



Aplicação de técnicas de mineração de dados como ferramenta para geração de informação: um caso sobre o uso de análise sistemática em periódicos da área de design de moda

Application of data mining techniques as a tool for generating information: a case on the use of systematic analysis in journals in the area of fashion design

Bibiana Silveira Horn

*UNISINOS – Universidade do Vale do Rio dos Sinos
e-mail: bibianash@unisinobr*

Vinicius Gadis Ribeiro

*Universidade Federal do Rio Grande do Sul – UFRGS – Campus Litoral Norte
e-mail: vinicius.gadis@ufrgs.br*

Jorge Zabadal

*Universidade Federal do Rio Grande do Sul – UFRGS – Campus Litoral Norte
e-mail: jorge.zabadal@ufrgs.br*

Sidnei Renato Silveira

*Universidade Federal de Santa Maria – UFSM – Campus Frederico Westphalen/RS
e-mail: sidneirenato.silveira@gmail.com*

Cristiano Bertolini

*Universidade Federal de Santa Maria – UFSM – Campus Frederico Westphalen/RS
e-mail: cristiano.bertolini@ufsm.br*

Recebido em 04/11/2019. Aceito em 29/06/2020.

Resumo

O presente trabalho descreve a aplicação da técnica de mineração de dados em um estudo que utilizou a análise sistemática como metodologia de pesquisa. O objetivo deste estudo é o de descrever o uso da mineração de dados em uma área de pesquisa que não costuma utilizar este tipo de técnica. Como prova de conceito, foi aplicada essa técnica na análise de 276 artigos de três periódicos qualificados na área específica. Verificou-se que a estatística descritiva e algumas técnicas de mineração de dados, como apoio à análise sistemática, proporcionaram resultados relevantes para o estudo. Considerando os resultados do presente trabalho, pesquisadores podem empregar o processo aqui proposto em suas investigações, possibilitando fácil visualização de dados nem sempre evidentes.

Palavras chave: Mineração de dados. Análise Sistemática. Análise de Agrupamentos. Produção Científica.

Abstract

This paper describes the application of data mining technique in a study that used a systematic analysis as a research methodology. The purpose of this article is present the use of data mining in an area of research that seldom uses this type of technique. As proof of concept, this technique was applied in the analysis of 276 articles from three qualified journals in specific area. At the end, it is possible to verify that data mining, combined with systematic analysis and descriptive statistics provided a number of results and graphs of extreme importance for research in which it was applied.

Keywords: Data Mining. Systematic analysis. Clustering analysis. Design journals.

1. Introdução.

Pode-se afirmar que a produção científica em design de moda no Brasil - enquanto área de pesquisa e disseminação de conhecimento - é contemporânea (PIRES, 2002; CALDAS, 2006). A maioria dos cursos de nível superior são relativamente recentes, tendo havido a difusão de cursos de design de moda no país. Esse fenômeno se deve a vários fatores, tais como: demanda de empresas por profissionais capacitados, altos valores econômicos gerados pelo mercado de moda, relação empresa e universidade, por exemplo, o que, conseqüentemente, desencadeia a necessidade de profissionais docentes para atuarem nestes cursos.

No Brasil, o desenvolvimento do Design de Moda como campo de pesquisa científica deve-se em grande parte a trabalhos originados a partir de dissertações e teses, eventos e um número ainda significativamente pequeno de periódicos. No Brasil, algumas pesquisas já foram conduzidas, com o intuito de mapear a produção acadêmica na área de moda. Andrade Neto (2012) analisou publicações de artigos relacionados à moda em eventos - no caso, cinco edições do P&D (Congresso Brasileiro de Pesquisa em Design) e quatro edições do CIPED (Congresso Internacional de Pesquisa em Design). Outro trabalho que aborda a moda como área de disseminação de conhecimento foi conduzido por Bonadio (2010), a partir de teses e dissertações referentes à moda.

Neste contexto, o presente artigo tem, como objetivo, descrever de que forma pode-se utilizar a análise sistemática para verificação de produção científica em periódicos. Este trabalho apresenta o processo, desde a coleta e codificação dos dados, até a aplicação das técnicas de mineração de dados utilizadas como apoio à análise sistemática para obtenção dos resultados - iniciando a coleta com a leitura analítica de cada artigo, e posteriormente passando por diversas etapas, como a categorização dos trabalhos, pela criação de uma lista de categorias temáticas. As interpretações das leituras também foram empregadas para identificação de dados pertencentes às variáveis relacionadas à metodologia de pesquisa presente nos artigos. Em alguns casos, foi necessário o emprego de estatística descritiva - sobretudo em análises onde quantidades eram relevantes, complementando os resultados obtidos através da mineração de dados.

Como prova de conceito foram utilizados alguns dos resultados obtidos em Horn (2014), estudo anterior que analisou o emprego de metodologias de pesquisa baseados em produções de periódicos brasileiros qualificados da área de design. Entende-se por periódico qualificado o veículo acadêmico de disseminação de informação integrante do Qualis-Periódicos¹ na área de Design na CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior². Nesse contexto, buscou-se verificar de que forma a metodologia com esse aporte pode ser empregada e, como prova de conceito, apresenta alguns resultados referentes ao emprego dessa técnica na revisão sistemática.

O presente trabalho está estruturado da seguinte forma: na seção dois é apresentada a contextualização de Análise Sistemática e Mineração de Dados. O processo e as ferramentas de pesquisa são descritos na seção três, a qual apresenta o processo de preparação e coleta de dados, a mineração de dados utilizando a ferramenta de software *Rapidminer*³, e também o processo que foi necessário no pós-processamento. Na seção quatro, são exibidas algumas análises como: análise descritiva das instituições que mais publicaram artigos, a abordagem de metodologia de pesquisa científica nestes artigos e regras gerais dos periódicos analisados. Fazendo-se, nas considerações finais, uma análise do emprego dessa metodologia.

2. Análise sistemática e mineração de dados.

O método de pesquisa escolhido foi a análise sistemática – ou, como também é conhecida, revisão sistemática que, segundo Petticrew e Roberts (2006), é mais adequada para o efeito de responder a perguntas específicas e testar hipóteses do que outras avaliações tradicionais. A técnica empregada foi mineração de dados. A análise sistemática é utilizada com frequência em pesquisas da área médica e de áreas da saúde em geral. Para Sampaio e Mancini (2007), tais tipos de análises são metódicos e explícitos, sendo necessário que se possa reproduzir o estudo posteriormente. “Este tipo de estudo serve para nortear o desenvolvimento de projetos, indicando novos rumos para futuras investigações e identificando quais métodos de pesquisa foram utilizadas em uma área” (SAMPAIO e MANCINI, 2007). A identificação de métodos de pesquisa

¹ O Qualis-Periódicos é um sistema usado para classificar **a produção científica dos programas de pós-graduação** no que se refere **aos artigos publicados em periódicos científicos**. Tal processo foi concebido para atender as necessidades específicas do sistema de avaliação e é baseado nas informações fornecidas por meio do aplicativo Coleta de Dados. Como resultado, disponibiliza uma lista com a classificação dos veículos utilizados pelos programas de pós-graduação para a divulgação da sua produção.

² Órgão do Ministério da Educação responsável pelo reconhecimento e a avaliação de cursos de pós-graduação stricto-sensu (mestrado profissional, mestrado acadêmico e doutorado) em âmbito nacional.

³Disponível em <http://rapid-i.com/>

foi um dos objetivos do trabalho em qual foram aplicados o método e técnica citados anteriormente.

Também Akobeng (2005) caracteriza a análise sistemática como uma forma de pesquisa que utiliza métodos explícitos; além disso, avalia de forma crítica e sintetiza os estudos de forma sistemática. Já Linde e Willich (2003) afirmam que a revisão sistemática utiliza dados de outros estudos de determinados temas, assim como outros tipos de estudos de revisão. Esse tipo de análise torna possível uma gama maior de resultados relevantes, não limitando as conclusões há poucos trabalhos.

Embora a literatura tradicionalmente empregue técnicas de estatística descritiva para realização de análise sistemática, o presente trabalho empregou técnicas de mineração de dados. A mineração de dados trata-se de uma das etapas do processo conhecido como Descoberta de Conhecimento em Bases de Dados. Segundo Goldschmidt e Passos (2005), a descoberta do conhecimento em Bases de Dados também chamada de *Knowledge Discovery in Data Bases* (KDD) é um processo composto por três fases: Pré-processamento, Mineração de Dados e Pós Processamento. Os autores afirmam que essa área vem despertando o interesse das comunidades científica e industrial. Muitas vezes alguns autores podem se referir à KDD e Mineração de Dados como sinônimos; porém, para os autores em questão, a Mineração de Dados está inserida no KDD, constituindo-se em uma de suas etapas.

A primeira etapa da descoberta de conhecimento, segundo Goldschmidt e Passos (2005), é o Pré-processamento, onde acontece a captação, organização e o tratamento de dados. Essa etapa é muito importante no processo de KDD, sendo incluídas atividades de correção de dados errados, até o ajuste de formatos dos dados para mineração de dados.

Segunda etapa é a fase da mineração de dados; essa é a principal etapa do KDD e envolve a aplicação de algoritmos sobre os dados em busca de conhecimento implícitos e úteis. A forma como esse conhecimento será representado, em um modelo que será o resultado de todo processo depende diretamente do(s) algoritmo(s) utilizados na mineração de dados.

A fase de Pós-processamento engloba o tratamento do conhecimento obtido na mineração de dados; as principais funções são elaboração e organização, também podendo incluir simplificação de gráficos, diagramas, ou relatórios demonstrativos.

Entende-se a descoberta de conhecimento como “KDD é a descoberta de novos conhecimentos, seja padrões, tendências, associações, probabilidades ou fatos, que não são óbvios ou de fácil identificação” (AMARAL, 2001, p.13). Nisbet, Elder e Miner (2009) apresentam o KDD conforme a figura 1. Trata-se de todo processo, que engloba o acesso, exploração e preparação dos dados, depois a modelagem, implementação e monitoramento do modelo. E, nesse processo, são aplicadas as atividades de mineração de dados.

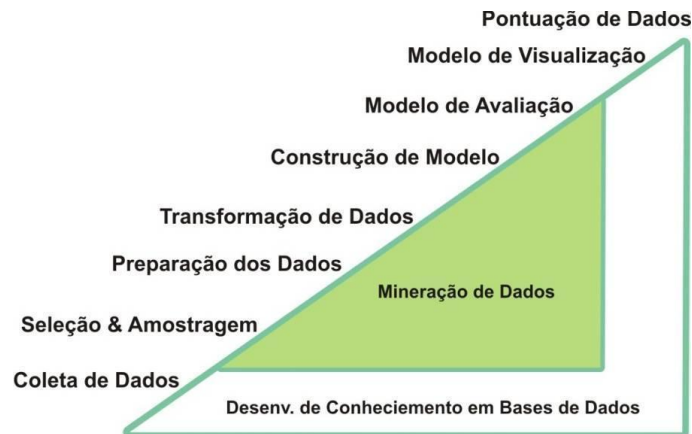


Figura 1: Relação entre mineração de dados e KDD
 Fonte: Nisbet, Elder e Miner (2009, p.17)

Uma unidade que pode ser empregada em mineração de dados é a variável. Pode-se definir variável como: “Os tipos de dados indicam a forma em que eles estão armazenados. Os tipos de variáveis expressam a natureza com que a informação deve ser interpretada” (GOLDSCHIMITD & PASSOS, 2005, p.24). Os autores classificam as variáveis em três tipos: variáveis nominais ou categóricas - utilizadas para nomear objetos, podem pertencer a um conjunto finito e pequeno de estados possíveis. Como exemplo, podemos citar estado civil de uma pessoa. Nesse tipo de variável, não existe ordenamento de valores, e eles, podem ser representados por tipos de dados alfanuméricos. Dessa forma, não se pode dizer que viúvo, por exemplo, é menor ou maior do que casado. As variáveis discretas são semelhantes às nominais, porém os seus valores podem possuir ordenamento, e esse terá algum significado. Os dias da semana são um exemplo, onde a terça feira, vem depois da segunda e antes da quarta. Já as variáveis contínuas são quantitativas e seus valores possuem uma relação de ordem. Neste caso o conjunto de valores pode ser finito ou infinito, e normalmente os valores são representados por algum tipo de dado numérico, por exemplo, idade ou renda.

A mineração de dados trabalha com uma série de tarefas, e com uma considerável quantidade de técnicas distintas. No presente trabalho, foram empregadas as técnicas de *clusterização* – ou agrupamento - e a associação por regras, pois são as tarefas mais clássicas da mineração de dados.

Nisbet, Elder e Miner (2009) definem a Clusterização/Agrupamento, como a detecção de subgrupos semelhantes entre uma grande variedade de casos e atribui essas observações aos seus subgrupos ou clusters. “Tão importante quanto identificar tais grupos é a necessidade de determinar como esses grupos são diferentes” (NISBET, ELDER & MINER, 2009, p. 147). Dias (2001) define a clusterização ou segmentação como a divisão de um grupo heterogêneo em subgrupos ou grupos mais homogêneos, e cita como exemplo agrupar clientes por região do país ou por similaridade de comportamento de compra. A figura 2 ilustra a mineração de dados inserida no processo de descoberta de conhecimento, apresentando destaque às tarefas e técnicas empregadas no presente trabalho.

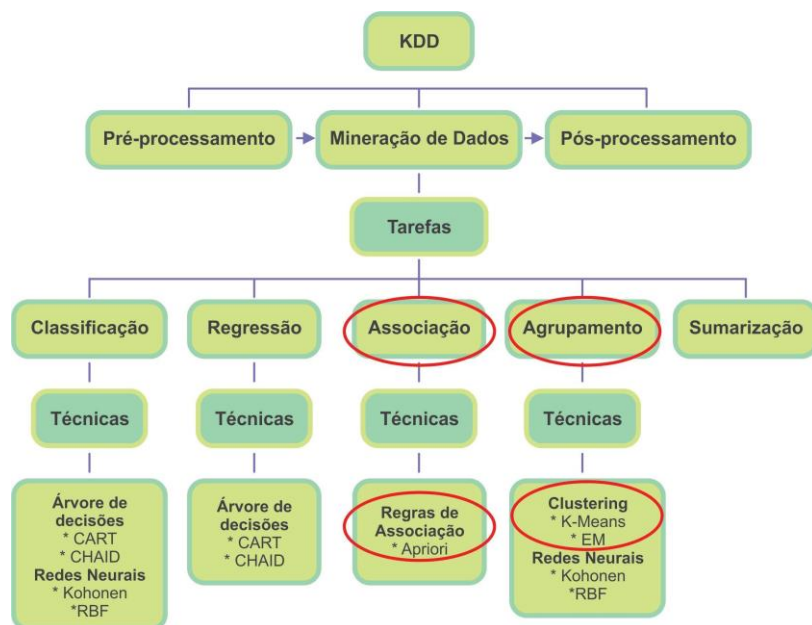


Figura 2: Processo de KDD e Mineração de Dados
Fonte: Elaborado pelos autores

Podem-se visualizar as três etapas citadas anteriormente, em seguida as tarefas de mineração de dados e, depois, algumas técnicas ou algoritmos básicos de mineração de dados. Deve-se destacar que as tarefas e técnicas presentes no mapa não englobam todas as funções da mineração de dados.

3. Ferramentas e processos.

Foram definidos como objeto de pesquisa artigos científicos publicados em três periódicos de moda brasileiros. Os critérios para escolha destas revistas foram: a) ser periódico brasileiro; b) específico de moda; c) ativo (ou em circulação) e d) constante na base *WebQualis* da área 29 da CAPES – que compreende as áreas de Arquitetura, Urbanismo e Design. O período analisado compreendeu o primeiro semestre desde 2007 até o primeiro semestre de 2013.

As revistas analisadas foram Dobras⁴, IARA: Revista de Moda, Cultura e Arte⁵ e Modapalavra⁶. Definido o período, foram considerados 276 artigos. Empregou-se uma planilha utilizando o *Microsoft Excel* para tabulação de dados, onde primeiramente foram inseridos dados referentes aos artigos - independente se esses artigos eram específicos de moda ou não -, como periódico, volume, mês e ano de publicação, instituição de ensino, título, palavras chave e se o artigo era da área da moda ou de uma área correlata. Dessa forma, todos os artigos participam da contagem e verificação inicial.

3.1 Preparação e coleta de dados.

⁴http://www.estacaolettras.com.br/revista_dobras.php

⁵ <http://www.iararevista.sp.senac.br/>

⁶ <http://www.ceart.udesc.br/modapalavra/edicao11/>

As leituras dos artigos foram realizadas com base na seguinte definição de moda: Moda é um fenômeno social relacionado à temporalidade, que em determinada época da história pode ter sido percebido como diferenciador de classes - mas que, antes de qualquer premissa, é um fator de individualidade diretamente relacionado às aparências e também ligado a questões culturais, sociais, políticas, antropológicas e econômicas. Criou-se uma lista de categorias⁷ temáticas para definir artigos de áreas correlatas e artigos de moda. Já com relação à categorização do método de pesquisa empregado, foi considerado o proposto por Ribeiro *et. al* (2011).

Desconsiderando-se os artigos de áreas correlatas, permaneceram duzentos e vinte e quatro artigos. Desses, foram analisados os dados, tais como o tipo de pesquisa, o tipo de coleta de dados utilizados, o caráter qualitativo ou quantitativo da pesquisa, o objeto de pesquisa e os autores e suas instituições de origem – esse último campo, com o objetivo de se verificar a possibilidade de grupos de pesquisa surgindo em determinadas universidades.

É importante destacar que algumas pesquisas não mencionavam a instituição de ensino a qual pertencia – tendo sido classificadas como “outras”. Muitos artigos não apresentavam dados de metodologia de pesquisa científica ou coleta de dados. Para que fosse possível completar os dados, foi necessário - através de leitura e interpretação dos artigos com base em Collado, Lucio e Sampieri (1997), Rea e Parker (2002) e outros autores específicos de cada metodologia de pesquisa -, definir os dados de metodologia e coleta de dados. É importante ressaltar área de moda é uma área de recente de investigação acadêmica no Brasil, de forma que a análise buscou considerar a realidade de cada universidade – visto que universidades com maior tempo de existência, com sólidos grupos de investigação tendem a gerar resultados com maior rigor metodológico.

⁷História da Moda, História da Moda no Brasil, Moda e Comportamento, Moda e Sociedade, Filosofia da Moda, Moda e Arte, Moda e Cultura, Moda e Literatura, Moda e Artesanato, Semiótica na Moda, Relação corpo e moda, Consumo de moda, Moda e Comunicação, Marketing de Moda, Mercado de Moda, Indústria da Moda, Ensino de Moda, Projeto de Moda, Terminologias de Moda, Ergonomia na Moda, Moda e design de superfície, Moda e Sustentabilidade, Calçados, Análise de artefatos de Moda, Epistemologia do Design de Moda, Moda e Tecnologia, Acessórios de Moda, Moda e Criatividade e Moda e Fetichismo.

1	A	B	C	D	E	F	G	H	I	J	K	L	M	
1	Periódico	Ano	IE	Título	Categoria	Técnica	Métodos de Pesqu	Técnica de Coleta de Dados	Autor	Autor	Autor	Autor	Metodologia científica	
199	Modapala	2009	PUC-MINAS	MODA E ALTERIDADE: Moda	Mercado de moda	Quali	Bibliográfica	Coleta Bibliográfica	Geanr	0	0	0	Não aborda	
200	Modapala	2009	FEEVALE	e*ALTA COSTURA NACIONAL História da moda; Terminologias de Mo		Quali	Estudo de Caso	Análise de fotografia; Análise sem	Thisa	Claudi	Denis	0	Abordagem detalhada	
201	Modapala	2009	UDESC	e*VESTIDO DE 1910: memóri História da moda		Quali	Dissertação projet	Análise de Artefato; Entrevistas	Valdir	Lucian	José	A	0	Abordagem detalhada
202	Modapala	2009	UDESC	e*Muito Além da Superfície Semiótica na Moda		Quali	Bibliográfica	Coleta Bibliográfica	Renat	Sandr	0	0	Não aborda	
203	Modapala	2009	UFSC;UDESC	e*ILUSTRAÇÃO DE MODA NA Ensino de moda		Quali	Dissertação projet	Coleta Bibliográfica	Caroli	Anton	0	0	Não aborda	
204	Modapala	2009	UFSC;UDESC	e*ATELÉ E EDUCAÇÃO A DIS Ensino de moda		Quali	Dissertação projet	Coleta Bibliográfica	Lucian	Leah	Marili	Eliana	0	Não aborda
205	Modapala	2009	FEEVALE	O CONSUMIDOR DE MODA N Moda e Cultura; Consumo de Moda		Quali	Estudo comparati	Entrevistas	Ana	Si	Claudi	Denis	0	Abordagem detalhada
206	Modapala	2009	Outras	A PUBLICIDADE E A LINGUAG Marketing de moda		Quali	Estudo de Caso	Análise de fotografia	Roche	0	0	0	Não aborda	
207	Modapala	2009	Outras	DESABILIDADE, DESIGN DE M Projeto de moda; Ergonomia na Moda		Quali	Dissertação projet	Entrevistas	Sheila	0	0	0	Abordagem detalhada	
208	Modapala	2009	UCAM	O CARÁTER ARTESANAL DO F Moda e artesanato		Quali	Reflexiva	Coleta Bibliográfica	Tatian	0	0	0	Não aborda	
209	Modapala	2009	UDESC	e*RELAÇÕES DE GÊNERO E N Semiótica na Moda; Marketing de moda		Quali	Estudo de Caso	Análise de fotografia	Danie	0	0	0	Não aborda	
210	Modapala	2009	UDESC	e*O CONSUMO E A CELEBRAI Consumo de moda		Quali	Histórica	Coleta Bibliográfica	Gisele	0	0	0	Abordagem superficial	
211	Modapala	2009	PUC-RIO;UDE	e*ECODESIGN, SUSTENTABIL Moda e Sustentabilidade		Quali	Dissertação projet	*Manufatura/Resultados	Janair	Neide	0	0	Não aborda	
212	Modapala	2008	UNIBAN	FAST FASHION: ESTRATÉGIA Consumo de moda		Quali	Bibliográfica	Coleta Bibliográfica	Danie	0	0	0	Não aborda	
213	Modapala	2008	Outras	MODA E SUBJETIVIDADE: COI Moda e Sociedade		Quali	Reflexiva	Coleta Bibliográfica	Maria	0	0	0	Não aborda	
214	Modapala	2008	UDESC	SUSTENTABILIDADE AMBIEN Moda e Sustentabilidade		Quali	Reflexiva	Coleta Bibliográfica	Neide	Lucian	0	0	Não aborda	
215	Modapala	2008	Outras	GAROTAS DE SÃO PAULO: IM História da moda no BR		Quali	Histórica	Análise de Ilustração	Rosan	0	0	0	Não aborda	
216	Modapala	2008	Outras	e*O ESTILO ENQUANTO LÓGI História da moda; Moda e sociedade		Quali	Reflexiva	Coleta Bibliográfica	Flávia	0	0	0	Não aborda	
217	Modapala	2008	Outras	e*GERAÇÕES JUVENIS E A M Moda e comportamento		Quali	Bibliográfica	Coleta Bibliográfica	Maria	0	0	0	Não aborda	
218	Modapala	2008	Outras	e*MODA E ARTE SOB O OLH# Moda e comportamento		Quali	Bibliográfica	Coleta Bibliográfica	Silvia	0	0	0	Não aborda	
219	Modapala	2008	UDESC	ESTRUTURA DA CADEIA PROI Indústria da moda		Quali	Bibliográfica	Coleta Bibliográfica	Sandr	0	0	0	Abordagem superficial	
220	Modapala	2008	UDESC	Usabilidade do Vestuário: F# Ergonomia na Moda		Quali	Bibliográfica	Coleta Bibliográfica	Icléia	0	0	0	Não aborda	

Figura 3: Planilha Excel com os dados brutos tabulados, por categorias definidas
 Fonte: Elaborado pelos autores

Com estas interpretações, considerou-se um campo que identificava o nível de profundidade sobre a metodologia no artigo. A figura 3 apresenta a planilha Excel que constituiu a fonte para a tabulação dos dados e serviu como base de dados brutos. Em alguns casos, o estudo não mencionava a metodologia; em outros casos, abordava esse tema de forma superficial, relatando um ou dois dados no resumo ou na introdução. Outros abordavam a metodologia de forma mais detalhada, fornecendo mais dados sobre a metodologia, um resumo mais completo e, em alguns casos, constava a seção específica de metodologia.

Após a fase de coleta de dados, foram realizadas correções e os ajustes no formato dos dados. No caso do software escolhido, foi necessário realizar codificações substituindo dados por códigos numéricos, e em alguns casos por palavras. Estas substituições são feitas conforme a tarefa e a técnica de mineração escolhidas dentro do *software*.

3.2 Mineração de dados utilizando *software*.

Um exemplo de verificação obtida por meio da clusterização, foram os métodos de pesquisa empregados nos artigos presentes nos três periódicos analisados. A partir dos dados o programa gera uma opção de gráfico, como o apresentado na Figura 4. Nesse caso optou-se pelo gráfico 3D colorido, tendo sido selecionados um tipo de dado para cada eixo - x, y, z.

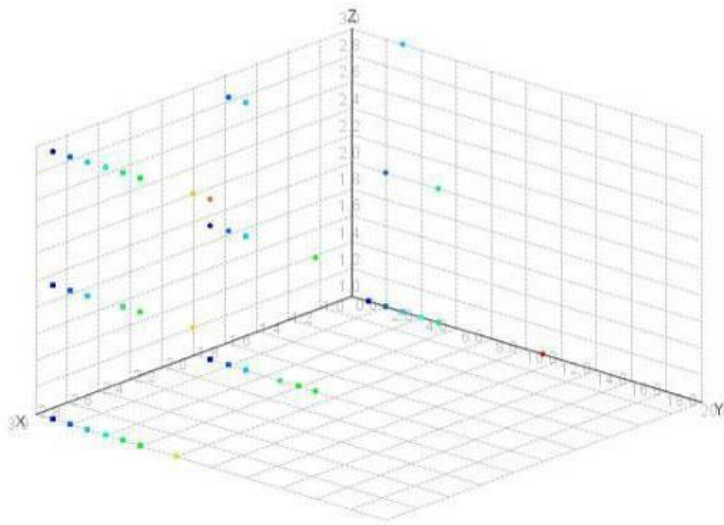


Figura 4: Exemplo de gráfico da técnica de clusterização
 Fonte: Elaborado pelos autores

Neste caso os periódicos estão no eixo X, o método de pesquisa utilizado para Y e abordagem de metodologia científica para Z. O campo cor foi marcado para método de pesquisa utilizado, que foram onze encontrados nos artigos: Pesquisa Histórica, Pesquisa Bibliográfica, Estudo de Caso, Pesquisa Reflexiva (neste caso artigos de Reflexão), Dissertação Projeto, Estudo Comparativo, Revisão Sistemática, Pesquisa-ação, *Survey*, Pré-experimento e Estudo de Caso somado a pesquisa Histórica.

Associação é a relação de itens que ocorrem com frequência em conjuntos de dados. Goldschmidt e Passos (2005) definem da seguinte maneira: “Intuitivamente essa tarefa consiste em encontrar conjuntos de itens que ocorram simultaneamente e de forma frequente em um banco de dados” (GOLDSCHMITD e PASSOS, 2005, p. 59). Um exemplo clássico citado por Dias (2001) são as compras ou produtos que são colocados juntos no carrinho de supermercado. Para Tan, Steinbach e Kumar (2005), a força de uma regra de associação pode ser medida em termos do seu suporte e confiança. Suporte determina a frequência com que uma regra é aplicável a um determinado conjunto de dados, enquanto confiança determina a frequência com que os itens em Y aparecem em transações que contêm X. Um suporte alto é importante: se ele for baixo, a regra pode ocorrer apenas por acaso. Da mesma forma, a confiança também, quanto maior melhor, pois ela mede a confiabilidade da interferência feita por uma regra.

Já a Associação por Regras trabalha com dados binários. Neste caso utilizamos 0 e 1 para codificar os dados, referente a deter ou não alguma propriedade de interesse. Foi necessário um novo tipo de planilha de dados, na qual cada linha representa um tipo de transação. No caso dessa pesquisa, cada linha do gráfico em questão representa uma publicação, e cada coluna, um item ou atributo. Nesse tipo de análise, há opção de gerar gráfico e de se fazer leitura e interpretação das regras. Foi descartada a opção do uso do gráfico gerado no programa por resultar em um gráfico com muitas informações, no caso da análise desses dados.

As leituras das regras podem ser feitas através da interface do *Rapidminer*. A Figura 5 ilustra o resultado gerado pela aplicação, no formato de planilha. Percebe-se que a mesma fornece,

entre outras informações – como o valor de suporte e de confiança –, coeficientes relevantes para análise de regras por associação. A força de uma regra de associação pode ser medida em termos do seu suporte e confiança. É importante um suporte elevado: se ele for baixo, é indicativo de que a regra pode ocorrer apenas por acaso. O mesmo ocorre com a confiança – quanto maior, melhor –, pois ela mede a confiabilidade da interferência feita por uma regra.

No.	Premises	Conclusion	Sup.	Confid.	LaPl.	Gain	p-s	Lift	Confi.
674	Quali	Nacional	0.924	0.937	0.959	-1.045	-0.001	0.999	0.987
737	Nacional	Quali	0.924	0.986	0.993	-0.951	-0.001	0.999	0.938
499	Quali	Não aborda metodologia científica	0.741	0.751	0.876	-1.232	0.006	1.008	1.022
749	Não aborda metodologia científica	Quali	0.741	0.994	0.997	-0.750	0.006	1.008	2.237
465	Quali	Único Autor	0.710	0.719	0.861	-1.263	0.001	1.001	1.003

Figura 5: Planilha análise de regras

Fonte: Elaborado pelos autores

No caso, um suporte e uma confiança com valores altos são bons, mas nem suficientes: o número de dados na premissa e na conclusão também deve ser maior. Caso contrário, geram-se regras simples - até mesmo óbvias – com baixa relevância.

Na etapa de pós-processamento, optou-se por simplificar e redesenhar alguns gráficos. Dessa forma, ficaram mais claros, apresentando os dados com seus nomes e não apenas com dados numéricos. Também foram empregadas as funções de estatística descritiva do Microsoft Excel – gerando, assim dados, quantitativos para alguns pontos que eram relevantes, como por exemplo: categorias temáticas mais pesquisadas, *ranking* de instituições que mais publicaram artigos e métodos de coletas mais utilizados.

4. Análise dos resultados.

Os principais resultados foram aqui selecionados como alguns exemplos do emprego das técnicas – sendo apresentada, ao final dessa seção, a síntese de todos os resultados.

Sobre as Universidades pode-se, na Figura 6, identificar quais são as que mais geraram publicações. A maioria dos registros não registrou filiação a qualquer instituição – mas, no caso dos trabalhos com Instituições identificadas, verifica-se que as que detêm maior número de publicações foram UDESC, PUC-SP, UFSC e SENAC-SP.



Figura 6: Instituições de ensino com maior número de publicações
 Fonte: Elaborado pelos autores

Um dos pontos mais relevantes para o problema da pesquisa foi a verificação da presença/ausência de uma seção de metodologia de pesquisa utilizada nos artigos analisados - por se tratar de um levantamento voltado para epistemologia do design de moda. Como já foi mencionado em seção anterior, criou-se uma classificação para abordagem científica, que resultou em 167 artigos que não mencionam metodologia, 30 que mencionam a metodologia, porém não dão informações completas e 27 artigos que mencionam a metodologia detalhadamente.

No exemplo da Figura 7, o periódico que mais utilizou detalhadamente metodologias de pesquisa foi Modapalavra; na contagem da abordagem “emprega/apresenta metodologia de modo superficial”, esse mesmo periódico também se destacou, e na contagem dos que não abordavam metodologias, os números foram similares entre os três periódicos analisados.

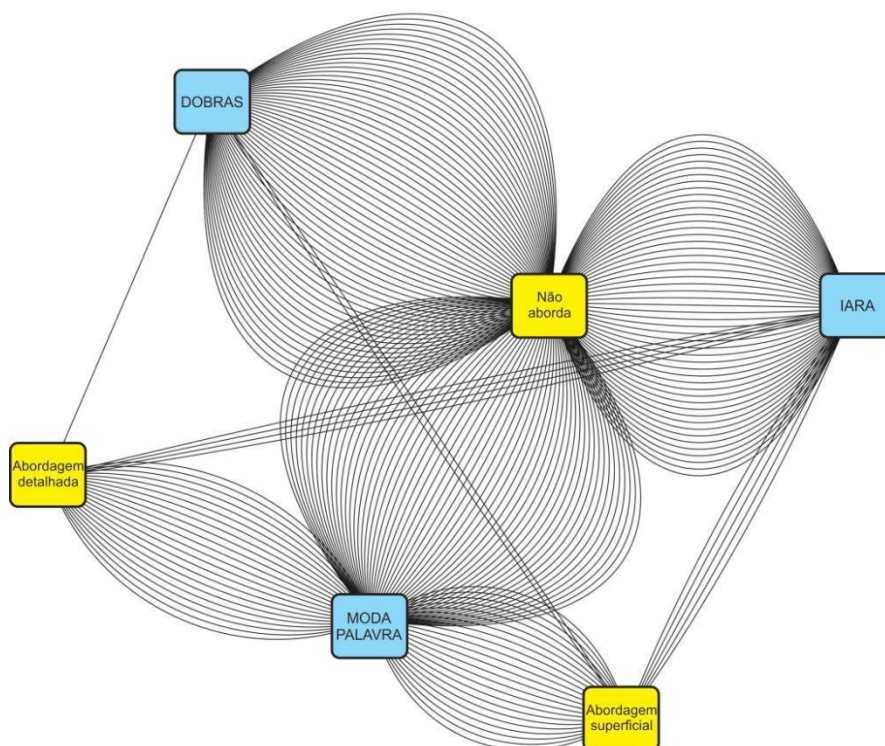


Figura 7: Abordagem de metodologia científica

Fonte: Elaborado pelos autores

Conforme mostra a Figura 8, pode se visualizar a regra na linha 455. Esta regra afirma que 99,5 das pesquisas que não abordam metodologia de pesquisa, também são de caráter qualitativo, e esta regra ocorre em 74% de todos artigos analisados - nesse caso, considerando os artigos desses periódicos. E na regra da linha 467, pode-se verificar a regra mostra que 100% dos artigos que não abordam metodologia e utilizam coleta bibliográfica são de caráter qualitativo: esta regra aparece em 38% de todos os artigos que estão na base de dados - esses totalizando 224 artigos.

Tais análises foram obtidas com o intuito de identificar as mais diversas informações - tais como temáticas mais pesquisadas, métodos de pesquisa e coletas de dados mais utilizados, instituições com maior número de publicações, periódicos que mais aborda a metodologia de pesquisa, caráter das pesquisas.

Além do panorama geral da pesquisa em design de moda, foi possível ainda construir um perfil de tendência de cada periódico através dos resultados obtidos. Maiores detalhamentos são disponibilizados em Horn (2014).

No.	Premises	Conclusion	Sup...	Confide.	LaPla.	Gain	p-s	Lift	Convi.
374	Quali	Não aborda metodologia científica	0.741	0.751	0.876	-1.232	0.006	1.008	1.022
455	Não aborda metodologia científica	Quali	0.741	0.994	0.997	-0.750	0.006	1.008	2.237
287	Quali	Coleta Bibliográfica	0.469	0.475	0.739	-1.504	0.002	1.004	1.004
454	Coleta Bibliográfica	Quali	0.469	0.991	0.997	-0.478	0.002	1.004	1.420
275	Quali	Modapalavra	0.451	0.457	0.730	-1.522	0.002	1.004	1.003
453	Modapalavra	Quali	0.451	0.990	0.997	-0.460	0.002	1.004	1.366
268	Quali	Bibliográfica	0.429	0.434	0.719	-1.545	0.006	1.014	1.010
456	Bibliográfica	Quali	0.429	1	1	-0.429	0.006	1.014	∞
251	Quali	Não aborda metodologia científica, Coleta Bibliogr	0.388	0.394	0.699	-1.585	0.005	1.014	1.009
298	Não aborda metodologia científica	Coleta Bibliográfica	0.388	0.521	0.795	-1.103	0.036	1.101	1.100
299	Não aborda metodologia científica	Quali, Coleta Bibliográfica	0.388	0.521	0.795	-1.103	0.039	1.111	1.109
300	Quali, Não aborda metodologia científica	Coleta Bibliográfica	0.388	0.524	0.797	-1.094	0.038	1.108	1.107
396	Coleta Bibliográfica	Não aborda metodologia científica	0.388	0.821	0.942	-0.558	0.036	1.101	1.420
397	Coleta Bibliográfica	Quali, Não aborda metodologia científica	0.388	0.821	0.942	-0.558	0.038	1.108	1.445
398	Quali, Coleta Bibliográfica	Não aborda metodologia científica	0.388	0.829	0.945	-0.549	0.039	1.111	1.484
467	Não aborda metodologia científica, Coleta Bibliográfica	Quali	0.388	1	1	-0.388	0.005	1.014	∞
212	Quali	Coleta Bibliográfica, Bibliográfica	0.357	0.362	0.683	-1.616	0.005	1.014	1.008
375	Coleta Bibliográfica	Bibliográfica	0.357	0.755	0.921	-0.589	0.154	1.761	2.330
376	Coleta Bibliográfica	Quali, Bibliográfica	0.357	0.755	0.921	-0.589	0.154	1.761	2.330

Figura 8: Regras gerais de todos os periódicos
Fonte: Elaborado pelos autores

5. Considerações finais.

Assumindo-se que a publicização de resultados representa a etapa final de um processo de pesquisa, pode-se verificar que a análise dos métodos e as técnicas de pesquisa empregados nos artigos desses periódicos forneceram uma série de informações relevantes para o problema da pesquisa realizada. Também poderiam ser utilizadas outras técnicas de mineração de dados e mesmo de mineração de textos. Contudo, para maior eficiência, essas exigem grande volume de dados e/ou textos e, nesse caso, foram analisados apenas 224 artigos dos 270 inicialmente previstos, embora com maior detalhamento. Apesar da limitação da quantidade de artigos, percebe-se que a mineração de dados é uma técnica que pode ser empregada neste tipo de estudo, e oferece uma variedade grande de ferramentas. Nesse sentido, ressalta-se que o presente trabalho buscou tão somente apresentar uma prova de conceito. Deve ser considerado que, para o emprego de todo o potencial dessas técnicas, a quantidade deve ser consideravelmente maior.

Em alguns pontos pesquisados foi necessária utilização de estatística descritiva com auxílio do *Microsoft Excel* e o redesenho de alguns gráficos para melhor visualização. Uma limitação do presente trabalho se refere ao intervalo temporal de dados empregado. Um possível trabalho futuro seria a ampliação dos dados brutos, abrangendo maior intervalo temporal. Outro possível trabalho futuro seria o emprego de técnicas de mineração de textos, buscando identificar conceitos e ideias embutidas nos textos, e verificando possíveis razões sobre o porquê tão pouca preocupação metodológica, em termos de empregar robustos métodos de pesquisa - considerando que a quantidade de produção acadêmica sobre a Moda em periódicos qualificados é restrita, talvez resultados mais ricos possam ser obtidos em periódicos mais tradicionais e de outros países.

Considerando-se o processo de pesquisa empregado, pesquisadores de outras áreas que desejarem realizar análise sistemática devem considerar empregar os seguintes passos: visualização dos dados que estão sendo extraídos e combinados, definição de uma pergunta e adoção de um método objetivo, avaliar e selecionar os resultados que são de fato relevantes para a resposta da pergunta da primeira fase.

Pode-se afirmar que o método de revisão sistemática pode ser enriquecido com o emprego de técnicas quantitativas, como teoria dos grafos (RIBEIRO, ZABADAL & SILVEIRA, 2016), mineração de dados e de textos. Simples tabulação de determinadas variáveis em uma planilha Excel já possibilita a identificação de informações de âmbito geral.

Além do simples emprego da revisão sistemática apenas como filtragem de artigos por meio de seleção de palavras-chave, pode-se identificar temas emergentes (por exemplo, por meio do indicativo se determinados temas estão aumentando a frequência através das datas de publicação), se há aumento da quantidade de grupos de pesquisa, etc. Regras podem ser geradas a partir da base, e agrupamentos podem ser identificados: tais informações seriam dificilmente identificadas a partir da simples análise individual de cada artigo. Ademais, os programas de mineração de dados têm, de modo geral, ferramentas de visualização que facilitam o processo de análise da produção de interesse. A recomendação, para o sucesso nesses casos, seria o emprego de grande quantidade de artigos – o que constituiu uma limitação do presente trabalho, cujo compromisso era verificar a viabilidade por meio de uma prova de conceito com os dados – entenda-se artigos - disponíveis.

Com relação aos resultados obtidos no trabalho, no que tange ao tema que constituiu o problema em questão – emprego de metodologia de pesquisa na produção científica em uma área iniciante em pesquisa acadêmica -, foi possível obter o panorama geral das pesquisas acadêmicas na área em questão - design de moda em periódicos brasileiros qualificados -, revelando as temáticas mais pesquisadas, os métodos de pesquisa e coletas de dados mais utilizados, instituições com mais publicações. É mais significativo, obteve-se a informação de que a maioria dos trabalhos pesquisados - referentes aos projetos de pesquisa que os originou -, não considerou metodologia de pesquisa com dados primários. Esse fato constitui oportunidades acadêmicas para o aprofundamento da investigação no Brasil. Além desses dados, também foi possível definir o perfil de cada um dos periódicos analisados, o que ajuda na orientação aos pesquisadores brasileiros da área em questão, quanto à submissão de trabalhos.

Agradecimentos.

A primeira autora agradece à CAPES pelo apoio financeiro para realização deste trabalho.

Referências.

AKOBENG, A. K. Understanding systematic reviews and meta-analysis. *Archives Disease Childhood*, 90:845-8. 2005. Disponível em: <http://adc.bmj.com/content/90/8/845>. Acesso em 10 de mar. 2019.

ANDRADE NETO, M. L. A produção científica de Design de Moda no Brasil: um estudo bibliométrico. *DAMT: Design, Arte, Moda e Tecnologia*. São Paulo, 2012. Disponível em: <http://sitios.anhembibrasil.br/damt/arquivos/1.pdf>. Acesso em: 9 jun. 2019.

AMARAL, F. C. N. do. *Data Mining - Técnicas e aplicações para marketing direto*. São Paulo: Berkley, 2001.

- BONADIO, M. C. A produção acadêmica sobre moda na pós-graduação *Stricto Sensu* no Brasil. *IARA: revista de Moda, Cultura e Arte*. São Paulo, v.3, n.3, 2010.
- CALDAS, D. *Observatório de sinais: Teoria e prática da pesquisa de tendências*. Rio de Janeiro: Senac, 2006.
- GOLDSCHMIDT, R.; PASSOS, E. *Data Mining – um guia prático*. Rio de Janeiro: Elsevier, 2005.
- HORN, B. S. *Uma análise da pesquisa em design de moda no Brasil a partir de periódicos da área*. Dissertação (Mestrado em Design). UniRitter, Porto Alegre, 2014.
- LINDE, K.; WILLICH SN. How objective are systematic reviews? Differences between reviews on complementary medicine. *Journal of the Royal Society of Medicine*, 90:845-8. 2003. Disponível em: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC539366/> Acesso em 10 de mar. 2019.
- NISBET, R.; ELDER, J.; MINER, G. *Handbook Statistical Analysis Data Mining Applications*. Oxford: Elsevier, 2009.
- PETTICREW, C.; ROBERTS, H. *Systematic Reviews in the Social Sciences: A practical guide*. Oxford, UK: Blackwell Publishing, 2006.
- PIRES, D. B. A história dos cursos de design de moda no Brasil. *Revista Nexos: Estudos em Comunicação e Educação*. São Paulo, n.9, ano VI, 2002.
- REA, L. M.; PARKER, R. A. *Designing and Conducting Survey Research: A Comprehensive Guide*. New Jersey: Wiley/Jossey-Bass, 2005.
- RIBEIRO, V. G.; SILVEIRA, S. R.; BISCONSIN, E. P.; ZABADAL, J. G. Uma análise conceitual sobre métodos de pesquisa utilizados em Design. *Revista D (UniRitter)*, v. 3, p. 97-112, 2011.
- RIBEIRO, V. G.; ZABADAL, J. G. ; SILVEIRA, S. R. O Emprego da Teoria dos Grafos para Análise de Dados de Publicações em Design: Design Studies (2001-2011). In: Semana de Extensão, Pesquisa e Pós-graduação do Centro Universitário Ritter dos Reis, 2016, Porto Alegre. Anais da Semana de Extensão, Pesquisa e Pós-graduação do Centro Universitário Ritter dos Reis. Porto Alegre: Editora UniRitter, 2016. v. 12. p. 10-16.
- SAMPAIO, R.F.; MANCINI, M.C. Estudos de revisão sistemática: Um guia para síntese criteriosa da evidência científica. *Revista Brasileira de Fisioterapia*. São Carlos, v.11, n.1, p. 83-89. 2007.
- SAMPIERI, C. R. H. ;COLLADO, C. F.; LUCIO, P.B; *Metodología de la investigación*. Colombia: Panamericana Formas e Impresos S.A., 1997.
- TAN, P.; STEINBACH, M.; KUMAR, V. *Introduction to Data Mining*. London: Addison Wesley, 2005.