

INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL

EDOARDO EUGENIO SIGAUD GONZALES

***TRADE DRESS: UMA ANÁLISE DE JURIMETRIA COM FERRAMENTAS DE
INTELIGÊNCIA ARTIFICIAL***

Rio de Janeiro

2019

Edoardo Eugenio Sigaud Gonzales

Trade dress: uma análise de jurimetria com ferramentas de inteligência artificial

Dissertação apresentada ao Programa de Mestrado Profissional em Propriedade Intelectual e Inovação, da Academia de Propriedade Intelectual, Inovação e Desenvolvimento - Divisão de Pós-graduação e Pesquisa, Instituto Nacional da Propriedade Industrial – INPI, como requisito parcial para a obtenção do título de Mestre em Propriedade Intelectual e Inovação.

Orientador: Prof. Dr. Vinícius Bógea Câmara

Coorientador: Prof. Dr. Sérgio Branco

Rio de Janeiro

2019

Ficha catalográfica elaborada pela Biblioteca de Propriedade Intelectual e Inovação “Economista Cláudio Treiguer” – INPI

Bibliotecário Evanildo Vieira dos Santos CRB7-4861

G643t Gonzales, Edoardo Eugenio Sigaud.

Trade dress: uma análise de jurimetria com ferramentas de inteligência artificial. /
Edoardo Eugenio Sigaud Gonzales, Rio de Janeiro, 2019. Dissertação (Mestrado Profissional em Propriedade Intelectual e Inovação) – Academia de Propriedade Intelectual Inovação e Desenvolvimento, Divisão de Programas de Pós-Graduação e Pesquisa, Instituto Nacional da Propriedade Industrial – INPI, Rio de Janeiro, 2019.

123 f.; fig.

Orientador: Vinícius Bogéa Câmara.

Coorientador: Sérgio Branco.

1. Propriedade industrial – *Trade dress*. 2. *Trade dress* – Contrafação. 3. *Trade dress* – Jurimetria. I. Instituto Nacional da Propriedade Industrial (Brasil).

CDU: 347.772(81)

Autorizo, para fins acadêmicos e científicos, a reprodução total ou parcial desta dissertação, desde que citada a fonte.

Assinatura

Data

Edoardo Eugenio Sigaud Gonzales

Trade dress: uma análise de jurimetria com ferramentas de inteligência artificial

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-Graduação em Propriedade Intelectual e Inovação, do Instituto Nacional da Propriedade Industrial

Aprovado em ____ de _____ do ____

Orientador: Prof. Dr. Vinícius Bógea Câmara

Instituto Nacional da Propriedade Industrial

Coorientador: Prof. Dr. Sérgio Branco

Instituto de Tecnologia e Sociedade do Rio de Janeiro

Banca Examinadora: Dr. Sergio Medeiros Paulino de Carvalho, do Instituto Nacional da Propriedade Industrial; Dra. Patrícia Pereira Peralta, do Instituto Nacional da Propriedade Industrial; e Dr. Mário Ferreira de Pragmácia Telles; da Pontifícia Universidade Católica do Rio de Janeiro.

A Ata de Defesa com as respectivas assinaturas dos membros encontra-se no processo de vida acadêmica do aluno.

Rio de Janeiro

2019

À minha esposa, Julia, e à minha filha, Catarina

À minha família e amigos, que incentivaram cada passo desta longa jornada

AGRADECIMENTOS

Agradeço a meus pais, Angela e Astyr, que incutiram em mim o amor pelo conhecimento, pela leitura e pelo diálogo. A eles devo tudo, inclusive minha própria existência. Nenhum agradecimento posto em palavras poderá exprimir adequadamente a gratidão e o amor que lhes devoto.

Agradeço a minha esposa, Prof. Dra. Julia Zardo, que tanto me apoiou e suportou as ausências do convívio familiar; e à minha filha, Catarina, que tantas vezes se chateou com a ausência do papai. O amor delas é meu norte, a inspiração última de tudo que faço.

Agradeço ao meu orientador, Prof. Dr. Vinicius Bógea Câmara, e meu co-orientador, Prof. Dr. Sérgio Branco, por suas ajudas incansáveis, sua compreensão com os desafios do caminho, e seu incentivo constante. Sem eles, este trabalho não teria sido possível.

Agradeço a todo corpo docente e administrativo da Academia do Instituto Nacional da Propriedade Industrial, sempre solícitos e buscando auxiliar os alunos em suas jornadas acadêmicas.

Em especial agradeço ao membros titulares e suplentes da banca de defesa, que de modo tão gentil doaram seu tempo e conhecimento para análise deste humilde trabalho; Prof. Dr. Sergio Medeiros Paulino de Carvalho (titular); Prof^a Dra^a. Patricia Peralta (titular); Prof. Dr. Mário Pragmácia (titular); Prof^a. Dr^a. Elizabeth Ferreira da Silva (suplente); Prof^a. Dr^a Bianca Xavier Gomes (suplente); agradeço ainda à Coordenação da Academia, na pessoa da sra. Patricia Trotte, por sua paciência e solicitude com minhas infindáveis dúvidas e pedidos.

Agradeço aos inúmeros colegas de trabalho e de academia, que contribuíram com ideias, orientações e debates, ajudando a enriquecer este trabalho. São tantos os nomes que não os poderia listar a todos neste breve espaço, mas em especial agradeço a Luiz Sá Lucas e Maria Elisa Tulipano, que me deram valiosas opiniões em relação aos complexos meandros da estatística, e a todos os meus colegas de turma do Mestrado 2016 da Academia do INPI, sempre dispostos a um enriquecedor debate de ideias, com civilidade e amizade.

O Direito não jaz na letra morta das leis: vive na tradição judiciária, que as atrofia, ou
desenvolve.

Rui Barbosa

RESUMO

GONZALES, Edoardo Eugenio Sigaud. *Trade dress: uma análise de jurimetria com ferramentas de inteligência artificial.* 2019. 123 f. Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-Graduação em Propriedade Intelectual e Inovação, do Instituto Nacional da Propriedade Industrial, Rio de Janeiro, 2019.

O presente estudo busca delinear o instituto do *trade dress* através de uma revisão teórica da literatura, conjugada posteriormente a uma análise de jurimetria acerca da jurisprudência selecionada, com uso de ferramentas automatizadas (software) de processamento de linguagem natural. O estudo foi conduzido através da seleção de julgados e aplicação de métodos de análise automatizada e testes de hipóteses estatísticas, no intuito de proporcionar melhor compreensão, não apenas de seus aspectos teóricos, mas principalmente de sua efetiva aplicação pelo Judiciário, de modo empírico.

Palavras-chave: *Trade dress*. Jurimetria. Processamento de linguagem natural.

ABSTRACT

GONZALES, Edoardo Eugenio Sigaud. *Trade dress: uma análise de jurimetría com ferramentas de inteligência artificial.* 2019. 110 f. Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-Graduação em Propriedade Intelectual e Inovação, do Instituto Nacional da Propriedade Industrial, Rio de Janeiro, 2019.

The present study seeks to probe the legal construct of *trade dress* through a theoretical revision of the literature, later complemented to a jurimetrics analysis on the selected legal precedents, with the use of natural language processing automated tools. The study was conducted through the selection of rulings and the application of automated analysis methods and statistical hypothesis tests, in order to provide a better understanding, not only of its theoretical aspects, but mainly of its effective application by the Judiciary, in an empirical way.

Keywords: *Trade dress*. Jurimetrics. Natural language processing.

LISTA DE FIGURAS

Fig. 1 – Metodologia de identificação de tokens proposta por KONCHADY (2006)	p.51
Fig. 2 – Processo de tokenização seguido de remoção de stopwords	p.52
Fig. 3 – Índice invertido elaborado com PLN	p.55
Fig. 4 – Dinâmica de Indexação temática	p.56
Fig. 5 - Representação matemática do teste de Shapiro-Wilk	p.55
Fig. 6 – Formula matemática aplicável ao test t de Welch	p.58
Fig. 7 – Wordcloud representativo do <i>Corpus</i> após pré-processamento de termos	p.62
Fig. 8 – Resultados quantitativos do TJSP	p.65
Fig. 9 – Resultados quantitativos do TJRJ	p.66
Fig. 10 – Resultados comparativos entre os tribunais	p.67
Fig. 11 – Perfil dos julgados do TJSP	p.68
Fig. 12 – Perfil dos julgados do TJRJ	p.69
Fig. 13 – Probabilidade de êxito comparada entre Tribunais	p.69
Fig. 14 – Resultados de primeira instância agrupados por decisão final	p.71
Fig. 15 – Histograma de Frequência Relativa de COR no <i>corpus</i>	p.73
Fig. 16 – distribuição de frequências de COR por tipo de decisão final	p.74
Fig. 17 – Histograma da Frequência relativa de MARCA	p.76

Sumário

1	INTRODUÇÃO	12
	Problema de Pesquisa.....	12
	Objetivo Geral.....	12
	Objetivos Específicos.....	12
2	REFERENCIAL TEÓRICO	14
2.1	Inovação, propriedade intelectual e concorrência	14
2.2	A Concorrência Desleal	15
	2.2.1 Previsão legal no Brasil: aspecto criminal	16
	2.2.2 Previsão Legal no Brasil: Aspecto cível	16
2.3	A Doutrina da Concorrência Desleal	17
	2.3.1 O <i>Trade dress</i> e a concorrência desleal	20
2.4	A Teoria do <i>Trade dress</i>	21
	2.4.1 <i>Trade dress</i> como signo distintivo: sua relação com a Marca	24
	2.4.2 Significação secundária (<i>Secondary Meaning</i>).....	28
	2.4.3 A Cor e o <i>Trade dress</i>	32
	2.4.4 A Prova Pericial nas Ações de <i>trade dress</i>	35
	2.4.5 A Importância da Jurisprudência para o Instituto do <i>trade dress</i>	36
3	METODOLOGIA.....	37
3.1	Jurimetria	37
	3.1.1 Conceito e Aplicação	37
	3.1.2 A Jurimetria no Brasil.....	38
3.2	Conceito de Inteligência Artificial e evolução histórica	41
3.3	Recuperação de Informação e Mineração de Textos	43
3.4	Processamento de Linguagem Natural.....	44
3.5	Coleta de textos.....	45

3.6	Pré-processamento	49
3.6.1	Modelo de Espaço Vetorial (<i>Bag-of-words</i>)	49
3.6.2	Tokenização (Tokenization)	50
3.6.3	Remoção de Stop-words	51
3.6.4	Stemming	52
3.7	Indexação	53
3.7.1	Indexação textual	54
3.7.2	Indexação temática.....	55
3.8	Teste de Shapiro-Wilk	56
3.9	Teste de Wilcoxon	57
3.10	Welch's t-Test	57
3.11	Teste de Chi-Square de Pearson.....	58
3.12	Análise manual dos julgados.....	58
3.13	Análise automatizada dos julgados	59
3.13.1	Importação dos arquivos em PDF e transformação em formato TXT	59
3.13.2	Inspeção visual.....	60
3.13.3	Remoção de Stopwords e limpeza do texto	60
3.13.4	Tokenização (Atomização)	61
3.13.5	Inspeção visual de tokens.....	61
3.13.6	Criação de wordcloud específico aos termos relevantes.....	61
3.13.7	Criação do <i>Thesaurus</i>	62
3.13.8	Indexação textual e temática conforme elementos selecionados	63
3.14	Aplicação dos testes de normalidade e de validação de hipótese	63
3.15	Limitações metodológicas.....	64
4	ANÁLISE JURIMÉTRICA DA JURISPRUDÊNCIA.....	65
4.1	Número total de ações julgadas.....	65
4.1.1	Resultados do Tribunal de Justiça de São Paulo.....	65

4.1.2	Resultados do Tribunal de Justiça do Rio de Janeiro.....	65
4.1.3	Resultados Comparativos.....	66
4.2	Perfil de decisões	67
4.2.1	Resultados TJ/SP	68
4.2.2	Resultados TJ/RJ.....	68
4.2.3	Resultado Comparativo.....	69
4.3	Perfil das decisões de primeira instância	70
4.3.1	Objetivo	70
4.3.2	Resultados agregados.....	70
4.4	Análise quantitativa e qualitativa do elemento “cor” nos julgados de <i>Trade dress</i> ..	71
4.4.1	Objetivo	71
4.4.2	Resultados agregados.....	73
4.5	Análise quantitativa e qualitativa do elemento “marca” nos julgados de <i>Trade dress</i> ..	75
4.5.1	Objetivo	75
4.5.2	Resultados agregados.....	75
4.6	Análise quantitativa e qualitativa do elemento “perícia” nos julgados de <i>Trade dress</i> .	77
4.6.1	Objetivo	77
4.6.2	Resultados agregados.....	77
5	CONCLUSÃO	79
5.1	Discussão dos resultados.....	79
5.2	Possíveis problemas de pesquisa futuros	81
5.2.1	Problemas metodológicos	81
5.2.2	Possíveis problemas de pesquisa futuros	81
6	Referências	83
7	APÊNDICE 1 – CODIGO-FONTE DO SOFTWARE DE JURIMETRIA UTILIZADO NO PROCESSAMENTO DOS DADOS	93

1 INTRODUÇÃO

Problema de Pesquisa

Objetivo Geral

O objetivo geral deste trabalho é analisar o instituto jurídico do *trade dress*, inicialmente traçando um panorama de sua doutrina, para em seguida realizar uma análise quantitativa e qualitativa em relação à sua efetiva expressão na jurisprudência selecionada, assim almejando uma melhor compreensão do instituto a partir de sua externalização jurisprudencial.

Deste modo, serão buscadas inicialmente as fontes doutrinárias que delineiam seus elementos norteadores, em busca de uma compreensão teórica do instituto e seu escopo de aplicação. A partir deste arcabouço teórico, será feita uma análise de jurimetria da jurisprudência brasileira, dentro de um recorte temporal e geográfico específico.

Espera-se que o resultado de tal estudo propicie uma maior compreensão do instituto jurídico do *trade dress* na forma como se manifesta no mundo do Direito, notadamente na forma como os tribunais pátrios aplicam o instituto.

Objetivos Específicos

Constituem objetivos específicos deste trabalho: (i) Examinar a doutrina acerca do instituto do *trade dress*, em busca de sua conceituação, origem e elementos constitutivos; (ii) Efetuar um levantamento quantitativo acerca das lides que versam sobre violações de *trade dress*, dentro da jurisprudência selecionada, aplicando conceitos de jurimetria a esta análise; (iii) Efetuar uma análise quantitativa e qualitativa da aplicação do conceito de “Cor” como elemento do *trade dress* na jurisprudência, relacionando sua presença nos julgados com o sucesso ou insucesso nas lides examinadas; (iv) Efetuar uma análise quantitativa e qualitativa da aplicação do conceito de “Marca” como elemento do *trade dress* na jurisprudência, relacionando sua presença nos julgados com o sucesso ou insucesso nas lides examinadas; (v) Efetuar uma análise quantitativa e qualitativa da aplicação do conceito de “Perícia” como elemento do *trade dress* na jurisprudência, relacionando sua presença nos julgados com o sucesso ou insucesso nas lides examinadas.

Uma parte dos objetivos específicos será atingido através da revisão visual dos julgados pertinentes, efetuada por esforço humano. Outra parcela dos resultados será obtida com aplicação de um software desenvolvido especialmente para estes objetivos, que permitirá a revisão e o cálculo estatístico de diversos parâmetros atinentes à jurimetria do instituto jurídico; a aplicação de meios de análise automatizada servirá para efetuar a revisão de um número maior de julgados, que de outra maneira seria deveras onerosa para elaboração manual individualizada. Ademais, o software proporciona a capacidade de aplicar conceitos estatísticos capazes de embasar inferências acerca da situação do instituto nos tribunais estudados.

2 REFERENCIAL TEÓRICO

2.1 Inovação, propriedade intelectual e concorrência

Desde há muito já se reconhece o papel do empreendedor e da inovação no crescimento e desenvolvimento econômico, e na geração de valor. Joseph Schumpeter (1984) foi um dos precursores dos estudos no campo das ciências econômicas, associando seus estudos às teorias de desenvolvimento. Em suas primeiras obras, Shumpeter (1911) destaca o papel dos indivíduos empreendedores como sendo os responsáveis pelo ato da inovação e de sua introdução nos mercados. Em um segundo momento (idem, 1942), houve certa mudança quanto à identificação do principal responsável pela inovação, que passou a ser apontada como sendo as grandes sociedades empresariais e suas áreas de Pesquisa e Desenvolvimento (P&D). A inovação, para este economista, é elemento central nas políticas industriais, sendo considerada muito além do mero enfoque de produtos inovadores:

O impulso fundamental que inicia e mantém o movimento da máquina capitalista decorre dos novos bens de consumo, dos novos métodos de produção ou transporte, de novos mercados, das novas formas de organização industrial que a empresa capitalista cria (idem, 1984, p. 72).

Mais recentemente, tem-se observado uma crescente importância dos ativos intangíveis para geração de valor na atividade empresarial. A relação entre ativos intangíveis e vantagem competitiva teve sua origem no trabalho pioneiro de Edith Penrose sobre a teoria baseada em recursos, e hoje pode-se afirmar que poucas sociedades empresárias podem sustentar algum grau de competitividade no cenário nacional ou internacional sem apoio de valiosos ativos intangíveis (SANTOS, 2006).

Dentro deste contexto, a propriedade intelectual desponta como elemento primordial de organização da atividade empresarial, capaz de assegurar ao seu detentor um direito de exclusiva, e assim constituir ativo valioso e até vital ao sucesso empresarial. Entretanto, o alcance da proteção por direitos de propriedade intelectual mostrou-se insuficiente para tutelar a exploração efetiva dos aspectos intangíveis da organização empresarial, e sua proteção contra usos ilícitos por terceiros (SOARES, 2004).

Em decorrência desta limitação, surge uma visão de proteção calcada não na propriedade em si, mas na vedação aos atos desleais de concorrência. Uma gama de atos de

deslealdade que não viola propriamente o direito de propriedade, mas que com eles transita numa ‘zona cinzenta’ de associação e confusão, suscita especial interesse; eis que buscando valer-se do reconhecimento, da expertise, da imagem e da credibilidade alcançada pelo concorrente, para dela se apropriar, o ator econômico desleal não viola sua propriedade diretamente, mas por via transversa almeja confundir o público, ou sugerir, ainda que subconscientemente, uma associação entre produtos e serviços que não existe de fato (DANNEMAN; MOREIRA, 2005; BITTAR FILHO, 2002).

O reconhecimento de tais práticas desonestas – e da necessidade de coibi-las – é algo antigo na história da humanidade. Hermano Duval (1976) aponta que já havia regramento protetivo da concorrência desonesta desde o ano 1410 na Inglaterra, sob o regime da *common law*, e mais recentemente por positivação na lei alemã, datada de 27 de maio de 1896. A partir do século XX, em âmbito internacional, teve primeiro assento na vedação ao que se convencionou chamar de “concorrência desleal”, positivada internacionalmente no texto da Convenção da União de Paris.

2.2 A Concorrência Desleal

Na ordem jurídica internacional, a Concorrência Desleal encontra assento na Convenção da União de Paris (BRASIL, 1975), cujo artigo primeiro preceitua:

A proteção da propriedade industrial tem por objeto as patentes de invenção, os modelos de utilidade, os desenhos ou modelos industriais, as marcas de serviço, o nome comercial e as indicações de procedência ou denominações de origem, bem como a repressão da **concorrência desleal** (grifo nosso).

Prossegue afirmando que a concorrência desleal consiste em atos contrários aos “usos honestos”, em seu artigo 10:

- (1) Os países da União obrigam-se a assegurar aos nacionais dos países da União proteção efetiva contra a concorrência desleal.
- (2) Constitui ato de concorrência desleal qualquer ato de concorrência contrária aos **usos honestos** em matéria industrial ou comercial
- (3) Deverão proibir-se especialmente:
 - 1º Todos os atos suscetíveis de, por qualquer meio, estabelecer confusão com o estabelecimento, os produtos ou a atividade industrial ou comercial de um concorrente.
 - 2º As falsas afirmações no exercício do comércio, suscetíveis de desacreditar o estabelecimento, os produtos ou a atividade industrial ou comercial de um concorrente.
 - 3º As indicações ou afirmações cuja utilização no exercício do comércio seja suscetível de induzir o público em erro sobre a natureza, modo de fabrico, características, possibilidades de utilização ou quantidade das mercadorias (grifo nosso).

2.2.1 Previsão legal no Brasil: aspecto criminal

A Concorrência Desleal é configurada como uma concorrência indireta, ilícita, na qual os atos de empresa praticados por uma pessoa natural ou jurídica possuem o objetivo precípua de atingir os interesses comerciais de outrem, também este empresário.

Segundo preceitua Hermano Duval, é “a agressão à atividade concorrente em violação aos preceitos éticos da correção profissional” (DUVAL, 1976, p. 126).

Diversas condutas afeitas à prática de concorrência desleal encontram-se tipificadas como crime, no ordenamento brasileiro, na Lei de Propriedade Industrial (BRASIL, 1996) em seu Capítulo 6º, “Dos Crimes de Concorrência Desleal”, Artigo 195. Deste artigo, alguns incisos tipificadores de conduta são afeitos ao conceito de *trade dress*:

Art. 195. Comete crime de concorrência desleal quem:

-
 - III - emprega meio fraudulento, para desviar, em proveito próprio ou alheio, clientela de outrem;
 - IV - usa expressão ou sinal de propaganda alheios, ou os imita, de modo a criar confusão entre os produtos ou estabelecimentos;
 - V - usa, indevidamente, nome comercial, título de estabelecimento ou insígnia alheios ou vende, expõe ou oferece à venda ou tem em estoque produto com essas referências;
 - VI - substitui, pelo seu próprio nome ou razão social, em produto de outrem, o nome ou razão social deste, sem o seu consentimento;
- Pena - detenção, de 3 (três) meses a 1 (um) ano, ou multa.

Pode-se verificar que tais incisos guardam estreita relação com o conceito de *trade dress* e sua violação. Porém, é importante ressaltar que as condutas tipificadas nos incisos não são exaustivas, no condão de determinar o escopo de atos violadores de *trade dress*, como se depreende da tutela cível da matéria.

2.2.2 Previsão Legal no Brasil: Aspecto cível

Em seu aspecto cível, encontra-se previsão legal a embasar da proteção ao *trade dress*, no artigo 209 da Lei da Propriedade Industrial (LPI), in verbis:

Art. 209. Fica ressalvado ao prejudicado o direito de haver perdas e danos em resarcimento de prejuízos causados por atos de violação de direitos de propriedade industrial e **atos de concorrência desleal não previstos nesta Lei, tendentes a prejudicar a reputação ou os negócios alheios, a criar**

confusão entre estabelecimentos comerciais, industriais ou prestadores de serviço, ou entre os produtos e serviços postos no comércio (BRASIL, 1996, grifo nosso)

É importante frisar que o artigo 209 consiste em adição ao rol de condutas do artigo 195, pelo uso da expressão “não previstos nesta lei”. Esta é a melhor inteligência do artigo em cotejo, uma vez que já existe a previsão na lei, qual seja, do artigo 195 (BARBOSA, 2001; MARTINS, 2006). Dannemann e Moreira esclarecem que legislador não teve o condão de excluir do art. 209 os atos elencados no artigo 195 da LPI (DANNEMAN; MOREIRA, 2005 p. 436).

Esta concorrência possui em seu cerne os atos de desonestade, deslealdade, desídia na prática comercial. Pode se consubstanciar em desrespeito ao direito do consumidor, desvio de clientela, violação do segredo de empresa, entre outros. Porém, não são todos estes atos que representam violação ao *trade dress*, mas sim uma minoria deles, como se verá adiante neste estudo.

2.3 A Doutrina da Concorrência Desleal

Para o eminentíssimo jurista Pontes de Miranda, no atinente à concorrência desleal, seria ilícito aquilo que, por reprodução ou imitação, origina confusão; na ausência de confusão, há necessariamente distinguibilidade, e portanto ausente a deslealdade (MIRANDA, 1983).

O Prof. Fábio Ulhoa Coelho (2007) frisa a diferença entre as formas existentes de concorrência ilícita, ou seja, entre a concorrência desleal e a infração da ordem econômica, de modo que esta última se diferencia no quanto sua prática atenta contra as estruturas da economia de mercado. Enquanto a primeira atenta contra a lealdade entre competidores, a segunda teria um campo de atuação mais abrangente, quando comparado à concorrência desleal. A concorrência desleal teria o condão de atingir os interesses dos rivais, causando-lhes perdas de clientela, e com isto, obtendo ganhos.

Haveria, segundo esse jurista, uma classificação da concorrência desleal em duas espécies: a específica e a genérica; na primeira, a tipificação penal está presente nas condutas lesivas, como por exemplo, os direitos sobre patentes ou marcas, ou seja, este modo de concorrência é sancionado não só no âmbito civil, mas também no penal, conforme Lei da Propriedade Industrial, Artigo 195 (BRASIL, 1996); já a segunda se relaciona com a responsabilidade extracontratual, que se configura na sanção somente cível - a repressão

assegura ao empresário, que fora vítima a devida composição dos danos sofridos, além da abstenção de comportamentos similares, ou sua cessação imediata, quando cabível, como resultado de obrigação de fazer ou não fazer, consoante art. 209 da Lei de Propriedade Industrial. Ainda seguindo a classificação acima, a específica se consuma pelos tipos penais elencados, que se revelam facilmente como subsunção da conduta à norma penal tipificadora; e a genérica, de mais difícil definição em razão da norma jurídica aberta (COELHO, 2007, p. 91).

O Prof. Rubens Requião (2010) inicia seus estudos acerca do assunto salientando a desigualdade (entendida como noção moral) entre concorrência ilícita e desleal, de modo que é a deslealdade na concorrência que vem a produzir, consequentemente, a denominação de concorrência ilícita. A deslealdade seria elemento determinante nesta conduta. A repressão desta espécie de concorrência, segundo preceitua o autor, possui como objetivo central a proteção da clientela (do empresário), contra atos denegridores de concorrentes, ou seja, haveria necessidade de existir uma relação concorrencial de fato, e não apenas potencial.

Entretanto, esta posição não é unânime, e diversos autores entendem que o aproveitamento parasitário, mesmo no caso de inexistência de concorrência direta, seria justificativa adequada e bastante à repressão dos atos de violação ao *trade dress* (BARBOSA, 2011; SOARES, 2004).

Não configurariam atos de concorrência desleal aqueles praticados de forma honesta, sem afronta à moral ou à boa-fé, como os que os negociantes utilizam para elevar a boa reputação de seus produtos, ou até mesmo as expressões exageradas com que os exaltem. Os atos de concorrência desleal estão ordenados, segundo Requião (2010) em três categorias, quais sejam: 1) atos tendentes a criar confusão; 2) atos tendentes a desviar clientela de outrem por processos desonestos ou ‘artificiais’; e 3) atos contrários à moralidade comercial.

A primeira categoria consistiria nos mais comuns. Quem os realiza se determina a angariar vantagens da confusão provocada dolosamente entre a empresa e seus produtos, e a empresa ou os produtos de um competidor, aproveitando-se da confusão para proveito próprio. Estão previstos nos Artigos 193 e 195 (incisos IV, VI e VIII) da LPI. Os atos da segunda possuem como objetivo a desonra do competidor e/ou de seus produtos, seja com a publicação de notícias, informações falsas, insinuações nocivas, deste modo desviando a clientela do competidor. Podem ser seguidas de modo indireto, comparativo. Estão previstos no Artigo 195 (incisos I, II e III) da LPI. Os da terceira seriam configurados como sendo de corrupção, suborno de empregados ou de outros elementos da coletividade da empresa, com o propósito de que divulguem segredos, e também a atribuição de qualidades que a empresa não

possui. Estão previstos no Artigo 195 (incisos VII, IX, X, XI, XII, XIII e XIV) da LPI. O autor cita ainda os atos que não configurariam, *in re ipsa*, concorrência desleal, tais como a contrafação de marcas, e a violação da patente de invenção (REQUIÃO, 2010).

Para Hermano Duval (1976), todos sentem ou percebem o fenômeno da concorrência desleal, mas poucos a conceituam de forma precisa. Haveria uma dificuldade semântica, posto que a palavra “deslealdade” possui várias acepções de acordo com a subjetividade da linguagem e os costumes dos locais. A primeira repressão a esta espécie de concorrência ocorreu na Inglaterra, no ano de 1410, com base na *common law*, e a primeira lei que sistematizou o instituto fora a alemã, datada de 27 de maio de 1896. A concorrência desleal é relacionada por esse autor a conceitos éticos e morais.

Neste contexto, as violações de patentes e marcas não são entidades autônomas, porém, meros delitos-meio, para assim, conceber o delito-fim de concorrência desleal; na prática desta concorrência, há um bem jurídico maior que se eleva, e domina o ato, o que, segundo ele, desejam os concorrentes, é atingir a reputação, a fama, o bom conceito, a notoriedade dos produtos de outrem, para com isso angariar sua clientela, violando portanto, como já dito, as normas da ética, da lealdade e da honestidade (DUVAL, 1976, p. 365).

Duval (1976) argumenta que, pela extensa variedade de ações delituosas que integram a concorrência desleal, não é possível à doutrina reduzi-las a um tipo único, com características igualitárias, como ocorre, por exemplo, no homicídio, no furto, na lesão corporal, entre outros crimes regidos pelo direito penal; e ainda, não é possível exaurir sua definição legislativa no âmbito cível, sendo mais adequada a norma aberta. A exclusiva designação que a doutrina lhe atribui é o do dolo genérico, pois o concorrente tem consciência de seu ato (*animus dolandi*), e mesmo assim o realiza; mas ausente esta concepção, a análise da concorrência desleal deve ser apreciada segundo as circunstâncias de cada caso concreto.

Carlos Alberto Bittar Filho (1989) preceitua que esta modalidade de concorrência repercute por ações de captação de clientela alheia, ações estas que, quando praticadas, devem estar dentro do mercado correspondente ao de seu concorrente e referir-se à criação já existente e em uso anterior pelo mesmo. Existiria concorrência desleal em todos os atos de outrem que se aproveitem de modo indevido de criação ou de avultamento alheio para angariar, sem esforço, a recíproca clientela. O autor ressalta alguns pressupostos elencados pela doutrina e pela jurisprudência para a identificação da concorrência desleal no caso concreto, quais sejam: desnecessidade de dolo ou de fraude, bastando a culpa do agente (responde o mesmo por simples negligência), desnecessidade de verificação de dano em concreto (basta a possibilidade ou perigo de sua superveniência), necessidade de existência de colisão

(verificada por campos colidentes de interesses, mesma atividade negocial, por exemplo), necessidade de existência de clientela (pois é o objetivo precípua visado, em sua totalidade ou parcialmente, pelo concorrente) e por fim, o ato ou procedimento suscetível de repreensão (o mesmo deve ir contra as práticas normais dos negócios, qualificado por ausência – desrespeito à ordens de direito ou moral).

A distinção entre os atos de concorrência desleal e leais é imprescindível pois, segundo o autor, servem para sancionar determinadas ações que, quando praticadas, desrespeitam as regras impostas, prejudicando o livre funcionamento da concorrência e da obtenção de clientela, e acarretando prejuízos para os titulares do direito violado – não só prejuízos patrimoniais, como também morais; surgem, portanto, diferentes procedimentos que são capazes de ‘confundir’ este livre funcionamento, como por exemplo, o aproveitamento indevido de marca, de embalagem, de publicidades alheias, entre outros procedimentos que representam abuso aos costumes da indústria e do comércio (BITTAR FILHO, 1989, p. 290).

Para Fran Martins (2006) o exercício da atividade empresarial acarreta como consequência necessária a concorrência, que se manifesta por vários aspectos, a exemplo: na apresentação de produtos, no sistema de vendas, na publicidade, na reputação, etc. Esta rivalidade, por assim dizer, deve ser leal, honesta, transparente, de modo exibir a boa-fé como símbolo fundamental da relação empresarial. A concorrência desleal se caracterizaria, portanto, pela ausência destes elementos, quando empresários buscam, de maneira desonesta, prejudicar os negócios alheios em benefício próprio.

Fran Martins (2006) ressalta a existência de uma miríade de atos que não estão mencionados na lei (Artigo 195 da LPI), mas que podem causar prejuízos aos comerciantes, não só à sua reputação, mas atingindo também seus negócios, ou ainda, causando confusão entre estabelecimentos comerciais ou industriais, ou entre mercadorias e produtos colocados no comércio (MARTINS, 2006, p. 128).

2.3.1 O *trade dress* e a concorrência desleal

Dentre todos os atos passíveis de configurar a concorrência desleal, de especial interesse a este trabalho estão aqueles atos que a doutrina convencionou chamar de violação ao “*trade-dress*” ou, no Brasil, ao “conjunto-imagem”.

Trata-se de atos que consistem em uma cópia ou imitação de elementos perceptíveis, que não violam propriamente um registro ou depósito de propriedade industrial, mas ainda assim se aproveitam de uma característica distintiva de um produto, serviço ou

estabelecimento, com a finalidade de causar confusão no público, ou de se aproveitar da reputação e da aceitação que tais produtos, serviços ou estabelecimentos tenham angariado junto ao consumidor (BARBOSA, 2012).

Existe aqui, muitas vezes, um aproveitamento da percepção subconsciente do público acerca de um produto, serviço ou estabelecimento, utilizando ou fazendo referência às suas cores, formatos, posições ou arranjos, embalagens, invólucros, fontes tipográficas, e até mesmo características como organização dos elementos de um website¹ (MONTAURY et al., 2009), entre outros elementos (MORO, 2005).

De fato, são tão variados os elementos passíveis de proteção pela teoria do *Trade dress* que não existe um elenco exaustivo na doutrina; a lei, por sua vez, sequer tenta elencá-los, ao menos no ordenamento brasileiro. Resta, portanto, a análise da jurisprudência em busca de um prumo que possa orientar os empresários em suas práticas, aumentando assim a segurança jurídica dos atos de empresa (SOARES, 2004).

Abordaremos, no próximo capítulo, a teoria do *trade dress*, relacionando-a com a concorrência desleal, para melhor compreender seu escopo e aplicação, e posteriormente possibilitar o exame objeto deste trabalho, qual seja, a análise dos julgados selecionados recentes acerca do tema.

2.4 A Teoria do *trade dress*

O instituto do *trade dress* é de criação relativamente recente: surgiu pela primeira vez em um caso emblemático da Suprema Corte do Estados Unidos da América, datado de 1992. Tratava-se de lide envolvendo duas empresas que operavam cadeias de fast-food, Taco Cabana, Inc. e Two Pesos, Inc. A primeira operava uma rede de restaurantes de comida mexicana no estado do Texas, cujo primeiro estabelecimento foi inaugurado em 1978, sendo depois seguido por outros cinco, abertos em 1985.

Os restaurantes da rede Taco Cabana se caracterizavam, conforme descrito na demanda judicial, por ambientes decorados com pinturas, murais e cores brilhantes específicas. Apresentavam ainda ambientes internos e externos adornados com toldos e guarda-sóis (UNITED STATES, 1992).

¹ Em artigo da Revista da ABPI nº 100, de mai/jun de 2009, Luiz Edgard Montaury Pimenta e Marianna Furtado de Mendonça aventaram a possibilidade de se tutelar o conteúdo de websites por meio da proteção ao *trade dress*, desde que o projeto gráfico não seja funcional e apresente suficiente distintividade (originária ou adquirida).

Por outro lado, a rede de restaurantes Two Pesos surgiu em Houston em 1985, apresentando um conjunto visual muito semelhante ao da rede Taco Cabana, que àquela época já era uma cadeia consolidada no mercado. Os restaurantes da Two Pesos rapidamente se expandiram, abrindo diversas outras lojas em diferentes localidades.

A enorme semelhança com os novos estabelecimentos da Two Pesos levou a Taco Cabana a recorrer ao judiciário, alegando prática de atos de concorrência desleal, e logrando êxito em sua demanda. A decisão da Suprema Corte reconheceu a ilicitude da conduta, conceituando então o instituto do *trade dress*:

Trade dress é a imagem total do negócio (...) Pode incluir um perfil e uma aparência geral da parte externa do restaurante, o símbolo que o identifica, o plano interior da cozinha a decoração, o ‘menu’, o equipamento para servir a cozinha, o uniforme dos atendentes e outros detalhes que refletem a imagem global do restaurante (SOARES, 2004, p.86).

O caso foi solucionado mediante aplicação do Lanham Act, legislação norte-americana que regula o sistema nacional de registro e proteção de marcas, mais especificamente em sua seção 43 (a), que veda o uso de sinais ou signos distintivos com intuito de causar confusão ou induzir o consumidor a erro em relação às informações dos produtos ou serviços, in verbis:

Lanham Act, seção 43 (a), False designations of origin, false descriptions, and dilution forbidden:

(1) Any person who, on or in connection with any goods or services, or any container for goods, uses in commerce any word, term, name, symbol, or device, or any combination thereof, or any false designation of origin, false or misleading description of fact, or false or misleading representation of fact, which— (A) is likely to cause confusion, or to cause mistake, or to deceive as to the affiliation, connection, or association of such person with another person, or as to the origin, sponsorship, or approval of his or her goods, services, or commercial activities by another person (UNITED STATES, 1946).

Tal julgado foi de grande importância na evolução jurisprudencial do conceito, cristalizando o entendimento de que a proteção visual de estabelecimentos comerciais pode ocorrer por meio de *trade dress*. Neste sentido, a Corte norte-americana decidiu:

“[T]rade dress” is the total image of the business. Taco Cabana’s *trade dress* may include the shape and general appearance of the exterior of the restaurant, the identifying sign, the interior kitchen floor plan, the decor, the menu, the equipment used to serve food, the servers’ uniforms, and other features reflecting on the total image of the restaurant. 1 App. 83-84. The Court of Appeals accepted this definition and quoted from Blue Bell Bio-Medical v. Cin-Bad, Inc., 864 F.2d 1253, 1256 (CA5 1989): “The ‘*trade dress*’ of a product is essentially its total image and overall appearance.” See 932 F.2d 1113, 1118 (CA5 1991). It “involves the total image of a product,

and may include features such as size, shape, color or color combinations, texture, graphics, or even particular sales techniques.” John H. Harland Co. vs Clarke Checks, Inc., 711 F.2d 966, 980 (CA11 1983). Restatement (Third) of Unfair Competition 16, Comment a (Tent. Draft No. 2, Mar. 23, 1990)

A expressão ‘conjunto-imagem’, cunhada como tradução do original ‘*Trade dress*’ para o português, surgiu com a evolução do conceito de proteção da vestimenta de uma marca, passando então a abarcar a imagem total de um produto ou serviço (BARBOSA, 2011).

O Prof. José Carlos Tinoco Soares (2004) conceitua o *trade dress* como a imagem total pela qual o produto ou serviço é apresentado ao mercado; sua aparência geral, que inclui, exemplificativamente, características da embalagem, do recipiente, cores e elementos decorativos, entre outras. Representaria, portanto, uma combinação própria de elementos ou figuras que, em razão de seu uso, tornam-se associadas e funcionam como um indicador de origem de um produto ou serviço.

Com efeito, o conjunto de elementos que pode configurar *trade dress* é sempre apontado como exemplificativo, dada a variedade de sua externalização. Podem incluir, entre outros, uma cor ou esquema de cores, forma ou formato, embalagem, configuração do produto, sinais, frases, disposição, estilização e tamanho de letras, gráficos, desenhos, emblemas, brasões, texturas e efeitos ou ornamentos em geral, sempre que capazes de identificar produto ou serviço e diferenciá-lo dos demais (DANIEL, 2003).

Stephen Ladas (1975) entendia os elementos do *trade dress* - a aparência, formato, modelo, detalhe, de um produto – como passível de proteção pelo direito autoral, quando a criação for artística e original; através do desenho industrial, caso a criação possua as características para ser protegida por este instituto; através da proteção marcária, se dentre as marcas passíveis de registro em determinado país abrange marcas mistas e tridimensionais; através da concorrência desleal, quando a imitação gerar engano ou confusão ao público.

É importante ressaltar que o conjunto-imagem não se refere apenas aos produtos e suas características. Com muita propriedade acrescenta Silva (2013, p. 150) que “o *trade dress* engloba não só a aparência exterior dos produtos como também o aspecto distintivo interno e/ou externo de estabelecimentos industriais, comerciais ou prestações de serviços”. Esta, aliás, sua acepção original, na lide entre Two Pesos e Taco Cabana, onde se visava justamente a proteção dos elementos de decoração e organização dos estabelecimentos comerciais.

O *trade dress* consistiria, em resumo, no “aspecto visual característico de como um produto ou serviço é apresentado ao público” (OLAVO, 2008, p. 430).

Tal externalidade visual tem a intenção precípua de tornar o produto ou serviço mais atrativo ao público, ao mesmo tempo que o associa a uma determinada origem. Neste condão, Gustavo Piva de Andrade (2011) entende a função estética como de suma importância, quando considera que tais elementos são cuidadosamente elaborados para realçar distintividade, e transmitir um ar de sofisticação, assim tornando o respectivo produto ou serviço cada vez mais atrativo em termos estéticos.

A conceituação do instituto como essencialmente ligado a elementos visuais o aproximaria da proteção marcária. O conjunto-imagem seria assim compreendido como uma extensão da marca, emanando do direito marcário (PINTO; MONTENEGRO; PINHEIRO, 2012).

Entretanto, à similaridade das discussões doutrinárias acerca das marcas olfativas e sonoras, o *trade dress* também deve ser entendido como englobando tais sinais perceptíveis olfativos ou auditivos, ainda que não visualmente. De fato, nada impediria, em face da positivação aberta do instituto, que cheiros ou sons associados a produtos, serviços e estabelecimentos comerciais gozem de efetiva proteção contra atos desleais de concorrência, com base na construção lógica de que estes, igualmente, podem representar elementos de associação do público-consumidor com a origem e qualidade dos mesmos (ANDRADE, 2011).

2.4.1 *Trade dress* como signo distintivo: sua relação com a Marca

Com efeito, há uma relação lógico-jurídica muito próxima entre o *trade dress* e a Marca. Ambos representam um conjunto de sinais perceptíveis, aptos e tendentes a identificar um produto ou serviço em relação à sua origem ou fabricação, ou de identificá-los na mente dos consumidores. Alguns autores entendem que o *trade dress* teria natureza jurídica de signo distintivo (ANTONIAZZI, 2010).

Ademais, é de interesse a relação do fenômeno jurídico das marcas com a semiologia, e como afetam – ou deveriam afetar – a percepção do *trade dress* e sua aplicação ao caso concreto. A Semiologia surge do trabalho autônomo mas contemporâneo de Charles Sanders Peirce (1934) e Ferdinand de Saussure (1916), que adotam a expressão semiologia (do grego *semeîon*) para nomear a ciência do estudo dos signos dentro da sociedade. Assim leciona o saudoso Prof. Denis Barbosa:

O signo seria, funcionalmente, a representação de algo, em face de uma pessoa determinada; a representação presume um sistema, e o liame entre signo e objeto se define por oposição – define-se o signo pelo que ele não é, pelo encadeamento estrutural de diferenças. Enfatiza-se aqui o estrutural – espaço em que a causalidade não é mecânica, ponto a ponto, mas sistêmica. Tal sistema não é estático. Sistema e evento significativo (“langue” e “parole”, no dizer de Saussure) são noções inter-atuantes, de forma que a representação em cada caso resulta da ação de todo sistema e cada evento de significação potencialmente pode afetar o sistema. Essa noção de sistema simbólico – a significação nasce da estrutura e não só de uma oposição binária – justifica, no âmbito do direito de marcas, a análise da chamada teoria de diferença, objeto de análise quando tratarmos de confundibilidade das marcas. **Não se pode eficazmente comparar uma marca e outra para se determinar a eventual violação de direitos, sem levar em conta o sistema das representações no mercado pertinente, organizado como uma subestrutura diferenciada.** Um postulado da semiologia é que as estruturas da sociedade são constituídas à maneira da linguagem, o Direito não menos do que as demais. Como indica Beebe, um objeto primordial da semiologia seria o direito de marcas (BEEBE, Barton, 2004 apud BARBOSA, 2006, pag. 36, grifo nosso)

Embora atinente especificamente ao fenômeno marcário, as considerações semiológicas têm igual pertinência na análise da violação de *Trade dress*, que possui conteúdo mais subjetivo e difuso. Outros institutos ou fenômenos jurídicos, por terem função análoga à marca – o desenho industrial, o *trade dress* e as construções publicitárias (inclusive sinais de propaganda) se aproveitariam dos mesmos pressupostos semiológicos das marcas em sua interpretação e aplicação prática. Com efeito, se a avaliação de infração marcária enfrenta a necessidade de avaliar “***o sistema das representações no mercado pertinente, organizado como uma subestrutura diferenciada***” (BARBOSA, 2006, p. 36), da mesma forma a avaliação de violação de *trade dress* estaria sujeita ao mesmo requisito semiológico.

As marcas, porém, são disciplinadas no ordenamento jurídico brasileiro por meio da Lei 9.279/96 – Lei de Propriedade Industrial (BRASIL, 1996). O referido diploma legal conceitua a marca por meio de seu artigo 122: “São suscetíveis de registro como marca os sinais distintivos visualmente perceptíveis, não compreendidos nas proibições legais”.

Deste modo, marca é um signo ou sinal, com intenção distintiva ou com caráter de distintividade, visualmente perceptível, com ressalva das restrições legais, que se encontram elencadas no artigo 124 da mesma Lei. A lei ainda prevê, em seu artigo 123, três espécies de marcas: as de produto ou serviço, usadas para distinguir produtos ou serviços de outros similares; as marcas de certificação, usadas para atestar conformidade com normas ou especificações técnicas; e as marcas coletivas, usadas para identificar produtos ou serviços de membros de uma determinada entidade (BRASIL, 1996).

Voltamos a buscar albergue nas lições do Prof. Denis Borges Barbosa:

Marca é o sinal visualmente representado, que é configurado para o fim específico de distinguir a origem dos produtos e serviços. Símbolo voltado a um fim, sua existência fáctica depende da presença destes dois requisitos: capacidade de simbolizar, e capacidade de indicar uma origem específica, sem confundir o destinatário do processo de comunicação em que se insere: o consumidor. Sua proteção jurídica depende de um fator a mais: a apropriabilidade, ou seja, a possibilidade de se tornar um símbolo exclusivo, ou legalmente unívoco, em face do objeto simbolizado (BARBOSA, 2001, p. 700).

Conforme determina o Instituto Nacional da Propriedade Industrial (INPI), a marca pode se apresentar como: a) nominativa: aquela formada pela combinação de letras e números; b) figurativa: aquela que inclui imagens e letras de outros alfabetos; c) mista: formada por palavras e imagens; e d) tridimensional: forma particular de um produto ou recipiente, quando capaz de distingui-lo de outro semelhante (INSTITUTO NACIONAL DE PROPRIEDADE INDUSTRIAL, 2015).

O *trade dress*, embora guarde relação com a marca na medida em que possui função de identificar e distinguir, com ela não se confunde. O *trade dress*, como já visto, não possui positivação específica no ordenamento brasileiro, admitindo elementos e componentes que a proteção marcária rechaça.

Por oportuno, colacionamos lição do Prof. José de Oliveira Ascensão (2002):

É corrente colocar como função originária da marca a de indicar a origem do produto (que mais tarde, se estenderia ao serviço também). A marca asseguraria que os produtos a que se aplicava tinham a mesma origem. Supomos que hoje este entendimento é insustentável. Desde que se admite a transmissão da marca independentemente do estabelecimento (quebrando assim o princípio chamado de adesão) já a explicação deixava de servir, porque a origem passava a ser diferente e a marca continuava a mesma. Mas sobretudo, se se admite que se concedam licenças de utilização da marca a várias entidades, que têm condições de exploração diversas, a marca deixou de dizer seja o que for sobre a origem do produto. Por isso sob a mesma marca circulam produtos bons e produtos maus. As reformulações que têm sido tentadas não são convincentes. Entre nós, Couto Gonçalves procura encontrar, como sub-rogado da unidade de origem empresarial, uma unidade de origem pessoal: há sempre uma pessoa a quem se atribui o ônus do uso não enganoso dos produtos ou serviços marcados. É não dizer nada: exprime-se apenas que a marca tem um titular, e que esse titular, como todos, tem poderes e deveres. Nada se diz ao público sobre a origem dos produtos ou serviços. Na realidade, há que distinguir função de determinação de origem e função distintiva. Esgotada a função de determinação de origem, só resta a função distintiva (pag. 17).

Pode-se ainda afirmar que o *trade dress* oferece uma proteção mais ampla e mais genérica do que a marca, englobando elementos que, inclusive, estariam vedados ao registro

de marca, tais como expressões de propaganda, o acondicionamento de produtos e cores isoladas. A marca, em verdade, constitui um elemento do *trade dress*, na medida em que contribui parcialmente com a caracterização da “vestimenta” total de um negócio, mas não exclusivamente; o conjunto-imagem compreende a marca, mas não se esgota nela (BARBOSA, 2006). Neste sentido, Cerqueira (1952) preconiza:

Entre os meios de criar confusão com os produtos ou artigos de um concorrente, acode à idéia, desde logo, o uso de sinais distintivos idênticos ou semelhantes aos que os assinalam. Quando êsses sinais distintivos se acham registrados como marcas, a hipótese cai sob o domínio da lei respectiva, sendo punível como contrafação (Código, art. 175, II). Tratando-se de sinais não registrados, é condição essencial à ação baseada em concorrência desleal que êsses sinais, além de reunirem os requisitos intrínsecos que os tornem distintivos e suscetíveis de constituir marcas, estejam em uso prolongado, de modo a se tornarem conhecidos como marcas dos produtos concorrentes. Aplica-se o mesmo princípio quando se trata de imitação ou reprodução do aspecto característico do produto ou à forma de sua apresentação (embalagens, envoltórios, recipientes, etc.), quando não sejam vulgares ou pertencentes ao domínio público. Equiparam-se aos produtos, para o efeito de sua proteção contra a concorrência desleal, como já foi dito, os serviços oferecidos ao público por uma emprêsa ou estabelecimento.

Também não se pode afirmar que a marca tridimensional corresponda ao *trade dress*, uma vez que esta remete tão-somente ao “formato característico, não funcional e particular que é dado ao próprio produto ou ao seu recipiente” (BERTOLDI; RIBEIRO, 2013. p. 127).

[Nas marcas tridimensionais] o elemento essencial é a forma em três dimensões, sem a qual não se pode vislumbrar tal marca, deixando-se avaliação das figuras ou elementos nominativos a ela apostos em plano secundário (OLAVO apud MORO, 2009, p. 260).

Por fim, cabe ressaltar que as Marcas são passíveis de registro, e que de fato a propriedade e exclusividade (observados os princípios da territorialidade e da especialidade) somente se obtém com o registro válido perante o Instituto Nacional de Propriedade Industrial, nos termos do art. 129 da Lei 9279/1996 (BRASIL, 1996).

O *trade dress*, por outro lado, e por óbvio, independe de registro, visto que é construção jurisprudencial, e não se pode sequer falar em direito de propriedade ou exclusividade erga omnes em relação aos seus elementos, uma vez que sua proteção se dá sempre no conteúdo de uma lide fundada em concorrência desleal, mediante cotejo de elementos probatórios específicos ao caso, e que os efeitos da sentença são produzidos apenas inter partes (PINTO; MONTENEGRO; PINHEIRO, 2012).

No âmbito deste estudo, buscar-se-á relacionar a presença do elemento “marca” nos julgados de *trade dress*, no intuito de avaliar se tal associação tende a alterar as chances de

êxito em uma lide, conforme detalhado no capítulo de Metodologia e no capítulo de Resultados.

2.4.2 Significação secundária (*Secondary Meaning*)

O fenômeno da significação secundária pode ser definido como a aquisição, pelo uso ou aplicação repetida de um signo ou característica, em um contexto de práticas de mercado, de distintividade absoluta, de forma que torna-se possível obter o registro de tais signos que, originalmente, pela falta de distintividade, pertenciam ao domínio comum (BARBOSA, 2011).

Seguimos com as pertinentes palavras do Prof. Denis Barbosa (2011, p. 23):

O fenômeno da significação secundária tem em comum com o instituto do *trade dress* o fato de sua natureza juscriadora resultar de práticas simbólicas da atividade de mercado. No entanto, não se cuida de exigir, para a proteção do *trade dress* pela concorrência desleal, que se comprove que tal elemento tenha adquirido significação secundária. No direito brasileiro, porém, ainda que não exista tal exigência, cabe sempre a imposição de que o elemento protegido através do art. 195 da Lei 9.279/96 (que é a base da rejeição à cópia do *Trade dress*) seja capaz de confundir a clientela quanto à origem dos produtos e serviço. Para a determinação dessa confusão em concreto, a existência de uma percepção pelo público de que o *trade dress* tenha adquirido significação secundária é útil, mas não exigível.

O Acordo TRIPs – Acordo Sobre Aspectos dos Direitos de Propriedade Intelectual Relacionados ao Comércio traz uma previsão acerca da significação secundária:

15.1 Qualquer sinal, ou combinação de sinais, capaz de distinguir bens e serviços de um empreendimento daqueles de outro empreendimento, poderá constituir marca. Estes sinais, em particular palavras, inclusive nomes próprios, letras, numerais, elementos figurativos e combinação de cores, bem como qualquer combinação destes sinais, serão registráveis como marcas. Quando estes sinais não forem intrinsecamente capazes de distinguir bens e serviços pertinentes, os Membros poderão condicionar a possibilidade do registro ao caráter distintivo que tenham adquirido pelo seu uso. [...] (BRASIL, 1994)

Cita-se ainda o caso Qualitex Co. vs. Jacobson Products Co. (UNITED STATES, 1995) onde restou determinado que as cores isoladamente, não poderiam ser registradas como marcas ou protegidas como *trade dress*, exceto se devidamente comprovado que estas adquiriram significado secundário ou *secondary meaning*, através de seu emprego efetivo e continuado, vindo a ser percebidas pelos consumidores como um forte indicativo da origem daqueles produtos ou serviços.

As decisões neste condão consolidaram a interpretação de que o *trade dress*, para se valer de proteção, apresente o requisito da distintividade, seja ela originária ou secundária. A distintividade originária se apresenta quando os elementos do *trade dress* possuem uma originalidade inerente a si; equivale dizer, quando desde o seu uso inicial eles já possibilitavam que o consumidor reconhecesse a origem do produto ou do serviço respectivo, pois é suficientemente distinto dos demais conjuntos imagens dos concorrentes e/ou não detém nenhuma característica comum no ramo de mercado em que se aplica.

Alternativamente, tem-se a distintividade secundária quando um determinado conjunto-imagem não possui, de início, uma distintividade inerente a si, por apresentar caracteres comuns aos demais conjuntos imagens daquele segmento de atividade econômica. Entretanto, com o uso continuado, tal conjunto imagem conquista reconhecimento do consumidor, que passa a associá-lo fortemente com os produtos ou serviços respectivos, a despeito de sua falta de distintividade original. Tem-se neste caso o fenômeno do *secondary meaning* em ação (BARBOSA, 2011).

Ainda de destaque é o caso Samara Brothers vs. Wall Mart, julgado em 2010 pela Suprema Corte norte-americana.

WAL-MART STORES, INC. V. SAMARA BROTHERS, INC., Decided March 22, 2000. [...] Trademark law has been held to embrace not just word marks and symbol marks, but also “*trade dress*” - a category that originally included only the packaging, or “dressing, “of a product, but in recent years has been expanded by many courts of appeals to encompass the product’s design. (...) Design, like color, is not inherently distinctive. The attribution of inherent distinctiveness to certain categories of word marks and product packaging derives from the fact that the very purpose of attaching a particular word to a product, or encasing it in a distinctive package, is most often to identify the product’s source. Where it is not reasonable to assume consumer predisposition to take an affixed word or packaging as indication of source, inherent distinctiveness will not be found. With product design, as with color, consumers are aware of the reality that, almost invariably, that feature is intended not to identify the source, but to render the product itself more useful or more appealing. (UNITED STATES, 2010).

Neste caso, esposou-se o entendimento de que a proteção do *design* de produtos (incluindo cor e embalagem) deveria estar condicionada à sua comprovada distintividade, adquirida pelo uso, uma vez que estas característica de *design* não seriam “*inherently distinctive*”².

Em tribunais pátrios, temos o *leading case* do “polvilho antisséptico”, in verbis:

A denominação necessária, segundo a doutrina (Tratado da Propriedade, Industrial, Gama Cerqueira, 2º ed., pág. 1151), implica na ausência de outra

² Inerentemente distintivas, em tradução livre.

expressão pela qual se possa designar o produto. Não é o caso do Polvilho Antisséptico, porquanto o produto é um pó, também geralmente conhecido como talco, ou substância medicamentosa em forma de pó, ao passo que polvilho pode designar várias outras coisas, mas principalmente (como é de todos sabido) a farinha amilácea finíssima, obtida da mandioca, amido ou goma, mencionada ao Novo Dicionário Aurélio, citado pelo réu. Não é portanto, a expressão necessária para designar o produto. Todavia, conforme indica o próprio texto legal, mesmo os nomes de uso necessário, comum ou vulgar, e que tenham relação com o produto a distinguir, poderão ser registrados, "quando se revestirem de suficiente forma distintiva. Assim, embora "polvilho" e "Antisséptico" possam ser consideradas expressões de uso comum ou vulgar, quando empregadas isoladamente, deixam de o ser quando utilizadas em conjunto, como Polvilho Antisséptico, caso em que apenas se poderá dizer que se trata de um produto popular, tradicional o que, porém, não o impede, mas até recomenda, o registro da marca. A autora comprovou que vem usando a marca Polvilho Antisséptico desde o inicio do século ou seja, desde o ano de 1903, quando o produto correspondente foi licenciado pela Saúde Pública (fls. 26). Além disso, também vem sendo usada em Portugal, onde está devidamente registrada (fls. .12/11). As declarações de fls. 28/31, passadas por firmas tradicionais do ramo, assim como os depoimentos de fls. 62/64, prestados por dirigentes de antigas firmas do setor farmacêutico, demonstram a veracidade das afirmações da autora nesse respeito. Como se isso não bastasse, difícil que haja alguém, no Brasil, com mais de cinquenta anos de idade, que não se lembre que desde sua infância o Polvilho Antisséptico, da Casa Granado, era o produto sempre usado para curar frieiras e brotoejas. Muito significativo o caso relatado pelas testemunhas, relativo ao uso da marca Polvilho Antisséptico pela multinacional Johnson & Johnson, cerca de quinze anos atrás, tendo a referida firma desistido do mesmo, por reconhecer tratar-se de patrimônio da Casa Granado. Embora inaplicável ao caso o art., 6º quinques da Convenção de Paris, na versão citada pela autora (decorrente da Revisão de Estocolmo, a que o Brasil não aderiu), certo é que o Item 2º do referido artigo, na revisão de Haia (fl. 44), estabelece que na apreciação do caráter distintivo de uma marca, dever-se-ão ter em conta todas as circunstâncias de fato, especialmente as de duração do uso da marca. Este dispositivo se torna aplicável à espécie, por força do art. 64 do Código da Propriedade Industrial. Não ficou demonstrado o argumento do réu, no sentido de que tal ressalva só prevalecia nos países em que a propriedade da marca se adquire por mera ocupação uma vez que não foi citado e parece inexistir qualquer dispositivo convencional ou legal que estabeleça tal restrição, quanto à aplicação do citado texto da Convenção de Paris. Decisão de 1985 do Tribunal Federal de Recursos na Apelação Cível Nº 102.635 - RJ Registro nº 5.796J56. (BARBOSA, 2011, p. 1)

E ainda, digno de citação o mui salutar precedente do TJ-SP:

O ponto nuclear da lide repousa numa questão de fato, consistente em saber se o termo preponderante Alpargatas, da denominação social da autora, São Paulo Alpargatas S/A, constitui ou não palavra genérica, que, designativa de determinado gênero de produto industrial, ou mercadoria, seja ou não suscetível de representar marca de indústria e merecer consequente tutela jurídica, no confronto com a denominação da ré, Cervera Alpargatas Ltda. E a resposta é francamente positiva. Ninguém questiona que o vocábulo alpargata é substantivo que, na significação original, denota espécie de sapato, ou calçado, com algumas particularidades que não vêm ao propósito.

Por demonstrá-lo, escusa recorrer aos dicionários. Mas a questão fática, que está a desafiar a qualificação jurídica neste processo, é outra, e está em indagar se, difundindo seu uso como expressão nominativa de singular espécie de calçado, por ela fabricado há muitos anos, não acabou a autora por emprestar-lhe conteúdo semiológico translato e singular, que entrou, no universo da concorrência mercantil, a identificar não só o produto em si, mas o comerciante mesmo, e todas as suas atividades industriais e comerciais. Foi o que, como fato público e notório, aconteceu. Por um fenômeno semântico não incomum e já esmiuçado por estudos pertencentes às áreas da Semasiologia e da Gramática Histórica, a palavra alpargatas perdeu o sentido primitivo corrente e passou a traduzir, nas relações comerciais e até nas comunicações cotidianas o específico calçado então fabricado pela autora e, por contaminação ideológica, a sua própria identidade, no mundo empresarial e, sobretudo, no plano da concorrência mercantil. De tal arte que o termo, cuja preponderância na denominação é intuitiva e manifesta (mot vedette dos franceses), ganhou a capacidade de suscitar imediata associação de idéias com aquele produto industrial característico e com a individualidade objetiva da autora e do seu grupo econômico, que hoje, não obstante interrompida ou terminada a fabricação, se identificam por essa única palavra, no relacionamento com o público consumidor (cf. fls. 105-128). A autora é individuada apenas como Alpargatas, assim apresentando-se perante o mercado, como sendo por este reconhecida! Este fato vistoso, explicável à luz de múltiplos fatores de ordem históricocomercial e lingüística, é que possibilitou à autora obter o registro do vocábulo, como marca notória (cf. fls. 31 e 130), objeto de previsão e especial tutela do art. 67, caput e parágrafo único, da Lei 5.772, de 21.12.71, além de vários outros registros concernentes a mercadorias distintas (cf. fls. 131-142), nenhum dos quais foi, até agora, desconstituído por ação própria (arts. 98 e 99 da mesma Lei 5.772/71). De modo que, até sobrevir o termo preclusivo da ação de nulidade de cada registro, se tem de reconhecer-lhes a todos a eficácia própria, que está em atribuir "direito real sobre bem incorpóreo a quem se disse titular do direito ao registro" (Pontes de Miranda, Tratado de Direito Privado, São Paulo, Ed. RT, 4.^a ed., 1977, t. XVII/70, § 2.018, n. 1). Mas o certo é que nulidade alguma se pode encontrar nos registros, que atenderam àquele peculiar fenômeno semântico, capaz de subtrair à palavra os caracteres de vulgaridade e necessariedade que a fariam irregistrável (art. 65, inc. 20, da Lei 5.772). E é justa a proteção concedida, porque foram as atividades permanentes e o conceito público da autora que operaram a translação de sentido, pela qual Alpargatas deixou de significar uma espécie de sapato, passando a simbolizar e nominar um produto determinado e seu fabricante TJSP - Ap 82.301-1 - 2.^a Câmara - j. 10/2/1987 - rel. Antonio Cezar Peluso (BARBOSA, 2011, p. 2).

Do quanto exposto, pode-se afirmar que o fenômeno da significação secundária pode ter relevância e, indubitavelmente, contribuir para a configuração de violação de *Trade dress*, mas que de modo algum é imprescindível para sua caracterização (BARBOSA, 2011; CRUZ, 2009).

2.4.3 A cor e o *trade dress*

A proteção da cor como elemento do *trade dress* – e sua associação com práticas desleais de comércio – não é recente. Já no longínquo ano de 1891, lecionava Marafy em comentário jurisprudencial sobre o uso de cores não suscetíveis de proteção como marca nas embalagens:

Mas a partir do momento em que a cor, seja pela adoção de disposições especiais, seja pela aplicação da dita cor a certas partes do produto, qualquer que seja a simplicidade das ditas combinações, chega a constituir um signo distintivo suscetível de atrair o olhar do comprador, ela entra dentro da categoria dos signos não enumerados do artigo primeiro da lei de 23 de junho de 1857, que podem servir de marca de fábrica; Nós devemos entretanto chamar a atenção que, mesmo não sendo a cor considerada como signo distintivo passível de ser validamente depositado, a imitação das cores, a partir do momento em que elas sejam claramente características, é sempre considerada pelo juiz, em caso de imitação fraudulenta de uma marca, como uma circunstância muito agravante, demonstrando e evidenciando a intenção do imitador de estabelecer confusão certamente dolosa³ (apud BARBOSA, 2011, p. 148).

Já naquela época se admitia a proteção à cor pela via da repressão à concorrência desleal, que neste caso se referiu como ‘imitação fraudulenta’ com intuito de “estabelecer confusão [...] dolosa”.

A proteção de cores como marca há muito é aceita em boa parte dos ordenamentos jurídicos pelas chamadas “marcas cromáticas”. Estas nada mais seriam do que um sinal composto por uma cor ou conjunto de cores, dispostas ou não de certo modo, e tendentes às funções da marca, qual sejam, identificar produtos e serviços e distingui-los dos concorrentes. A marca cromática pode existir em diversas configurações ou tipos, como a marca que reveste um produto ou embalagem (ou presente no serviço), de maneira isolada; como disposição de cores aplicada ao produto ou embalagem (ou presente no serviço); ou como a própria cor da marca figurativa ou mista. Pode também assumir a forma de cor única, também referida como *de per se*, que contém uma única cor e tonalidade. Também pode se apresentar como composição ou padrão com ordem específica (CASTRO, 2017).

³ No original : “Mais attendu que, du moment que la couleur, soit par l'adoption de dispositions spéciales, soit par l'application de ladite couleur à certaines parties du produit, quelle que soit d'ailleurs la simplicité desdites combinaisons, arrive à constituer un signe distinctif susceptible de frapper l'oeil de l'acheteur, elle rentre dans la catégorie des signes non énumérés de l'article premier de la loi du 23 juin 1857 qui peuvent servir de marque de fabrique;” “Nos devons cependant faire remarquer que, même lorsque la couleur n'entre pas en ligne de compte comme signe distinctif pouvant être valablement déposé, l'imitation des couleurs, lorsqu'elles sont franchement caractéristiques, est toujours considérée par le juge, em cas d'imitation fraudulente de la marque, comme une circonstance très aggravante démontrant avec évidence l'intention de l'imitateur d'établir une confusion tout au moins blâmable en équité”.

A International Trademark Association - INTA (1997) já se posicionou no sentido de que:

[...] fica resolvido que é posição da INTA que as cores, seja como uma combinação de cores ou uma única cor, podem servir como marca e, dessa forma, em circunstâncias apropriadas, poderão ser passíveis de reconhecimento marcário, proteção e registro (tradução nossa).

Alguns dos problemas levantados em desfavor do registro de marcas são atinentes à dificuldade de estabelecer padrões ou códigos absolutos para as cores e de distinguir suas tonalidades em casos de aplicação efetiva (*shade confusion theory*), e também no eventual esgotamento das cores perceptíveis ao olho humano, impossibilitando novos registros (*depletion theory*) (CASTRO, 2017).

Ademais, colacionamos alguns julgados selecionados acerca do tema:

As marcas ‘Nescau’ e ‘Toddy’ são renomadas e se distinguem em todas as classes sociais, por conseguinte, o alegado pela apelante não está caracterizado, até porque as partes competem amplamente no mercado e reiteradamente têm modificado as embalagens dos produtos que fabricam; **a utilização das cores branca, amarela, vermelha e azul, por si só, não pode dar origem à concorrência desleal.** O que efetivamente existe é que os produtos fabricados pelas partes são muitos próximos e originários de um vegetal, qual seja, o cacau, dissolvido num produto natural, que é o leite, assim sendo, há pouco a distinguir um do outro no que envolve a apresentação, contudo, é público e notório, portanto, independe de provas, que ‘Toddy’ e ‘Nescau’ são marcas distintas, não se podendo subestimar a opção do consumidor na aquisição de um produto por outro, já que se trata de preferência pessoal por esta ou aquela marca, levando em consideração o sabor e até mesmo o preço. **Desta forma, a pretensão da apelante de que estaria havendo cópia, ainda que parcial, da embalagem utilizada no produto ‘Nescau’ não tem consistência, haja vista que a autora não tem exclusividade de uso das cores,** e nem mesmo da forma como o chocolate líquido se apresenta, ou do invólucro que poderia conter o mencionado líquido, o que, inclusive, pode ser comparado com marcas distintas de automóveis, pois o consumidor mesmo tendo veículos com a mesma cor e eventual semelhança no modelo, também levará em consideração em caso de aquisição do produto as peculiaridades técnicas e outros itens correlatos, como o preço, no entanto, sempre terá realce a marca escolhida, o mesmo acontecendo no caso em exame. (...) Desta forma, o que efetivamente busca a apelante é excluir a ré da concorrência de mercado, o que não pode ocorrer, pois, inclusive, é benéfico ao consumidor a escolha do produto levando em consideração as marcas existentes, bem como contribui para que o preço não se eleve demasiadamente, haja vista que se trata de um componente da alimentação principalmente na infância (TJSP Ap. Cív. N° 252.756-4/8-00, 7^a Cam. Dir. Priv., rel. Des. Natan Zelinschi de Arruda, j. 25.06.08, grifo nosso).

TUTELA ANTECIPADA - Marca – Concorrência desleal - Pretensão à abstenção de uso de embalagens nas cores amarelo e rosa em produtos do gênero alimentício voltados a público infantil - Indeferimento pelo MM.

Juízo - Admissibilidade dessa decisão - Ausência de sérios indicativos acerca do risco de possível confusão pelo consumidor e efetivo prejuízo à empresa recorrente - Inexistência de prova inequívoca a convencer da verossimilhança da alegação – Recurso não provido (TJSP, AI nº 517.443-4/0-00, 6ª Cam. Dr. Priv., rel. Des. Encinas Manfre, j. 02.08.07, grifo nosso).

Nesse passo, a cor de fundo amarela ou de letras em verde, que sequer possuem idêntica tonalidade ao do novo produto (fls. 291, 1005 e 1006, 1233 e 1234), são elementos insuficientes para amparar a tutela almejada, mormente face das outras cores, grafismos é nitidamente distintos. É preciso conviver com uma margem de proximidade, quando incapaz de consubstanciar uma associação imitativa à marca que já foi de maior relevância no mercado consumidor. Contrariamente, estaríamos diante da monopolização do amarelo como coloração de fundo para embalagens de dentífricos, ainda que não houvesse a combinação estampada no ato registrário junto ao INPI (fls. 152) ou a elaboração de forma visualmente imitadora. (...) Vale dizer, a semelhança quanto à cor não atinge a similaridade passível de reprimenda pela via judicial, porque eventual escopo impeditivo demandaria a comprovação de uma identificação ou uniformidade visível, como já dito, face à apreciação do todo e capaz de gerar confusão junto ao mercado consumidor (...) De qualquer modo, em face da nossa legislação não admitir a apropriação de cores no âmbito marcário, resta necessário tolerar uma certa semelhança quanto à coloração preponderante, desde que não leve à confusão do mercado consumidor, segundo a visão do homem médio, no caso, concebida pela percepção amplamente majoritária dos consumidores avaliados.” (TJSP, Ap. Cív. N° 245.538-4/7-00, 5ª Câm. Dir. Priv., rel. Des. Oscarlino Moelle, j. 11.06.08, grifo nosso).

Da coleção de julgados acima, depreende-se que a proteção da cor como elemento do *trade dress* não tem aceitação pacífica, ou até mesmo majoritária, nos tribunais pátrios. Entretanto, decisões em sentido contrário podem ser encontradas:

No tocante ao uso da embalagem, entendeu a sentença de rejeitar o pedido da Apelante dada a semelhança de formato e cores com a embalagem da marca KOLESTON e, como uma consequência lógica desse entendimento não poderia declarar inexistência de violação da marca KOLESTON, como pretendido pela Apelante. A sentença foi categórica ao afirmar que é evidente a imitação da embalagem do produto daquelas, utilizando-se a autora de denominação COR E TON em lugar daquela cujo registro detém, repita-se ainda uma vez COR & TON, tornando-se patente a imitação, mormente em caixa do mesmo tamanho e **coloração assemelhada**. A possibilidade enfatizada pela sentença de induzimento do consumidor a erro, dúvida, ou confusão é manifesta. Seria pura perda de tempo realizar uma perícia para a constatação de algo tão óbvio. A **alegação de que outros fabricantes usam embalagens com as mesmas cores é obviamente irrelevante para o julgamento desta causa onde se discutem a questão entre apelante e apelada** (TJRJ, AC 2003.001.08938, Des. Ruyz Athayde Alcântara de Carvalho, 9ª Câmara Cível, DJ 29/01/04, grifo nosso).

No âmbito deste estudo, buscar-se-á, através de análise estatística da jurisprudência selecionada, determinar se o elemento “cor” influencia positiva ou negativamente os resultados das lides de *trade dress*, conforme detalhado no capítulo de metodologia.

2.4.4 A Prova Pericial nas Ações de *trade dress*

No exercício de suas funções, o magistrado pode necessitar do auxílio, constante ou eventual, de outras pessoas que, tal como ele, devem atuar com diligência e imparcialidade, consoante o art. 149 do Código de Processo Civil - CPC (BRASIL, 2015). Quando a matéria arguida demandar conhecimentos técnicos ou científicos afeitos a determinadas áreas de conhecimento técnico, o julgador será assistido por perito ou órgão, segundo o artigo 156, §1º do Código de Processo Civil (idem). A lei nº. 13.105/2015 do novo Código de Processo Civil (idem) trouxe a possibilidade do juiz também ser assistido por “órgãos técnicos ou científicos”, não estando limitado apenas a pessoas físicas na condição de “profissionais de nível universitário”, tal como dispunha o código revogado.

Nem sempre a prova pericial é imprescindível para o deslinde da questão litigiosa. A prova pericial se dá por exame, vistoria ou avaliação, e poderá ser determinada de ofício ou a requerimento das partes. Será indeferida quando: a) não houver a necessidade de conhecimento especial de técnico para prova do fato; b) o fato já estiver comprovado por outros meios de prova; e; c) a verificação for impraticável, segundo artigo 464, §1º do Código de Processo Civil (BRASIL, 2015).

A prova pericial poderá ser dispensada, quando as partes apresentarem, sobre as questões de fato, pareceres técnicos ou documentos elucidativos que forem considerados suficientes pelo magistrado, tal como consta no artigo 472 do Código de Processo Civil (idem).

No caso específico deste estudo, tratamos da necessidade de prova pericial em ações de *trade dress*. O levantamento de jurimetria aplicado ao corpus selecionado indica que a prova pericial não é prevalente, como se verá no capítulo de resultados. Ademais, alguns autores criticam sua necessidade, em consonância com o entendimento exarado no julgado abaixo transcreto:

Se se tratar de colidência de forma ou de figura, sobretudo naqueles processos em que se alega a possibilidade de indução do consumidor a confusão ou erro, a determinação de prova pericial seria elemento infirmador ontológico dessa alegação. **Muito simples: o que se buscaria seria a**

percepção de um consumidor médio. Nomear perito para aferir colidência “visual” é quase como dizer que a colidência é quase tão imperceptível que só mesmo um experto no assunto poderia aferi-la (Tribunal Regional Federal da 2ª Região, 1ª Turma Especializada, JC. Márcia Helena Nunes, AC 2008.51.01.805451-9, DJ 25.09.2009, grifo nosso).

A despeito de posições conflitantes sobre a questão procura-se, neste estudo, determinar quantitivamente a aplicação da prova pericial nas ações de *trade dress* e, assim, buscar relacioná-las com o êxito, determinando seu efetivo impacto na decisão final do magistrado.

2.4.5 A Importância da Jurisprudência para o Instituto do *trade dress*

Pode-se arguir que o reconhecimento da jurisprudência como fonte do direito restou pacificado, em face da Emenda Constitucional nº 45, que acrescentou o artigo 103-A à Constituição Federal Brasileira, e introduziu em nosso ordenamento jurídico a edição de Súmulas com caráter vinculante (BRASIL, 2004).

Especificamente em relação ao instituto do *trade dress* que surgiu e permanece em nosso ordenamento como uma construção puramente jurisprudencial, é ainda mais clara a importância do estudo da jurisprudência para delimitar suas características e delinejar seu escopo. É o estudo da jurisprudência que pode, em maior medida, trazer luz a este instituto. Decerto, em face da ausência de positivação, o julgador possui enorme discricionariedade para determinar o alcance do instituto em sua aplicação prática, qual seja, no deslinde das situações de concorrência desleal levadas ao crivo judicial (MELLO, 2012).

No âmbito deste estudo, tendo em vista a aplicação da metodologia de Jurimetria, a jurisprudência ganha contornos ainda mais vitais à compreensão buscada do instituto jurídico, como se depreenderá da explanação pertinente do capítulo de metodologia, e da análise e discussão dos resultados da análise estatística descritiva e inferencial.

3 METODOLOGIA

3.1 Jurimetria

3.1.1 Conceito e Aplicação

As ciências de todos os ramos sempre se valeram de métodos matemáticos para formular hipóteses e experimentos. Decerto, em alguns ramos de investigação científica, como nas ciências ditas exatas, os métodos de análise quantitativa são o principal instrumento de observação e aferição de fenômenos, e consequentemente, de falseamento de hipóteses.

Por outro lado, as ciências humanas e sociais sempre tiveram aplicação mais restrita de ferramentas quantitativas e métodos matemáticos ou estatísticos. Indubitavelmente, os fenômenos que busca deslindar são muitas vezes carregados de grande subjetividade, ou apresentam-se em um espectro de variações e combinações que pode ser até mesmo infinito, o que dificulta – mas não inviabiliza – qualquer tentativa de modelagem matemática dos mesmos. A ciência do Direito é um exemplo de tal caso (ZABALA; SILVEIRA, 2014).

Entretanto, a ideia de utilizar métodos quantitativos para estudar os fenômenos jurídicos não é algo novo. O primeiro trabalho conhecido sobre o tema data de 1709, da lavra do grande matemático Nicolau Bernoulli, que escreveu sobre temas relacionados à probabilidade de confiança em testemunhas, probabilidade de inocência de um acusado, e precificação de seguros, em sua tese de doutoramento intitulada “*De usu artis conjectandi in jure*”.

A aplicação da matemática, e mais especificamente dos métodos estatísticos, à ciência do Direito e à prática jurídica, entretanto, não teve grande relevo na literatura até o momento. Mais recentemente, em meados do século XX, Loewinger (1948) cunhou a expressão *Jurimetrics* para se referir a uma conjugação de teoria jurídica, métodos computacionais e estatística, com fulcro de analisar a jurisprudência e os fenômenos jurídicos em geral, e tornar a prática do direito mais previsível.

Em 1959 foi criado o primeiro periódico referente ao tema, editado pela *American Bar Association* em periodicidade trimestral, onde trata-se de temas relacionados ao Direito, à computação e à estatística, intitulado “*Jurimetrics Journal*”⁴.

Joseph Kadane, pesquisador da Universidade Carnegie Mellon é uma das referências na área, tratando da aplicação estatística em questões como seleção de jurados, auditoria de impostos, lides de direito empresarial, e mesmo ensaios sobre a ética na apresentação de métodos quantitativos em tribunais de júri (KADANE; LEHOCZKY, 1976).

3.1.2 A Jurimetria no Brasil

O presente trabalho se utilizará de conhecimentos e paradigmas de uma relativamente nova disciplina da ciência do Direito, a jurimetria. A jurimetria pode ser definida como a disciplina do Direito que utiliza a metodologia estatística para estudar o funcionamento do ordenamento jurídico, de modo empírico.

A jurimetria pretende determinar a realidade do Direito na sua prática diária, observando o funcionamento das instituições jurídicas e das sentenças, autos, contratos e demais populações de normas individuais produzidas pelo sistema jurídico diariamente. A jurimetria busca investigar como o Direito influencia a vida cotidiana das pessoas de modo empírico.

Como exemplo de jurimetria, temos uma pesquisa sobre a dissolução de sociedades nos tribunais estaduais brasileiros entre 1997 e 2011. A pesquisa identificou que as ações de dissolução de sociedade sugem em um momento de crise política interna, no qual os sócios discordam a respeito da gestão da sociedade. Ademais, a pesquisa aponta que a retirada e a expulsão por falta grave são utilizadas de forma deturpada, indicando que a lei deveria prever mecanismos de separação baseado em leilões ou propostas para auxiliar a sociedade a sobreviver à lide (NUNES, 2012).

No Brasil, a produção científica sobre o tema vem crescendo. Bertran (2006, 2007) produziu um trabalho para avaliação de contratos de arrendamento mercantil indexados em dólar, de caráter bastante similar aos objetivos deste trabalho, qual seja, a aplicação de métodos matemáticos, por meio de software, para determinar o comportamento dos tribunais em relação ao tema estudado. Puschel (2011) coordenou uma investigação sobre a suposta

⁴ site: www.law.asu.edu/jurimetrics

‘indústria do dano moral’ no judiciário brasileiro, terminando por concluir que tal ‘indústria’ não existe, a despeito de alguma críticas metodológicas.

Um dos problemas ainda enfrentados na realidade nacional é o pouco acesso a dados de qualidade, normalmente com grandes restrições, erros, discrepâncias de formato, entre outros problemas. Estes problemas certamente não se restringem ao Brasil. Especificamente em relação ao direito, porém, trabalhamos ainda com a questão da complexidade da informação. Conforme exposto anteriormente neste trabalho, a dicotomia entre dados e informações é ainda mais aguda no tratamento de informações codificadas em linguagem natural (ou linguagem humana), não apenas devido à enorme variedade sintática, gramatical e semântica possível para exprimir os mesmos conceitos, mas também à dificuldade de estabelecer parâmetros objetivos de classificação destes conceitos. Esta é uma dificuldade inerente a qualquer estudo que se proponha aplicar métodos matemáticos, quantitativos, a fenômenos precipuamente subjetivos, como os jurídicos. Pode-se afirmar que os fenômenos jurídicos são eminentemente subjetivos, uma vez que a norma jurídica, elemento básico da ciência do direito, admite interpretação, e dá azo inclusive à disciplina da hermenêutica (CAVALCANTI, 2004).

Um dos objetivos da estatística, como ramo de estudo da matemática, é extrair informação a respeito de quantidades de interesse, porém desconhecidas, tais como valores, medidas de tempo ou espaço, probabilidades ou qualquer outra grandeza de interesse, passível de metrificação. Neste condão, ela busca armar o observador com ferramentas capazes de analisar as informações quantificadas, medir incertezas e assim auxiliar na tomada de decisões. Também é definida como o ramo da matemática que lida com coleta de dados, organização, análise, interpretação e apresentação (DODGE, 2006).

Esta dinâmica possui correspondência imediata no Direito, como exemplificativamente, no instituto da tutela de urgência, antecipada ou cautelar. O próprio Código de Processo Civil de 2015 usa, inclusive, a expressão “probabilidade”, em seu art. 300, in verbis: “A tutela de urgência será concedida quando houver elementos que evidenciem a **probabilidade do direito e o perigo de dano ou o risco** ao resultado útil do processo” (BRASIL, 2015, grifo nosso).

Em tais tutelas, cabe ao julgador verificar a “probabilidade do direito” (ou verossimilhança) e o “perigo de dano ou risco ao resultado útil do processo”. A primeira pode ser entendida como um juízo acerca da probabilidade de um ato ou fato jurídico ser verdadeiro, formal ou materialmente; e a segunda, como um juízo acerca da probabilidade de ocorrência de um dano, e, adicionalmente, de tal dano ser de impossível reparação.

É imprescindível acrescentar que a verossimilhança refere-se não só à matéria de fato, como também à plausibilidade da subsunção dos fatos à norma invocada, conduzindo aos efeitos pretendidos. O magistrado precisa avaliar se há probabilidade de ter acontecido o que foi narrado e quais as chances de êxito do demandante (DIDIER et al., 2008, p. 627).

Ademais, o próprio instituto da “prova” pode ser comparado aos dados de que o estatístico se vale para elaborar suas hipóteses; juízos subjetivos acerca de sua confiabilidade são muito comuns na prática judiciária, tais como, por exemplo, na confiabilidade da prova testemunhal (THEODORO, 2016). Neste campo, muito valeria a aplicação de métodos formais para aglutinar a informação extraída dos dados à opinião subjetiva de um especialista ou perito, aplicando-se a chamada teoria da decisão (DEGROOT, 2005; SOUZA, 2007).

A despeito da reconhecida importância dos cálculos probabilísticos na aplicação do Direito, a disciplina da Jurimetria ainda tem farto terreno para avançar. O maior entrave, sem dúvida, se relaciona com a necessidade de aprimorar técnicas de análise de informação em grande volume, codificadas em linguagem natural. A estatística se vale de observações mensuráveis de subconjuntos amostrais, e quanto mais numerosos forem tais unidades de observação em relação ao seu universo total de interesse, tão maior será sua precisão e consequente validade (BASU; GHOSH, 1988).

Considerando que a jurisprudência se apresenta sempre codificada em linguagem natural, a análise humana de grandes amostras se torna deveras onerosa, ou mesmo impossível, além de se sujeitar ao erro humano. Uma ferramenta de análise automatizada, ainda que certamente sujeita a margem de erro, pode propiciar a capacidade de escrutínio de enormes volumes amostrais. Sua margem de erro pode ser estimada, e progressivamente diminuída, com otimizações de algoritmos e outras técnicas computacionais (CARRILHO, 2007). Tal proceder contribuiria sobremaneira para a precisão e validade estatística das hipóteses.

Com o surgimento dos computadores modernos, e mais recentemente, com os avanços tecnológicos e científicos relacionados à Inteligência Artificial, este tipo de estudo torna-se mais acessível economicamente (dispensando, por exemplo, o uso de supercomputadores), e mais confiável em termos teóricos (posto que apoiado em métodos e algoritmos bem descritos e revisados na literatura, e fartamente validados empiricamente). Ademais, as técnicas de aprendizado de máquina (*machine learning*) e os algoritmos preditivos que as possibilitam já se encontram em estado avançado de desenvolvimento (ARANHA, 2007), e podem propiciar novos ganhos de eficiência e precisão nos estudos de jurimetria, abrindo até mesmo

possibilidades antes impensáveis dentro da prática jurídica, como a estimativa de chances de acolhimento de teses jurídicas específicas no bojo de uma ação judicial.

No Brasil, tivemos um marco importante no estudo da Jurimetria com a fundação da Associação Brasileira de Jurimetria – ABJ, entidade sem fins econômicos, constituída na forma de associação civil, com os seguintes propósitos:

A Associação Brasileira de Jurimetria (ABJ) busca contribuir para o aperfeiçoamento da sociedade brasileira através do levantamento de subsídios empíricos e quantitativos para elaboração de políticas públicas, sempre tendo em vista três missões principais: (i) Reunir e incentivar pesquisadores com preocupação em investigar e descrever os processos de decisão em que são criadas as normas individuais e concretas; (ii) Disciplinar a jurimetria como um ramo do conhecimento jurídico, definindo suas premissas, seus fundamentos, seus conceitos e relações essenciais; (iii) Colaborar com entidades públicas e privadas no esforço estratégico de aperfeiçoar os mecanismos de prestação jurisdicional através da elaboração de leis e da administração dos tribunais (ASSOCIAÇÃO BRASILEIRA DE JURIMETRIA, 2018).

3.2 Conceito de Inteligência Artificial e evolução histórica

O conceito de Inteligência Artificial (IA) se contrapõe à noção de inteligência natural, ou seja, aquela exibida por humanos e animais. Inteligência artificial adquiriu status de disciplina acadêmica em 1956, com o Congresso de Dartmouth (Dartmouth Summer Research Project on Artificial Intelligence), considerado amplamente na literatura como marca de fundação dos estudos acadêmicos sobre o tema (MOOR, 2006, pp. 87-9). Este encontro reuniu um grupo de matemáticos e cientistas, com base na proposta formalmente apresentada por McCarthy et al. (1955), onde se encontra a primeira menção à expressão ‘Inteligência Artificial’:

We propose that a 2-month, 10-man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer (p. 2)

Convencionou-se chamar de Inteligência Artificial todas aquelas técnicas ou métodos pelos quais um computador pode modelar os processos normalmente associados à inteligência natural, tais como reconhecimento de padrões, compreensão de linguagem, aprendizado e

solução de problemas não-padrонizados. A longo prazo, o objetivo dos estudos neste campo seriam a modelagem e criação de uma inteligência geral, capaz de realizar as mesmas tarefas ou objetivos que seres dotados de inteligência natural. Hodernamente, o campo de IA comprehende subdivisões que muitas vezes não se comunicam entre si, indicando uma dificuldade em integrar os avanços em modelos específicos ao um todo coeso que atinja o objetivo de uma inteligência geral (RUSSELL; NORVIG, 2003).

Na passado, o estudo da IA experimentou fases de grande otimismo, alternadas com fases de desinteresse e pessimismo, face às dificuldades de avanço. Recentemente, com o advento de poder de processamento elevado, aliado a um baixo custo, grande disponibilidade de dados (*big data*), e aprimoramentos na compreensão teórica dos modelos, o campo da IA experimentou uma ‘renascença’, recebendo muita atenção na mídia e investimentos em pesquisa (BASTOS, 2006).

As ferramentas de IA hoje ocupam inquestionável, e por vezes essencial, papel na indústria de tecnologia, auxiliando na solução de problemas de altíssima complexidade, que até pouco tempo eram computacionalmente muito onerosos ou teoricamente insanáveis. Um exemplo é o advento das assistentes pessoais virtuais, que comprehendem linguagem natural como falada por humanos, executam tarefas solicitadas, e aprendem com seus erros (AUGELLO et al., 2016).

Com efeito, o campo da linguagem é um dos pilares dos estudos de IA, notadamente nas capacidades de reconhecimento da fala (*speech recognition*), e na compreensão do conteúdo da linguagem (*natural language processing*) (ARANHA, 2007).

Além disso, com a disponibilidade de grande volume de dados e avanços nos modelos matemáticos, observou-se significativo aprimoramento no campo dos modelos preditivos, também chamados de modelos de aprendizagem de máquina (*machine learning models*) através dos quais os computadores processam exemplos, reconhecem padrões, e aprendem a executar tarefas, predição de eventos futuros ou classificações similares, sem supervisão adicional. Modelos preditivos de maior ou menor complexidade estão presentes em praticamente toda a indústria e ciência modernas, desde previsões meteorológicas e simulações de engenharia, até o prosaico preenchimento automático de formulários eletrônicos variados (GEISSER, 2016).

A taxonomia das técnicas ou ferramentas de IA, entretanto, não é pacífica na literatura. Muitas técnicas que eram consideradas como afeitas à IA, com sua popularização e simplificação, são removidas do campo da IA e passam a ser consideradas como tecnologia rotineira, em um fenômeno chamado “Efeito IA” (BAILEY, 2016).

No âmbito deste trabalho, serão aplicadas técnicas de IA relativas à Mineração de Textos e Processamento de Linguagem Natural, e Modelos Preditivos. O objetivo destas técnicas é obter uma compreensão de determinados assuntos ou tópicos através do processamento automatizado de textos a eles referentes, com baixo grau de interação ou supervisão humana, e resultando em grandezas quantificáveis.

A vantagem obtida é a possibilidade de recuperar ou extrair informações relevantes de um volume grande de dados, com maior grau de precisão e/ou abstração, e menor incidência de erro humano (FINLAY, 2014).

3.3 Recuperação de Informação e Mineração de Textos

Os termos “dados” e “informação”, embora utilizados muitas vezes como sinônimos, possuem significações distintas. Dado é um fenômeno desprovido de significado e contexto (LAUDON; LAUDON, 2002). Informação resulta do processamento, manipulação e organização de dados, interpretados, compreendidos e contextualizados, que adquirem assim um significado (GOLDSCHMIDT; PASSOS, 2005). Dados tem natureza objetiva, enquanto informações tem natureza subjetiva; agentes diversos podem atribuir contexto e significado diverso ao mesmo conjunto de dados. A maioria das organizações não sofre de carência de dados, mas sim de um excesso; o principal problema é a inconsistência e redundância dos dados (SINGH, 2001).

Informações e dados podem ser armazenadas de modos variados. O armazenamento digital, utilizando representações binárias registradas como cargas eletromagnéticas em dispositivos computacionais, é relativamente novo na história da humanidade, tendo surgido com o advento dos computadores. Historicamente, o registro mais comum de informações é através da linguagem escrita, registrada em textos impressos.

Entretanto, enquanto o acesso a este tipo de informação é lento, complexo, e de baixo rendimento. Para efetuar todas as etapas de manipulação desta informação, é imprescindível a intervenção humana, que constitui o principal gargalo da eficiência e rapidez deste processo, devido à limitação natural do ser humano em processar grandes volumes de informação e reter o conhecimento para acesso futuro (MANDEL; SIMON; DELYRA, 1997). Métodos tradicionais de recuperação de informação baseados em atuação humana não permitem o tratamento e manipulação de grande volume de dados (SINGH, 2001).

Técnicas de mineração de dados, embora bem conhecidas e provadas, são inadequadas para recuperação de informações. Recentemente comprovou-se que oitenta e cinco por cento

de todos os dados do mundo estão armazenados na forma de documentos de texto, o que representa um desafio de acesso à informação neles codificada (GDS PUBLISHING, 2008; IBM, 2008), uma vez que precluem a utilização de técnicas simples de mineração de dados.

Neste contexto, a utilização de técnicas de recuperação de informação ganha enorme relevância, em especial para o armazenamento de informações na forma de texto e a recuperação da informação a eles associada (BAEZA-YATES; BERTIER, 1999; MANNING; RAGHAVAN; SCHUTZE, 2007).

Técnicas de mineração de textos (*Text Mining*) são necessárias para recuperar e tratar informações textuais, obtendo informações úteis ao usuário a partir de uma base de dados em linguagem escrita. O principal objetivo de tais técnicas é a análise automatizada de documentos textuais, com o fulcro de recuperar ou obter o conhecimento implícito neles registrado (ARANHA; PASSOS, 2006).

As técnicas atuais de mineração de textos ainda estão em estágio de estudos e aperfeiçoamentos. O processamento total e isento de erros da linguagem natural ainda está fora do alcance da tecnologia atual. Porém, existem diversas técnicas capazes de extrair informações valiosas e adequadas da linguagem natural presente nos documentos textuais (NEVES, 2012).

3.4 Processamento de Linguagem Natural

O Processamento de Linguagem Natural (PLN) consiste em um agregado de métodos formais utilizados para análise de textos, geração automatizada de textos e apreensão de informações neles codificada. Pode ser entendido como uma etapa ou técnica utilizada no processo de mineração de textos. O uso de técnicas de PLN na mineração de textos, embora não seja essencial, resulta em um ganho na qualidade da recuperação de informações, o que pode justificar o esforço computacional adicional. O processo de Mineração de Textos consiste em cinco etapas: Coleta de documentos, pré-processamento, indexação, mineração e análise (ARANHA, 2007).

As linguagens de programação são facilmente interpretadas por computadores, devido à sua natureza objetiva, precisa, de regras fixas e bem determinadas. Com efeito, a omissão de um simples caractere torna uma instrução ininteligível para a máquina, que necessita de um conjunto lógico preciso e isento de erros ou omissões para sua execução. Por outro lado, a linguagem humana é extremamente complexa para a interpretação por computadores, devido ao seu caráter simbólico, muitas vezes ambíguo e sujeito a interpretações de contexto, de

regras gramaticais, e de arcabouço cultural e abstrato no qual se insere ou foi criada (INSITE, 2001).

Deste modo, o Processamento de Linguagem Natural tem por objetivo final capacitar os computadores a compreender e compor textos, através de reconhecimento de contexto, análise sintática e semântica, análise léxica e morfológica, e até mesmo apreensão de conceitos abstratos (ARANHA, 2007).

No âmbito deste trabalho, utilizar-se-á um subconjunto das técnicas que compõem a Mineração de Textos, com aplicação de técnicas específicas de Processamento de Linguagem Natural, com intuito de obter informações relevantes a partir do *corpus* composto pela jurisprudência selecionada acerca do instituto do *trade dress*, de acordo com os objetivos de pesquisa pretendidos.

3.5 Coleta de textos

A primeira etapa apontada por Aranha (2007) é a coleta dos textos. Neste condão, em linha com o objeto de estudo deste trabalho, procedeu-se à seleção e coleta de textos decisões judiciais de segunda instância (jurisprudência) que versassem precipuamente sobre o instituto do *trade dress*.

A coleta foi efetuada através de buscas simples por palavras-chave (*keyword search*), considerado um dos métodos mais conhecidos e simples de seleção e coleta de informações (CARRILHO J., 2007). Através deste procedimento, construiu-se o universo amostral do estudo em epígrafe.

O tipo de estudo adotado foi exploratório. Esse tipo de estudo permite ao pesquisador coletar dados em busca de padrões, e através da formulação de questões e do refinamento da matéria, proceder a um estudo mais sistemático (PATTON, 1990). O objetivo do estudo é trazer informação rica, comprehensiva e sistematizada através dos casos de interesse, seja por métodos qualitativos ou quantitativos (BARAÑANO, 2008).

O escopo deste trabalho comprehende, outrossim, uma análise documental e bibliográfica exploratória, de natureza quantitativa e qualitativa, com fulcro no exame da doutrina e das decisões judiciais acerca do instituto jurídico do *trade dress*.

Para efetuar a análise da jurisprudência, é necessária a escolha de um recorte temporal e geográfico específico, uma vez que o exame de toda a jurisprudência pátria se mostraria inviável.

Optou-se, destarte, por um recorte geográfico limitado a dois Tribunais de Justiça, dos Estados do Rio de Janeiro e de São Paulo. O recorte temporal selecionado foi de três anos, compreendendo os anos-calendário de 2015, 2016 e 2017.

A escolha dos Tribunais de Justiça obedeceu a dois critérios norteadores: importância em volume de ações, e importância em termos de influência junto aos operadores do direito.

O Tribunal de São Paulo é o mais importante do Brasil em termos numéricos. Sob sua competência territorial, julgou-se em segundo grau um total de sessenta e oito ações versando precipuamente sobre violações de *trade dress*, no triênio em escrutínio.

Com efeito, segundo dados de 2014, o Estado de São Paulo tem um Produto Interno Bruto de R\$1,858 trilhões, que corresponde a 32,2% do Produto Interno Bruto nacional (INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA, 2015). São Paulo concentra parte considerável da atividade empresarial brasileira. Natural que tal Estado concentre também o maior número de ações judiciais relacionadas ao *trade dress*, uma vez que o instituto é afeito, por sua natureza, à atividade empresarial⁵.

Entretanto, o estudo de um Tribunal apenas não permitiria a comparação com resultados semelhantes de outros tribunais. Eventuais idiossincrasias do TJ/SP restariam despercebidas, uma vez que ausente uma contraposição com os julgados de outro Tribunal de Justiça. O resultado assim obtido necessariamente teria menor valor estatístico.

Da mesma forma, o estudo estatisticamente relevante da jurisprudência de qualquer Tribunal carece de um volume mínimo de documentos que lhe assegure um grau aceitável de validade. Este volume não é desprezível, e embora se tenha optado pela análise parcialmente automatizada dos documentos, ainda assim um exame e validação manual se fazem necessários, o que preclui, dentro do âmbito limitado de uma dissertação, a escolha de um número grande de tribunais para cotejo.

Optou-se, portanto, pela análise adicional de um Tribunal de Justiça estadual, o do Estado do Rio de Janeiro. A escolha foi motivada, como acima descrito, pelo critério de relevância e influência para os operadores do direito.

O Estado do Rio de Janeiro, antiga capital do país e sede de uma das primeiras faculdades de direito do Brasil, mantém ainda uma grande relevância no meio jurídico, a despeito de sua decadência econômica em tempos recentes. Historicamente, sempre albergou grande número de doutrinadores e bancas de causídicos de grande renome, servindo de baliza

⁵ “*trade dress*” significa, em tradução literal, “vestido do negócio”.

à aplicação do direito em todo o país, notadamente em matérias afeitas à Propriedade Intelectual.

O Rio de Janeiro é sede do Tribunal Regional Federal da 2^a Região e do Instituto Nacional da Propriedade Industrial - INPI, autarquia responsável pelos registros e depósitos de propriedade industrial. Sob a égide do TRF2, são concentrados no Rio de Janeiro todos os julgamentos relacionados à competência de atos administrativos do INPI, o que lhe confere grande importância na aplicação do direito da propriedade industrial.

Em seu meio acadêmico, florescem muitas das publicações mais relevantes no âmbito da propriedade industrial, e lecionam muitos dos doutrinadores mais respeitados do país na matéria. Destarte, foi o Tribunal escolhido como contraponto à análise da jurisprudência do TJ/SP.

Em relação ao recorte temporal, igual ponderação foi feita. Um recorte deveras extenso produziria uma gama muito ampla de documentos, de difícil análise nas possibilidades limitadas de uma dissertação de mestrado. Por outro lado, um recorte muito curto limitaria a validade estatística do estudo em questão.

Outra questão metodológica a ser considerada refere-se à validade de um estudo de caráter eminentemente documental, focado na revisão do resultado prático da prestação jurisdicional. Como tal, debruça-se sobre a externalidade da ponderação das lides pelo julgador, a jurisprudência, que é fonte do direito⁶. A escolha de um intervalo temporal muito extenso arriscaria a seleção de decisões já ultrapassadas, seja pela natural evolução dos entendimentos relacionados ao tema, como pela igualmente natural sucessão de magistrados. Equivale dizer, se pretende um entendimento válido e relevante ao tempo deste trabalho, e, portanto, impõe-se estudar o passado recente.

A partir desta ponderação, optou-se por um interregno trienal, que já fornece número adequado de documentos, simultaneamente provendo um entendimento prático relevante ao operador do direito, nesta data. Foi considerada na busca a data de publicação no Diário Oficial da decisão de segundo grau; a data de publicação é o termo inicial da contagem dos prazos processuais, e o momento em que a decisão passa a produzir efeitos no mundo do direito, genericamente falando, vis-à-vis a redação do art. 224, CPC (BRASIL, 2015) o que justifica sua escolha como marco temporal da seleção.

Art. 224. Salvo disposição em contrário, os prazos serão contados excluindo o dia do começo e incluindo o dia do vencimento.

⁶ Para Arnaldo Rizzato, as fontes do direito originam de fato o próprio direito, de modo que seu estudo é essencial à sua compreensão; “*as fontes do direito são os elementos geradores do direito, ou os elementos de onde deriva o direito*” (2003, p. 37).

§ 1º Os dias do começo e do vencimento do prazo serão protraídos para o primeiro dia útil seguinte, se coincidirem com dia em que o expediente forense for encerrado antes ou iniciado depois da hora normal ou houver indisponibilidade da comunicação eletrônica.

§ 2º Considera-se como data de publicação o primeiro dia útil seguinte ao da disponibilização da informação no Diário da Justiça eletrônico.

§ 3º A contagem do prazo terá início no primeiro dia útil que seguir ao da publicação.

A coleta e seleção documental relativa à jurisprudência foi feita pela Internet, uma vez que ambos os Tribunais oferecem uma interface *online* de busca e obtenção do texto integral das decisões que compõem a jurisprudência. Neste sentido, foram utilizadas as seguintes páginas de internet:

Tribunal de Justiça do Estado de São Paulo	https://esaj.tjsp.jus.br/cjsg/consultaCompleta.do
Tribunal de Justiça do Estado do Rio de Janeiro	http://www4.tjrj.jus.br/ejuris/ConsultarJurisprudencia.aspx

Para efeitos deste trabalho, buscou-se apenas coletar as decisões que versam precipuamente sobre *trade dress*. O instituto é arguido como elemento da tese jurídica em diversos tipos de ações judiciais fundadas em concorrência desleal e à violação de direitos de propriedade industrial, notadamente à violação marcária, porém sua preponderância como elemento principal da lide é questionável nestes casos; é possível arguir que sempre se trata de elemento acessório e de importância secundária à tese jurídica nestas ações, uma vez que a alegação principal nestes casos pode ser a prática diversa e concomitante de outros atos de concorrência desleal.

Uma vez que o objetivo deste estudo é analisar a jurisprudência específica sobre o *trade dress*, buscou-se focar a seleção e coleta exclusivamente nos julgados que tenham o instituto do *trade dress* como claramente preponderante. Para efetuar esta seleção, optou-se por uma busca por palavra-chave somente na ementa da decisão judicial, já que tal elemento contém os termos preponderantes à lide.

A busca foi feita utilizando as palavras-chave “*trade dress*” e sua expressão correspondente no vernáculo, “conjunto-imagem”. Foram buscadas também suas grafias variantes “*trade-dress*”, “*tradedress*”, “conjunto imagem” e “conjuntoimagem”, para assegurar que eventuais variações relativas à presença ou ausência de hífen ou espaço não resultassem em erro de busca. Os resultados das buscas foram cotejados para garantir a inclusão da totalidade da jurisprudência aplicável.

Foram examinados apenas acórdãos de apelação, sendo descartados os julgados referentes a Agravos e Embargos de qualquer natureza. Agravos dizem respeito a questões incidentais no processo, e via de regra não atacam diretamente a matéria arguida. No escopo deste estudo, os Agravos foram suprimidos, sem prejuízo maior aos objetivos almejados. Embargos dizem respeito a questões acessórias relacionadas ao julgado, como falta de clareza ou contradições no texto do acórdão, e via de regra não atacam diretamente a matéria arguida. No escopo deste estudo, os Embargos foram igualmente suprimidos, por igual motivo (THEODORO, 2016).

Por fim, restringiu-se a busca às decisões da Justiça Estadual dos Estados selecionados, omitindo-se propositalmente eventual recurso aos Tribunais superiores (Superior Tribunal de Justiça e Supremo Tribunal Federal). Esta omissão deve-se ao fato de que o presente trabalho busca deslindar questões de direito material afeitas ao instituto do *trade dress* na jurisprudência, e não se propõe estudar questões de direito processual, como a prescrição, por exemplo. As questões atinentes ao *trade dress* são eminentemente probatórias, uma vez que envolvem a coleção de elementos fáticos de convencimento do julgador relativos à conduta comissiva desleal. Os Tribunais Superiores estão impedidos de reexaminar as questões materiais em face da Súmula 7 do Superior Tribunal de Justiça, que veda o reexame de provas. Deste modo, a supressão desta análise dos Tribunais Superiores não resultará em empecilho à formulação de uma conclusão válida no âmbito e escopo desta dissertação.

3.6 Pré-processamento

Após a coleta dos dados, estruturados em texto no caso em tela, faz-se necessário aplicar uma série de passos ou tratamentos ao conjunto de textos, de maneira a torna-los adequados ao processamento e manipulação por algoritmos de mineração de textos e/ou de aprendizado de máquina (GONÇALVES et. al, 2006) Os sistemas de mineração de textos devem submeter aos seus algoritmos as coleções textuais de modo pré-processado, com uma representação do texto mais estruturada, de modo a obter maior eficiência e menor taxa de erro nas etapas subsequentes (GOMES, 2008).

3.6.1 Modelo de Espaço Vetorial (*Bag-of-words*)

Nesta etapa de Pré-processamento é criado um modelo de representação dos documentos, ou seja, trata-se de transformar os textos de linguagem natural em dados

estruturados e matematicamente computáveis. A literatura de Recuperação de Informações propõe diversos modelos para efetuar esta representação estruturada de documentos textuais. O modelo mais utilizado, entretanto, é o Modelo de Espaço Vetorial (SILVA, 2007).

Uma vez criado este modelo, é feita uma etapa de seleção de características que melhor expressam o conteúdo e significado do texto, tornando-o computacionalmente manejável para os objetivos colimados. No caso em estudo, a seleção de características será feita em consonância com o referencial teórico acerca do instituto do *trade dress*, notadamente em relação ao papel de seus elementos e da prova pericial no resultado dos julgados.

3.6.2 Tokenização (Tokenization)

O processo de tokenização (*tokenization*) é considerado a primeira etapa de Pré-processamento de textos, e tem como finalidade separar o texto em unidades mínimas (*tokens*) que contudo representem a mesma semântica original. Alguns autores preferem o termo **atomização** para fazer referência ao processo de *tokenização* (FINATTO, 2005).

O termo *token*, usado para referir estas unidades mínimas, geralmente corresponde a uma palavra do texto tratado, embora isso nem sempre ocorra; muitas vezes o *token* é, por exemplo, uma série de algarismos ou numerais (e.g. “21/10/2017”, “R\$100,00”) ou um conjunto de palavras que possui sentido único (e.g. “chave-mestra”).

O processo de *tokenização* se baseia, em grande medida, no reconhecimento da separação de palavras por caracteres de controle de arquivo ou de formatação, como espaços ou sinais de pontuação. Estes indicadores são então considerados ***tokens delimitadores*** (FELDMAN; SANGER, 2007). Embora relativamente simples, para seres humanos, pode ser exigido um considerável poder computacional durante a tarefa de tokenização. Isto ocorre devido à multiplicidade de funções que os delimitadores podem exercer: um ponto pode significar o final de uma frase (e.g. “A ação foi arquivada.”) ou então um elemento-chave na formação de um token (e.g. “A ação foi arquivada sob o número 008763.98-2018.”). A despeito de algumas limitações, a tokenização utilizando delimitadores costuma apresentar bons resultados no Pré-processamento (CARRILHO, 2007). Em alguns casos de Mineração de Textos com objetivos que requeiram maior grau de abstração ou menores taxas de erros específicos, pode ser necessário adotar metodologias de *tokenização* mais complexas, como a proposta por Konchady (2006), que utiliza dicionários de dados e regras de formação de

palavras para manter o mesmo nível semântico dos *tokens* antes do processo, e evitar certos erros decorrentes do processamento mais simples.

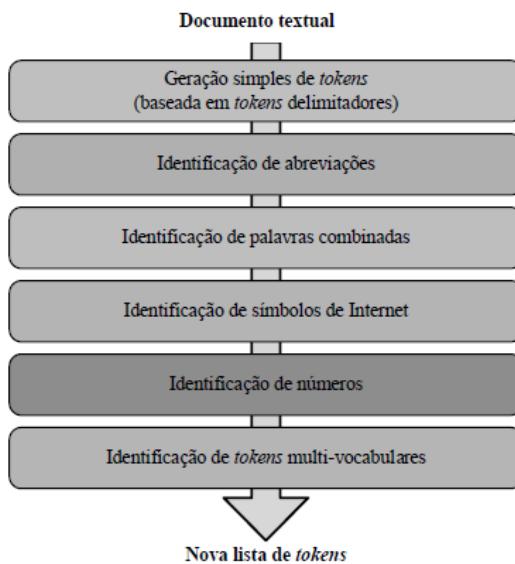


Fig. 1 – Metodologia de identificação de *tokens* proposta por Konchady (2006).

No escopo deste trabalho, utilizaremos o algoritmo de *tokenização* disponível no pacote denominado “*tm*” para linguagem R (FEINERER et al., 2018), instalado no software Rstudio⁷.

3.6.3 Remoção de Stop-words

A remoção de palavras de pouco ou nenhum valor semântico é o passo seguinte no Pré-processamento computacional do *corpus*. Em qualquer documento, existe um número considerável de palavras que não contribuem com valor semântico relevante, sendo úteis tão-somente para a compreensão geral do texto. Após a *tokenização*, estes *tokens* são então eliminados da representação vetorial do texto, em busca de uma redução da dimensionalidade do *corpus*, e consequentemente maior rapidez e otimização na extração do conhecimento ali contido. Estes *tokens* são chamados *stopwords*, e integram uma de *stoplist* de um sistema de Mineração de Textos. A *stoplist* costuma conter elementos como conjunções, preposições, pronomes e artigos, considerados de baixa ou nenhuma relevância semântica para a representação de textos, no modelo de espaço vetorial. Uma *stoplist* adequada permite a eliminação de termos irrelevantes para o processo de Mineração de Textos, e representa em

⁷ A instalação dos pacotes usados é efetuada automaticamente pelo script constante do APENDICE 1.

média uma redução de 50% a 60% do total de *tokens* de um *corpus* (BASTOS, 2006; SILVA, 2007).

A figura XXX ilustra um exemplo de tokenização seguido de remoção de stopwords.

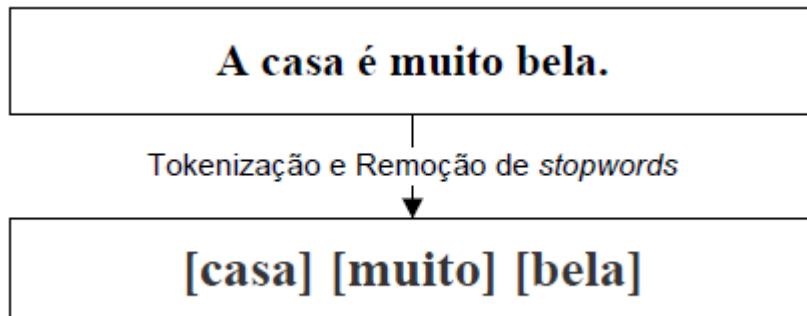


Fig. 2 – Processo de *tokenização* seguido de remoção de *stopwords*

No âmbito deste trabalho, será utilizado o dicionário de *stopwords* de língua portuguesa disponível no pacote “tm” da linguagem R, instalado no RStudio (FEINERER, 2018).

3.6.4 Stemming

Qualquer língua apresenta flexões de palavras nas mais diversas formas. Na língua portuguesa, um substantivo pode apresentar flexões de gênero, número e grau, mantendo, porém, o mesmo valor semântico. O processo de formação de vocábulos é, via de regra, realizado pela derivação de radicais, mas mantendo essencialmente a significação, em sentido amplo (CEGALLA, 2005).

No tratamento semântico de grandes volumes de dados, seria deveras oneroso computacionalmente considerar todas as flexões como vocábulos distintos. Ademais, não há ganho de eficiência no processo de Mineração de Textos com tal proceder. Destarte, aplica-se o processo de *stemming*, que consiste numa redução das palavras ao seu radical original (*stem*, do inglês ‘caule’), eliminando suas flexões ou derivações. Objetiva-se, com este processo, diminuir a dimensionalidade dos textos tratados, considerando como um único token a representação de diversas palavras semanticamente idênticas ou similares.

Algoritmos de *stemming* costumam depender de regras linguísticas, o que os torna mais ou menos eficientes e adequados em relação ao idioma em que são aplicados. Porém, tais algoritmos não visam aplicar as regras linguísticas básicas, mas tão-somente otimizar o

desempenho computacional dos sistemas de Mineração de Textos. Em razão de sua aplicação por vezes divorciada das regras linguísticas escorreitas, dois tipos principais de erros podem resultar desta técnica: *Overstemming e Understemming*.

Ocorre o *Overstemming* quando a remoção de caracteres é excessiva, atingindo não apenas a derivação ou flexão do vocábulo, mas sim o seu radical. Este erro ocasiona o registro de um único radical para vocábulos distintos. Um exemplo deste erro seria a redução da palavra “gramática” para o radical “grama”, quando o correto seria “gramát”.

Já no *Understemming*, o problema é oposto: os caracteres retirados são insuficientes, e o conjunto resultante ainda contém caracteres parte de uma derivação ou flexão. Este erro resulta em radicais distintos para palavras da mesma origem. Um exemplo deste erro seria a redução da palavra “conferência” ao radical “conferênc”, quando o correto seria “confer” (SOARES, 2013).

Para evitar erros de stemming, recomenda-se análise manual dos resultados, ainda que por amostragem, ou focada em termos mais relevantes ao objetivo pretendido.

O algoritmo de stemming mais utilizado na língua portuguesa é o de (PORTER, 1980), embora aplique-se também o ‘Método do Stemmer S’, e o algoritmo de (LOVINS, 1968).

No escopo deste trabalho, não foi utilizado nenhum algoritmo ou método de stemming, em função da baixa complexidade da análise, e do pequeno esforço computacional envolvido com este volume de processamento. Deste modo, não há preocupações com erros de *stemming*; porém, especial cuidado deve ser tomado na construção do *Thesaurus* para garantir a presença de todas as flexões dos termos cuja análise é pretendida.

3.7 Indexação

Uma vez concluído o Pré-processamento, o texto está apto a ser indexado. Antes fracoamente estruturado e com excesso de informação de baixo ou nenhum valor, agora a coleção textual encontra-se bem estruturada e, na medida do possível, livre de informações redundantes ou errôneas. A indexação pode ser comparada ao processo de geração de um sumário de um livro, onde seu conteúdo é elencado através de tópicos ou palavras de maior relevância.

A indexação de documentos data de mais de 4 mil anos, através de técnicas de catalogação manual por índices (BAEZA-YATES; BERTIER, 1999). O processo de indexação resulta em um ganho de tempo e esforço computacional na consulta à informação

armazenada, e a única penalidade é o custo de espaço computacional para armazenamento dos índices. Como *esforço* computacional (ciclos de processamento de um microprocessador) é mais caro e lento do que *espaço* computacional (armazenamento e acesso a memória eletrônica), a troca é obviamente vantajosa. Em muitos casos, a indexação é essencial, pois de outro modo o acesso a grandes quantidades de texto seria tecnicamente e/ou economicamente inviável (MANNING; RAGHAVAN; SCHUTZE, 2007).

A etapa de indexação sofre direta influência do processo de pré-processamento que lhe precede. Caso o *corpus* tenha sido tratado com técnicas de Processamento de Linguagem Natural (PLN), o conteúdo possuirá uma riqueza adicional de informações e características linguísticas e semânticas, que poderá ser utilizado na construção do índice. Ademais, a própria indexação poderá fazer uso de PLN, originando duas abordagens distintas ao processo de indexação: **Indexação Textual** ou **Indexação Temática**.

3.7.1 Indexação textual

A **Indexação Textual** é efetuada pela indexação de termos presentes no documento. É um processo automático e autocontido, que não utiliza informações externas, como um dicionário de palavras, exemplificativamente. Dependendo do algoritmo utilizado na indexação, pode-se realizar consultas com operadores booleanos ou de proximidade. Atualmente a técnica de indexação mais utilizada é a de *índices invertidos* (KONCHADY, 2006; FELDMAN; SANGER, 2007). Um índice invertido é uma estrutura de dados na forma de uma lista ordenada contendo todas as palavras distintas encontradas no *corpus*, e os documentos que o compõem. Recebem essa denominação por inverter a hierarquia da informação; em lugar de uma lista de documentos contendo termos, gera-se uma lista de termos que referenciam documentos. Eventualmente, outras informações como posição e frequência dos termos também podem ser armazenadas. A figura 3 apresenta um exemplo de índice invertido elaborado (SOARES, 2013) com técnicas de PLN, com todas as palavras em caixa baixa e na sua forma singular.

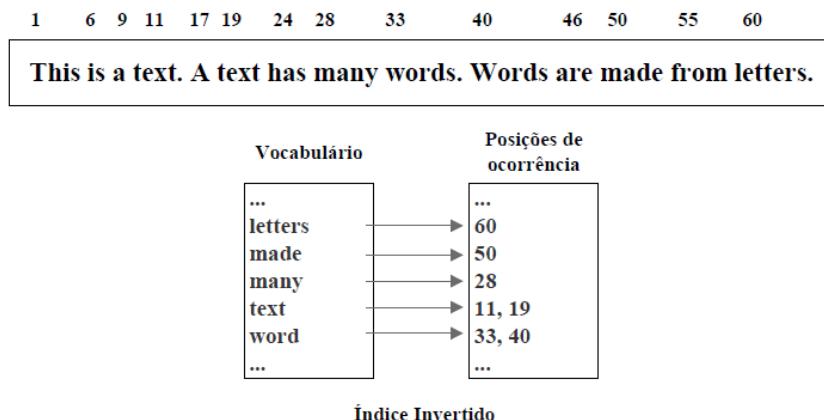


Fig. 3 – Índice invertido elaborado com PLN

3.7.2 Indexação temática

A **Indexação Temática** é efetuada mediante constante consulta a uma fonte externa de informações, que neste caso consiste em um dicionário comumente referido como *Thesaurus*, contendo uma listagem que mapeia em um único termo (denominado termo preferido) um conjunto de sinônimos (denominados termos não preferenciais). Assim, uma palavra é automaticamente associada aos seus sinônimos. Deste procedimento resultam índices mais compactos, além de possibilitar a identificação e recuperação de informações grafadas de formas diversas, porém com mesmo valor semântico. A maior dificuldade desta etapa é a criação de dicionários adequados (SOARES, 2013). A indexação temática está diretamente relacionada ao processo de *lemmatization*, posto que é através deste que torna-se possível agregar palavras do mesmo campo lexical, equivalentes em sentido, no âmbito pretendido.

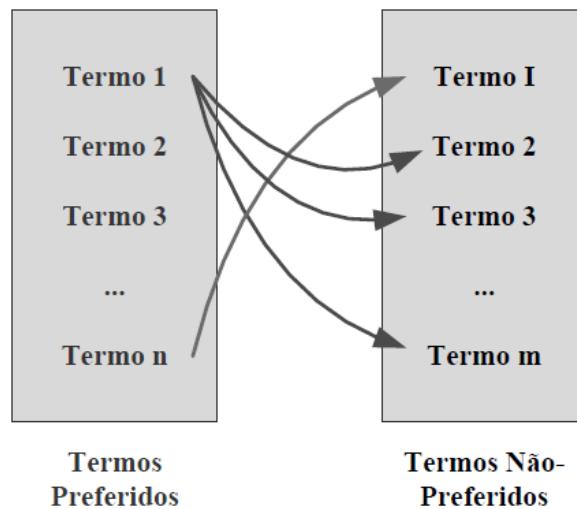


Fig. 4 – Dinâmica de Indexação temática

No âmbito deste trabalho, serão executadas tanto a indexação textual quanto a temática. Na indexação textual, os termos ou palavras (*tokens*) que compõe o *corpus* serão identificados e mapeados; na indexação temática, serão aglutinados em uma única entidade semântica, na qual os *tokens* encontram-se agrupados pelo seu conceito subjacente. A indexação temática permitirá a aplicação de técnicas de estatística inferencial para validar ou refutar as hipóteses objeto de teste, através dos algoritmos referidos no software.

3.8 Teste de Shapiro-Wilk

O teste de Shapiro-Wilk (*Shapiro-Wilk test*) foi publicado originalmente por Samuel Sanford Shapiro e Martin Wilk, em 1965. É um teste de normalidade estatística, ou seja, se propõe a determinar matematicamente se um conjunto amostral deriva de uma distribuição normal. Esta informação é relevante na ciência estatística, pois influencia as inferências que serão possíveis a partir do conjunto amostral, bem como a validade dos demais testes estatísticos de hipótese (SHAPIRO; WILK, 1965).

No caso de populações com distribuição normal, pode-se aplicar o teste t de Welch, que consiste em uma variação do Test t de Student (WELCH, 1947); em caso contrário, o teste de Wilcoxon é a abordagem considerada correta para validação de hipóteses.

$$W = \frac{\left(\sum_{i=1}^n a_i x_{(i)}\right)^2}{\sum_{i=1}^n (x_i - \bar{x})^2},$$

Fig. 5 - Representação matemática do teste de Shapiro-Wilk

3.9 Teste de Wilcoxon

O teste de Wilcoxon (*Wilcoxon signed-rank Test*) foi desenvolvido inicialmente por Frank Wilcoxon, em 1945, como um teste não paramétrico de hipóteses estatísticas, usado para comparar duas amostras relacionadas, casadas, ou repetidas medidas de uma amostra única, com o intuito de determinar se as médias de suas respectivas populações diferem de modo estatisticamente significativamente. A escolha do Teste de Wilcoxon se dá como melhor alternativa quando não é possível assumir uma distribuição normal da amostra em exame. É considerado um dos mais tradicionais e reconhecidos testes de hipótese no ramo da estatística (DEGROOT, 2005).

A distribuição normal pode ser verificada pela aplicação do Teste de Shapiro-Wilk, acima referido. No caso deste estudo, as amostras objeto de teste não apresentaram distribuição normal, o que indica a aplicação do Teste de Wilcoxon para avaliar as hipóteses nulas descritas nos resultados, acerca de influência dos conceitos de “cor”, “marca” e “perícia” no resultado das lides amostrais. (WILCOXON, 1945).

3.10 Welch's t-Test

O teste de Welch (Welch's t-test) é utilizado em estatística para determinar se duas populações, ou dois conjuntos amostrais tem médias similares. Consiste em uma adaptação do teste t de Student (Student's t-test). É muitas vezes referido como “unpaired” ou “independent samples” test. O teste t de Student assume que ambas as populações ou amostras possuem distribuição normal com igual variância, enquanto o teste de Welch mantém a presunção de distribuição normal, porém adaptado para variância desiguais.

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}}$$

Fig. 6 – Formula matemática aplicável ao test t de Welch

No âmbito deste trabalho, será o teste mais adequado para determinar a validade da hipótese nula de relação entre frequência relativa de “Marca” e resultado final da lide (GRAYBILL et al., 1998).

3.11 Teste de Chi-Square de Pearson

O teste de Pearson chi-squared consiste em um teste de validação de hipótese estatística, usado para avaliar se dois conjuntos amostrais de variáveis categóricas possuem divergências aleatórias ou dependentes. É aplicado precipuamente em dados não pareados de amostras volumosas.

Existem diversos testes de ‘chi-squared’ (Yates, likelihood ratio, portmanteau time series, etc.) em estatística, porém o teste de Pearson é considerado o mais amplamente utilizado. Resulta do trabalho de Karl Pearson no início do século XX (PEARSON, 1900).

3.12 Análise manual dos julgados

Inicialmente, após a coleta dos documentos relevantes conforme a metodologia acima descrita, procedeu-se à análise inicial do perfil dos mesmos. Esta análise, embora bastante trabalhosa, necessita de processamento humano, tendo em vista que não existem ferramentas computacionais disponíveis para obter estes dados de forma automatizada, com confiabilidade. Nesta análise exordial, buscou-se determinar o seguinte:

- (i) número total de decisões coletadas;
- (i) número total de decisões coletadas por ano;
- (iii) número total de decisões coletadas por tribunal;
- (iv) resultado de primeira instância para cada uma das decisões, segundo a taxonomia de ‘êxito’, e ‘sucumbência’;

(v) resultado final de cada ação, classificado segundo a taxonomia de ‘êxito’, e ‘sucumbência’.

(vi) decisão unânime de segunda instância, classificada em ‘sim’ ou ‘não’;

(vii) existência de perícia ou não nas lides examinadas, segundo a taxonomia de ‘sim’ ou ‘não’

Os resultados foram tabulados em uma planilha em .XLS (*excel spreadsheets*), que posteriormente foi introduzida no script desenvolvido na forma de uma matriz de dados (*data-matrix*), apta a receber o tratamento computacional.

3.13 Análise automatizada dos julgados

A análise manual dos julgados foi sucedida pelo tratamento automatizado dos textos que compõem a jurisprudência selecionada, conforme metodologia acima descrita.

Para tanto, foi desenvolvido um software (script) na linguagem R, que pode ser executado com uso da plataforma *open-source* RSTUDIO⁸.

Utilizando este software, procedeu-se à aplicação das ferramentas de Mineração de Textos detalhadas, da seguinte forma:

3.13.1 Importação dos arquivos em PDF e transformação em formato TXT

Considerando o formato dos julgados sempre disponíveis em PDF, foi necessário efetuar a transformação para o formato TXT (texto puro, sem formatação), que pode então ser processado pelas bibliotecas disponíveis. Para tanto, foi utilizada a biblioteca “pdftools”, para a linguagem R (OOMS, 2018). O script de transformação encontra-se descrito no APENDICE 1.

Esta conversão não é isenta de problemas, muitos deles inerentes ao próprio formato PDF. Para examinar os erros possíveis e buscar mitigações, foi feita a seguir uma inspeção visual nos resultados, através da revisão dos arquivos TXT gerados e armazenados no diretório de trabalho.

Nesta etapa, não foi identificado nenhum erro ou falha na conversão, restando perfeitos e íntegros os textos originais, sem qualquer alteração gramatical ou sintática. Em um dos documentos, foi identificada uma alteração de codificação de caracteres, que foi corrigida

⁸ Disponível em www.rstudio.com.

na fase de pré-processamento com eliminação dos caracteres errôneos. Também não foi detectado nenhum erro relacionado à falta de páginas ou trechos de decisões.

Uma nova inspeção visual é necessária na etapa de conversão dos arquivos TXT para o formato de um *corpus* e posteriormente de um *data frame ou matriz*.

3.13.2 Inspeção visual

A inspeção visual nos resultados, embora trabalhosa, é absolutamente necessária para garantir a mínima correção e utilidade das informações resultantes. Essa inspeção é efetuada novamente após a conversão dos arquivos TXT para o formato de um *corpus* dentro do ambiente do Rstudio. Nesta etapa, inúmeros problemas de conversão foram detectados.

Como exemplos, podemos citar a inserção de “\r\n” como forma de codificar mudanças de linha e parágrafo no texto resultante do PDF. O texto resultante da transformação, sem tratamento, restaria assim grafado⁹:

PODER	JUDICIÁRIO\r\n
TRIBUNAL DE JUSTIÇA DO ESTADO DE SÃO PAULO\r\n	
Este documento é cópia do original que recebeu as seguintes assinaturas digitais:\r\nPg. inicial Pg. final Categóri	
Nome do assinante Confirmação\r\n11 17 Acórdãos	
RICARDO JOSE NEGRAO NOGUEIRA 47A7F5D\r\n	
Eletrônicos\r\n18 26 Declarações de CARLOS ALBERTO	
GARBI 499B435\r\n	Votos\r\nPara conferir o
original	acesse o
	site:\r\nhttps://esaj.tjsp.jus.br/pastadigital/sg/abrirConferenciaDocumento.do, informando o processo\r\n1010406-06.2013.8.26.0100 e o código de confirmação da tabela acima.\r\n

Outros erros também foram encontrados na inspeção visual, como excesso de espaço entre palavras, caracteres esparsos, e caracteres irrelevantes à análise almejada (tais como sinal de ‘arroba’, ‘cifrão’, etc.). Questões relativas à codificação de caracteres também devem ser detectadas e corrigidas nesta etapa.

3.13.3 Remoção de Stopwords e limpeza do texto

⁹ Texto exemplificativo obtido de trecho de acórdão do processo CNJ Nº 1010406-06.2013.8.26.0100, de relatoria do Desembargador Carlos Garbi, TJ/SP.

Com base nestas observações, foi elaborado uma rotina simples implementada no código-fonte do software (APENDICE 1), capaz de ‘limpar’ o resultado de inserções indevidas, equivocadas ou inúteis, não apenas os códigos “\r\n” mas também outros caracteres (inclusive espaços) excessivos, inúteis ou errôneos. Foram também removidas as chamadas ‘stopwords’, conforme detalhado acima.

3.13.4 Tokenização (Atomização)

Após estas etapas iniciais, foi efetuada a tokenização (ou atomização) do texto, conforme metodologia acima, implementada no código-fonte do Software (APENDICE 1).

Após a tokenização, o *corpus* está pronto para tratamentos seguintes, uma vez concluído seu processo de criação do espaço vetorial.

3.13.5 Inspeção visual de tokens

Após o pré-processamento do *corpus*, foi realizada uma nova inspeção visual nos tokens (ou *features*) identificados, conforme tratados pelo pacote “tm”, da linguagem R.

Esta segunda inspeção permite identificar erros ou falhas na codificação de caracteres, delimitação de tokens, e outros problemas que podem originar da transformação e tratamento do *corpus*.

3.13.6 Criação de wordcloud específico aos termos relevantes

Uma vez selecionados os termos mais significativos em relação ao conteúdo semântico, foi possível criar uma nuvem de palavras (wordcloud) contendo apenas os termos e suas respectivas frequencias representadas na forma de ‘nuvem’. O tamanho do termo e sua proximidade com o centro indicam maior frequencia. Esta representação gráfica permite um exame muito facilitado e intuitivo da frequencia e importância dos termos no *corpus* examinado.



Fig. 7 – Wordcloud representativo do *Corpus* após pré-processamento de termos

3.13.7 Criação do *Thesaurus*

A partir da seleção dos termos relevantes, é possível asseverar que alguns termos, ainda que não necessariamente compõem o mesmo campo lexical, correspondem a um campo semântico comum, e podem portanto ser unificados para o processo de *indexação temática* descrito na seção 3.7.2 supra. Exemplificativamente, o termo ‘cor’ pode ser considerado sinônimo (ou *lemma*) de ‘tonalidade’, ‘coloração’ ou ‘tom’. Deste modo, ao se buscar pela importância do elemento ‘cor’ nos julgados de *trade dress*, é salutar, e mesmo necessário, que se aglutine os demais elementos que podem ser reduzidos ao *lemma* de ‘cor’, produzindo assim um entendimento mais fidedigno da importância da cor, como elemento do *trade dress*, nos julgados, mesmo quando a palavra ‘cor’ não é utilizada em favor de um sinônimo.

A construção do *Thesaurus* pode ser uma das etapas mais complexas e onerosas do processamento de linguagem natural, pois envolve necessariamente uma avaliação manual dos termos, e um juízo acerca dos objetivos almejados e da metodologia aplicada.

Em alguns casos, pode-se agrupar todos os sinônimos; em outros casos, a presença de sinônimos deve ser preservada para fins de análise. Além disso, um simples dicionário de sinônimos muitas vezes não é capaz de conter todas as expressões possíveis dentro de um contexto estudado, ou pode omitir palavras compostas, expressões gramaticais, orações e outros elementos sintáticos que possuem algum nível de correspondência ou relevância semântica com o escopo do estudo.

Há ainda uma complexidade adicional quando se trabalha com traduções, pois muitas vezes os sinônimos escolhidos pelo tradutor correspondem a significações minimamente diferentes, porém de impacto ao resultado pretendido (SOARES, 2013).

No âmbito deste estudo, buscou-se apenas um *Thesaurus* simples, com agrupamento de sinônimos dos termos identificados na inspeção visual, e acréscimo de outros de acordo com o dicionário de sinônimos disponível em www.sinonimos.com.br.

A lista completa de sinônimos usada para a elaboração do *Thesaurus* foi implementada no código fonte conforme comentário presente no código. Para COR, foram relacionadas as palavras "cor", "cores", "matiz", "matizes", "tom", "tons", "tonalidade", "tonalidades", "colorido", "coloridos", "colorida", "coloridas", "pigmentação", "tingimento". Para MARCA, foram selecionados os seguintes sinônimos: "marca", "marcas", "signo", "signos"¹⁰.

A presença ou não de prova pericial foi tabulada por análise manual, com valores categóricos “SIM” ou “NÃO”, conforme presente ou não a perícia técnica em cada lide, não se aplicando ao caso a construção de um *Thesaurus*.

3.13.8 Indexação textual e temática conforme elementos selecionados

Após a conclusão destas etapas de pré-processamento, foi feita uma indexação temática derivada da indexação textual, para cada documento, levando em consideração a frequência dos termos agrupados conforme seus *lemmas* no *Thesaurus*, assim criando uma matriz de dados (*data matrix*) na qual cada linha corresponde a um documento (no caso, a uma julgado de segunda instância relativo ao um processo), e cada coluna corresponde a um fator (ou índice) relativo àquele documento. Esta metodologia de cálculo de frequencia relativa é amplamente utilizada há décadas, e fartamente documentada na literatura (MÄNTYLÄ et. al., 2018; GAHG; SHAH, 2014).

3.14 Aplicação dos testes de normalidade e de validação de hipótese

Após o tratamento que permite a análise de linguagem natural do *corpus*, foi possível o cálculo das frequencias relativas dos termos do *Thesaurus*, ou das ocorrências de ‘perícia’, e o teste de sua relação estatística significante através da aplicação dos testes de Shapiro-Wilk

¹⁰ Fonte: <https://www.sinonimos.com.br/cor/> e <https://www.sinonimos.com.br/marca/>. Acessado em 15/01/2019.

(teste de normalidade), teste de Wilcoxon, teste Chi-Squared de Pearson, ou Welch's t-test (testes de hipótese), conforme descritos neste capítulo. Os resultados serão discutidos no Capítulo 4.

3.15 Limitações metodológicas

Um problema metodológico relevante à busca refere-se à existência de dois graus de jurisdição em sede de justiça comum: a primeira e a segunda instância. Via de regra, as lides iniciam-se na primeira instância e terminam após o trânsito em julgado da decisão final de segunda instância, porém nem todos os processos seguem até a segunda instância; muitos cessam sua vida ainda na primeira instância, por homologação de acordo entre as Partes, ou por sentença não recorrida, seja ela formadora de coisa julgada formal ou material (THEODORO, 2016). Assim, uma busca pelo intervalo temporal nos julgados de segunda instância necessariamente vai omitir os resultados de primeira instância que não tivessem sido objeto de recurso. Entretanto, como o objetivo deste estudo é uma análise exploratória do entendimento dos Tribunais, pode-se prescindir do exame de tais decisões de primeira instância, que via de regra não representam jurisprudência relevante.

Adicionalmente, outras possíveis limitações metodológicas serão discutidas no capítulo de conclusões.

4 ANÁLISE JURIMÉTRICA DA JURISPRUDÊNCIA

4.1 Número total de ações julgadas

4.1.1 Resultados do Tribunal de Justiça de São Paulo

Inicialmente, pretende-se determinar em números absolutos a incidência de ações de *Trade dress* julgadas nos tribunais em análise, com decisão de segunda instância publicada durante o triênio 2015 a 2017.

No triênio em análise, houve um total de **sessenta e nove** julgados, distribuídos da seguinte forma:

- a. No ano de 2015, 15 (quinze) julgados;
- b. No ano de 2015, 27 (vinte e sete) julgados;
- c. No ano de 2015, 27 (vinte e sete) julgados;

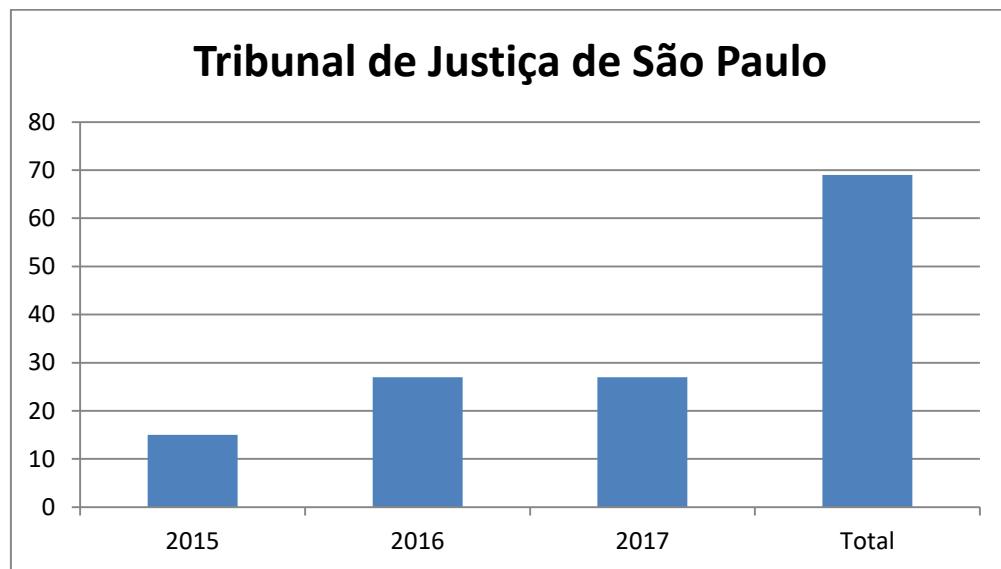


Fig. 8 – Resultados quantitativos do TJSP

4.1.2 Resultados do Tribunal de Justiça do Rio de Janeiro

Em sequência à análise quantitativa genérica do Tribunal de Justiça de São Paulo, o mesmo procedimento será aplicado aos julgados do Tribunal de Justiça do Rio de Janeiro.

No triênio em análise, houve um total de **vinte e dois** julgados, distribuídos da seguinte forma:

- a. No ano de 2015, 9 (nove) julgados;
- b. No ano de 2015, 5 (cinco) julgados;
- c. No ano de 2015, 8 (oito) julgados;

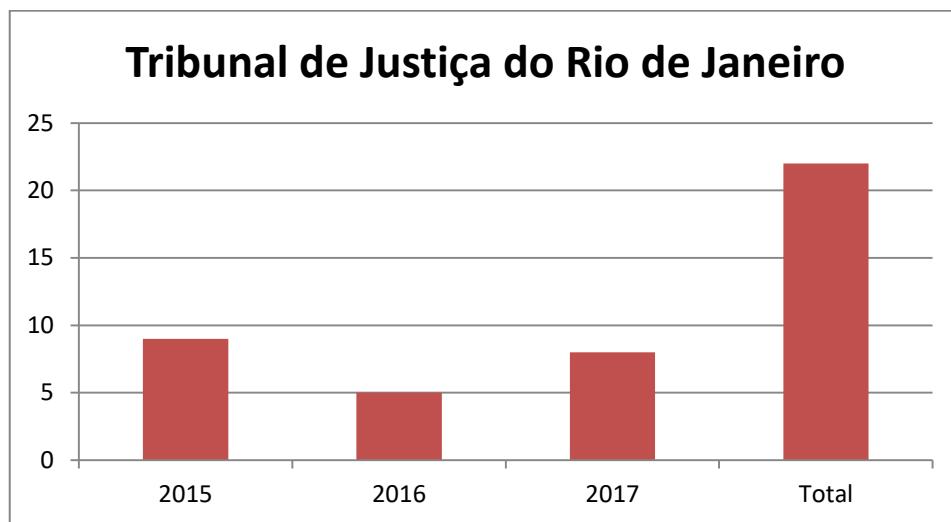


Fig. 9 – Resultados quantitativos do TJRJ

4.1.3 Resultados Comparativos

Uma análise comparativa dos numeros absolutos de julgados de cada tribunal revela a enorme discrepância entre o volume de julgados do Tribunal de São Paulo em relação ao Tribunal do Rio de Janeiro. O total de ações julgadas no triênio em São Paulo foi mais do que o triplo daquelas julgadas no Tribunal fluminense. Estes números corroboram a posição de dominância do Estado de São Paulo em relação à atividade econômica, e demonstram que o referido Tribunal teve um corpo de jurisprudência muito mais volumoso no período, acerca do *Trade dress*.

Também se pode perceber que o Tribunal paulista teve um aumento expressivo no número de julgados a partir do ano de 2015, aumento este que se manteve até 2017, último ano de análise. Já o Tribunal fluminense teve um decréscimo no número de julgados, indicando uma tendência de diminuição das ações de *trade dress* ao longo da série histórica examinada.

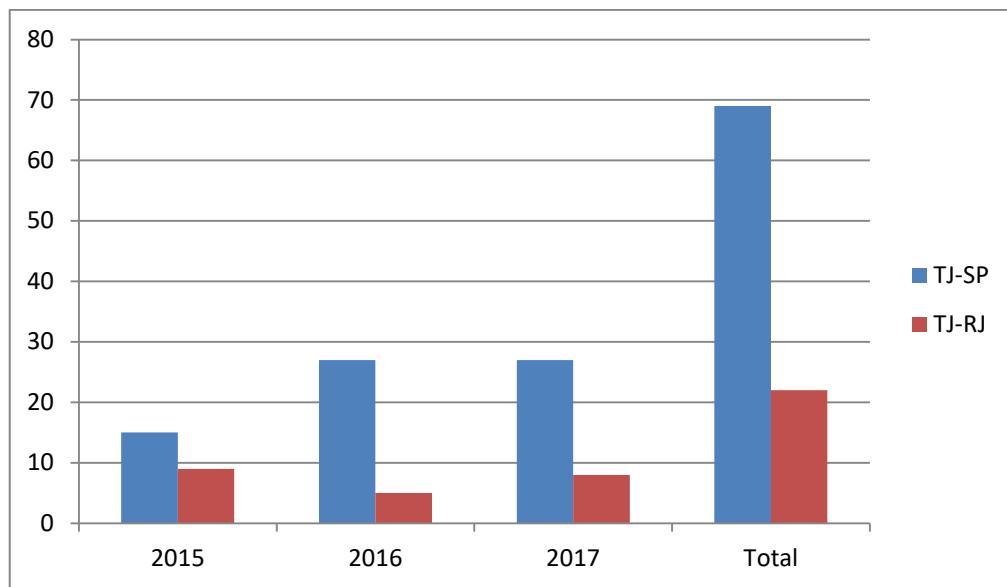


Fig. 10 – Resultados comparativos entre os tribunais

4.2 Perfil de decisões

Após estabelecer os números absolutos de julgados, e efetuar a comparação entre os tribunais, proceder-se-á a uma avaliação quantitativa do resultado das lides em análise, tanto em primeira instância como em segunda instância.

Esta análise foi feita por inspeção visual de cada decisão e de seu conteúdo respectivo, buscando-se determinar: (i) o resultado da lide em primeira instância; (ii) o resultado da lide em segunda instância; e (iii) o número de decisões de primeira instância reformadas em segunda instância.

Nesta etapa, as lides a cujas decisões se referem serão classificadas como **ÊXITO** ou **SUCUMBÊNCIA**. Para fins de avaliação dos resultados de primeira instância, foram consideradas exitosas as lides cujo pedido principal de reconhecimento de violação de *Trade dress* foi acolhido pela sentença do juiz de primeiro grau. Igualmente, serão consideradas como ‘exitosas’ as ações cujo pedido principal de reconhecimento de violação de *Trade dress* foi acolhido ao final da lide, em grau de recurso, seja com manutenção ou reforma da sentença de primeiro grau.

Decisões de êxito parcial foram consideradas como êxito, sempre que reconhecerem a existência de violação de *Trade dress*, ou seja, o pedido principal da lide. Não será feita qualquer análise ou levantamento em relação a valores de condenação, honorários de sucumbência ou outras questões processuais. A classificação como **ÊXITO** depende tão-

somente do reconhecimento judicial da existência de uma violação de *trade dress* na relação jurídica litigiosa.

4.2.1 Resultados TJ/SP

Após análise dos resultados das lides, constata-se que o Tribunal paulista tem reconhecido a violação de *trade dress* em sessenta e seis por cento das lides julgadas no período.

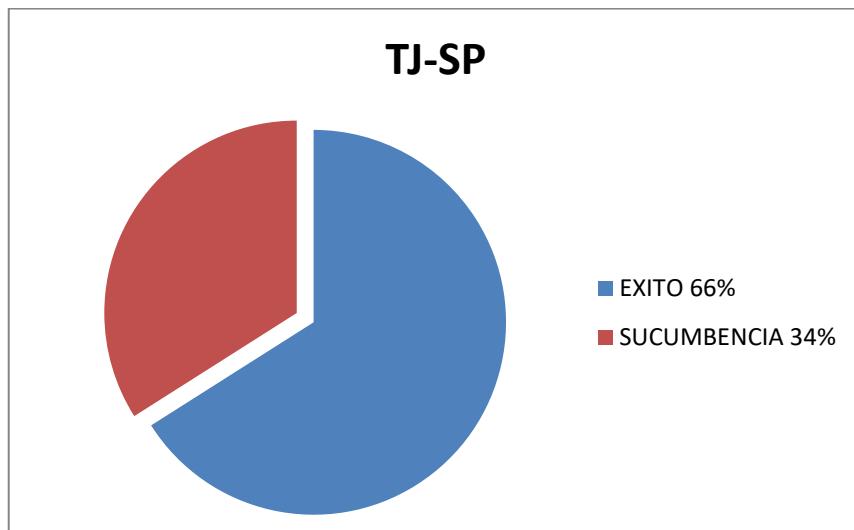


Fig. 11 – Perfil dos julgados do TJSP

4.2.2 Resultados TJ/RJ

Após análise dos resultados das lides, observa-se que o Tribunal fluminense tem reconhecido a violação de *trade dress* em oitenta e oito por cento das lides julgadas no período.

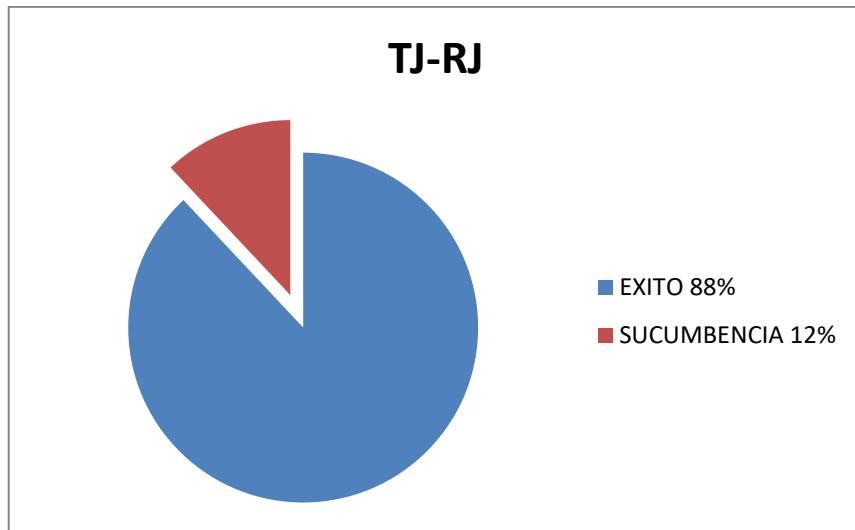


Fig. 12 – Perfil dos julgados do TJRJ

4.2.3 Resultado Comparativo

A análise dos resultados das lides permite afirmar que a probabilidade de obter êxito em uma ação de violação de *trade dress*, em um caso hipotético, é bem mais alta no Tribunal do Rio de Janeiro. Havendo possibilidade de escolher o foro de litígio, os autores aumentariam muito suas chances de êxito escolhendo litigar no Tribunal do Rio de Janeiro.

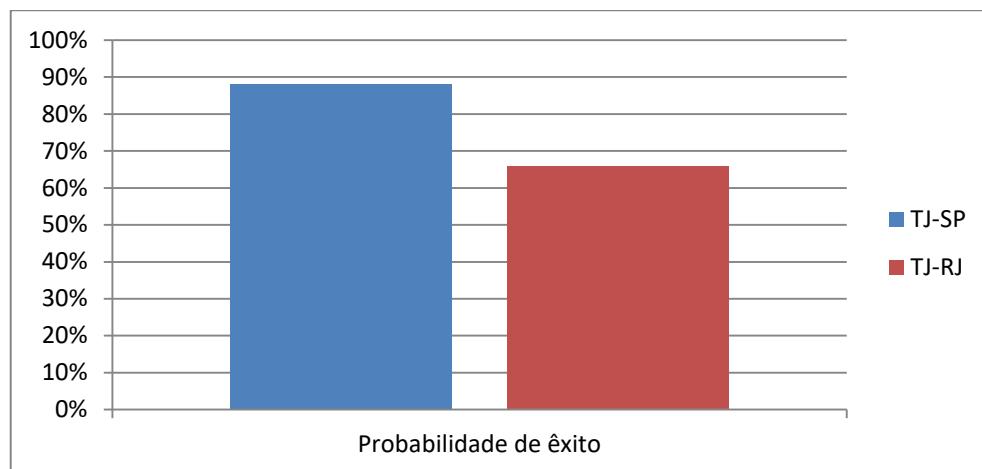


Fig. 13 – Probabilidade de êxito comparada entre Tribunais

Do mesmo modo, é possível afirmar que a probabilidade de êxito, ainda que parcial, é consideravelmente maior do que a probabilidade de sucumbência, em ambos os tribunais estudados, no interregno temporal em análise. Isto significa que o autor de uma ação fundada

em violação de *trade dress* tem mais chances de êxito do que o réu, no caso abstrato, de acordo com este subconjunto amostral.

4.3 Perfil das decisões de primeira instância

4.3.1 Objetivo

O objetivo desta análise é determinar se as decisões de primeira instância são significativamente relacionadas, do ponto de vista estatístico, com o resultado final da lide. Para este fim, foram tabuladas as decisões de primeira instância, seguindo a convenção de E para EXITO, S para SUCUMBENCIA, e considerando-se Exito Parcial como EXITO. A decisão final da lide, como já descrito, foi tabulada como 1 (um) para EXITO e 0 (zero) para SUCUMBENCIA. As decisões de ambos os tribunais foram agregadas, devido ao baixo número de observações (amostras) existentes no TJ-RJ, de modo a conferir maior precisão e validade estatística ao tratamento.

4.3.2 Resultados agregados

A análise de resultados de primeira instância segregadas por tipo de decisão final retornou o seguinte gráfico, onde 0 (zero) são as decisões finais de EXITO e 1 (um) são as decisões finais de sucumbência.

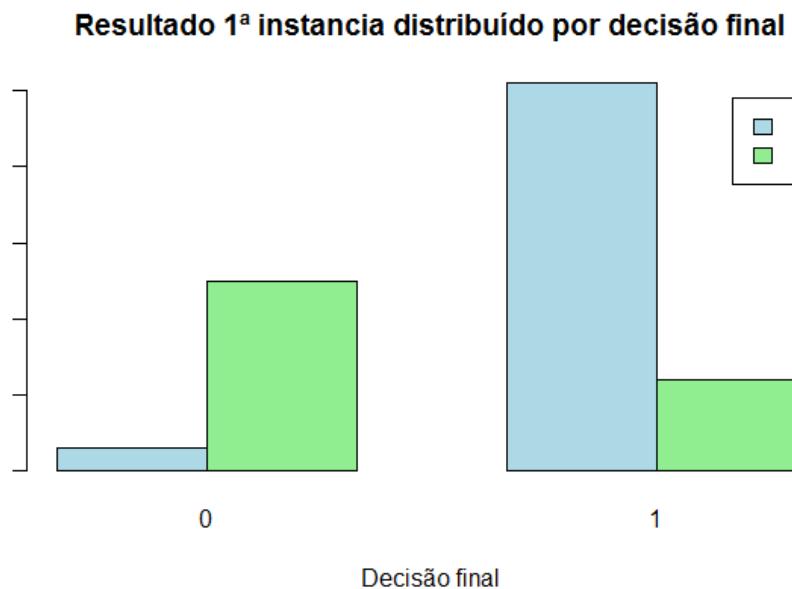


Fig. 14 – Resultados de primeira instância agrupados por decisão final

A análise acima nos permite inferir que os resultados desfavoráveis de primeira instância encontram-se muito mais presentes nas ações cuja resultado final foi a sucumbência do Autor (Decisão ZERO). A variável “resultado de primeira instância” está estatisticamente relacionada com a solução final da lide.

Em outras palavras, nos casos em que os resultados de primeira instância foram desfavoráveis ao Autor, sua chances de obter resultado final favorável (i.e., a reforma da decisão de primeiro grau) são significantemente menores. De modo inverso, nos casos em que o Autor obteve êxito em primeira instância, sua chances de manter a decisão favorável em grau recursal (i.e., manutenção da decisão de primeiro grau) são significantivamente maiores.

4.4 Análise quantitativa e qualitativa do elemento “cor” nos julgados de *trade dress*

4.4.1 Objetivo

O objetivo da análise desta etapa é determinar, primeiramente, a frequência relativa da palavra “cor” (e seus sinônimos e flexões), determinando o quanto frequentemente ocorrem, e assim indicando a importância relativa do conceito de “cor” em cada julgado analisados. Esta abordagem de cálculo de frequência relativa é amplamente utilizada em técnicas de

processamento de linguagem natural, especialmente no modelo de espaço vetorial, em aplicações de análise de sentimento (GHAG; SHAH, 2014; MÄNTYLÄ et al., 2018).

Em seguida, através da aplicação dos testes de normalidade (Shapiro-Wilk Test) e do Teste de hipótese não paramétrico (Wilcoxon Test), determinar se existe uma relação estatisticamente significante entre a frequência do conceito de “cor” no julgado, e o êxito final na lide.

Caso tal relação exista, seria indicativa de que julgados com forte fundamentação em violações de *Trade dress* baseadas ou associadas fortemente com “cor” possuem maior ou menor probabilidade de êxito.

A hipótese nula (H_0) do teste de Wilcoxon é de que as variáveis são independentes; equivale dizer, que a frequencia relativa de COR não tem relação estatística significativa com o resultado final da lide, enquanto a hipótese alternativa (H_1) seria da relação significativa de “cor” com êxito final na lide.

Para determinar se a hipótese nula é válida ou não, calcula-se um valor denominado valor-p (d inglês, *p-value*). Este valor indica se as variáveis comparadas são independentes (ou seja, se não têm relação estatisticamente significante entre si), ou se são dependentes (equivale dizer, possuem relação estatisticamente significante entre si). O nível de significância é tradicionalmente fixado em 0.05 (cinco por cento). Deste modo, se o valor-p na comparação entre duas variáveis é maior do que o valor de significância 0.05 ($p\text{-value} > 0.05$), pode-se aceitar a hipótese nula como válida, e concluir que as variáveis estudadas são independentes.

A escolha de qual teste de hipótese (também chamado de teste de significância ou teste estatístico) aplicar vai depender da distribuição das variáveis do conjunto amostral: caso apresentem distribuição normal, aplica-se uma classe de testes – caso não presentem distribuição normal, aplica-se outra classe de testes.

Para elaboração deste teste, não foi possível criar subconjuntos referentes aos julgados do TJ-RJ e do TJ-SP, devido ao baixo número de amostras do TJ-RJ, que poderia causar distorções significativas. Optou-se, portanto, nesta etapa, pela análise dos resultados agregados de ambos os tribunais, considerados como uma única entidade.

4.4.2 Resultados agregados

Conforme detalhado no capítulo de Metodologia, aplicou-se nesta etapa dois testes sequenciais. Primeiramente, aplicou-se o Teste de Shapiro-Wilk para verificar a normalidade, implementado com a seguinte linha de código, devidamente comentada no APENDICE 1:

```
#summary
group_by(color_decision, decision) %>%
summarise(
  count = n(),
  mean = mean(color_relfreq, na.rm = TRUE),
  sd = sd(color_relfreq, na.rm = TRUE))
shapiro.test(color_decision$color_relfreq)
```

Resultando em:

```
Shapiro-Wilk normality test
data: color_decision$color_relfreq
W = 0.77517, p-value = 1.764e-10
```

Observe-se que o *p-value* é bastante inferior ao valor de significância (*valor de significância* é 0.05), o que informa **a rejeição da hipótese nula de distribuição normal**. Isto está em consonância com a análise visual do histograma de distribuição de frequência relativa para COR:

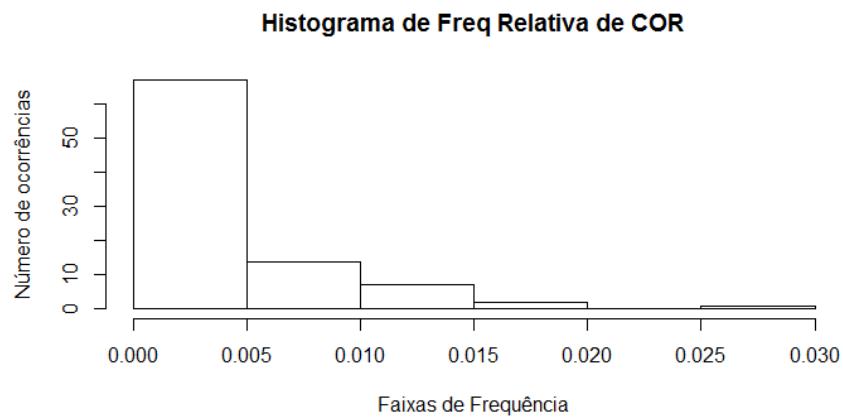


Fig. 15 – Histograma de Frequência Relativa de COR no *corpus*

Uma vez asseverado que a amostra não apresenta distribuição normal, deve-se optar pelo teste de hipótese de Wilcoxon, implementado no código da seguinte forma:

```
#Hipótese nula = as variáveis são independentes e não relacionadas
x=color_decision$color_relfreq[color_decision$decision==1]
y=color_decision$color_relfreq[color_decision$decision==0]
```

```
res <- wilcox.test(x,y,alternative="l")
```

Resultando em:

```
Wilcoxon rank sum test with continuity correction
data: x and y
W = 755, p-value = 0.1376
```

A partir do resultado, pode-se aferir que valor *p-value* de significância ficou bastante acima do valor mínimo (*p-value* > 0.05), indicando que a hipótese nula não pode ser refutada; em outras palavras, **não existe relação estatisticamente significante entre a frequência relativa do conceito COR nos julgados analisados e o resultado final da lide.**

Pode-se formular esta asserção, de modo alternativo, da seguinte forma: as médias das frequências relativas do conceito de COR nos julgados do tipo 1 (um) não diferem significativamente das médias das frequências relativas nos julgados de tipo 0 (zero).

Ao plotar as frequências relativas no gráfico abaixo pode-se averiguar que as linhas de frequência média, de fato, apresentam-se bastante similares.

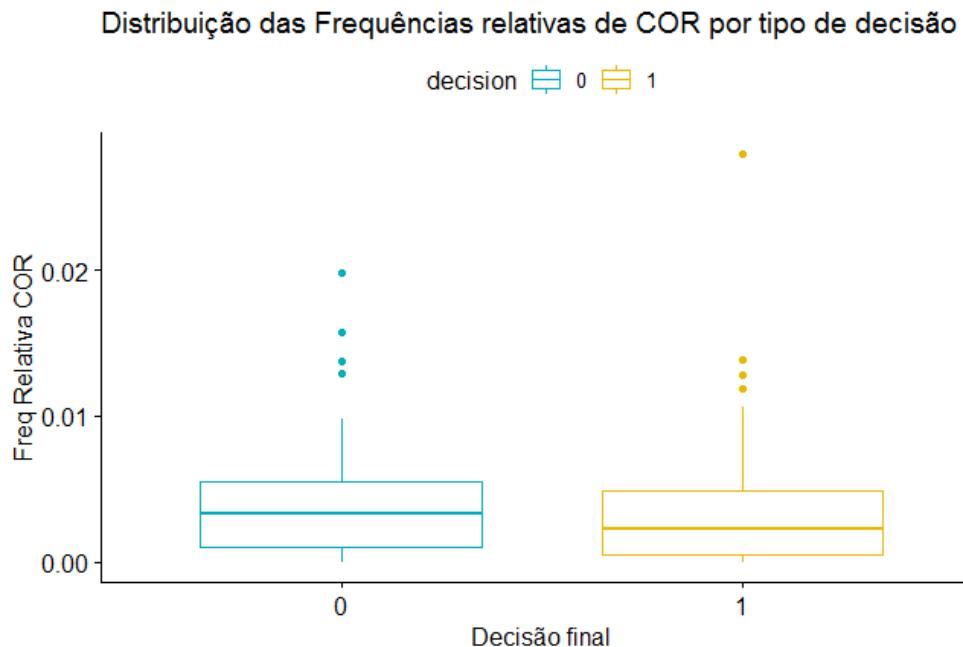


Fig. 16 – Distribuição de frequências de COR por tipo de decisão final

4.5 Análise quantitativa e qualitativa do elemento “marca” nos julgados de *Trade dress*

4.5.1 Objetivo

O objetivo da análise desta etapa é determinar, primeiramente, a frequência relativa da palavra “marca” (e seus sinônimos e flexões), determinando o quanto frequentemente ocorrem, e assim indicando a importância relativa do conceito de “marca” em cada julgado analisado (GHAG; SHAH, 2014). Em seguida, através da aplicação dos testes de normalidade (Shapiro-Wilk Test) e do Teste de hipótese não paramétrico (Wilcoxon Test), determinar se existe uma relação estatisticamente significante entre a frequência do conceito de “marca” no julgado, e o êxito final na lide.

Caso tal relação exista, seria indicativa de que julgados com forte fundamentação em violações de *Trade dress* baseadas ou associadas fortemente com “marca” possuem maior ou menor probabilidade de êxito.

Para elaboração deste teste, não foi possível criar subconjuntos referentes aos julgados do TJ-RJ e do TJ-SP, devido ao baixo número de amostras do TJ-RJ, que poderia causar distorções significativas. Optou-se, portanto, nesta etapa, pela análise dos resultados agregados de ambos os tribunais, considerados como uma única entidade.

4.5.2 Resultados agregados

Primeiramente, aplicou-se o teste de Shapiro-Wilk para determinar a normalidade. Porém, ao contrário da frequência relativa de COR, mencionada e testada acima, o teste indicou que a distribuição da frequência relativa de MARCA segue uma distribuição normal. Em face deste resultado, não se pode aplicar o teste de hipótese de Wilcoxon, e deve-se aplicar o ‘teste *t de Student*’, que é o correto para teste de hipóteses em amostras normais (GRAYBILL et al., 1998).

O teste de Shapiro-Wilk foi implementado pelas seguintes linhas de código, devidamente comentadas no APENDICE 1:

```
#marca
#Compute unpaired two-samples Wilcoxon test
x=color_decision$marca_relfreq[color_decision$decision==1]
y=color_decision$marca_relfreq[color_decision$decision==0]
res <- wilcox.test(x,y,alternative="g")
```

Resultando em:

```
shapiro.test(color_decision$marca_relfreq)
Shapiro-Wilk normality test
data: color_decision$marca_relfreq
W = 0.97441, p-value = 0.06974
```

Como se pode verificar, p-value é maior do que o valor de significância (p-value > 0.05), o que indica que a hipótese nula não pode ser refutada. Equivale dizer, a distribuição de frequência relativa de MARCA nos julgados segue uma curva normal¹¹, o que impede a aplicação do teste de Wilcoxon e demanda a aplicação do ‘teste t de Student’.

Com efeito, a plotagem do histograma de distribuição da frequência relativa de MARCA sugere a distribuição normal, por simples análise visual:

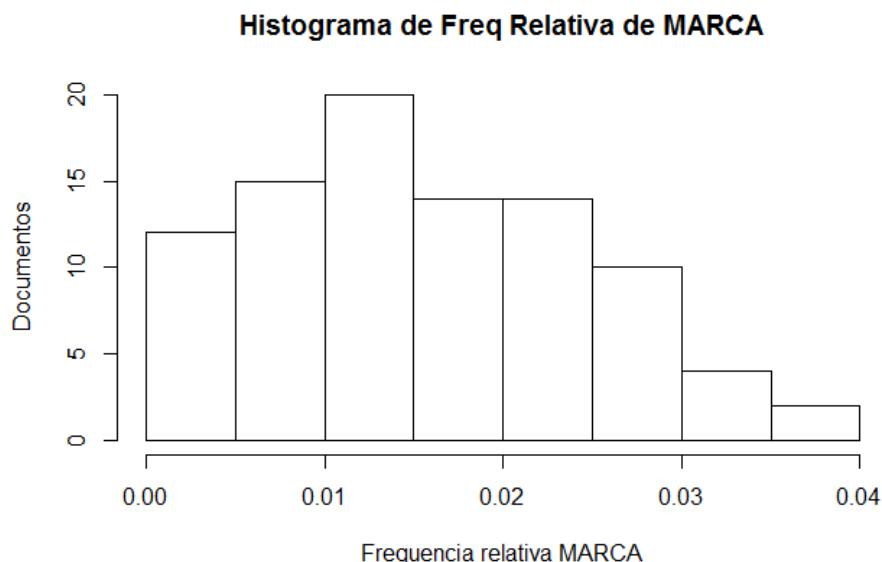


Fig. 17 – Histograma da Frequênci a relativa de MARCA

A partir desta constatação, aplicou-se o teste ‘Welch’s t-test’, implementado através das seguintes linhas de código devidamente comentadas no APENDICE 1:

```
t.test(marca_relfreq ~ decision, data = color_decision)
```

Resultando em:

```
Welch Two Sample t-test
data: marca_relfreq by decision
t = 0.86626, df = 44.099, p-value = 0.391
```

¹¹ Curva normal também é referida como distribuição Gaussiana, ou Curva Bell, pela sua semelhança com a forma de um sino (HECKMAN, 1995).

```

alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-0.002519426 0.006318374
sample estimates:
mean in group 0 mean in group 1
0.01697692 0.01507744

```

Novamente, deve-se observar o *p-value* resultando do teste, que neste caso também foi bastante superior ao valor de significância (*p-value* > 0.05). Disto resulta que a hipótese nula (H_0) não pode ser refutada, e não há relevância estatística entre a frequencia relativa de MARCA e o êxito final nas lides, de acordo com a população estudada.

4.6 Análise quantitativa e qualitativa do elemento “perícia” nos julgados de *trade dress*

4.6.1 Objetivo

O objetivo da análise desta etapa é determinar, primeiramente, a prevalência da perícia nos julgados analisados, determinando o quanto frequentemente ocorrem. Em seguida, através da aplicação dos testes de normalidade (Shapiro-Wilk Test) e do Teste de Pearson Chi-Squared (*Pearson's Chi-Squared Test*), determinar se existe uma relação estatisticamente significante entre a existência de prova pericial no julgado e o êxito final na lide.

Caso tal relação exista, seria indicativa de que julgados onde existe perícia técnica em violações de *Trade dress* possuem maior ou menor probabilidade de êxito. A hipótese nula (H_0) do teste de Chi-Squared seria da independencia das variáveis, enquanto a hipótese alternativa (H_1) seria da relação (ou dependência) entre existência de perícia e êxito final da lide.

Para elaboração deste teste, não foi possível criar subconjuntos referentes aos julgados do TJ-RJ e do TJ-SP, devido ao baixo número de amostras do TJ-RJ, que poderia causar distorções significativas. Optou-se, portanto, nesta etapa, pela análise dos resultados agregados de ambos os tribunais, considerados como uma única entidade.

4.6.2 Resultados agregados

Após aplicação do teste de Pearson Chi-Squared, obteve-se o seguinte resultado:

```

chisq.test(counts)
Pearson's Chi-squared test with Yates' continuity correction
data: counts
X-squared = 3.0197e-31, df = 1, p-value = 1

```

Observa-se que o valor p é bastante superior ao valor de significância ($p\text{-value} > 0.05$), o que indica que a hipótese nula (H_0) não pode ser refutada; equivale dizer, não se pode afirmar que existe algum tipo de relação estatisticamente válida entre a existência de perícia e o resultado final da lide.

5 CONCLUSÃO

5.1 Discussão dos resultados

O presente estudo balizou-seem duas frentes: a análise da doutrina do *trade dress*, e posteriormente sua análise do ponto de vista da jurisprudência, com utilização da metodologia de jurimetria.

A revisão teórica acerca do instituto não encontrou grandes surpresas. Trata-se de instituto fartamente discutido pelos doutrinadores. As dissensões doutrinárias são de menor monta, e não representam uma cisão abrupta entre correntes; de modo geral, e em relação aos seus caracteres principais, há um consenso substancial no tratamento teórico do instituto. Existe, porém, um reconhecido caráter discricionário na sua aplicação prática, no cotidiano das empresas e operadores do direito: a avaliação de infração, de ‘atos confusórios’, de ‘possibilidade de confusão ou associação’, são todos critérios norteadores que muitas vezes falham em proporcionar um norte absoluto, ou sequer um norte confiável. Há muita divergência entre julgados, e casos aparentemente semelhantes muitas vezes recebem decisões diametralmente opostas.

Neste contexto, procurou-se uma ferramenta capaz de se contrapor a tal discricionariedade, e o grau de insegurança jurídica que lhe acompanha, com base no estudo empírico do fenômeno, tomando com base a jurisprudência selecionada. Esta ferramenta foi a jurimetria, conjugada com a potencialidade de processamento massivo de dados proporcionada pela tecnologia computacional moderna, notadamente as ferramentas de processamento de linguagem natural que compõem o ferramental dito ‘inteligência artificial’.

Indubitavelmente, as ferramentas aplicadas neste trabalho apresentam grau de simplicidade abaixo do estado da técnica, tendo em vista os avanços quase diários nos estudos de ciência da computação, em especial de inteligência computacional aplicada.

A despeito de limitações técnicas, metodológicas e até epistemológicas deste trabalho, ainda assim os resultados apontam em direções de grande interesse não apenas aos especialistas em propriedade intelectual, mas aos operadores do direito como um todo.

Inicialmente, o tratamento jurimétrico apontou algumas grandezas numéricas que dispensam interpretação – os números sozinhos são deveras eloquentes. Pode-se verificar que o Tribunal de São Paulo julga mais ações de *trade dress* do que o Rio de Janeiro, e que ao

contrário do tribunal fluminense, o tribunal paulista vem observando um crescimento no número de julgados anuais.

Tal crescimento, ainda que relativamente relevante, ainda assim aponta que as ações de *trade dress* são de pouquíssima frequência no judiciário, e representam um percentual ínfimo do total de ações distribuídas, para a amostra estudada.

Em seguida, buscou-se determinar o perfil destas decisões em números. Restou claro que há um viés de êxito ao pleito autoral, ou seja, na maior parte dos casos, em ambos os tribunais estudados, o êxito é mais frequente que a sucumbência, ao final da lide.

Porém, tal conclusão foi complementada com uma análise do perfil de decisões de primeira instância, em busca de um entendimento maior sobre a possibilidade efetiva de reverter condenações em primeira instância, ou vice versa. A figura 14 acima nos indica que a reversão de sentenças de condenação é menos comum – ou estatisticamente menos provável – do que a reversão de sentenças de absolvição, conforme descrito no item 4.3.1 supra.

Em seguida, buscando sucedâneo nas informações numéricas (estatística descritiva), e aplicando metodologias de processamento de linguagem natural, buscou-se obter uma conclusão qualitativa (estatística inferencial) acerca de alguns tópicos ou elementos específicos.

Primeiramente, utilizando-se de algoritmos de modelo de espaço vetorial e frequência relativa de termos, procurou-se determinar se uma decisão teria preponderância de algum “elemento” ou “característica”, seja pelo cálculo da frequência relativa das palavras associadas ao conceito tratado (no que se denomina um “*Thesaurus*” em processamento de linguagem natural), seja pela revisão manual dos julgados para tabulação de informações relevantes (como no caso da perícia).

Uma vez determinada em um espaço objetivo – numérico – que tal frequência ou associação semântica estava presente, buscou-se em seguida determinar se a frequência de tais elementos teria influência com o resultado da lide. Equivale dizer, se há sugestão estatisticamente válida de que a ocorrência de tais conceitos nos julgados apontaria uma predisposição (ou melhor dizendo, uma maior probabilidade) ao sucesso ou insucesso nas lides.

Para isso, foram realizados diversos testes estatísticos na amostra, onde restou demonstrado que: (i) a preponderância do elemento COR em um julgado de *trade dress* não influencia estatisticamente seu resultado final; (ii) a preponderância do elemento MARCA em um julgado de *trade dress* não influencia estatisticamente seu resultado final; e (iii) a existência de prova pericial em um julgado não influencia estatisticamente seu resultado final.

É importante ressaltar que as conclusões são limitadas ao escopo que nos fornece a ciência estatística e a matemática. Não é possível determinar que tais elementos não influenciem no resultado da lide no caso concreto; mas é possível afirmar que, para o universo amostral trabalhado, não existe relação estatisticamente significante entre as variáveis elencadas. Equivale dizer, os resultados apontam para uma indiferença em relação a estes conceitos estarem ou não presentes na argumentação jurídica da lide, e seu eventual sucesso.

5.2 Possíveis problemas de pesquisa futuros

5.2.1 Problemas metodológicos

Alguns possíveis problemas metodológicos merecem menção.

Uma das premissas adotadas foi de que a frequência relativa de um termo e seus sinônimos ou equivalentes semânticos é um indicador adequado de sua relevância e importância no conteúdo total do texto. Esta premissa baseou-se em décadas de estudo de processamento de linguagem natural, e é uma das técnicas mais adotadas hoje em campos que necessitem de tais tratamentos, como análise de sentimento, classificação de textos, etc. Esta premissa, embora robusta teoricamente, pode não ser totalmente válida, ou apresentar deficiências quando aplicada ao estudo da jurisprudência. Outros fatores de aprimoramento deste tratamento poderiam oferecer resultados mais precisos, ou até mesmo contraditórios a este trabalho.

Além desta limitação, podemos ainda ressaltar algumas outras: que o universo amostral estudado pode ser insuficiente para extrapolar as conclusões para o restante dos tribunais pátrios; que a falta de análise de tribunais superiores pode omitir reversões de julgados de relevância estatística; que seria necessário obter uma amostra genérica de ações de diversos tipos para verificar se o viés de decisão se aplica somente a esta classe de ações, ou está em consonância com o percentual de êxito em outras classes de ações.

As limitações metodológicas foram minimizadas ou mitigadas quando possível no âmbito deste trabalho.

Os resultados apresentados aplicam-se somente à amostra estudada.

5.2.2 Possíveis problemas de pesquisa futuros

Como possíveis temas a serem desenvolvidos no futuro, pode-se cogitar em um estudo mais aprofundado que englobe outros tribunais, e um intervalo temporal ampliado, para comparar os resultados e determinar sua validade no caso de extrapolados para um universo maior.

Pode-se também determinar o perfil de decisões por magistrado e por câmara julgadora, em busca de determinar se há uma uniformidade de decisão, ou se algum julgador ou órgão possui um viés estatístico relevante.

Pode-se buscar por outros termos, ou conjuntos de termos, em uma tentativa de refinar os resultados do ponto de vista matemático; pode-se ainda buscar aplicar uma metodologia mais complexa de processamento de linguagem natural, talvez resultando em extração de informações mais detalhadas acerca de temas específicos dentro da jurisprudência, como acolhimento de teses jurídicas específicas.

De modo geral, o campo de estudo da jurimetria está em sua infância, e muito ainda poderá contribuir para um entendimento mais completo, mais embasado e mais significativo do fenômeno jurídico, em quaisquer de suas áreas de concentração.

6 Referências

- ABJ - ASSOCIAÇÃO BRASILEIRA DE JURIMETRIA. **Estatuto Social.** 2018a. Disponível em <https://abj.org.br/estatuto/>. Acessado em 27/01/2019.
- _____. **Propósitos e valores.** 2018b. Disponível em <https://abj.org.br/proposito/>. Acessado em 27/01/2019.
- ANDRADE, Gustavo Piva de. *O trade dress e a proteção da identidade visual de produtos e serviços*. Rio de Janeiro: **Revista da ABPI**, n. 112, mai/jun 2011.
- ARANHA, C. N. **Uma abordagem de pré-processamento automático para mineração de textos em português: sob o enfoque da inteligência computacional.** Tese de Doutorado, Departamento de Engenharia Elétrica, Pontifícia Universidade Católica, Rio de Janeiro, 2007.
- ARANHA, Christian; PASSOS, Emmanuel. A Tecnologia de Mineração de Textos. **Revista Eletrônica de Sistemas de Informação**, [S.l.], v. 5, n. 2, aug. 2006. ISSN 1677-3071. Disponível em: <<http://www.periodicosibepes.org.br/index.php/reinfo/article/view/171>>. Acesso em: 12 feb. 2019. doi: <https://doi.org/10.21529/RESI.2006.0502001>.
- ASCENSÃO, José de Oliveira. **As funções das marcas e os descritores.** Rio de Janeiro: **Revista da ABPI**, nº 61, nov/dez 2002.
- ASSEF JUNIOR, Eduardo João. **A proteção do trade dress à luz da concorrência desleal. Direito dos negócios em debate.** São Paulo: Saraiva, 2011. p.439-458.
- AUGELLO A. et. al. A Model of a Social Chatbot. In: GALLO, Pietro G., HOWLETT R., JAIN L. (eds) Intelligent Interactive Multimedia Systems and Services 2016. **Smart Innovation, Systems and Technologies**, vol 55. Springer, Cham, 2016. DOI https://doi.org/10.1007/978-3-319-39345-2_57.
- BAEZA-YATES, R.; BERTIER, R. N. **Modern Information Retrieval.** Harlow: Addison-Wesley, 1999.
- BAILEY, Katherine. **Reframing the “AI Effect”.** 2016. Disponível em <https://medium.com/@katherinebailey/reframing-the-ai-effect-c445f87ea98b> Acessado em 02/09/2018.
- BARAÑANO, A.M. **Métodos e técnicas de investigação: manual de apoio à realização de trabalhos de investigação.** Lisboa: Edition Sílabo, 2008.
- BARBOSA, Denis Borges. **Uma introdução à propriedade intelectual.** Rio de Janeiro: Lumen Juris. 2001, 700 p. Disponível em: < <http://www.denisbarbosa.addr.com/arquivos/livros/umaintro2.pdf>> Acesso em: 02 de agosto de 2018.

_____. **O Fator semiológico na construção do signo marcário.** Tese apresentada ao Programa de Pós-graduação da Faculdade de Direito da Universidade do Estado do Rio de Janeiro, como requisito para obtenção do título de Doutor em Direito. Rio de Janeiro, 2006.

_____. **Do trade dress e suas relações com a significação secundária.** 2011a. Disponível em:

<http://www.denisbarbosa.addr.com/arquivos/200/propriedade/trade_dress.pdf>
Acessado em 02 de outubro de 2015.

_____. **Revistando o tema da significação secundária.** Novembro de 2011b.

Disponível em:

http://www.denisbarbosa.addr.com/arquivos/200/propriedade/significacao_secundaria.pdf.
Acessado em 13/01/2019.

BARBOSA, Pedro Marcos Nunes. Uma Sucinta Análise da Teoria dos Sistemas para com a Propriedade Intelectual, Rio de Janeiro: **Revista Criação do IBPI**, 3º. Vol. III, Ed. Lumen Juris, 2012.

BASTOS, V. M. **Ambiente de Descoberta de Conhecimento na Web para a Língua Portuguesa.** 2006. Tese de Doutorado. Departamento de Engenharia Civil, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2006.

BASU, Debabrata; GHOSH, Jayanta Kumar. **Statistical information and likelihood: a collection of critical essays.** Berlin: Springer-Verlag, 1988.

BERTOLDI, Marcelo M.; RIBEIRO, Marcia Carla Pereira. **Curso avançado de direito comercial.** 7ª. Ed., São Paulo: Editora Revista dos Tribunais, 2013.

BERTRAN, Maria Paula Costa. **Análise econômica como critério orientador de decisão judicial: aplicações e limites – estudo a partir do caso de revisão dos contratos de arrendamento mercantil com paridade cambial.** Dissertação de Mestrado, Universidade de São Paulo, São Paulo, 2006.

_____. **Leve um e pague dois: inusitadas consequências jurídicas da desvalorização monetária.** Relato da revisão dos contratos de arrendamento mercantil indexados ao dólar. Rio de Janeiro: Fundação Getúlio Vargas, 2007.

BITTAR FILHO, Carlos Alberto. **Teoria e Prática da Concorrência Desleal.** São Paulo: Editora Saraiva, 1989.

_____. **Tutela dos direitos da personalidade e dos direitos autorais nas atividades empresariais.** 2ª. ed, São Paulo: Editora Revista dos Tribunais, 2002.

_____. **A concorrência desleal e a confusão entre produtos.** Rio de Janeiro: Revista dos Tribunais, v. 550, 2003.

BRASIL. Decreto nº. 75.572, de 08 de abril de 1975. Convenção da União de Paris para a Proteção da Propriedade Industrial, de 20 de março de 1883, promulgada no Brasil pela DAI – Divisão de Atos Internacionais. Disponível em: www.inpi.gov.br/images/stories/CUP.pdf. Acesso em 10/01/2014.

_____. Decreto lei nº. 1.355, de 30 de dezembro de 1994. Promulga a Ata Final que Incorpora os Resultados da Rodada Uruguai de Negociações Comerciais Multilaterais do GATT. Disponível em <http://www.inpi.gov.br/legislacao-1/27-trips-portugues1.pdf>. Acessado em 29/01/2019.

_____. Emenda Constitucional nº 45, de 30 de dezembro de 2004. Altera dispositivos dos arts. 5º, 36, 52, 92, 93, 95, 98, 99, 102, 103, 104, 105, 107, 109, 111, 112, 114, 115, 125, 126, 127, 128, 129, 134 e 168 da Constituição Federal e acrescenta os arts. 103-A, 103B, 111-A e 130-A, e dá outras providências. Disponível em http://www.planalto.gov.br/ccivil_03/Constituicao/Emendas/Emc/emc45.htm. Acessado em 01/09/2018.

_____. Lei nº. 13.105, de 16 de março de 2015. Institui o Código de Processo Civil. **Diário Oficial da União**, Brasília, DF, 17 de março de 2015. Disponível em http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/Lei/L13105.htm

CARRILHO, J. **Desenvolvimento de uma metodologia para mineração de textos**. Dissertação de Mestrado, Departamento de Engenharia Elétrica, Pontifícia Universidade Católica, Rio de Janeiro, 2007.

CARVALHO, Patrícia Luciane de. A marca como elemento de propaganda e publicidade: a controvérsia entre o CONAR e o INPI. Rio de Janeiro: **Revista da ABPI**, nº 118, pp. 37-47, 2012.

_____. *Trade dress*: a proteção ao conjunto informativo do Banco Itaú S.A. Rio de Janeiro: **Revista da ABPI**, nº. 114, 2011.

CASTRO, Beatriz Vergaça. Sinais distintivos não tradicionais: o caso Louboutin. Rio de Janeiro: **Revista da ABPI**, nº. 148, 2017.

CEGALLA, D. P. **Novíssima Gramática da Língua Portuguesa**, 46ª. ed., São Paulo: IBEP, 2005.

CERQUEIRA, João da Gama. **Tratado da propriedade industrial**. Volume II, Tomo I, Parte II, Rio de Janeiro: Editora Forense, 1952.

CERQUEIRA, João da Gama. **Tratado da propriedade industrial: das marcas de fábrica e de comércio, do nome comercial, das insígnias, das frases de propaganda e das recompensas industriais, da concorrência desleal**. Volume II, Tomo II, Rio de Janeiro: Lumen Juris, 2010.

COELHO, Fábio Ulhôa. **Curso de Direito Comercial**. 11ª edição, São Paulo: Editora Saraiva, 2007.

CONSELHO NACIONAL DE JUSTIÇA. **100 maiores litigantes.** Disponível em: <www.cnj.jus.br>. Acesso em: 12 mar. 2011.

CRUZ, Alexandre. Pirataria e contrafação: da propriedade intelectual ao *trade dress*. In: **XXVIII Seminário Nacional da Propriedade Intelectual: inovação e desenvolvimento**. Organizado pela Associação Brasileira da Propriedade Intelectual. Rio de Janeiro: ABPI, 2009.

DANIEL, Denis Allan. **Litígios envolvendo conjunto-imagem (“Trade dress”) no Brasil.** 2003. Disponível em:
http://www.daniel.adv.br/port/articlespublications/denisdaniel/trade_dress.pdf.> Acessado em 02 de agosto de 2018.

DANNEMAN, Siemensen Bigler; MOREIRA, Ipanema. **Comentários à lei de propriedade industrial e correlatos.** Rio de Janeiro, São Paulo: Ed. Renovar, 2005.

DE MULDER, Richard; VAN NOORTWIJK, Kees; COMBRINKKUITERS, Lia. **Jurimetrics please! A history of legal informatics, lefis series**, v. 9, n. 5, p. 147-178, 2010.

DEGROOT, Morris H. **Optimal statistical decisions**. Wiley-Interscience, 2005.

DIDIER JR. et. al. **Curso de Direito Processual Civil, direito probatório, decisão judicial, cumprimento da sentença e coisa julgada.** Volume 2, 2^a ed., Salvador: Jus Podium, 2008.

DODGE, Y. **The Oxford Dictionary of Statistical Terms**. Oxford University Press, 2006.

DUVAL, Hermano. **Concorrência Desleal**. São Paulo, Editora Saraiva, 1976.

FEINERER et al. **Package ‘tm’**. 2018. Disponível em <https://cran.r-project.org/web/packages/tm/tm.pdf>. Acessado em 02/09/2018.

FELDMAN, R.; SANGER, J. **The Text Mining Handbook – Advanced Approaches in Analyzing Unstructured Data**. Cambridge University Press, 2007.

FINATTO, M. J. Análise Textual Assistida por computador: Reconhecimento Lingüístico-Terminológico do texto técnico-científico de Química em Português - Da coesão à Enunciação (TextQuim). **Relatório Final de Atividades**. CNPQ, Instituto de Letras, UFRGS. Porto Alegre, 2005.

FINLAY, Steven. **Predictive Analytics, Data Mining and Big Data: Myths, Misconceptions and Methods** (1st.). Basingstoke: Palgrave Macmillan, 2014.

FONSECA, B. M.; REIS, D. C. **O fantástico mundo da distância de edição**. 2002. Disponível em:
<http://homepages.dcc.ufmg.br/~nivio/cursos/pa03/seminarios/seminario4/seminario4.pdf>. Acessado em 02/09/2018.

FONSECA, F.; FIDALGO, R. **Gerenciamento de Dados e Informação**. Centro de Informática, Universidade Federal de Pernambuco, 2002.

GRAYBILL, Franklin; HARIHARAN K.; BURDICK, Richard K. **Applied Statistics: A First Course in Inference.** Prentice Hall, 1998.

GDS PUBLISHING. Managing the Data Explosion. **Business Management.** 2008.

GEISSER, Seymour. **Predictive Inference: An Introduction.** New York: Chapman & Hall, 2016.

GHAG, Kranti Vithal; SHAH, Ketan. SentiTFIDF – Sentiment Classification using Relative Term Frequency Inverse Document Frequency. **International Journal of Advanced Computer Science and Applications**, Vol. 5, n.1, 2014. Disponível em <https://pdfs.semanticscholar.org/0940/96af2bdd341fc04dc0814bb17a503f41333.pdf>. Acessado em 12/01/2019.

GOLDSCHMIDT, R.; PASSOS, E. **Data Mining: um guia prático.** Rio de Janeiro: Campus, 2005.

GOMES, R. M. **Mineração de Textos na Desambiguação.** Dissertação de Mestrado. Departamento de Engenharia Elétrica, Pontifícia Universidade Católica, Rio de Janeiro, 2008.

GONÇALVES, T. et al. Analyzing Part-of-Speech for Portuguese Text Classification. **Proceedings of Computational Linguistics and Intelligent Text Processing (CICLing)**, pp. 551- 562, Mexico City, Mexico, 2006.

HALD, Anders. **A history of mathematical statistics from 1750 to 1930.** New York: Wiley, 1998.

HECKMAN, James J. Lessons from the Bell Curve. **Journal of Political Economy.** 103 (5): 1091–120, 1995. doi:10.1086/262014.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Comunicação Social - Sala de Imprensa, (eds). **Contas Regionais: de 2010 a 2013, PIB do Mato Grosso acumula a maior alta (21,9%) entre todos os estados.** Em 19 de novembro de 2015. Disponível em <https://www.ibge.gov.br/>. Acesso em 25 de dezembro de 2015.

IBM. **Unstructured Information – The Knowledge Rush.** Disponível em The Knowledge Rush: http://domino.research.ibm.com/comm/research_projects.nsf/pages/uima.knowled Acesso em 03 de Maio de 2008.

INPI - INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL. Ministério da Economia. **Marca – Mais informações.** 2015. Disponível em <http://www.inpi.gov.br/menu-servicos/marcas/marca-2013-mais-informacoes>. Acessado em 10 de janeiro de 2019.

INSITE. **Grupo de Lingüística da Insite Processamento de Linguagem Natural.** 2001. Disponível em Processamento de Linguagem Natural: <http://linguistica.insite.com.br/nlp.phtml> Acesso em 01 de abril de 2008.

INTA - INTERNATIONAL TRADEMARK ASSOCIATION. **Guidelines for trademark examination.** 07/05/1997.

Disponível em <https://www.inta.org/Advocacy/Pages/ModelLawsandGuidelines.aspx>. Acessado em 28/01/2019.

KADANE, Joseph B.; FIENBERG, Stephen E.; DEGROOT, Morris H. **Statistics in the law.** United Kingdom: Oxford University Press, 2008.

KONCHADY, M. **Text Mining Application Programming.** 1a. ed., Charles River Media, 2006.

LADAS, Stephen. **Patents, Trademarks and Related Rights.** Cambridge: Harvard University Press, 1975.

LAUDON, K. C.; LAUDON, J. P. **Management Information Systems: managing the digital firm.** 7^a ed. New Jersey: Prentice-Hal, 2002.

LIMA, Maria Alicia. Pesquisa: propriedade intelectual. In: **XXVIII Seminário Nacional da Propriedade Intelectual: inovação e desenvolvimento.** Associação Brasileira da Propriedade Intelectual. Rio de Janeiro: ABPI, 2009.

LOEVINGER, Lee. **Jurimetrics: the next step forward.** Minnesota Law Review, v. 33, n. 5, p. 455-493, 1948.

_____. **Jurimetrics: the methodology of legal inquiry.** Law & Contemporary Problems, v. 28, n. 1, p. 5-35, 1963.

LOVINS, J. B. Development of a stemming algorithm. **Mechanical Translation and Computacional Linguistics**, Vol. 11, 22-31, 1968.

MANDEL, A.; SIMON, I.; DELYRA, J. Informação: computação e comunicação. **Revista da USP**, V. 35, pp. 11-45, 1997.

MANNING, C. D.; RAGHAVAN, P.; SCHÜTZE, H. **Introduction to Information Retrieval.** Cambridge University Press, 2007.

MÄNTYLÄ, Mika V.; GRAZIOTIN, Daniel; KUUTILA, Miikka. The evolution of sentiment analysis: A review of research topics, venues, and top cited papers. **Computer Science Review**, Volume 27, February 2018, Pages 16-32. Disponível em <https://doi.org/10.1016/j.cosrev.2017.10.002>.

MARTINS, Fran. **Curso de Direito Comercial.** São Paulo: Editora Forense. 30^a edição, 2006.

McCARTHY, J. et. al. **A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence.** 1955. Disponível em: <http://raysolomonoff.com/dartmouth/boxa/dart564props.pdf>. Acesso em 02 de setembro de 2018.

MELLO, Erika Farah de. *Trade dress: análise comparativa das decisões dos Tribunais Nacionais de 1996 a 2011.* Rio de Janeiro: **Revista da ABPI**, n. 121, pp.3-23, 2012.

MENEZES, Paula Bezerra de. Perícia em conflitos de marcas e trade-dress: transferência da decisão de mérito para um terceiro não togado? Rio de Janeiro: **Revista da ABPI**, n. 103. pp.46-50, 2009.

MIRANDA, Pontes de. **Tratado de direito privado – Tomo XVII. Propriedade intelectual. Propriedade industrial**, vol II. 4^a edição. São Paulo: RT, 1983.

MOHR, S. F.; MITCHELL, G.; WADIKA Jr., S. **US Trade dress Law: Exploring the Boundaries**. New York: INTA, 1997.

MONTAURY, Luiz Edgard; MENDONÇA, Marianna Furtado de. *Trade dress* e a Tutela dos Websites. Rio de Janeiro: **Revista da ABPI**, nº 100, 2009.

MOOR, J., The Dartmouth College Artificial Intelligence Conference: The Next Fifty years,, **AI Magazine**, Vol 27, N°, 4, pp. 87-9, 2006.

MORO, Maitê Cecília Fabri. **A Marca tridimensional, sua proteção e os aparentes conflitos com a proteção outorgada com outros institutos da propriedade intelectual**. Tese de Doutorado. São Paulo: Pontifícia Universidade Católica, 2005.

. **Marcas Tridimensionais**. Rio de Janeiro: Ed. Saraiva, 2009.

NEVES, Daniel Amorim Assumpção. **Manual de direito processual civil**. São Paulo: Editora Juspodivm, 2017.

NEVES, P. I. **Uma estratégia para apoiar a decisão baseada em Mineração de Textos Livres**. Dissertação de Mestrado. Departamento de Ciência e Tecnologia, Instituto Militar de Engenharia. 2012.

NUNES, Marcelo Guedes. **Jurimetria aplicada ao direito societário: um estudo estatístico da dissolução de sociedade no Brasil**. Tese de doutorado. Pontifícia Universidade Católica de São Paulo, São Paulo, 2012.

OLAVO, Carlos. A protecção do “trade dress” In: **Direito Industrial**, Vol. V, Associação Portuguesa de Direito Intelectual, 2008.

OLAVO, Carlos. A protecção do “trade dress” no direito português e no direito comunitário. Rio de Janeiro: **Revista da ABPI**, n. 82, 2006.

OOMS, Jeroen. **Package ‘pdftools’**. Disponível em <https://cran.r-project.org/web/packages/pdftools/pdftools.pdf>. Acessado em 02 de setembro de 2018.

PATTON, M.Q. **Qualitative evaluation and research methods**. 2nd ed., Newbury Park, CA: Sage publications, 1990.

PEARSON, Karl. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. **Philosophical Magazine**. Series 5. 50 (302): 157–175, 1900. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/14786440009463897>. Acessado em 02 de fevereiro de 2019.

PEIRCE, Charles Sanders. **Collected Papers Of Charles Sanders Peirce.** Charles Hartshorne & Paul Weiss eds., 1934.

PINTO, André Almeida Matos de Oliveira; MONTENEGRO, Diego; PINHEIRO, Lauro Augusto Vieira Santos. *O trade dress e sua aplicação no Brasil*. Rio de Janeiro: **Revista da ABPI**, n. 21, 2012.

PORTER, M. An algorithm for suffixing stripping. **Program: electronic library and information systems**. Vol. 14 (3), 130-137, 1980.

PÜSCHEL, Flavia Portella (coordenação) et. al. Dano moral. **Série Pensando o Direito**, n. 37. Brasília, DF: Secretaria de Assuntos Legislativos do Ministério da Justiça, 2010. Disponível em: <http://participacao.mj.gov.br/pensandoodireito/wp-content/uploads/2012/11/37Pensando_Direito.pdf> Acesso em: 29/12/2018.

REQUIÃO, Rubens. **Curso de Direito Comercial**. São Paulo: Editora Saraiva, 29^a edição, 2010.

RIZZARDO, Arnaldo. Parte Geral do Código Civil. 2^a ed. Forense, 2003, p. 37.

RODRIGUES, Maria Lúcia de Barros. *Trade dress ou conjunto-imagem e sua proteção jurídica*. In: **Estado e economia: estudos em homenagem a Ademar Pereira**. São Paulo: Saraiva, 2011.

RUSSEL, Stuart J.; NORVIG, Peter. **Artificial Intelligence: A Modern Approach**. Pearson Education. 2003.

SANTOS, J. L. et al. Ativos intangíveis: fonte de vantagem competitiva. **ConTexto**, Porto Alegre, v. 6, n. 10, 2º semestre, 2006.

SAUSSURE, Ferdinand de. **Course in General Linguistics**. Charles Bally & Albert Sechehaye eds., Wade Baskin trans. 1959 – 1916.

SCHUMPETER, Joseph A. **Capitalismo, Socialismo e Democracia**. Rio: Zahar. 1984.

The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest and the Business Cycl. Translated from the German by Redvers Opie, New Brunswick (U.S.A) and London (U.K.): Transaction Publishers. [1911], 2008.

SHAPIRO, S. S.; WILK, M. B. An analysis of variance test for normality (complete samples). **Biometrika**, 52 (3–4): 591–611. 1965. Disponível em <https://academic.oup.com/biomet/article-abstract/52/3-4/591/336553>. Acessado em 28 de janeiro de 2019.

SILVA, A. A. **Aíuri: Um portal para Mineração de Textos integrado a Grids**. Dissertação de Mestrado. Engenharia Civil, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2007.

SILVA, Alberto Luís Camelier da. **Concorrência desleal: atos de confusão**. São Paulo: Saraiva, 2013.

SINGH, H. S. **DATA WAREHOUSE**: Conceitos, tecnologias, implementação e gerenciamento. São Paulo: Makron Books, 2001.

SOARES, Fábio de Azevedo. **Categorização automática de textos baseada em mineração de textos**. Tese de doutorado. Departamento de Engenharia Elétrica, Pontifícia Universidade Católica, Rio de Janeiro, 2013.

SOARES, José Carlos Tinoco. **Tratado da propriedade industrial: patentes e seus sucedâneos**. São Paulo: Jurídica Brasileira, 1998.

_____. **Tratado da propriedade intelectual: marcas e congêneres**. São Paulo, Editora Jurídica Brasileira, v. 1, 2003.

_____. “Concorrência Desleal” vs. “*Trade dress*” e/ou “conjunto-imagem”. São Paulo, 2004.

SOUZA, Fernando Menezes Campello de. **Decisões racionais em situações de incerteza**. 2^a. ed. revista. São Paulo: Revista dos Tribunais, 2007.

TEIXEIRA, Cassiano Ricardo Golos. Concorrência desleal: *Trade dress*. São Paulo: **Revista Eletrônica do IBPI**, 2009.

THEODORO, Júnior Humberto. **Curso de Direito Processual Civil**: Teoria geral do direito processual civil, processo de conhecimento e procedimento comum, Vol. I, 57^a edição, Rio de Janeiro: Editora Forense, 2016.

UNITED STATES. **The Lanham (Trademark) Act**. 1946. (Pub.L. 79-489, 60 Stat. 427, enacted July 5, 1946, codified at 15 U.S.C. § 1051 et seq. 15 U.S.C. ch. 22)

_____. Supreme Court. **Two Pesos, Inc. v. Taco Cabana, Inc.** (91-971), 1992 - 505 U.S. 763. Disponível em: <<https://www.law.cornell.edu/supct/html/91-971.ZO.html>> Acessado em: 10/08/2018.

_____. Supreme Court. **Wall Mart vs. Samara Brothers**, 2000. Disponível no sítio: www.supremecourtus.gov – Wal-Mart Stores, Inc. vs. Samara Brothers, Inc. (99-150), 529 U.S. (2000). Acesso em 11/11/2018.

_____. Supreme Court. **John H. Harland Co. vs. Clarke Checks, Inc.**, 711 F.2d 966, 980 (CA11 1983). Restatement (Third) of Unfair Competition 16, Comment a Tent. Draft No. 2, Mar. 23, 1990

_____. Supreme Court. **Qualitex Co. vs. Jacobson Products Co.**, n. 514 U.S. 159, 162-163, 1995.

WECHSLER, Sergio et al. **Relatório de análise estatística sobre o projeto “Análise econômica do direito aplicada a decisões judiciais: o caso dos contratos de arrendamento mercantil para compra de veículos com cláusulas de reajuste associadas ao dólar”**. São Paulo: IME-USP, 2006. Disponível em: <http://www.numberscare.com/wp-content/uploads/2017/07/decisoes-judiciais.pdf>. Acessado em 30/01/2019.

WELCH, B. L. The generalization of "Student's problem when several different population variances are involved. **Biometrika**, 34 (1–2): 28–35, 1947.

WILCOXON, Frank. Individual Comparisons by Ranking Methods. **Biometrics Bulletin**, Vol. 1, nº 6. Dec. 1945.

Disponível em <https://sci2s.ugr.es/keel/pdf/algorithm/articulo/wilcoxon1945.pdf>. Acessado em 30/01/2019.

ZABALA, Filipe Jaeger; SILVEIRA, Fabiano Feijó. **Jurimetria: estatística aplicada ao direito**. Natal: Revista Direito e Liberdade, v. 16, n. 1, p. 87-103, 2014.

7 APÊNDICE 1 – CODIGO-FONTE DO SOFTWARE DE JURIMETRIA UTILIZADO NO PROCESSAMENTO DOS DADOS

```
#####
#  
# Início do Script  
#  
#####  
##### SECTION 1 - Installation and Loading #####  
###Instala os pacotes requeridos  
install.packages("tm")  
install.packages("readtext")  
install.packages("pdftools")  
install.packages("wordcloud")  
install.packages("RColorBrewer")  
install.packages("dplyr")  
install.packages("readxl")  
install.packages("ggpubr")  
install.packages("tibble")  
install.packages("ggplot2")  
install.packages("gtools")  
install.packages("reshape2")  
install.packages("tidyverse")  
install.packages("gplots")  
  
## Carrega os pacotes para uso no script
```

```
library("tm")
library("pdftools")
library("wordcloud")
library("RColorBrewer")
library("dplyr")
library("readxl")
library("ggpubr")
library("tibble")
library("ggplot2")
library("gtools")
library("reshape2")
library("widyr")
library("tidyverse")
library("gplots")

## END OF SECTION 1 ###

#####
#####

## SECTION 2 - Construction of Corpus ##

#Verifica o Diretorio de trabalho atual
getwd()

#
#for home machine

setwd("C://Users//ed087//OneDrive//Documentos//rulings")
```

```
#for work machine

setwd("C://Users//ESGusr//Documents//mestrado-docs//rulings")

# Dump content of 'rulings' directory (all TXT files) into your workind dir

## END OF SECTION 2 ##

#####
#####

## SECTION 3 - BUILD AND PRE-PROCESS THE CORPUS

## Transforma os arquivos TXT em um Corpus para tratamento

corpus <- Corpus(DirSource(getwd()))

## Examina o resultado; deve conter 91 documentos

summary(corpus)

##Cria o Thesaurus para o conceito de "cor" e "marca":

color <- c("cor", "cores", "matiz", "matizes", "tom", "tons",
         "tonalidade", "tonalidades",
         "colorido", "coloridos", "colorida", "coloridas",
         "pigmentação", "tingimento")

marca <- c("marca", "marcas", "signo", "signos")

# fonte: https://www.sinonimos.com.br/cor/
```

Funções de pré-processamento dos textos do *Corpus*

```
toSpace <- content_transformer(function (x , pattern ) gsub(pattern, " ", x))

corpus <- tm_map(corpus, toSpace, "/")

corpus <- tm_map(corpus, toSpace, "@")

corpus <- tm_map(corpus, toSpace, "\\|")

corpus <- tm_map(corpus, toSpace, "\r\n")

corpus <- tm_map(corpus, toSpace, "-")

corpus <- tm_map(corpus, toSpace, "nº")

corpus <- tm_map(corpus, toSpace, "ª")

corpus <- tm_map(corpus, toSpace, "º")

corpus <- tm_map(corpus, toSpace, "§")

corpus <- tm_map(corpus, toSpace, "»")

corpus <- tm_map(corpus, toSpace, "£")

corpus <- tm_map(corpus, toSpace, "iii")

corpus <- tm_map(corpus, toSpace, "«")

corpus <- tm_map(corpus, toSpace, "[\\]")

corpus <- tm_map(corpus, toSpace, "fls")

corpus <- tm_map(corpus, toSpace, "paulo")

corpus <- tm_map(corpus, tolower)

corpus <- tm_map(corpus, removeNumbers)

corpus <- tm_map(corpus, removePunctuation)

new_stops <- c("tribunal", "justiça", "desembargador", "voto",

               "poder", "judiciário", "b", "n", "r", stopwords("portuguese"))

corpus <- tm_map(corpus, removeWords, new_stops)

corpus <- tm_map(corpus, stripWhitespace)
```

```
## Inspeciona o resultado do primeiro documento para avaliação visual de erros
inspect(corpus[1])

## END OF SECTION 3 ##

#####
#####

## SECTION 4 - Transform corpus into DTM, TDM, DF

## Tranforma o corpus em Document-term matrix
corpus_dtm <- DocumentTermMatrix(corpus)
print(corpus_dtm)

corpus_m <- as.matrix(corpus_dtm)

## Transforma em Data Frame
corpus_df <- as.data.frame(corpus_m)

## Transforma o corpus em Term-document Matrix
corpus_tdm <- TermDocumentMatrix(corpus)
m <- as.matrix(corpus_tdm)
v <- sort(rowSums(m), decreasing = TRUE)
d <- data.frame(word=names(v),freq=v)

## END OF SECTION 4 ##

#####
```

```
#####
## SECTION 5 - Calculates all necessary values for treatment
#Calcula a percentagem global de "marca" e "cor"
d %>% filter(word %in% color) %>% summarise(n=sum(freq))/sum(d$freq)
d %>% filter(word %in% marca) %>% summarise(n=sum(freq))/sum(d$freq)

#generate wordlcoud
set.seed(1234)
wordcloud(words = d$word, freq = d$freq, min.freq = 1,
          max.words=200, random.order=FALSE, rot.per=0.35,
          colors=brewer.pal(8, "Dark2"))

#Determina quais elementos de "color" e "marca" encontram-se em corpus_df
color_available=color[color %in% names(corpus_df)]
marca_available=marca[marca %in% names(corpus_df)]

#seleciona as colunas correspondentes a estes elementos
corpus_color=corpus_df %>% select(color_available)
corpus_marca=corpus_df %>% select(marca_available)

#frequencia total das entidades "cor" e "marca"
corpus_color$color_total=rowSums(corpus_color)
corpus_marca$marca_total=rowSums(corpus_marca)

#total de palavras por documento
```

```

corpus_color$totalwords=rowSums(corpus_df)

summary(corpus_color$totalwords)

#relative frequency per file

corpus_color$relfreq=corpus_color$color_total/corpus_color$totalword

corpus_marca$relfreq=corpus_marca$marca_total/corpus_color$totalword

summary(corpus_color$relfreq)

summary(corpus_marca$relfreq)

corpus_color= rownames_to_column(corpus_color, "id")

corpus_marca= rownames_to_column(corpus_marca, "id")

corpus_color=full_join(corpus_color,corpus_marca,by="id")

## END OF SECTION 5 ##

#####
##### SECTION 6 - Join sheets with additional data #####
#####

getwd()

#for home machine:

setwd("C://Users//ed087//OneDrive//Documentos")

#for work machine:

setwd("C:/Users/ESGusr/Documents/mestrado-docs")

#for Maria's machine:

```

```

setwd("?")


Planilha_R <- read_excel("planilha_r.xlsx")
names(Planilha_R)[1]="id"


decisions <- read_excel("decisions.xls",col_names = FALSE)
decisions=data.frame(decisions)
names(decisions)=c("id","decision")


#join decisions and corpus_color - fix name to merge both data.frame - replace
.pdf.text for empty char

corpus_color[,1]=gsub(".pdf.txt","",corpus_color[,1])
names(corpus_color)[c(15,21)]=c("color_relfreq","marca_relfreq")

#ready to join, c

color_decision=full_join(corpus_color,decisions)
color_decision=full_join(color_decision,Planilha_R)
color_decision$decision=as.factor(color_decision$decision)

## END of SECTION 6 ##

#####
#####
```



```

#####
## SECTION 7 - Additional metrics and info
#Algumas métricas adicionais
#Media de palavras por documento para decision=1 (EXITO)
```

```

color_decision      %>%
filter(decision==1) %>%
%>%
summarize(n=sum(totalwords))/color_decision %>%
filter(decision==1) %>%
%>%
summarize(n=n())

```

```

#Media de palavras por documento para decision=0 (SUCUMBENCIA)

color_decision      %>%
filter(decision==0) %>%
%>%
summarize(n=sum(totalwords))/color_decision %>%
filter(decision==0) %>%
%>%
summarize(n=n())

```

#average of color and marca words per file decision=1, in average we have 7 words "color" per file, and 30 marca words per file

```

color_decision      %>%
filter(decision==1) %>%
%>%
summarize(n=sum(color_total))/color_decision %>%
filter(decision==1) %>%
%>%
summarize(n=n())

color_decision      %>%
filter(decision==1) %>%
%>%
summarize(n=sum(marca_total))/color_decision %>%
filter(decision==1) %>%
%>%
summarize(n=n())

```

```

#average of color words per file decision=0

#in average we have 6 words "color" per file, and 24 marca words per file

color_decision      %>%
filter(decision==0) %>%
%>%
summarize(n=sum(color_total))/color_decision %>%
filter(decision==0) %>%
%>%
summarize(n=n())

color_decision      %>%
filter(decision==0) %>%
%>%
summarize(n=sum(marca_total))/color_decision %>%
filter(decision==0) %>%
%>%
summarize(n=n())

```

```

#frequencies summary as percentage

summary(color_decision$color_relfreq*100)

```

```

summary(color_decision$marca_relfreq*100)

## PLOTS ##

#o boxplot nos mostra que a frequencia relativa de marca é maior que de cor
boxplot(color_decision$color_relfreq,color_decision$marca_relfreq,
        names=c("color", "marca"),main="Distribution of Relative Frequency \n of the
words per file,for color and marca")

#plot: distribuição de relfreq of color por documento por decision
ggboxplot(color_decision, x = "decision", y = "color_relfreq",
          color = "decision", palette = c("#00AFBB", "#E7B800"),
          ylab = "Freq Relativa COR", xlab = "Decisão final",main="Distribuição das
Frequências relativas de COR por tipo de decisão")

#new variables

color_decision$Resultado1inst[(color_decision$Resultado1inst!="S")]="E"
counts <- table(color_decision$Resultado1inst,color_decision$decision)
barplot(counts, main="Resultado 1ª instancia distribuído por decisão final",
        xlab="Decisão final", col=c("lightblue","lightgreen"),
        legend = rownames(counts),beside=T)###

##

table(color_decision$V.Unanime,color_decision$decision) #no plot

counts=table(color_decision$Perícia ,color_decision$decision)
barplot(counts, main="Pericia Distribution by Decision",

```

```

xlab="Decision", col=c("lightblue","lightgreen","yellow"),
legend = rownames(counts),beside=T)###


counts=table(color_decision$ANO,color_decision$decision)
barplot(counts, main="Year Distribution by Decision",
xlab="Decision", col=c("lightblue","lightgreen","yellow"),
legend = rownames(counts),beside=T)###


counts=table(color_decision$Tribunal,color_decision$decision)
barplot(counts, main="Tribunal Distribution by Decision",
xlab="Decision", col=c("lightblue","lightgreen"),
legend = rownames(counts),beside=T)###


## END OF SECTION 7 ##

#####
###


#####
## SECTION 8 - Join tables??
#join decisions and corpus_color
#fix name to merge both data.frame
#replace .pdf.text for emtpy char
corpus_df= rownames_to_column(corpus_df, "id")
corpus_df[,1]=gsub(".pdf.txt","",corpus_df[,1])

```

```

#ready to join,
color_df=full_join(corpus_df,decisions)
color_df=full_join(color_df,Planilha_R)
color_df$decision=as.factor(color_df$decision)

#number of files per type
color_decision %>% group_by(decision) %>% summarise(n = n())

#list of files with decision=0 (dec0), and 1 (dec1)
dec0=color_decision %>% filter(decision==0)%>% select(id)
dec1=color_decision %>% filter(decision==1)%>% select(id)

#delete.pdf.txt of the names
docs=names(corpus)
docs=gsub(".pdf.txt","",corpus_color[,1])

#create one corpus for each decision
corpus0=corpus[which(docs%in%dec0[,1])]
corpus1=corpus[which(docs%in%dec1[,1])]

#
corpus_tdm <- TermDocumentMatrix(corpus0)
m <- as.matrix(corpus_tdm)
v <- sort(rowSums(m), decreasing = TRUE)
d0 <- data.frame(word=names(v),freq=v)

```

```

#head(d0,10)

#findFreqTerms

#freqforcolor

#h0=head(d0,300)

wordcloud(words = d0$word, freq = d0$freq, min.freq = 1,
          max.words=200, random.order=FALSE, rot.per=0.35,
          colors=brewer.pal(8, "Dark2"))

#repeat for 1

corpus_tdm <- TermDocumentMatrix(corpus1)

m <- as.matrix(corpus_tdm)

v <- sort(rowSums(m), decreasing = TRUE)

d1 <- data.frame(word=names(v),freq=v)

head(d1,20)

#findFreqTerms

#freqforcolor

set.seed(1234)

wordcloud(words = d1$word, freq = d1$freq, min.freq = 1,
          max.words=200, random.order=FALSE, rot.per=0.35,
          colors=brewer.pal(8, "Dark2"))

## END OF SECTION 8 ##

#####

```

```
#####
## SECTION 9 - Naormality and Hypothesis Tests
## Testes de normalidade e hipótese

#summary
group_by(color_decision, decision) %>%
  summarise(
    count = n(),
    mean = mean(color_relfreq, na.rm = TRUE),
    sd = sd(color_relfreq, na.rm = TRUE))

#normality test para COR e MARCA
# Hipótese nula H0 = distribuição normal; Hipótese alternativa H1 = distribuição não é normal
#if the p-value > 0.05, the distribution of the data
#is not significantly different from normal distribution.
shapiro.test(color_decision$color_relfreq)
shapiro.test(color_decision$marca_relfreq)
hist(color_decision$color_relfreq)
hist(color_decision$marca_relfreq,
     main="Histograma de Freq Relativa de MARCA", ylab="Documentos",
     xlab="Frequencia relativa MARCA")

#Resultado: COR não tem distribuição normal; MARCA tem distribuição normal.

#Wilcoxon test para COR
```

```

#Hipótese nula = as variáveis são independentes e não relacionadas

x=color_decision$color_relfreq[color_decision$decision==1]

y=color_decision$color_relfreq[color_decision$decision==0]

res <- wilcox.test(x,y,alternative="l")

res

#res acima de 0.05 indica que a hipótese nula está correta


#t-TEST para MARCA

#Calculando hipótese com t.Test de Student

t.test(marca_relfreq ~ decision, data = color_decision)

summary(color_decision$marca_relfreq)

table(color_decision$Pericia)

#The p-value of the test is 0.1376, which is more than the significance level
#alpha = 0.05. We can conclude that ZEROS's means relfreq is not significantly
#different from ONES's means relfreq.

#Teste de hipótese para PERICIA

#Dois casos de Segredo de Justiça foram eliminados como estatisticamente irrelevantes

#Chi-square test examines whether rows and columns of a contingency table are
statistically significantly associated.

#Null hypothesis (H0): the row and the column variables of the contingency table are
independent.

#Alternative hypothesis (H1): row and column variables are dependent

```

```

pie(table(color_decision$Pericia))

counts=table(color_decision$decision[which(color_decision$Perícia!="Segredo")],
color_decision$Perícia[which(color_decision$Perícia!="Segredo")])

chisq.test(counts)

#p-value is higher than the 0.05 significance level.

#So we do not reject the null hypothesis

#and conclude that decision and pericia do not have a significant relationship.

## END OF SECTION 9 ##

#####
##### SECTION 10 - More Plots

#ploting for each file

#plot

ggboxplot(color_decision, x = "decision", y = "color_relfreq",
          color = "decision", palette = c("#00AFBB", "#E7B800"),
          ylab = "Freq Relativa COR", xlab = "Decisão final")

#plot

ggboxplot(color_decision, x = "decision", y = "marca_relfreq",
          color = "decision", palette = c("#00AFBB", "#E7B800"),
          ylab = "Freq Relativa MARCA", xlab = "Decisão final")

```

```
x=melt(color_decision[,c("totalwords")])  
color_decision %>% ggplot(aes(x=id,y=totalwords,fill=decision)) +  
  geom_bar(stat="identity") +  
  facet_wrap( ~ decision)
```

```
## END OF SECTION 10
```

```
#####
#####
```

```
#####
#####
```

```
##### END OF SCRIPT #####
```

```
#####
#####
```