

**UNIVERSIDADE DO VALE DO RIO DOS SINOS - UNISINOS  
UNIDADE ACADÊMICA DE PESQUISA E PÓS-GRADUAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM GESTÃO E NEGÓCIOS  
NÍVEL MESTRADO**

**MARCO ANTONIO TROIS ENDRES**

**APLICAÇÃO DA MINERAÇÃO DE OPINIÃO NO PLANEJAMENTO TURÍSTICO  
DO MUNICÍPIO DE GRAMADO/RS**

**São Leopoldo  
2016**

MARCO ANTONIO TROIS ENDRES

APLICAÇÃO DA MINERAÇÃO DE OPINIÃO NO PLANEJAMENTO  
TURÍSTICO DO MUNICÍPIO DE GRAMADO

Projeto de Qualificação apresentado como requisito parcial para obtenção do título de Mestre em Gestão e Negócios, pelo Programa de Pós-Graduação em Gestão e Negócios da Universidade do Vale do Rio dos Sinos - UNISINOS

Orientador: Prof. Guilherme Luís Roehe Vaccaro

São Leopoldo

2016

#### Dados Internacionais de Catalogação na Publicação (CIP)

E56a Endres, Marco Antonio Trois  
Aplicação da mineração de opinião no planejamento turístico do município de Gramado / Marco Antonio Trois Endres ; Orientador Guilherme Luís Roehe Vaccaro. -- 2016. 115 f.

Dissertação (Mestrado) -- Universidade do Vale do Rio dos Sinos, Programa de Pós-Graduação em Gestão e Negócios, Porto Alegre, 2016.

1. Turismo. 2. Planejamento. 3. Mineração de opinião. I. Vaccaro, Guilherme Luís Roehe, orient. II. Título.

CDU 338.48

MARCO ANTONIO TROIS ENDRES

APLICAÇÃO DA MINERAÇÃO DE OPINIÃO NO PLANEJAMENTO  
TURÍSTICO DO MUNICÍPIO DE GRAMADO/RS

Projeto de Qualificação apresentado como  
requisito parcial para obtenção do título de  
Mestre em Gestão e Negócios, pelo  
Programa de Pós-Graduação em Gestão  
e Negócios da Universidade do Vale do  
Rio dos Sinos - UNISINOS

Aprovado em 28 de Abril de 2016

BANCA EXAMINADORA

---

Prof. Dr. Guilherme Trez - UNISINOS

---

Prof. Dr. Sílvio César Cazella – UFCSPA

---

Prof. Dr. Fabrício Farias Tarouco - UNISINOS

Orientador: Prof. Guilherme Luís Roehe Vaccaro

Visto e permitida impressão

São Leopoldo, 01/06/2016

Prof. Dr. Guilherme Trez

Coordenador do PPG em Gestão e Negócios

*Dedico este trabalho aos meus pais, Júlio e Lilia, que me forneceram a base de valores que me guiam.*

## **AGRADECIMENTOS**

*A todos os meus colegas da turma do MPGN/2014, acima de tudo pela união. Esta conquista não teria sido possível sem vocês.*

*A Unisinos, na figura da coordenação do curso e seus professores, pelos ensinamentos que me foram passados, pelo convívio, pelo carinho e atenção dispensada, em especial, Guilherme Trez, Flávia Haas, Patrícia Fagundes, Marcelo Machado, Fonseca, Oscar Kronmeyer e Pierre Fayard. A experiência foi surpreendente.*

*A Rosa Helena Volk e Kalinka Silveira, pela disposição e interesse nesta pesquisa e à todos os demais que contribuíram de alguma forma com o desenvolvimento deste trabalho.*

*Ao grupo escoteiro Guia Lopes, chefias e lobinhos, que nos dias em que eu não estive presente, fizeram o melhor possível!*

*Ao meu orientador Guilherme Vaccaro, pela paciência ilimitada, pela humildade, pelos conhecimentos que me foram passados e pelo exemplo como profissional.*

*A minha esposa Clarissa e ao meu filho Bernardo, pelo apoio incondicional e por me acompanharem sempre, abdicando comigo dos momentos de lazer.*

*A DEUS, por sempre iluminar o meu caminho.*

*Muito obrigado à todos!*

*"Loucura? Sonho? Tudo é loucura ou sonho no começo. Nada do que o homem fez no mundo teve início de outra maneira, mas já tantos sonhos se realizaram que não temos o direito de duvidar de nenhum".*

*(Monteiro Lobato)*

## RESUMO

O propósito deste estudo é explorar o processo de descoberta de conhecimento e analisar as oportunidades geradas pela Mineração de Opinião como técnica para se obter um retorno sobre experiência do turista em relação aos produtos e serviços ofertados pelo destino turístico. Entender o turista quanto ao seu comportamento de compra e seus hábitos de viagem é fundamental para a ampliação do mercado turístico e melhoria da experiência turística do visitante. Usuários da web têm a oportunidade de registrar e divulgar suas ideias e opiniões através de comentários em redes sociais. Estas opiniões estão disponíveis e em grande volume para as organizações. Neste contexto perguntam-se, quais as contribuições da Mineração de Opinião na geração de informação útil para a gestão da atividade turística, como suporte ao processo de tomada de decisão no planejamento e no aprimoramento das suas ações? Este estudo teve como cenário de investigação o município de Gramado/RS e os comentários registrados em redes sociais pelos turistas que o visitam. Para alcançar o propósito deste estudo, foram extraídas opiniões do Twitter e Facebook e submetidas a uma técnica de análise de sentimentos. Como resultado do estudo, são apresentados e discutidos os resultados da aplicação da Mineração de Opinião consolidados de acordo com as dimensões de competitividade que o município é avaliado.

**Palavras-chave:** Turismo. Competitividade. Mineração de Opinião. Descoberta de Conhecimento

## **ABSTRACT**

The purpose of this study is to explore the knowledge discovery process and analyze the opportunities generated by the Opinion Mining as a technique to obtain a feedback on the tourist experience about products and services offered by the tourist destination. Understanding the tourist about their buying behavior and their travel habits is essential to the expansion of the tourist market and improvement of the tourist experience. Web users have the opportunity to register and show their ideas and opinions through posts on social networks. These opinions are available in high volume to organizations. In this context, what are the contributions of Opinion Mining to generate useful information for the management of tourism activities, to support the decision-making process in planning and improvement of their actions? This study analyses the comments registered on social networks by tourists who visit Gramado/RS. To achieve the purpose of this study, opinions were extracted from Twitter and Facebook and submitted to a sentiment analysis technique. As a result of the study are presented and discussed the results summarized according to the competitiveness of dimensions that the municipality is assessed.

**Keywords:** Tourism. Competitiveness. Opinion Mining. Knowledge Data Discovery.

## LISTA DE ILUSTRAÇÕES

Figura 1 - Problema de Pesquisa .....	18
Figura 2 - Balanço geral – índices de competitividade .....	19
Figura 3 - Sistema de Turismo (SISTUR) – Modelo Referencial .....	25
Figura 4 – Sistema turístico de Oliveira, Zouain e Barbosa .....	28
Figura 5 - Modelo conceitual de competitividade do destino turístico de Ritchie/Crouch.....	32
Figura 6 - A experiência turística no contexto sistêmico .....	38
Figura 7 - Representação do perfil do cientista de dados através do Diagrama de Venn.....	41
Figura 8 - Representação do processo de KDD.....	43
Figura 9 - Etapas da Mineração de Opinião.....	48
Figura 10 - Etapa da Coleta de Dados .....	61
Figura 11 - Extração de Comentários do Twitter com o NodeXL .....	62
Figura 12 - Etapa da Identificação do Sentimento.....	64
Figura 13 - Etapa de Classificação na Mineração de Opinião.....	68
Figura 14 – Polaridade final no nível do documento .....	68
Figura 15 - Etapa de Sumarização ou Agregação das Opiniões.....	69
Figura 16 - Categorização no Semantria.....	70
Figura 17 – Localizações Extraídas .....	77
Figura 18 – Hashtags Extraídos .....	78

## LISTA DE GRÁFICOS

Gráfico 1 - Índices gerais de competitividade – destino x Brasil x não capitais: 2008-2015 .....	16
Gráfico 2 - Índices gerais de competitividade – destino x Brasil x não capitais: 2008-2015 .....	57
Gráfico 3 - Índices do destino por dimensão, em ordem decrescente de desempenho .....	58
Gráfico 4 - Distribuição da Amostra Total.....	63
Gráfico 5 - Distribuição da Amostra do Facebook.....	64
Gráfico 6 - Distribuição da amostra conforme horário do post comentário.....	76
Gráfico 7 - Distribuição da amostra conforme dia do post comentário .....	77
Gráfico 8 - Frequências por Categoria ou Dimensão de Análise.....	84
Gráfico 9 - Frequências por Sentimento x Categoria ou Dimensão de Análise.....	85
Gráfico 10 - Índices Atrativos turísticos – destino x Brasil x não capitais: 2008-2015 .....	87
Gráfico 11 - Índices Monitoramento – destino x Brasil x não capitais: 2008-2015 ....	97

## LISTA DE TABELAS

Tabela 1 - Painel de Indicadores do Observatório de Turismo do RS.....	40
Tabela 2 - Modelo de Apresentação do Semantria .....	66
Tabela 3 - Pontuação final no nível do documento .....	67
Tabela 4 - Motivações – Fatores Pull .....	74
Tabela 5 - Frequências por Sentimento x Categoria ou Dimensão de Análise .....	86
Tabela 6 - Amostra da Dimensão de Análise Atrativos Turísticos.....	88
Tabela 7 - Amostra da Dimensão de Análise (In)segurança .....	89
Tabela 8 - Frequências para a Dimensão Condições Climáticas.....	90
Tabela 9 - Amostra da Dimensão de Análise Condições Climáticas.....	91
Tabela 10 - Comentários sem informação relevante.....	92

## LISTA DE QUADROS

Quadro 1 - Dimensões e Variáveis do Índice de Competitividade do Turismo Brasileiro .....	34
Quadro 2 - análise conceitual dos modelos de competitividade.....	36
Quadro 3 - Novas categorias com critério de avaliação baseadas no desempenho .	37
Quadro 4 - Técnicas, tarefas e algoritmos de Mineração de Dados.....	46
Quadro 5 - Fases, Artefatos e Objetivos .....	60
Quadro 6 - Configuração para Extração do Twitter .....	63
Quadro 7 - Categorização no Semantria.....	73
Quadro 8 - Dicionário criado e configurado para a categorização baseada na Matriz de Conceitos .....	80
Quadro 9 - Dicionário criado e configurado para a categorização baseada em queries com algoritmos booleanos.....	82

## LISTA DE ABREVIATURAS E SIGLAS

API	Application Program Interface
BI	Business Intelligence
CRISP-DM	Cross Industry Standard Process for Data Mining
CT	Classificação Automática de Textos
DCBD	Descoberta de Conhecimento em Base de Dados
DW	Data Warehouse
FGV	Fundação Getúlio Vargas
ICT	Índice de Competitividade do Turismo
KDD	Knowledge Discovery in Database (Mineração de Dados)
KDT	Knowledge Discovery Text (Mineração de Textos)
MTur	Ministério do Turismo do Brasil
NLTK	Natural Language Toolkit
OLAP	On-line Analytical Processing
OMT	Organização Mundial do Turismo
PIB	Produto Interno Bruto
PNL	Programação da Linguagem Natural
PNT	Plano Nacional de Turismo
SEBRAE	Serviço Brasileiro de Apoio a Micro e Pequenas Empresas
SETUR-RS	Secretaria Estadual de Turismo do RS
SMT	Secretaria Municipal de Turismo de Gramado

## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	<b>15</b>
1.1 TEMA e JUSTIFICATIVA DA PESQUISA .....	18
1.2 OBJETIVOS .....	22
<b>1.2.1 Objetivo Geral</b> .....	<b>22</b>
<b>1.2.2 Objetivos Específicos</b> .....	<b>22</b>
1.3 ESTRUTURA DA DISSERTAÇÃO .....	23
<b>2 REFERENCIAL</b> .....	<b>24</b>
2.1 O TURISMO .....	24
<b>2.1.1 Definição</b> .....	<b>24</b>
<b>2.1.2 A Abordagem sistêmica</b> .....	<b>25</b>
<b>2.1.3 O produto turístico como output do sistema turístico</b> .....	<b>26</b>
<b>2.1.4 A retroalimentação do sistema</b> .....	<b>27</b>
2.2 O PLANEJAMENTO DO TURISMO .....	29
<b>2.2.1 A abordagem da competitividade no planejamento turístico</b> .....	<b>31</b>
<b>2.2.2 A criação do Índice de Competitividade do Turismo (ICT)</b> .....	<b>33</b>
<b>2.2.3 A abordagem baseada na dimensão ex-post</b> .....	<b>35</b>
<b>2.2.4 Planejamento do turismo no RS</b> .....	<b>39</b>
2.3 A MINERAÇÃO DE DADOS E TEXTOS .....	41
<b>2.3.1 Conceitos e aplicações</b> .....	<b>41</b>
<b>2.3.2 A Mineração de Dados</b> .....	<b>42</b>
<b>2.3.3 A Mineração de Textos e o Processamento de Linguagem Natural (PNL)</b> .....	<b>46</b>
<b>2.3.4 A Mineração de Opiniões ou Análise de Sentimentos</b> .....	<b>47</b>
<b>2.3.5 A Classificação Automática de Documentos</b> .....	<b>49</b>
<b>2.3.6 A Descoberta de Conhecimento no Turismo</b> .....	<b>50</b>
<b>3 METODOLOGIA</b> .....	<b>54</b>
3.1 TIPO DE PESQUISA .....	54
3.2 UNIDADE DE ANÁLISE .....	55
<b>3.2.1 A Cidade de Gramado</b> .....	<b>55</b>
<b>3.2.2 A Secretaria Municipal de Turismo de Gramado</b> .....	<b>55</b>
<b>3.2.3 O município de Gramado como destino turístico</b> .....	<b>56</b>
3.3 MÉTODO DE TRABALHO .....	58
3.4 PROCEDIMENTOS PARA A MINERAÇÃO DE OPINIÃO .....	60

<b>3.4.1 Etapa da Coleta de Dados .....</b>	<b>61</b>
<b>3.4.2 A identificação da opinião ou sentimento.....</b>	<b>64</b>
<b>3.4.3 A classificação da opinião ou sentimento .....</b>	<b>67</b>
<b>3.4.4 A sumarização ou agregação.....</b>	<b>69</b>
<b>3.5 DELIMITAÇÃO DO ESTUDO.....</b>	<b>73</b>
<b>4 APRESENTAÇÃO DO ESTUDO .....</b>	<b>76</b>
<b>4.1 RESULTADOS DA ETAPA DA COLETA .....</b>	<b>76</b>
<b>4.2 RESULTADOS DA ETAPA DE IDENTIFICAÇÃO E CLASSIFICAÇÃO DO SENTIMENTO.....</b>	<b>78</b>
<b>4.3 RESULTADOS DA ETAPA DE SUMARIZAÇÃO OU AGREGAÇÃO.....</b>	<b>79</b>
<b>4.4 ANÁLISE DE RESULTADOS.....</b>	<b>Erro! Indicador não definido.</b>
<b>4.4.1 Dimensão de Análise: Atrativos Turísticos.....</b>	<b>87</b>
<b>4.4.2 Dimensão de Análise: (In)segurança.....</b>	<b>89</b>
<b>4.4.3 Dimensão de Análise: Condições Climáticas .....</b>	<b>90</b>
<b>4.5 DISCUSSÃO .....</b>	<b>92</b>
<b>5 CONSIDERAÇÕES FINAIS .....</b>	<b>96</b>
<b>5.1 LIMITAÇÕES DA PESQUISA .....</b>	<b>97</b>
<b>5.2 SUGESTÕES PARA TRABALHOS FUTUROS .....</b>	<b>98</b>
<b>REFERÊNCIAS.....</b>	<b>100</b>
<b>APÊNDICE A – CARTA CONVITE.....</b>	<b>108</b>
<b>APÊNDICE B – RELATÓRIO PRELIMINAR PARA A SMT.....</b>	<b>109</b>

## 1 INTRODUÇÃO

Segundo o Ministério do Turismo do Brasil (MTur), a participação do setor turístico na economia brasileira já representa 3,7% do PIB do nosso país, se consolidando como importante atividade econômica para geração de emprego, desenvolvimento social, investimentos em infraestrutura, sustentabilidade e modelagem do ambiente competitivo (BRASIL, 2012).

As políticas públicas, em âmbito nacional e regional, através de planos de desenvolvimento turísticos, avaliam o desempenho dos principais destinos turísticos em relação à um grande número de variáveis. O Plano Nacional de Turismo (PNT), elaborado pelo MTur, com diretrizes da Organização Mundial do Turismo (OMT), pode ser citado como exemplo de política pública de turismo.

O PNT segue uma estrutura lógica que interliga as diretrizes estratégicas, os objetivos e resultados esperados e a proposição e o desenvolvimento de programas, projetos e ações. O Índice de Competitividade do Turismo Nacional (ICT), medido anualmente pela Fundação Getúlio Vargas (FGV) e pelo Serviço Brasileiro de Apoio a Micro e Pequenas Empresas (SEBRAE), tem por objetivo o acompanhamento do desempenho dos destinos turísticos. Os principais destinos turísticos nacionais são avaliados pelo ICT, entre eles o município de Gramado/RS.

O município de Gramado/RS está em posição de destaque no cenário turístico do Brasil. Possui um índice geral de competitividade medido pelo MTur acima da média dos demais municípios do país, além da primeira posição no ranking do site de avaliações Tripadvisor (TRIPADVISOR, 2015). O índice geral alcançado por Gramado indica que o destino se posiciona no nível 4 de competitividade, e registrou estabilidade em relação ao último ano da pesquisa. O resultado é superior à média Brasil e à média das não capitais, conforme é possível observar no gráfico 1:

Gráfico 1 - Índices gerais de competitividade – destino x Brasil x não capitais: 2008-2015



Fonte: BRASIL. MTur; SEBRAE; FGV, 2015.

Segundo o MTur, a vantagem do município de Gramado/RS se deve a 3 fatores principalmente:

- Forte posicionamento turístico no cenário nacional, resultado das ações promocionais realizadas pelo destino e pelo enfoque na qualidade dos serviços prestados aos turistas;
- Realização de eventos tradicionais de forma inovadora, como o Natal Luz, referência no país por suas atrações exclusivas e pela estrutura disponível para recepção e entretenimento dos turistas e;
- Existência de dois órgãos locais de gestão do turismo – a Secretaria Municipal de Turismo e a autarquia Gramadotur –, que contam com orçamentos exclusivos e possuem atribuições distintas, contribuindo para o desenvolvimento turístico.

A Secretaria Municipal de Turismo de Gramado (SMT) necessita, para poder planejar as ações necessárias conforme diretrizes do PNT, de informações estratégicas atualizadas e integradas, muitas delas não disponíveis em simples consultas a bancos de dados.

Nos últimos anos, surgiram várias técnicas e de Descoberta de Conhecimento, tais como a Descoberta de Conhecimentos em Bancos de Dados (KDD) e a Descoberta de Conhecimento em Textos (KDT). Estas técnicas permitem a identificação de padrões de relacionamento entre dados que, adequadamente interpretados, podem levar à abstração de conhecimento inédito e relevante. Juntamente com a evolução destas tecnologias veio também o aumento da quantidade e variedade de informações registradas.

Os diversos tipos de dados existentes atualmente acrescentam mais complexidade ao processo de extração e análise. As análises necessitam de versatilidade para poderem considerar, além de dados estruturados na forma de tradicionais registros no nível de usuário, dados de imagens, dados oriundos de sistemas de geoprocessamento, de tweets, de posts, textos de e-mails, entre outros formatos. A definição da amostragem precisa levar em consideração o tipo de dado analisado.

A Mineração de Textos surgiu nesse contexto. A Mineração de Textos busca a descoberta de regras e padrões em dados que podem trazer conhecimento útil para apoiar os responsáveis pela tomada de decisões. Esses dados podem estar disponíveis em forma de opiniões dos consumidores, disponíveis em redes sociais.

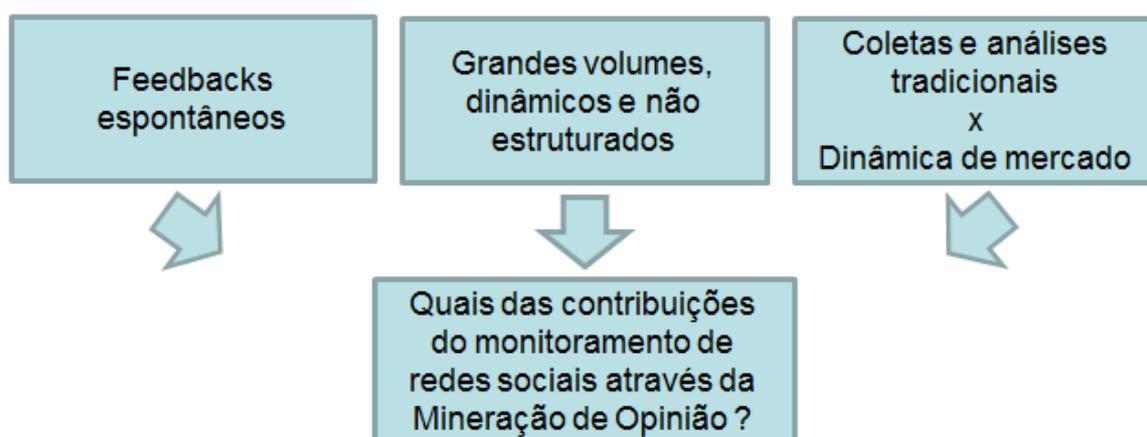
As opiniões possuem grande influência sobre o comportamento das pessoas. Decisões sobre o consumo de um produto ou serviço são frequentemente baseadas em opiniões de pessoas próximas, de especialistas, ou de estudos conduzidos por instituições especializadas. As organizações baseiam suas estratégias de negócio e investimentos na opinião de seus clientes sobre seus produtos ou serviços. Tradicionalmente, a resposta a questões envolvendo a opinião pública envolve técnicas como pesquisa de campo, telefonemas ou questionários escritos. Estas técnicas envolvem custos, são restritas, seu retorno é demorado e muitas vezes pouco eficaz.

O surgimento das redes sociais alterou este cenário, disponibilizando a indivíduos e organizações conteúdo de opinião diversificado e em grandes volumes. Usuários da web têm a oportunidade de registrar e divulgar suas ideias e opiniões

através de comentários, fóruns de discussão, blogs, redes sociais, entre outros. Para as organizações, isto significa oportunidades de ampliar as fontes de opinião quantitativa ou qualitativamente, tornar mais baratas as formas de coleta, e reduzir o tempo necessário para disponibilização da informação. Contudo, o grande volume de informação produzido diariamente implica a necessidade de métodos e ferramentas capazes de processar automaticamente não apenas o conteúdo das publicações, mas também a opinião e sentimento que expressam.

Desta forma, esta pesquisa tem como objetivo responder a seguinte questão ilustrada na figura 1: quais as contribuições da Mineração de Opinião na geração de informação útil para a gestão da atividade turística, como suporte ao processo de tomada de decisão no planejamento e no aprimoramento das suas ações?

Figura 1 - Problema de Pesquisa



Fonte: Elaborado pelo autor.

## 1.1 TEMA E JUSTIFICATIVA DA PESQUISA

O município de Gramado está em posição de destaque no cenário turístico do Brasil. Atualmente, o município sustenta a primeira posição no Relatório Índice de Competitividade do Turismo Nacional. Manter, aprimorar e inovar em seus produtos e serviços requer informação atualizada.

Segundo a SMT, a principal finalidade do ICT para a administração do município é a utilização das informações publicadas para planejar e desenvolver vantagens competitivas, norteando a elaboração de políticas públicas que eliminem, gradativamente, os entraves ao desenvolvimento sustentável da atividade turística do município.

Conforme demonstrado na figura 2, o índice geral diminuiu entre 2013 e 2015:

Figura 2 - Balanço geral – índices de competitividade

Dimensões	BRASIL			NÃO CAPITAIS			GRAMADO		
	2013	2014	2015	2013	2014	2015	2013	2014	2015
ÍNDICE GERAL	58,8	59,5	60,0	53,1	53,4	53,8	66,3	65,2	65,4
INFRAESTRUTURA GERAL	68,6	68,2	67,7	63,8	62,5	61,8	80,9	80,9	76,1
ACESSO	62,6	62,2	61,9	53,8	52,4	52,4	70,4	69,0	72,1
SERVIÇOS E EQUIPAMENTOS TURÍSTICOS	56,8	58,7	59,0	48,1	49,6	49,5	71,3	70,5	73,1
ATRATIVOS TURÍSTICOS	63,2	63,4	63,2	63,4	62,8	62,6	67,3	65,9	65,3
MARKETING E PROMOÇÃO DO DESTINO	46,8	48,4	48,5	44,4	45,7	45,0	56,2	56,5	54,3
POLÍTICAS PÚBLICAS	57,6	58,1	58,9	54,4	53,9	55,3	69,2	66,7	72,0
COOPERAÇÃO REGIONAL	44,6	48,3	50,0	44,9	49,3	51,7	34,1	46,6	38,9
MONITORAMENTO	37,4	36,2	36,3	31,9	30,7	30,4	42,1	33,6	26,3
ECONOMIA LOCAL	63,6	63,6	64,7	55,2	54,8	55,8	74,2	71,8	75,4
CAPACIDADE EMPRESARIAL	61,2	61,9	62,7	43,5	44,8	45,7	66,8	60,6	61,3
ASPECTOS SOCIAIS	59,4	59,7	60,5	56,7	56,8	57,9	70,7	66,7	68,6
ASPECTOS AMBIENTAIS	67,7	67,3	68,2	63,6	62,4	63,5	71,6	69,6	67,3
ASPECTOS CULTURAIS	58,2	62,0	64,0	52,4	55,6	57,6	58,7	62,5	65,9

Fonte: BRASIL. MTur; SEBRAE; FGV, 2015.

Em 2015, a cidade de Gramado atingiu uma nota dois décimos acima da nota de 2014, passando de 65,2% para 65,4%. Em 2016, a cidade figura em 25º lugar no ranking geral de competitividade, num total de 65 destinos avaliados. O índice geral alcançado pelo município indica que o destino se posiciona no nível 4 de competitividade, o resultado é superior à média Brasil e à média das não capitais. As dimensões com os maiores índices são Infraestrutura geral, Economia local, Serviços e equipamentos turísticos, Acesso, Políticas públicas, Aspectos sociais, Aspectos ambientais, Aspectos culturais, Atrativos turísticos e Capacidade empresarial.

Segundo o MTur, a metodologia utilizada no Índice de Competitividade se baseia no conceito de que o destino desenvolve e aprimora a capacidade de gerar negócios. Isto é, adota-se a compreensão de que os destinos podem se autoanalisar, visando a identificação dos pontos fortes e fracos de cada um e, assim,

contribuindo para o planejamento e o desenvolvimento de vantagens competitivas. Esta abordagem não leva em consideração a avaliação da demanda sobre os produtos e serviços ofertados, fator que impede que o destino turístico se adapte às mudanças de mercado.

Analisando os índices que compõem o índice geral, pode ser observado que os principais índices que colaboraram para a diminuição do índice geral do município foram:

- a) Acesso
- b) Serviços e equipamentos turísticos
- c) Atrativos turísticos
- d) Monitoramento

Quanto a este último em especial, diz respeito aos investimentos em (i) pesquisa de demanda; (ii) pesquisa de oferta; (iii) sistemas de estatísticas do turismo; (iv) medição dos impactos da atividade turística; e (v) setor específico de estudos e pesquisas. A dimensão Monitoramento refere-se, em parte, aos objetivos propostos por este trabalho que, por si só, colabora com a melhoria do índice nesta dimensão.

Conforme o ICT (BRASIL, 2015), a pesquisa em turismo proporciona um diagnóstico de determinada situação, com base no grau de conhecimento acerca da avaliação dos turistas sobre os serviços ofertados no local visitado, assim como satisfação, hábitos, atitudes e expectativas. Entender o turista - seu comportamento de compra e seus hábitos de viagem - é, pois, fundamental para a ampliação do mercado turístico e melhoria da experiência turística do visitante.

Dessa forma, nesta variável, Pesquisa de Demanda, se averigua a realização de pesquisas periódicas e contínuas de demanda no destino, considerando informações fundamentais: dados de perfil socioeconômico; informações sobre a viagem (motivo da viagem, forma de organização, tipo de alojamento utilizado etc); gasto médio do turista; permanência média do turista no destino; grau de satisfação em relação à viagem e/ou intenção de retorno e a aplicabilidade e o aproveitamento dos resultados desta pesquisa para elaboração de políticas públicas, planejamento, marketing e promoção do destino.

A pesquisa em turismo proporciona um diagnóstico de determinada situação, com base no grau de conhecimento a cerca da avaliação dos turistas sobre os serviços ofertados no local visitado, assim como satisfação, hábitos, atitudes e

expectativas. Entender o turista quanto ao seu comportamento de compra e seus hábitos de viagem é fundamental para a ampliação do mercado turístico e melhoria da experiência turística do visitante.

A Internet, por exemplo, tem reformulado a forma como as pessoas planejam e fazem suas viagens, afetando significativamente o sistema de turismo. Os consumidores geram várias formas de conteúdo em *websites* tais como redes sociais, blogs, comunidades virtuais, wikis, redes sociais e sites de compartilhamento de mídia como o Twitter e o Facebook. Estes sites são cada vez mais populares entre os viajantes ditos *on-line* da internet. Muitos desses *websites* sociais ajudam os consumidores a publicar e compartilhar suas experiências pessoais que, então, podem servir como informações para os outros (XIANG; GRETZEL, 2010).

Considerado este padrão de comunicação, a Mineração de Opiniões ou Análise de Sentimentos pode apoiar no processo de tomada de decisões fornecendo informações para a SMT, demais secretarias e demais atores envolvidos na atividade turística sobre a percepção do turista quanto aos serviços e produtos oferecidos em Gramado, pois:

- a) O turista deixa rastro de informações através do uso de tecnologias móveis, redes sociais e comércio eletrônico;
- b) Estes dados são multidimensionais, dinâmicos e volumosos. Necessitam de recursos computacionais para obtenção, análise e armazenamento;
- c) Entidades do governo buscam informações em fontes tais como companhias de transporte, pesquisas de opinião, operadores turísticos e fontes pulverizadas na web. Isso leva tempo e as ações decorrentes da análise destas informações não acompanham a velocidade com que o mercado de turismo se modifica.

As informações cuja existência o presente trabalho propõe investigar, podem também apoiar no estudo da concorrência. Segundo Porter (2004), muitas companhias não coletam informações sobre os concorrentes de forma sistemática, mas agem na base de impressões informais, de conjecturas e da intuição derivada de um pequeno número de informações que os gestores recebem dos concorrentes. Opiniões sobre os produtos e serviços ofertados pelos destinos turísticos concorrentes estão disponíveis em várias redes sociais.

Assim, esta pesquisa contribui com a literatura e com os gestores públicos do município de Gramado ao analisar as oportunidades geradas pelo uso da Mineração de Opiniões em redes sociais como ferramenta de apoio às secretarias de governo do município de Gramado/RS no planejamento das suas ações voltadas ao desenvolvimento do turismo, de forma geral e específicas quanto à dimensão Monitoramento.

O trabalho também colabora com o desenvolvimento do turismo de forma geral, uma vez em que as oportunidades oferecidas por este estudo podem ser estendidas para outros atores envolvidos com a atividade turística e para outros municípios da região, apoiando a gestão do turismo de forma integrada.

A seguir, são detalhados os objetivos e delimitações da pesquisa e estrutura do projeto.

## 1.2 OBJETIVOS

### 1.2.1 Objetivo Geral

Este estudo tem como objetivo geral analisar os benefícios e oportunidades do uso da Mineração de Opinião no apoio à tomada de decisão no processo de planejamento turístico do município de Gramado/RS.

### 1.2.2 Objetivos Específicos

- I. Elaborar um framework que justifique a utilização da Mineração de Opinião integrada à metodologia do ICT utilizada na avaliação do município de Gramado/RS e que apoie na consecução dos demais objetivos específicos estabelecidos;  
Realizar um estudo piloto com a finalidade de testar o framework através da configuração de uma ferramenta de que permita a mineração de opiniões de turistas do município de Gramado/RS organizadas de acordo com as dimensões sobre as quais o município é avaliado e;
- II. Analisar e discutir os resultados encontrados à luz da avaliação de referencial teórico sobre a dimensão analisada.

### 1.3 ESTRUTURA DA DISSERTAÇÃO

A apresentação deste estudo está estruturada em 5 capítulos. O capítulo 1 contém a introdução, a apresentação do tema e da justificativa da pesquisa, seguido dos objetivos e estrutura do projeto.

O capítulo 2 trata do referencial utilizado para a elaboração da pesquisa, abordando os temas turismo, planejamento turístico, Mineração de Dados e Textos e os casos de uso na área do turismo.

No capítulo 3 são tratados os procedimentos metodológicos que foram seguidos para se alcançar os objetivos propostos por este estudo e envolvem a aplicação da Mineração de Opiniões propriamente dita. No capítulo 4 é apresentada a análise dos resultados encontrados após a aplicação da técnica de Mineração de Opinião proposto.

Ao final, são apresentadas as conclusões, propostas de trabalhos futuros, as referências bibliográficas referenciadas ao logo deste projeto e um Apêndice contendo a carta-resposta formalizando a intenção de colaboração mútua entre a SMT, o Autor e a Universidade e a estrutura do questionário submetido à SMT para avaliação dos resultados encontrados.

## **2 REFERENCIAL**

As teorias podem ser entendidas como resultados das ações de observar, examinar e especular a respeito de algum fenômeno. A visão de alguns autores a respeito dos temas Mineração de Dados, Mineração de Textos e Opiniões e Turismo auxiliarão as análises da pesquisa que será desenvolvida, a formulação das hipóteses para pesquisa e a contextualização das futuras conclusões.

### **2.1 O TURISMO**

Esta seção trata da definição e a abordagem sistêmica do Turismo, seguido dos conceitos que envolvem o planejamento do turismo até o modelo de planejamento adotado pelo MTur para avaliar os principais destinos indutores da atividade turística no Brasil. A revisão destes conceitos é essencial para o entendimento da metodologia de avaliação do destino turístico através do ICT e sua interseção com a Mineração de Opinião, tema abordado na segunda parte deste referencial.

#### **2.1.1 Definição**

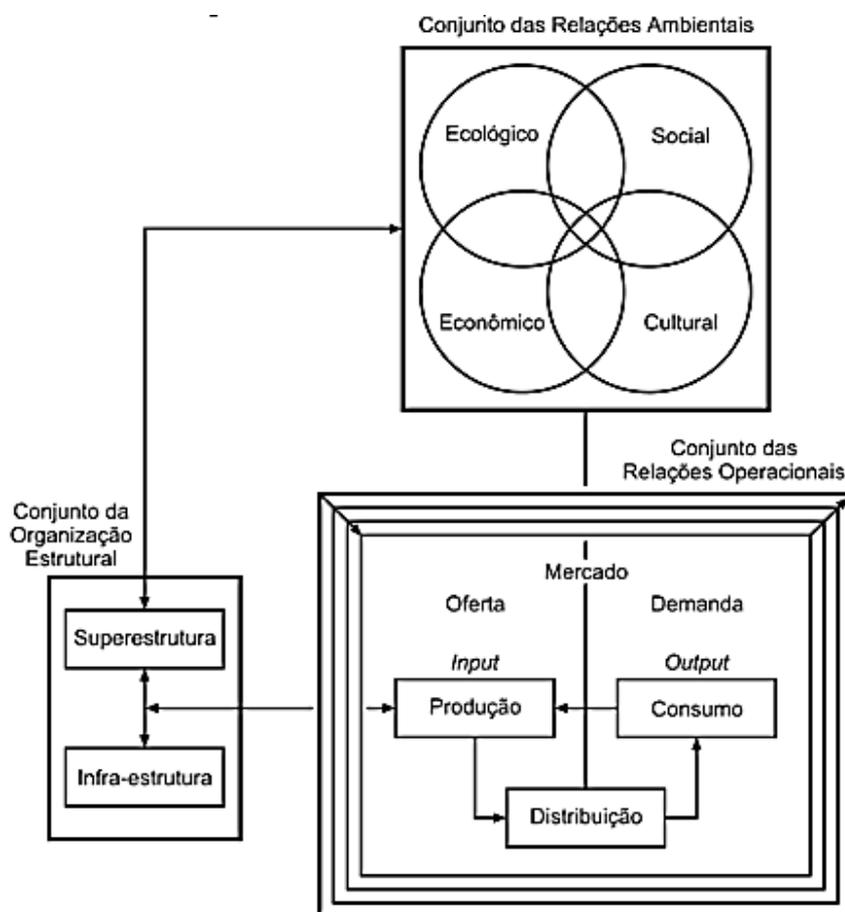
O turismo é um fenômeno social, cultural e econômico que se caracteriza pelo deslocamento de pessoas para lugares fora do seu ambiente habitual para fins pessoais ou profissionais. Estas pessoas são chamadas visitantes. O turismo tem a ver com as suas atividades, algumas das quais envolvem despesas turísticas (WORLD TOURISM ORGANIZATION, 2014). Segundo Beni (1990), o turismo é uma atividade que resulta do somatório de recursos naturais do meio ambiente, culturais, sociais e econômicos e, assim, o campo de seu estudo é abrangente, complexo e multicausal e, para organizar esse complexo de fatores, é preciso dispor de um quadro referencial dinâmico, flexível, adaptável, de fácil leitura e compreensão e integrado, o que pode ser instrumentalizado pela Teoria dos Sistemas.

### 2.1.2 A Abordagem sistêmica

O turismo pode ser visto pela ótica de Bertalanffy com a *Teoria Geral dos Sistemas*, que permite a sua análise como um todo ou dividido em partes para melhor compreensão do seu estudo. Existem várias abordagens sobre turismo e muitas delas são baseadas na teoria geral dos sistemas.

Entre as abordagens sistêmicas existentes e mais difundidas no Brasil está o sistema turístico SISTUR, proposto por Mário Carlos Beni (2001), ilustrado na figura 3:

Figura 3 - Sistema de Turismo (SISTUR) – Modelo Referencial



Fonte: BENI, 2001.

Beni (2001) identificou os elementos do SISTUR, dividindo-os em três grandes conjuntos: Relações Ambientais, Organização Estrutural e Ações Operacionais, quais sejam:

- Conjunto das Relações Ambientais: compreende os subsistemas ecológico, econômico, social e cultural;
- Conjunto da Organização Estrutural: engloba os subsistemas da superestrutura e da infraestrutura. A superestrutura é o conjunto de normas, regras e leis que regulam o funcionamento da atividade turística. Já a infraestrutura é composta pela estrutura de acesso, a estrutura urbana e outros serviços básicos, como o abastecimento de água e;
- Conjunto das Ações Operacionais: compreende a oferta e a demanda, o consumo e a distribuição de produtos turísticos.

### **2.1.3 O produto turístico como output do sistema turístico**

A Organização Mundial do Turismo (1998) utiliza o termo “produto turístico” para definir o que seria a oferta ou output do sistema turístico. Os atrativos e as atividades turísticas, as infraestruturas que servem aos atrativos, as acomodações, enfim esse conjunto de serviços e facilidades compõe o produto turístico. Em outras palavras, o processamento do produto turístico é objetivo final do sistema turístico.

Para Oliveira (2013), no processo de formação do produto turístico participam empresas de turismo e de suporte que se fazem uso de infraestruturas para fornecer ao turista uma experiência única. Por isso, quanto maior for a sintonia entre essas empresas, melhor será a qualidade da experiência turística. A sinergia entre os atores e organizações do destino, quando trabalham em rede, confere ao sistema turístico local um perfil dinâmico capaz de fazer frente a mudanças do ambiente externo.

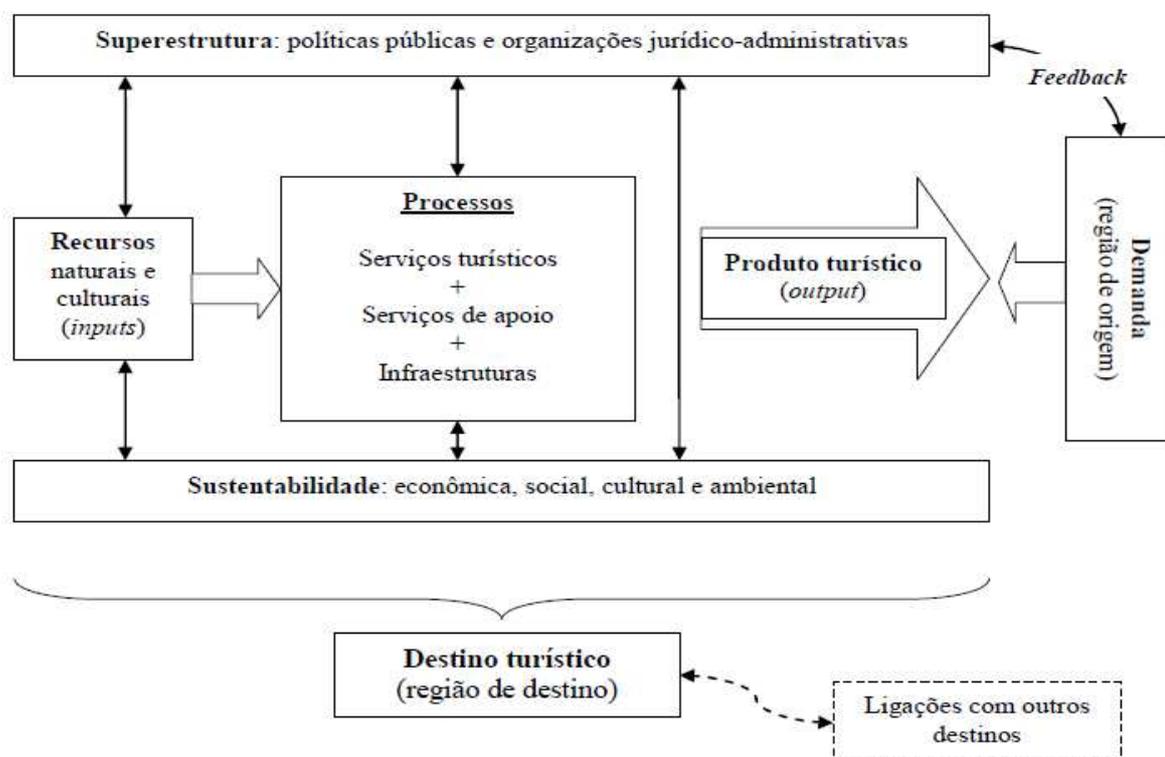
A OMT (ORGANIZAÇÃO MUNDIAL DO TURISMO, 1998) utiliza o termo “produto turístico” para definir o que seria a oferta ou o output do sistema turístico. Os atrativos e as atividades turísticas, as infraestruturas que servem aos atrativos, as acomodações, enfim esse conjunto de serviços e facilidades compõe o produto turístico. Em outras palavras, o processamento do produto turístico é objetivo final do sistema turístico.

#### **2.1.4 A retroalimentação do sistema**

Para Oliveira (2013), considerando que todo sistema deve possuir um mecanismo de realimentação (feedback) para que ele possa se ajustar às variações do ambiente, o sistema turístico pode basear-se em pesquisas de satisfação com o turista para avaliar a qualidade do produto turístico. A demanda é a razão da existência e do funcionamento de todo o sistema turístico, cujo output (produto turístico) deve satisfazer às suas necessidades. O equilíbrio do sistema é representado pelas dimensões da sustentabilidade: econômica, social, cultural e ambiental. O feedback é obtido pela pesquisa de satisfação com a demanda. O contato com o ambiente externo ao sistema é realizado por meio da demanda e por pontes ou redes com outros destinos turísticos.

Com base no modelo de sistema de turismo definido por Beni (2001), Leiper (1979), OMT (ORGANIZAÇÃO MUNDIAL DO TURISMO, 1998), Uhlmann (2002), Oliveira, Zouain e Barbosa (2013) elaboraram um novo sistema turístico tendo como unidade de análise o município e ressaltando os seguintes subsistemas: o produto turístico, os processos produtivos, os recursos, as organizações jurídico-administrativas, a sustentabilidade e a demanda. Esta última está no ambiente externo ao sistema turístico do destino e entrará em contato com ele pelo consumo do produto turístico. Segundo os autores, por essa visão sistêmica, a organização político-administrativa do destino está incumbida de coordenar as redes de cooperação, monitorar a qualidade do produto turístico (feedback) e a sustentabilidade do destino, conforme figura 4:

Figura 4 – Sistema turístico de Oliveira, Zouain e Barbosa



Fonte: OLIVEIRA; ZOUAIN; BARBOSA, 2012.

Do ponto de vista do turista, a experiência do turismo é vista como integral e está descrita na literatura como uma amálgama de serviços turísticos e instalações localizadas sobre um território compartilhado (OLIVEIRA, 2013). Assim, atrações turísticas e instalações são vistos como uma combinação de experiências de consumo que envolve ambos os atores públicos e privados (YOON; UYSAL, 2005).

A avaliação desta experiência pelo consumidor pode ser traduzida pela satisfação dele com a qualidade dos produtos e serviços prestados ou, simplesmente, o produto turístico. Para Oliver (2014), satisfação é a resposta da realização do consumidor, no sentido de avaliar se as características do bem ou serviço atendem às suas necessidades, proporcionando um nível prazeroso de consumo.

Por meio da visão sistêmica do destino percebe-se a complexidade de fatores que interagem para fornecer o output do sistema, ou o produto turístico. Esse produto pode ser traduzido pela experiência turística do visitante enquanto em contato com o destino e sua opinião (feedback), satisfatória ou insatisfatória, determinará o desempenho do sistema.

Para Oliveira (2013), a experiência de consumo no turismo é complexa, resultante da interação do indivíduo com um ou vários serviços (transporte, hospedagem, centros de visitantes, lazer, etc.), bem como com os elementos ambientais e públicos (clima, infraestruturas, população residente). Mostrou-se que muitos estudos exploram o desempenho de um destino por meio da análise de declaração de satisfação do turista em relação a diferentes aspectos do destino. A identificação dos fatores que fazem surgir a (In)satisfação do consumidor serviria para estabelecer dois objetivos simultâneos aos gestores dos destinos: minimizar a insatisfação e maximizar a satisfação.

A avaliação da demanda sobre a experiência turística pode ser considerada como uma medida de desempenho da competitividade do destino (análise ex-post). Por outro lado, a identificação dos fatores que induziu a satisfação ou insatisfação (análise ex-ante) auxiliaria na análise da competitividade sistêmica e estrutural.

Assim, se a perspectiva do turista sobre o destino fosse incorporada aos modelos de competitividade, haveria duas vantagens. Primeiro, uma medida de desempenho para ajustar o sistema às condições da demanda. Segundo, confrontar a importância dos fatores de competitividade, elaborado pelos estudiosos, com a opinião dos consumidores.

## 2.2 O PLANEJAMENTO DO TURISMO

Segundo Lohmann e Netto (2008), o planejamento turístico é um processo que visa, a partir de uma situação dada, a orientar o desenvolvimento turístico de um empreendimento, local, região, município, estado ou país, tendo como meta alcançar os objetivos propostos.

Independentemente do modelo de planejamento a ser seguido, algumas etapas básicas necessitam ser seguidas. Marques e Bissoli (2000), Barretto (2000), Rose (2002) e Ruschmann (1997, 2002) concordam com a necessidade das seguintes etapas:

- a) Análise do ambiente (externo/interno): é a etapa na qual se cria uma visão geral da área em que será desenvolvido o planejamento.
- b) Diagnóstico: etapa destinada a levantar o maior número de dados sobre o local a ser planejado.

- c) Prognóstico: etapa destinada para a análise das possibilidades de desenvolvimento do turismo, baseada em dados estatísticos anteriores, na política governamental para a região planejada e nas tendências de desenvolvimento do turismo para o destino em questão.
- d) Estratégia: é a forma como será conduzida a implantação da atividade turística para que os objetivos estabelecidos sejam alcançados.
- e) Implantação: é a fase destinada para a operacionalização da estratégia definida.
- f) Avaliação: destinada à avaliação dos resultados e aplicação de melhorias, e reformulação de estratégias.

O Planejamento, do ponto de vista governamental, consiste em um processo para estabelecer objetivos e definir linhas de ação e planos detalhados para atingi-los. É possível distinguir os seguintes pontos básicos do planejamento: estabelecer objetivos, definir cursos de ação e determinar as necessidades de recursos. (BENI, 1999).

Para Hall (2004), são sete as funções do poder público no turismo: coordenação, planejamento, legislação e regulamentação, o governo deve incentivar e difundir o turismo e ocupar-se de oferecer programas de turismo social. Segundo Beni (2006):

A Política Nacional de Turismo deve funcionar como o elo integrador entre as esferas de decisão governamental, representadas pelos segmentos oficiais do sistema e os beneficiários e usuários dessa política. No seu interior, encontram-se as diretrizes e políticas básicas que expressam os caminhos para atingir objetivos nacionais para o turismo [...]. (BENI, 2006, p.16).

Segundo Trigo (2002), a desenvolvimento permanente de programas regionais de turismo em todas as regiões do Brasil, articulados com programas de âmbito nacional, não deve ser de obrigação exclusiva dos governos. Os empresários, investidores e a sociedade civil em geral devem participar desses projetos para que o turismo se consolide como um setor dinâmico, lucrativo e com crescimento não apenas na quantidade de turistas ou no volume de divisas em circulação, mas na qualidade dos serviços prestados e da mão de obra utilizada.

Vale salientar que, apesar da relação próxima entre políticas públicas de turismo e planejamento turístico, trata-se de conceitos diferentes. A política pública

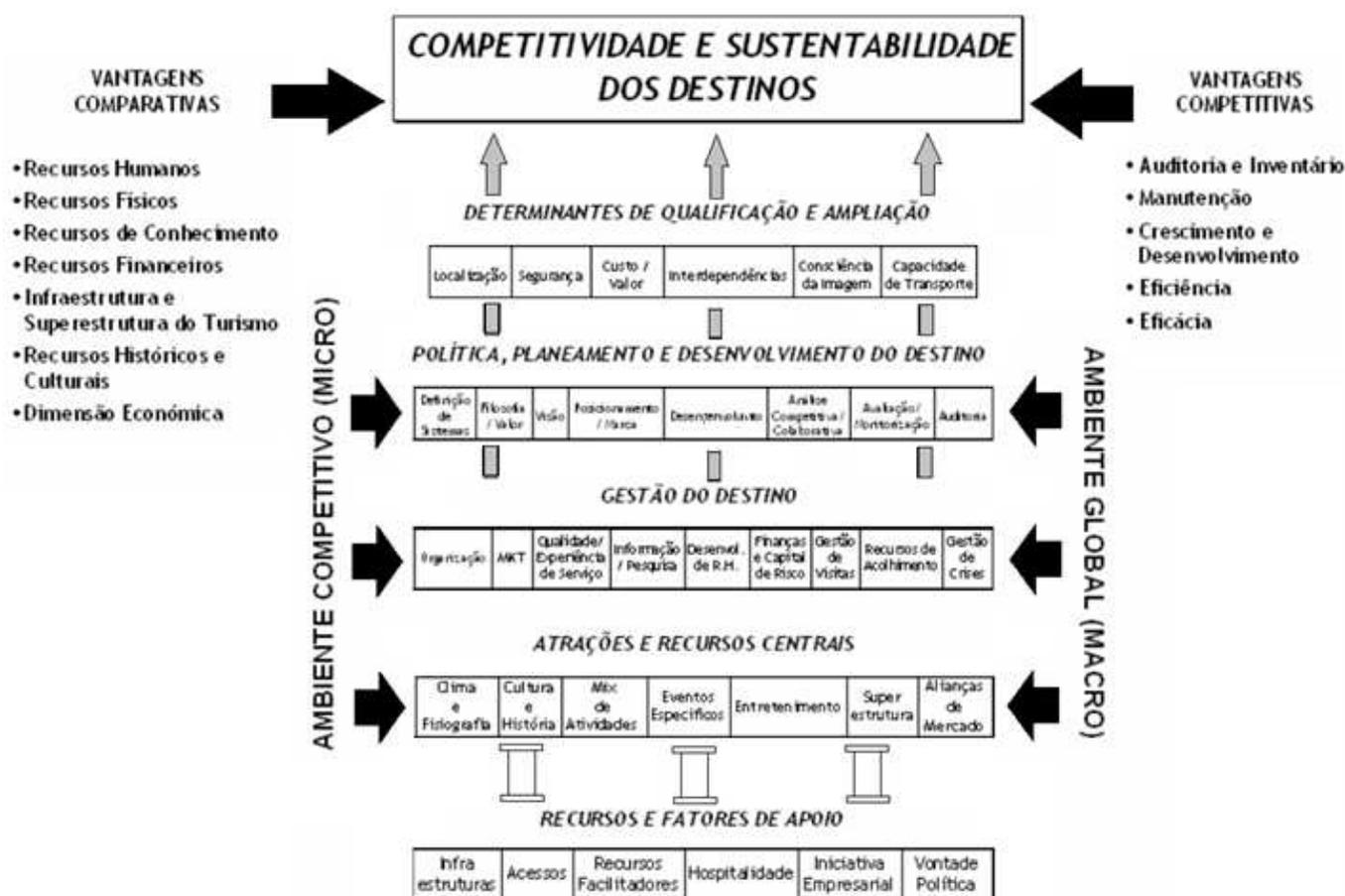
de turismo estabelece as diretrizes gerais de desenvolvimento turístico de um país, estado ou município, expressas em um plano de turismo, como é o caso do PNT. O planejamento deve consultar e se basear na política pública de turismo para ser desenvolvido, transformando-se assim num instrumento viabilizador de ações (LOHMANN; NETTO, 2008).

### **2.2.1 A abordagem da competitividade no planejamento turístico**

A competitividade vem sendo adotada como diretriz central em diversas análises econômicas nacionais e internacionais. O MTUR tem a competitividade como diretriz estratégica do Plano Nacional de Turismo e conceitua competitividade como a capacidade crescentede gerar negócios nas atividades econômicas relacionadas com o setor de turismo, de forma sustentável, proporcionando ao turista uma experiência positiva. Vários estudos acadêmicos propõe modelos para análise da competitividade no turismo.

Muitos dos modelos de competitividade adotados atualmente tem como base modelos como o de Ritchie e Crouch (2010) apresentado na figura 5:

Figura 5 - Modelo conceitual de competitividade do destino turístico de Ritchie/Crouch



Fonte: RITCHIE e CROUCH, 2010.

Segundo Alves e Nogueira (2014),

[...] esses modelos constituem a base teórica para o desenvolvimento dos modelos mais recentes e tratam-se essencialmente de modelos teóricos, conceituais e descritivos, portanto, que procuram definir os fatores que influenciam a competitividade dos destinos turísticos. São fortemente baseados no conceito de criação de vantagens competitivas proposto por Porter[...] (1990<sup>1</sup> apud ALVES; NOGUEIRA, 2014, p.4).

Para Patrucco e Barreto (2011), a competitividade de uma destinação turística exige que o setor público, o setor privado e a sociedade assumam uma atitude de cooperação para trabalhar de forma conjunta focalizando: a melhoria da atratividade do destino, a melhoria da eficácia do marketing, a melhoria da produtividade, o controle dos limites ambientais e culturais e a melhoria da gestão deste complexo e diverso fenômeno social.

<sup>1</sup> PORTER, Michael E. The competitive advantage of notions. *Harvard business review*, v. 68, n. 2, p. 73-93, 1990

### 2.2.2 A criação do Índice de Competitividade do Turismo (ICT)

O MTur (BRASIL, 2013) definiu sete áreas estratégicas, que vão da focalização e economicidade das obras de infraestrutura até a criação de um ambiente de negócios favorável, para alavancar as potencialidades do turismo no Brasil, com a finalidade de colocar o Brasil em lugar de destaque no cenário internacional e estimular o turismo doméstico.

Segundo o MTur (BRASIL, 2013), em 2030 a população do Brasil será de 223 milhões no Brasil. A expectativa de vida aumentará, juntamente com o número de idosos. Em 2030, haverá mais de um bilhão de idosos no mundo e cerca de 30 milhões no Brasil, representando 13% da população brasileira. Isso significa mais pessoas, com mais tempo e dinheiro para viagens e lazer. Para suportar o crescimento previsto, desenvolvimento do turismo deve ser apoiado por políticas públicas. A responsabilidade do poder público pelo desenvolvimento do turismo, muito embora deva considerar as peculiaridades regionais na promoção da articulação entre diversos setores e localidades, não pode deixar de fazer parte de um projeto nacional de desenvolvimento do turismo, pautado numa política central de planejamento tendente a subsidiar as diretrizes das políticas públicas de turismo traçadas em meio à heterogeneidade regional do país (KNUPP et al., 2012).

O MTur, através do PNT, traça diretrizes que norteiam o desenvolvimento do turismo brasileiro, entre as quais a regionalização como uma abordagem territorial e institucional para o planejamento, conforme recomendação da OMT. O plano, além de apoiar a gestão, a estruturação e a promoção do turismo no País, essa estratégia de estruturação de destinos com base na regionalização tem como objetivo melhorar a qualidade e aumentar a competitividade do turismo do Brasil (BRASIL, 2013). Knupp et al. (2012), observa que o PNT trata-se de uma política de desenvolvimento que comporta interações e engajamento com políticas de outras naturezas, sejam elas de maior ou menor abrangência.

Com base nesse objetivo, o MTur, juntamente com o SEBRAE e com a FGV, criou o Índice de Competitividade do Turismo Nacional. O índice é calculado a partir de uma coleta de dados baseada em entrevistas realizadas com gestores dos principais municípios destinos do Brasil (BRASIL, 2014). As dimensões e variáveis consideradas são apresentadas no quadro 1:

Quadro 1 - Dimensões e Variáveis do Índice de Competitividade do Turismo Brasileiro

ÍNDICE DE COMPETITIVIDADE DO TURISMO NACIONAL	DIMENSÃO	VARIÁVEIS						
	INFRAESTRUTURA GERAL	Capacidade de atendimento médico para o turista no destino	Fornecimento de energia	Serviço de proteção ao turista	Estrutura urbana nas áreas turísticas			
	ACESSO	Acesso aéreo	Acesso rodoviário	Acesso aquaviário	Acesso ferroviário	Sistema de transporte no destino	Proximidade de grandes centros emissores de turistas	
	SERVIÇOS E EQUIPAMENTOS TURÍSTICOS	Sinalização turística	Centro de atendimento ao turista	Espaço para eventos	Capacidade dos meios de hospedagem	Capacidade do turismo receptivo	Estrutura de qualificação para o turismo	Capacidade dos restaurantes
	ATRATIVOS TURÍSTICOS	Atrativos naturais	Atrativos culturais	Eventos programados	Realizações técnicas, científicas ou artísticas	Diversidade de atrativos, opções e equipamentos de lazer		
	MARKETING E PROMOÇÃO DO DESTINO	Plano de marketing	Participação em feiras e eventos	Promoção do destino	Estratégias de promoção digital			
	POLÍTICAS PÚBLICAS	Estrutura municipal para apoio ao turismo	Grau de cooperação com o governo estadual	Grau de cooperação com o governo federal	Planejamento para a cidade e para a atividade turística	Grau de cooperação público-privada		
	COOPERAÇÃO REGIONAL	Governança	Projetos de cooperação regional	Planejamento turístico regional	Roteirização	Promoção e apoio à comercialização de forma integrada		
	MONITORAMENTO	Pesquisas de demanda	Pesquisas de oferta	Sistema de estatísticas do turismo	Medição dos impactos da atividade turística	Setor específico de estudos e pesquisas		
	ECONOMIA LOCAL	Aspectos da economia local	Infraestrutura de comunicação	Infraestrutura e facilidades para negócios	Empreendimentos ou eventos alavancadores			
	CAPACIDADE EMPRESARIAL	Capacidade de qualificação e aproveitamento do pessoal local	Presença de grupos nacionais e internacionais do setor do turismo	Concorrência e barreiras de entrada	Geração de negócios e empreendedorismo			
	ASPECTOS SOCIAIS	Acesso à educação	Empregos gerados pelo turismo	Uso de atrativos e equipamentos turísticos pela população	Cidadania, sensibilização e participação na atividade turística	Política de enfrentamento e prevenção à exploração de crianças e adolescentes		
ASPECTOS AMBIENTAIS	Estrutura e legislação municipal de meio ambiente	Atividades em curso potencialmente poluidoras	Rede pública de distribuição de água	Rede pública de coleta e tratamento de esgoto	Coleta e destinação pública de resíduos	Patrimônio natural e unidades de conservação no território municipal		
ASPECTOS CULTURAIS	Produção cultural associada ao turismo	Patrimônio histórico cultural	Estrutura municipal para apoio à cultura					

Fonte: Adaptado de BRASIL. Ministério do Turismo; SERVIÇO BRASILEIRO DE APOIO ÀS MICRO E PEQUENAS EMPRESAS; FUNDAÇÃO GETULIO VARGAS, 2015.

O modelo de Barbosa (2012) vem sendo aplicado pelo Ministério do Turismo para a realidade brasileira. Este modelo foi concebido, tendo em vista a experiência do autor e de especialistas no setor, para a análise da competitividade de destinos brasileiros. Desde 2008, o Ministério do Turismo do Brasil (MTur) o utiliza para a mensuração da competitividade de 65 destinos que foram classificados como indutores do desenvolvimento turístico em suas respectivas regiões. A competitividade desses destinos vem sendo monitorada pelo MTur e o modelo de Barbosa é a metodologia adotada para essa medição.

Conforme abordado anteriormente, o MTur desenvolveu o índice de competitividade baseado no conceito de que o destino desenvolve e aprimora a

capacidade de gerar negócios e compreende que os destinos podem se autoanalisar, visando a identificação dos pontos fortes e fracos de cada um e, assim, contribuindo para o planejamento e o desenvolvimento de vantagens competitivas (BRASIL, 2015).

Segundo o MTur (BRASIL, 2015), as abordagens sob a perspectiva mais econômica tendem a considerar os resultados alcançados por ações já executadas, como os demonstrados por indicadores econômicos nacionais e regionais, de mercado, market share, número de turistas, registros de exportações, conquistas de investimentos, isto é, a competitividade é observada em uma dimensão ex-post, sob a ótica do desempenho de um setor, indústria, empresas ou países.

### **2.2.3 A abordagem baseada na dimensão ex-post**

O modelo adotado pelo MTur (BRASIL, 2013) não considera variáveis que possam captar o resultado da experiência turística pois, ao desenvolver tal metodologia para a composição do Índice, optou-se por explorar uma análise voltada para os recursos presentes no destino, ou seja, seguindo a abordagem da RBV (Visão Baseada em Recursos). O próprio autor do modelo sobre o qual foi baseada a metodologia do índice do MTur reconhece isso em seu trabalho:

[...] variáveis tais como a qualidade das experiências turísticas, atitudes, de atitudes de oficiais imigração / alfândega, valores e prioridades das partes interessadas e hospitalidade dos moradores para turistas, entre outros, não fazem parte dos fatores de competitividade do modelo proposto (BARBOSA, 2012, p. 88).

A RBV, cuja ótica é a da eficiência obtida através de formulação estratégica concebida de “dentro para fora”, sugere a análise dos recursos e capacidades da organização que podem ser fonte de vantagem competitiva. Barney (1991), em sua definição para a RBV, apresenta uma noção de recursos como sendo todos os ativos, capacidades, processos organizacionais, recursos financeiros, físicos, humanos, informações e conhecimentos controlados por uma empresa e que possibilitam a esta mesma empresa formular e seguir estratégias eficazes.

Para Oliveira (2013), fica claro que RBV é o arcabouço teórico que sustenta os constructos estudados. Os recursos são fáceis de serem observados documental ou empiricamente, porém as análises baseadas em recursos são, por definição, estáticas, conforme o autor demonstra o comparativo apresentado no quadro 2:

Quadro 2 - análise conceitual dos modelos de competitividade

Modelos de Mensuração de Competitividade de Destinos	Unidade de análise (país, região, cidade, indústria) / sistêmica, estrutural, negócio	Critério de avaliação (eficiência, desempenho)	Teoria predominante (econômica, organizacional)	Fonte dos dados	Fatores de competitividade (*)
Ritchie & Crouch (2003)	Genérico para: país, região, cidade, ou segmento específico do turismo / sistêmica e estrutural	Eficiência.	Organizacional (RBV, Novo Modelo de Organização Industrial)	Pesquisa com especialistas e dados secundários publicados.	- 35 fatores de oferta e 1 fator de demanda; - contribuições de especialistas.
Dwyer & Kim (2003) e seus seguidores	Países / sistêmica e estrutural	Eficiência.	Organizacional (RBV, Novo Modelo de Organização Industrial)	Pesquisa com especialistas e dados secundários publicados.	- 25 fatores de oferta, 3 fatores de macroambiente e 3 fatores de demanda. - contribuições da literatura e de especialistas.
Dwyer et al (2012) AID	Países / sistêmica e estrutural	Eficiência.	Organizacional (RBV, Novo Modelo de Organização Industrial)	Pesquisa com especialistas.	- 64 fatores de oferta, sendo 31 de gerenciamento de destinos e 33 de gerenciamento de empresas; - contribuição da literatura e dos próprios autores.
Gooroochurn & Sugiyarto (2004/7)	Países / estrutural	Eficiência.	Organizacional (RBV)	Dados secundários publicados.	- 23 fatores de oferta; - contribuições da literatura.
FEM (2007, 2009, 2011, 2013)	Países / estrutural	Eficiência.	Organizacional (RBV)	Pesquisa com especialistas e dados secundários publicados.	- 71 fatores de oferta. - contribuições de especialistas.
Barbosa (2012)	Cidades / estrutural	Eficiência.	Organizacional (RBV)	Pesquisa <i>in loco</i> e dados secundários publicados.	- 62 fatores de oferta. - contribuições de especialistas e da literatura.

Fonte: OLIVEIRA, 2013.

Embora o fenômeno da competitividade tenha sido objeto de discussão e produção acadêmica em turismo, pouco se discute sobre a competitividade dos destinos turísticos do ponto de vista da demanda, ou seja, competitividade sob a ótica dos visitantes dos destinos.

Oliveira (2013), em sua tese, defende que os modelos de mensuração da competitividade de destinos turísticos estão estruturados essencialmente com base na oferta e foram concebidos pelos estudiosos e profissionais do turismo, carecendo, portanto, de um instrumento que permitisse captar a perspectiva da demanda turística. Dito de outro modo, seria incorporar aos modelos de competitividade a análise da experiência do turista no destino, ou o feedback da demanda. Essa inovação poderia auxiliar gestores públicos e privados do turismo no refinamento das análises, diagnósticos e planejamentos da atividade.

Oliveira (2013) identificou nove novas subcategorias baseadas na percepção do turista (ex-post) sobre quais fatores devem ser os mais importantes para a competitividade no turismo. As Novas categorias com critério de avaliação baseadas no desempenho são apresentadas no quadro 3:

Quadro 3 - Novas categorias com critério de avaliação baseadas no desempenho

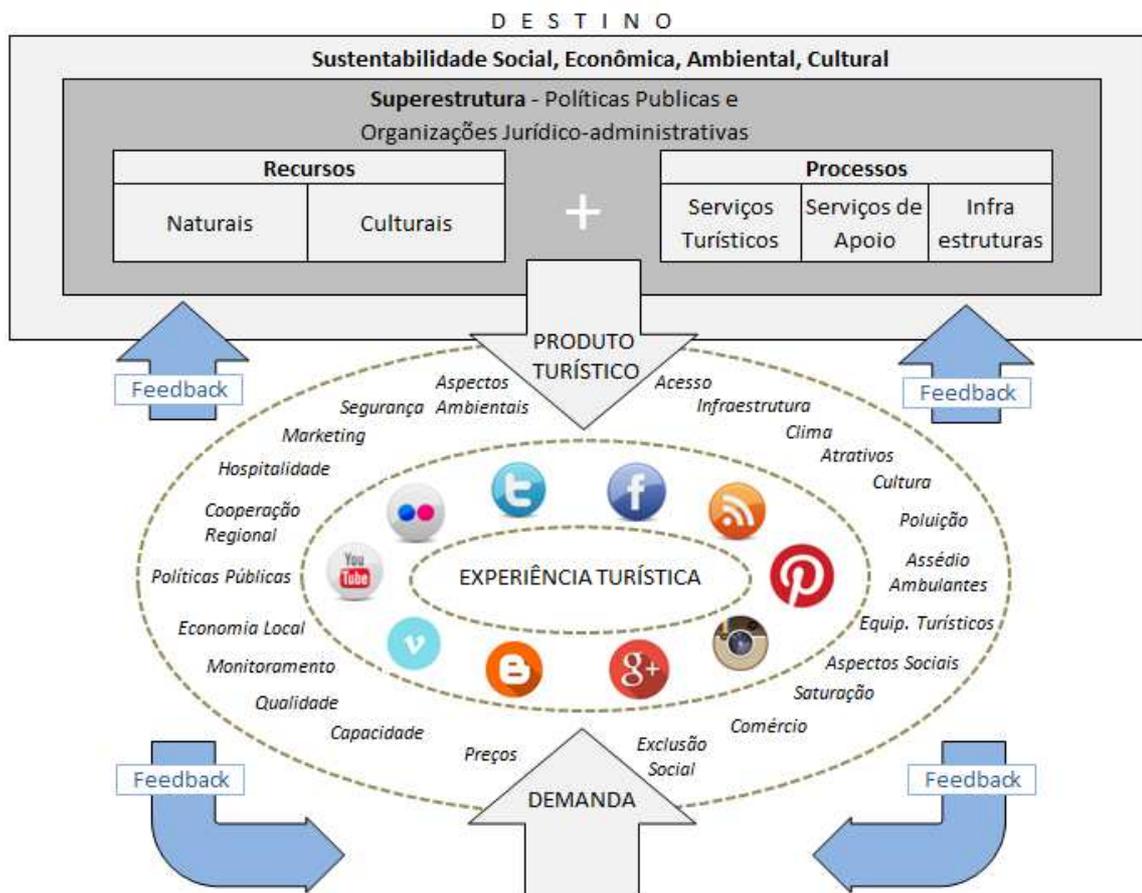
Categorias de análise	Subcategorias	Incidentes Críticos (IC)	Critério de avaliação	Origem teórica	Origem dos dados	Observação
Nova categoria	Hospitalidade dos residentes	21	Desempenho ( <i>ex-post</i> )	Visão de recursos (RBV)	Entrevista com o turista (MIC)	Os modelos do FEM e de Crouch e Ritchie apresentam categorias de hospitalidade
Nova categoria	Assédio de ambulantes	6	Desempenho ( <i>ex-post</i> )	Visão de recursos (RBV)	Entrevista com o turista (MIC)	Está ligado ao desequilíbrio socioeconômico e a exclusão social
Nova categoria	Desconforto causado pela exclusão social e prostitutas	10	Desempenho ( <i>ex-post</i> )	Visão de recursos (RBV)	Entrevista com o turista (MIC)	Está ligado ao desequilíbrio socioeconômico e a exclusão social
Nova categoria	Insegurança	22	Desempenho ( <i>ex-post</i> )	Visão de recursos (RBV)	Entrevista com o turista (MIC)	Está ligado ao desequilíbrio socioeconômico e a exclusão social
Nova categoria	Qualidade da prestação de serviços (turístico e outros)	32	Desempenho ( <i>ex-post</i> )	Visão de recursos (RBV)	Entrevista com o turista (MIC)	Juntamente com os atrativos formam a maior fonte de satisfação/insatisfação dos turistas
Nova categoria	Preços (serviços turísticos, refeições)	7	Desempenho ( <i>ex-post</i> )	Visão de recursos (RBV)	Entrevista com o turista (MIC)	Presente em todos os demais modelos de competitividade, podendo ser obtido por análises <i>ex-ante</i>
Nova categoria	Disponibilidade/locais para compras em geral	4	Desempenho ( <i>ex-post</i> )	Visão de recursos (RBV)	Entrevista com o turista (MIC)	As facilidades para compras podem ser obtidas também por análises <i>ex-ante</i>
Nova categoria	Excesso de utilização do atrativo	1	Desempenho ( <i>ex-post</i> )	Visão de recursos (RBV)	Entrevista com o turista (MIC)	Afeta a sustentabilidade do destino
Nova categoria	Poluição atmosférica	1	Desempenho ( <i>ex-post</i> )	Visão de recursos (RBV)	Entrevista com o turista (MIC)	Impacta na sustentabilidade do destino

Fonte: Adaptado de OLIVEIRA, 2013.

Por fim, Oliveira (2013) conclui que medir a experiência do turista acrescentaria aos modelos de competitividade não só um critério de avaliação de desempenho como também confrontaria a importância relativa dos fatores competitivos com a visão da demanda. A avaliação da demanda sobre a experiência turística pode ser considerada como uma medida de desempenho da competitividade do destino (análise *ex-post*). Por outro lado, a identificação dos fatores que induzem a satisfação ou insatisfação (análise *ex-ante*) auxiliaria na análise da competitividade sistêmica e estrutural. Oliveira (2013). A tradução da experiência do turista através das suas opiniões registradas em redes sociais pode auxiliar nesta análise.

A figura 6, adaptada pelo Autor com base no modelo de Oliveira, Zouain e Barbosa (2012), ilustra como a experiência do turista, registrada em redes sociais, interage com o sistema o sistema turístico:

Figura 6 - A experiência turística no contexto sistêmico



Fonte: Elaborado pelo autor.

O esquema apresentado na figura 5 demonstra um sistema turístico com os recursos naturais e culturais sendo explorados por processos comerciais, suportado por políticas públicas, organizações jurídico-administrativas e de acordo com diretrizes econômicas, sociais, ambientais e culturais sustentáveis. A soma de todos estes fatores representa o produto turístico que, em contato com o turista que se desloca até o destino, tem como resultado a experiência turística. Esta experiência, muitas vezes registrada em redes sociais, pode retroalimentar o sistema, fornecendo um feedback aos diversos atores envolvidos na formação do produto turístico. Este feedback pode ser obtido através da Mineração de Opinião, cujo referencial será abordado na seção 2.3 do estudo.

#### **2.2.4 Planejamento do turismo no RS**

Em nível regional, a Secretária do Turismo, Esporte e Lazer do Rio Grande do Sul (SETUR-RS), através do Observatório de Turismo do Rio Grande do Sul, apoia o processo de planejamento, estruturação e promoção da atividade turística, através da elaboração de estudos, pesquisas e análises para divulgação em publicações periódicas impressas e online que apresentem resultados sobre a oferta, a demanda e o consumo do turismo. Dois principais elementos da constituição do Observatório são o Painel de Indicadores e o Índice de Competitividade dos Destinos Turísticos (RIO GRANDE DO SUL, 2014).

O Painel de Indicadores é o instrumento de mensuração da atividade turística no Estado. O Índice é o resultado da análise de variáveis que permitem conhecer a conjuntura dos municípios a partir de 13 dimensões que demonstram a noção de eficiência para se promover e receber turistas, concomitante à geração de riquezas e benefícios sociais (RIO GRANDE DO SUL, 2014). Os indicadores supracitados são apresentados na tabela 1:

Tabela 1 - Painel de Indicadores do Observatório de Turismo do RS

Unid. Medida	Nome do indicador	Unid. Medida	Nome da fonte	Fonte	Períodicidade do dado	Divulgação	Escopo geográfico	Escopo de análise	Ano de referência	Ano de publicação	Início série histórica	Estratégia de utilização do dado	Comparativo no Sistema	Observação
<b>RECEITAS DO SETOR DO TURISMO</b>														
01	Participação Turismo no ISS	%	Secretarias Municipais (prim.) e Fibras(tec)	Fibras	Anual (Fibras) mensal (secretarias)	Anual (Fibras) trimestral (secretarias)	Municipal/Região/UF	Grupos e classes CNAE	2012 (para Fibras) 2014 (para secretarias)	2013 (para Fibras) e 2014 (para secretarias)	1989 (para Fibras) e 2014 (para secretarias)	Avaliar impacto e participação do turismo nas contas públicas locais	Não	Alternativamente, pode ser feito o cruzamento entre Fibras e PAS. Importante verificar licenças e alíquotas praticadas nos municípios do projeto-piloto.
02	Arrecadação de impostos	R\$	IBGE	PAS	Anual	Anual	Região/UF	Grupos e classes CNAE	2011	2013	2003	Demonstrar importância para economia do estado/região e sua contribuição em termos de tributos para o Governo Federal	Sim	Amostra de empresas
03	Valor pago de ISS pelas empresas de turismo: ICMS, ISS, SIMPLES, COFINS, PIS/PASEP, IPTU, IPVA, Taxas, IOF, Alvarás	R\$	Secretarias Municipais	Acordo	Trimestral	Trimestral	Municipal (projeto-piloto)	Grupos e classes CNAE	-	-	2014	Avaliar impacto do turismo sobre as contas públicas dos municípios indicadores do turismo	Não	(Projeto Piloto). Arrecadação de SETUR-RS com as Secret. Municipais de Fazenda dos municípios para levantamento da arrecadação de ISS do setor.
04	Arrecadação de ISS x receitas setor	%	Secretarias e STN	FINBRA e Secretarias	Anual	Anual	Região/UF para STN e municipal para secretarias	Grupos e classes CNAE	-	-	-	Apresentar a relação entre arrecadação ISS do setor e receitas de empresas de turismo	Sim	-
05	Lucro do setor privado	R\$	IBGE	PAS	Anual	Anual	Região/UF	Grupos e classes CNAE	2011	2013	2003	Monitoramento de desempenho das empresas de turismo.	Sim	Amostra de empresas
06	Receita bruta das empresas de turismo	R\$	IBGE	PAS	Anual	Anual	Região/UF	Grupos e classes CNAE	2011	2013	2003	Demonstrar importância do turismo para economia do estado/região e medir desempenho das empresas do Estado.	Sim	Amostra de empresas
07	Receitas do setor privado	R\$	IBGE	PAS	Anual	Anual	Região/UF	Grupos e classes CNAE	2011	2013	2003	Demonstrar importância do turismo para economia do estado. Pode também servir para mostrar eventual peso da carga tributária para as empresas do setor.	Sim	Amostra de empresas

Fonte: Adaptado de RIO GRANDE DO SUL, 2014.

## 2.3 A MINERAÇÃO DE DADOS E TEXTOS

Esta seção trata do referencial sobre Mineração de uma forma ampla, desde a conceituação, aplicações e metodologia até a abordagem de forma detalhada da Mineração de Opiniões, foco desta pesquisa.

### 2.3.1 Conceitos e aplicações

Atualmente, termos como “Ciência de Dados” e “*Big Data*” estão sendo utilizados para definir o que talvez seja muito mais um ofício do que ciência propriamente dita, sem considerar que a maioria das técnicas em uso não são novidade entre estatísticos e matemáticos. Segundo Schutt e O’neil (2013), “há profissionais com o título de Cientistas de Dados, mas há poucos professores de Ciência de Dados nas universidades”. Drew Conway (2010) define o perfil do Cientista de Dados (*Data Scientist*) apoiado em três pilares, basicamente formados por Conhecimento Computacional, Estatístico e Analítico, conforme figura 7:

Figura 7 - Representação do perfil do cientista de dados através do Diagrama de Venn



Fonte: Adaptado de CONWAY, 2010.

Isso pode estar mudando. Um cenário com crescente aumento da capacidade de armazenamento, do volume de dados, da quantidade de transações e da heterogeneidade entre as informações determinou a surgimento das técnicas elaboradas de descoberta de conhecimento.

De forma intencional ou não, em vários movimentos ou operações que realizados diariamente, deixa rastros de informações, tais como *posts* em redes sociais, imagens capturadas de câmeras, compras on-line, comentários em blogs bem como todo o tipo de informação, consentidas ou não, disponíveis para consulta na web. Segundo Larose (2014), a Mineração de Dados é uma das tecnologias mais promissoras da atualidade. Um dos fatores deste sucesso é o fato de dezenas, e muitas vezes centenas de milhões de reais serem gastos pelas companhias na coleta dos dados e, no entanto, nenhuma informação útil é identificada.

Na próxima seção serão abordados os conceitos para Mineração de Dados, o processo e fases para a sua aplicação, as técnicas mais utilizadas, seus benefícios e uma revisão dos trabalhos realizados de Mineração de Dados no Turismo.

### **2.3.2 A Mineração de Dados**

Em meio à grande quantidade e variedade de dados, é preciso analisar o que é relevante para a pesquisa de informações com valor. Para isso pode-se usar técnicas de Mineração de Dados.

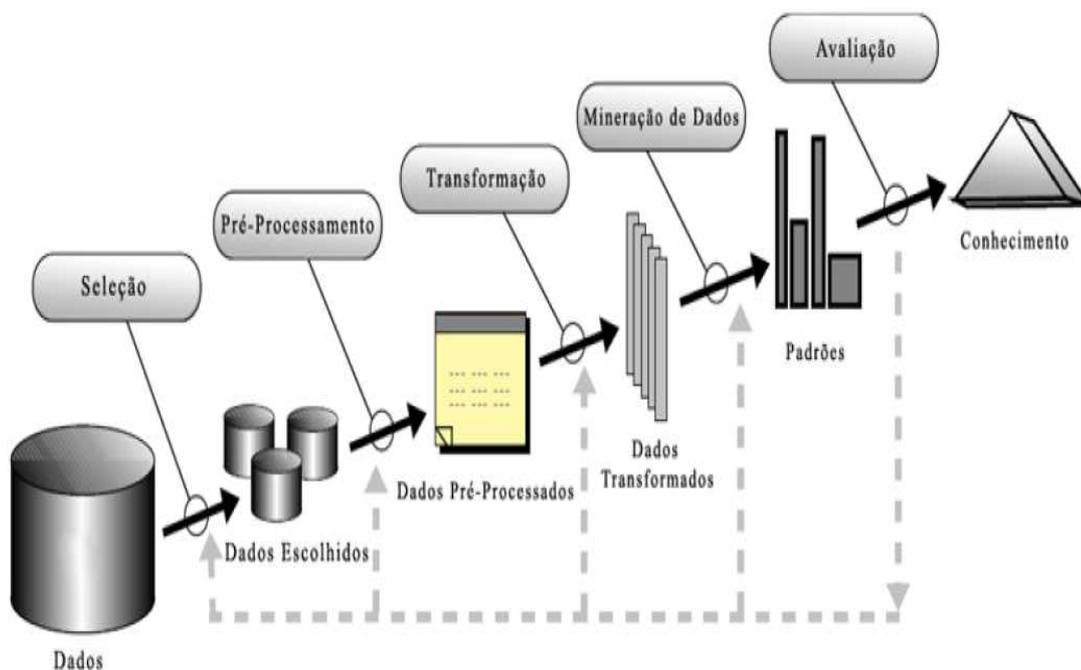
A Mineração de Dados é considerada uma etapa do processo de descoberta de conhecimento (Knowledge Discovery Database - KDD). Segundo Goebel e Gruenwald (1999), o termo KDD é usado para descrever o processo global de transformação de dados de baixo nível em conhecimento de alto nível. De forma simples, a descoberta de conhecimento em bases de dados é um processo não trivial de identificação de padrões válidos, úteis e compreensíveis de dados. Uma definição comum para Mineração de Dados a define como o processo de extração de padrões ou modelos de dados observados. Existem várias definições para o termo Mineração de Dados, mas algumas premissas são comuns entre eles. Segundo Braga (2005), Mineração de Dados é uma coleção de dados centrada no cliente.

Mineração de dados é um termo comum usado para um conjunto de técnicas e conceitos multidisciplinares. A Mineração de Dados envolve conceitos de bancos de dados, modelagem de informações e técnicas estatísticas.

A Mineração de Dados está inserida no processo de KDD, conforme pode ser visualizado através da Figura 8. Segundo Fayyad et al. (1996), o processo de KDD é composto de cinco etapas: seleção dos dados; pré-processamento e limpeza dos

dados; transformação dos dados; Mineração de Dados (Data Mining); e interpretação e avaliação dos resultados.

Figura 8 - Representação do processo de KDD



Fonte: FAYYAD et al., 1996.

Segundo Olson e Delen (2008), a Mineração de Dados pode ser útil para:

- Apoiar na melhoria de eventos com base no feedback dos turistas em eventos passados;
- Realizar campanhas de marketing direcionadas a partir da identificação de entidades como a Localização do turista que registrou um comentário e;
- Personalizar o atendimento de turistas cujo número de menções sobre uma determinada reclamação se propagaram pela rede.

As premissas em comum entre vários autores para que um processo seja definido como um “processo de Mineração de Dados” são as de que o conhecimento extraído necessita ser valioso e as técnicas utilizadas no processamento analítico dos dados podem ser automatizadas ou semiautomatizadas.

De forma geral, a Mineração de Dados permite filtrar informações relevantes e tendências para uso em processos de tomada de decisão e também a descoberta de padrões ou correlações invisíveis em bancos de dados originais ou construídos

(*Data warehouses*) e informações não observadas em análises do tipo *On-line Analytical Processing* (OLAP). Assim como as ferramentas ditas OLAP, a Mineração de Dados possui grande capacidade analítica, com a vantagem da predição ou da capacidade de prever eventos (BRAGA, 2005).

Atualmente, torna-se impraticável a gestão do conhecimento sem o uso de tecnologias de informação. O uso de tecnologias de informação associadas às tarefas de tomada de decisão consolidou o conceito de *Business Intelligence* (BI) nas empresas. Segundo Braga (2005), o BI utiliza conceitos e tecnologias que, fazendo o uso de acontecimentos (fatos) e sistemas baseados nos mesmos, apoia o processo de tomada de decisões em negócios. Segundo Angeloni (2003), informação e conhecimento são elementos fundamentais para a comunicação e a tomada de decisão nas organizações.

O crescimento da quantidade de regras necessárias para o tratamento de problemas como ambiguidade e reconhecimentos de entidades também pode ser considerado como um grande desafio para a evolução das técnicas e ferramentas de Mineração de Dados. Fontes de informações tais como blogs e redes sociais são reinventados constantemente, requerendo uma maior variedade de modelos.

Pode-se afirmar que não existe uma única técnica de Mineração de Dados que possa ser aplicada para qualquer problema. A escolha de uma ou mais técnicas adequadas bem como a escolha das ferramentas que as implementam dependem de cada situação. Segundo Harrison (1998), a escolha das técnicas de Mineração de Dados dependerá da tarefa específica a ser executada e dos dados disponíveis para análise. Para Larose (2014), as tarefas mais comuns são:

- a) **Sumarização ou Descrição (*Description*):** trata da descrição de padrões e tendências. Descrições de alta qualidade também podem ser obtidas através de análise exploratória de dados. A Análise Exploratória de Dados é o requisito inicial para a elaboração de modelos para representação da realidade que será analisada.
- b) **Estimativa ou Regressão (*Estimation*):** tarefas de regressão são similares às tarefas de classificação. O que difere é a forma de apresentação do resultado. Como exemplo pode-se citar o interesse em estimar a leitura da pressão sistólica de um paciente, com base na idade, sexo, peso e níveis de sódio. A relação entre estas variáveis

poderá fornecer um modelo de estimativa para aplicação em novos casos (LAROSE, 2014).

- c) **Predição (*Prediction*):** semelhante às tarefas de classificação e regressão, os resultados de uma tarefa de predição encontram-se no futuro. Qualquer um dos métodos e técnicas utilizados para Classificação e Estimativa pode também ser usado, dependendo da aplicação, para Predição. Como exemplo de aplicação deste tipo de tarefa pode-se citar a previsão do preço de uma determinada ação futuramente no mercado de ações ou prever o percentual de mortes no trânsito no próximo ano se o limite de velocidade for aumentado.
- d) **Classificação (*Classification*):** Na classificação existe uma variável-alvo que assume valores classificados ou de acordo com categorias, tais como faixa de renda, que pode ser dividida em 3 categorias (alta, média, baixa). Como exemplo de tarefa de Classificação, pode-se citar a classificação de um candidato em uma vaga que não esteja de acordo com o seu perfil ou qualificações ou diagnosticar se uma determinada doença está presente.
- e) **Segmentação ou Agrupamento (*Clustering*):** **Segmentação** refere-se ao agrupamento de objetos semelhantes. Um cluster é uma coleção de registros que são semelhantes entre si. A segmentação difere Classificação por não haver uma variável-alvo. Os algoritmos de Segmentação buscam a criação de grupos relativamente homogêneos buscarem através de iterações para maximizar e minimizar a similaridade dentro e fora do cluster, respectivamente. Um exemplo de aplicação prática da tarefa de Segmentação é a criação de um perfil sócio-demográfico a partir de um código postal de uma determinada região. A segmentação é frequentemente realizada como um passo preliminar em um processo de Mineração de Dados, com os aglomerados resultantes sendo utilizados como *inputs* para em uma outra técnica sucessivamente.
- f) **Associação (*Association*):** A tarefa de associação na Mineração de Dados tem como objetivo quantificar a relação de influência entre dois ou mais atributos, demonstrando o que ocorre com cada atributo se os demais forem modificados. Um exemplo prático da aplicação desta

tarefa é demonstrar que efeitos colaterais de uma droga podem surgir se a dose dela for aumentada.

Cada tarefa faz uso de um ou mais algoritmos ou técnicas estatísticas, conforme quadro 4 (DIAS, 2008):

Quadro 4 - Técnicas, tarefas e algoritmos de Mineração de Dados

Técnica	Descrição	Tarefas	Exemplos
Descoberta de Regras de Associação	Estabelece uma correlação estatística entre atributos de dados e conjuntos de dados	Associação	Apriori, AprioriTid, AprioriHybrid, AIS, SETM (Agrawal e Srikant, 1994) e DHP (Chen et al., 1996).
Árvores de Decisão	Hierarquização dos dados, baseada em estágios de decisão (nós) e na separação de classes e subconjuntos	Classificação Regressão	CART, CHAID, C5.0, Quest (Two Crows, 1999); ID-3 (Chen et al., 1996); SLIQ (Metha et al., 1996); SPRINT (Shafer et al., 1996).
Raciocínio Baseado em Casos ou MBR	Baseado no método do vizinho mais próximo, combina e compara atributos para estabelecer hierarquia de semelhança	Classificação Segmentação	BIRCH (Zhang et al., 1996); CLARANS (Chen et al., 1996); CLIQUE (Agrawal et al., 1998).
Algoritmos Genéticos	Métodos gerais de busca e otimização, inspirados na Teoria da Evolução, onde a cada nova geração, soluções melhores têm mais chance de ter "descendentes"	Classificação Segmentação	Algoritmo Genético Simples (Goldberg, 1989); Genitor, CHC (Whitley, 1993); Algoritmo de Hillis (Hillis, 1997); GA-Nuggets (Freitas, 1999); GA-PVMINER (Araújo et al., 1999).
Redes Neurais Artificiais	Modelos inspirados na fisiologia do cérebro, onde o conhecimento é fruto do mapa das conexões neuronais e dos pesos dessas conexões	Classificação Segmentação	Perceptron, Rede MLP, Redes de Kohonen, Rede Hopfield, Rede BAM, Redes ART, Rede IAC, Rede LVQ, Rede Counterpropagation, Rede RBF, Rede PNN, Rede Time Delay, Neocognitron, Rede BSB (Azevedo, 2000), (Braga et al., 2000), (Haykin, 2001)

Fonte: DIAS, 2008.

### 2.3.3 A Mineração de Textos e o Processamento de Linguagem Natural (PNL)

Segundo Kao e Potee (2007), a Mineração de Textos é a área da DM que trata da descoberta e extração de conhecimento interessante, não-trivial, em textos livres ou não estruturados, englobando desde recuperação de informações da web até tarefas supracitadas, como Classificação, Agrupamento (Clustering) e extração de conceitos para análise. A transformação do texto em dados para análise é feita por meio da aplicação do PNL.

Tipicamente, a PNL utiliza conceitos linguísticos, como partes do discurso e estrutura gramatical. A linguagem natural tem que lidar com situações complexas e de grande relevância, como ambiguidades. O PLN faz uso de diversas representações de conhecimento, tais como léxico, semântica, propriedades e regras gramaticais, ontologias ou dicionários de sinônimos e abreviaturas (KAO; POTEET, 2007).

#### **2.3.4 A Mineração de Opiniões ou Análise de Sentimentos**

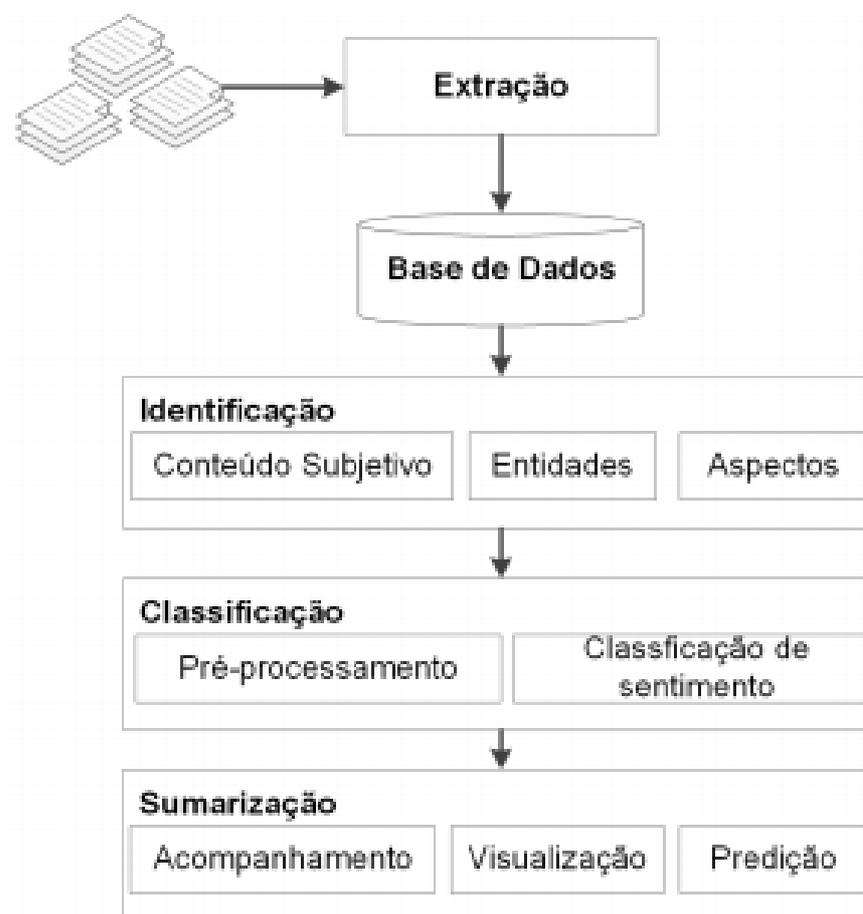
A Mineração de Opiniões pode ser vista como uma subárea da Extração de Informação que lida com a identificação de opiniões, avaliações e atitudes com relação a entidades como pessoas, produtos e organizações, expressas textualmente. Estimulada, sobretudo, pela crescente disponibilidade e diversidade de opiniões em blogs e redes sociais, a tarefa pode assumir diferentes aspectos: da caracterização de documentos (por exemplo, resenhas), como favoráveis ou desfavoráveis relativamente a um dado objeto, até à identificação de pontos de vista em debates políticos, trazendo desafios interessantes tanto do ponto de vista do processamento computacional quanto do ponto de vista linguístico (FREITAS, 2013).

Conforme Liu (2012), análise de sentimento e mineração opinião é o campo de estudo que analisa das pessoas opiniões, sentimentos, avaliações, atitudes e emoções da linguagem escrita. É uma das áreas de pesquisa mais ativas em processamento de linguagem natural e também é amplamente estudada em mineração de dados, mineração de Web, e mineração de texto. Na verdade, esta pesquisa se espalhou para fora da ciência da computação para as ciências da gestão e ciências sociais devido à sua importância para as empresas e sociedade como um todo. A crescente importância da análise de sentimento coincide com o crescimento das redes sociais, como comentários, fóruns de discussão, blogs, microblogs, Twitter e redes sociais.

Segundo Tsytarau e Palpanas (2012), o problema da mineração de opiniões pode ser estruturado em termos das seguintes tarefas genéricas:

- a) Identificar as opiniões expressas sobre determinado assunto ou alvo em um conjunto de documentos;
- b) Classificar a orientação ou polaridade desta opinião, isto é, se tende a positiva ou negativa e;
- c) Apresentar os resultados de forma agregada e resumida, conforme figura 9:

Figura 9 - Etapas da Mineração de Opinião



Fonte: Adaptado de TSYTSARAU; PALPANAS, 2012.

A polaridade da opinião define o sentimento, percepção ou atitude do público em relação ao alvo da opinião. A mineração de opiniões opera sobre porções de texto de quaisquer tamanho e formato, tais como páginas web, posts, comentários, tweets e revisões de produto (BECKER; TUMITAN, 2013). Toda opinião é composta de pelo menos dois elementos chave: um alvo e um sentimento sobre este alvo. Um

alvo pode ser uma entidade, aspecto de uma entidade, ou tópico, tais como produto, pessoa, organização, marca ou evento (LIU, 2012).

Um sentimento representa uma atitude, opinião ou emoção que o autor da opinião tem a respeito do alvo. A polaridade de um sentimento corresponde a um ponto em alguma escala que representa a avaliação positiva, neutra ou negativa do significado deste sentimento (TSYTSARAU; PALPANAS, 2012).

Segundo Liu (2012), a detecção do sentimento em um texto pode ocorrer em diferentes granularidades, sendo que a decisão do nível está sujeita ao contexto e aplicação. Já a análise pode ser em nível de documento, sentença ou frases. Ainda conforme o autor, opiniões referem-se a conteúdo subjetivo, escrito em linguagem natural.

A forma como as opiniões estão expressas influencia diretamente a habilidade de processá-las corretamente. As abordagens de classificação de polaridade podem ser divididas em quatro grandes grupos:

- a) Léxicas, com o uso de dicionários de sentimentos;
- b) Aprendizado de máquina, com o uso predominante de técnicas de classificação ou de regressão;
- c) Estatísticas, que utilizam técnicas para avaliar a recorrência de termos e;
- d) Semânticas, que definem a polaridade de palavras em função de sua proximidade semântica com outras de polaridade conhecidas. Técnicas destas diferentes abordagens podem ser combinadas para melhoria de resultados.

Técnicas destas diferentes abordagens podem ser combinadas para melhoria de resultados.

A análise de opiniões em grande escala não é prática sem agregação automática e sumarização, quando se busca a identificação de opiniões em maior nível do que a de um indivíduo, através de uma opinião média ou predominante um grupo de pessoas sobre algum tema, tema que é abordado na próxima seção.

### **2.3.5 A Classificação Automática de Documentos**

A Classificação Automática de Textos (CT) pode ajudar a solucionar problemas que lidam com grandes quantidades de documentos digitais, apoiando no

refinamento de resultados de buscas em sites como Yahoo! e o Google. Os classificadores automáticos ou semiautomáticos podem ser a chave para solucionar o problema de se lidar com enormes quantidades de textos e documentos digitais.

Segundo Langie e Lima (2003), a classificação automática consiste em classificar um documento digital em uma determinada categoria, de acordo com critérios estabelecidos a priori por programadores humanos. Essa classificação geralmente respeita uma estrutura hierárquica, utilizando uma árvore de categorias, permitindo que os documentos sejam classificados tanto nas folhas quanto nos nodos intermediários.

Dessa forma, uma palavra pode não ser muito esclarecedora para ajudar a classificar um texto entre as categorias Sintaxe, Semântica ou Fonologia, mas poderá ser um bom atributo para diferenciar textos entre categorias maiores, como História, Medicina ou Linguística Aplicada. No contexto da Mineração de Opinião, a Classificação Automática de Documento, categorização, classificação, ou ainda, clusterização de documentos consiste em organizar informações automaticamente em grupos de documentos similares, facilitando sua visualização e análise.

### **2.3.6 A Descoberta de Conhecimento no Turismo**

Macedo et al. (2011) analisam a contribuição da Mineração de Dados na mensuração da qualidade total dos serviços prestados. Segundo os autores, “a aplicação da Mineração de Dados para auxiliar a análise dos dados referentes aos indicadores de qualidade contribui para o processo de mensuração da qualidade de serviços”. O estudo analisou a contribuição das técnicas de Mineração de Dados para os modelos de gestão baseados em Modelos de Qualidade Total em geral aplicados ao segmento de prestação de serviços em geral.

Barcelos et al. (2014) utilizaram o VisMkt para demonstrar o desenvolvimento e a aplicação de softwares para a análise de dados, de modo a proporcionar novas diretrizes para o uso de grandes quantidades de texto não estruturado no contexto da gestão. O software foi utilizado na mineração de textos do site de avaliações de viagens Tripadvisor. Foram analisados 250 reviews no Tripadvisor, até dezembro de 2013, para os hotéis do município de Ouro Preto, Minas Gerais, Brasil: Solar do Rosário, Pousada Arcanjo, Luxor, Pousada Clássica, Boroni Palaci Hotel e Grande Hotel Ouro Preto.

Os autores criaram uma ferramenta de visualização e processamento de textos para análise de reviews do site. A implementação da solução computacional proposta no trabalho envolveu as etapas de coleta e pré-processamento de Mineração de Dados; armazenamento de informações em um banco de dados; a categorização e a classificação de termos; o processamento dos textos dos comentários de avaliações de hotéis aplicando índice dos termos categorizados; o processamento e a contagem dos termos agrupando-os como avaliações e a disponibilização dos dados em visualizações web. Através do trabalho realizado, os autores procuraram demonstrar novas formas de visualização das avaliações dos turistas. A técnica ou algoritmo utilizada através da ferramenta não foi divulgada.

Gazzana e Silveira (2009) desenvolveram um protótipo de sistema de recomendação para a indicação de pacotes turísticos a usuários. Segundo os autores do trabalho, essa ideia foi elaborada como o intuito de auxiliar usuários a efetivarem uma busca mais precisa por pacotes turísticos, evitando o recebimento de sugestões que não venham lhe interessar. Com a utilização de técnicas de filtragem baseada em conteúdo e informações coletadas de forma explícita e implícita, são sugeridos ao usuário pacotes turísticos que se encaixem no seu perfil. O sistema aplica o algoritmo de recomendação nos pacotes turísticos encontrados num destino selecionado e os classifica de acordo com as preferências do usuário. O sistema foi avaliado através de uma pesquisa qualitativa realizada com profissionais da área de turismo.

Reategui e Cazella (2005), no XXV Congresso da Sociedade Brasileira de Computação apresentaram trabalho onde descrevem as técnicas e estratégias de recomendação mais utilizadas, detalham a arquitetura dos sistemas e trazem exemplos de utilização destes tanto no meio acadêmico quanto na indústria. Na descrição das técnicas, abordaram Filtragem de Informações, Filtragem de Conteúdo, Filtragem Colaborativa, Filtragem Híbrida, Descoberta de Conhecimento e respectivas técnicas de mineração de dado. Por fim, discutiram a questão da privacidade e concluem que “[...] os sistemas de recomendação podem ser utilizados para que se possa conhecer melhor os hábitos de consumo e interesses dos clientes, e como este tipo de conhecimento pode ser empregado para personalizar o relacionamento entre a empresa e seus consumidores [...]”. (REATEGUI; CAZELLA, 2005, p.33).

Santos, Fernandes e Klein (2000), no VI Congresso Argentino de Ciências da Computação, apresentaram proposta de desenvolvimento de um sistema de apoio à decisão baseado em Mineração de Dados para as instituições envolvidas no “Projeto Segurança nas Praias”. O projeto tinha como objetivos principais a criação de uma base de informações sobre as características sócioambientais das praias catarinenses e a formulação de um programa de orientação ao público usuário, definindo o grau de periculosidade dos balneários. As técnicas de Mineração de Dados utilizadas não foram documentadas bem como os resultados alcançados.

Oliveira et al. (2012) em um estudo de caso sobre as Olimpíadas 2012 utilizando o Twitter, realizam a mineração de opinião com a finalidade de obterem informações que possibilitem a classificação de usuários dentro de grupos de interesses, análise de opiniões e aceitação de certos assuntos, com a finalidade de utilizar os resultados em análises de viabilidade de atletas, juntamente com uma classificação de opinião sobre os mesmos.

Thomaz (2014) investiga o monitoramento e mineração de conteúdos turísticos em redes sociais, as oportunidades e benefícios oferecidos e como podem ser utilizados estrategicamente para auxiliar na gestão de destinos turísticos. O objetivo principal foi o de avaliar a forma como os gestores de destinos turísticos do Brasil monitoram as redes sociais, além de comparar as principais funcionalidades das redes utilizadas. Inicialmente, foi realizada a mineração de termos conforme Estrutura para Coleta de Conteúdos do Twitter Relacionados aos Jogos Olímpicos de 2012, publicada por Neves e Marchiori (2014). Após a coleta e análise, foram utilizadas pesquisas quantitativas e qualitativas com os gestores dos destinos turísticos da região de Foz do Iguaçu-PR.

Agra et al. (2014), através de um estudo de caso realizado durante o período da Copa do Mundo da FIFA Brasil 2014, investiga como as redes sociais podem ser um instrumento facilitador para aproximar governo e cidadãos, ajudando a identificar e analisar a opinião pública e possibilitando, aos gestores públicos, atender melhor às necessidades da sociedade. O estudo analisou o sentimento baseado em hashtags publicados em redes sociais Twitter e Facebook associados aos temas “Governo” e “Copa” e conclui, após as análises sobre os dados coletados, que é possível obter insights significativos sobre determinado tema e aponta, dentre as dificuldades encontradas na condução do trabalho, a dificuldade na calibração das

palavras-chave com a finalidade de maximizar a cobertura dos comentários relevantes e, ao mesmo tempo, pré-filtrar publicações não relacionadas.

Nascimento (2012) analisa o sentimento de tweets relacionados aos fatos divulgados na mídia e que foram amplamente propagados dentro do Twitter. O estudo propôs uma classificação automática dos tweets com base em 3 classificadores de linguagem e o resultado obtido foi validado junto à classificação manual das mensagens, realizada previamente e conclui que o resultado obtido é satisfatório, pois a acurácia dos classificadores se manteve dentro do percentual de acertos corresponde à capacidade humana de classificar a subjetividade de um texto.

Outras contribuições encontradas para o setor turístico, embora realizados em outros países, foram:

- a) Pekar e Ou (2008) utilizaram técnica de análise de sentimento para avaliar 268 comentários sobre hotéis publicadas no site “epinions.com”.
- b) He, Zha e Li (2013) utilizaram a mineração de texto para analisar conteúdos e informações nos perfis do Facebook e Twitter das três maiores cadeias de pizza dos Estados Unidos.
- c) Neves e Marchiori (2014) monitoraram conteúdos do Twitter para avaliar a qualidade percebida de produtos e serviços turísticos durante os Jogos Olímpicos de Londres de 2012.

Entre as dimensões analisadas para a composição do índice de competitividade do turismo citados anteriormente, a dimensão “Serviços e Equipamentos Turísticos” envolve o uso de ferramentas de apoio ao Turista.

Os resultados documentados por Thomaz (2014), Pekar e Ou (2008), He, Zha e Li (2013) e Neves e Marchiori (2014) demonstram a viabilidade do uso de técnicas de Mineração de Textos em redes sociais na avaliação da qualidade de produtos e serviços.

Observa-se que foram identificados poucos estudos da aplicação da Mineração de Dados e Textos no contexto específico do planejamento da atividade turística. Porém, os estudos encontrados abordam todos os temas do referencial de forma adjacente e contribuíram sobremaneira para o desenvolvimento deste trabalho.

### 3 METODOLOGIA

Este capítulo aborda os procedimentos metodológicos utilizados para se alcançar os objetivos propostos neste trabalho, abordando o tipo de pesquisa, técnica de coleta e análise e uma síntese do planejamento realizado para a condução da pesquisa.

#### 3.1 TIPO DE PESQUISA

Para alcançar os objetivos propostos neste estudo, foi realizada uma pesquisa aplicada e exploratória com abordagem quantitativa, contemplando:

- a) Pesquisa bibliográfica sobre Sistema Turístico, Planejamento do Turismo e sobre as técnicas de Mineração de Opinião em redes sociais, documentada no capítulo 2 deste estudo;
- b) Aplicação da Mineração de Opinião sobre o município de Gramado/RS em dados de redes sociais e;
- c) Entrevista com gestor público para validação das informações encontradas.

Quanto à natureza, a pesquisa é classificada como Aplicada por que se trata de uma investigação que objetiva a geração de conhecimentos para a aplicação prática e voltada à solução de problemas. Segundo Gil (2002), a pesquisa aplicada abrange estudos elaborados com a finalidade de resolver problemas identificados no âmbito das sociedades em que os pesquisadores vivem.

Quanto aos objetivos gerais, a pesquisa em questão pode ser classificada como exploratória, uma vez os resultados esperados não são conhecidos de antemão e o objetivo maior deste estudo é, justamente, analisar as oportunidades caso elas se confirmem. Segundo Andrade (2003), a pesquisa exploratória tem como finalidade proporcionar mais informações sobre o assunto a ser investigado, possibilitando sua definição e seu delineamento.

## 3.2 UNIDADE DE ANÁLISE

A Unidade de Análise desta pesquisa foi selecionada com base no seu volume de informações ou comentários disponíveis em redes sociais, dada a importância do município de Gramado/RS no cenário turístico brasileiro. Também foram levados em consideração critérios de acessibilidade e conveniência, uma vez que foi necessário entrevistar os gestores da SMT em várias oportunidades. Assim, o objeto de análise deste estudo são os comentários que os turistas registram em redes sociais sobre suas percepções em relação aos serviços e produtos turísticos ofertados pelo município de Gramado/RS. A aplicação de uma técnica de coleta e Mineração de Opinião tem a finalidade de obter um feedback sobre a satisfação do turista. As opiniões on-line são de grande valor para os fornecedores de produtos e serviços, uma vez que representam o feedback do cliente de forma espontânea e genuína (BROSS, 2013).

A seguir, será apresentada a cidade sobre a qual as percepções dos turistas serão estudadas neste trabalho, bem como a Secretaria Municipal de Turismo, para quem o relatório final é destinado.

### 3.2.1 A Cidade de Gramado

Gramado tem a melhor estrutura turística do Rio Grande do Sul e é hoje o principal destino turístico do Estado e o quarto do Brasil. Com uma economia voltada ao turismo, o município de pouco mais de 33 mil habitantes está localizado na Serra Gaúcha e recebe anualmente cerca de 2,5 milhões de turistas.

Localizado a 126 km de Porto Alegre, Gramado é o sexto destino mais desejado pelos turistas brasileiros no país, perdendo apenas para Fernando de Noronha, Fortaleza, Rio de Janeiro, Salvador e Natal, todos destinos de praia. Os dados são do Ministério do Turismo (BRASIL, 2015).

### 3.2.2 A Secretaria Municipal de Turismo de Gramado

A Secretaria de Turismo de Gramado está situada no prédio sede da Prefeitura e tem como objetivo promover a cidade como destino e definir políticas de

turismo juntamente com as demais entidades municipais e/ou regionais (Sindicato Hotéis e Restaurantes, Convention, entre outros) (GRAMADO, 2015).

Através de diferentes ações, atende ao setor turístico profissional (agências) e ao turista individual que visita a cidade, prestando informações turísticas no Posto de Informações e nos Pórticos da cidade. Além disso, promove grandes eventos como o Carnaval, a Festa da Colônia, o Festival de Cinema e o Natal Luz. (GRAMADO, 2015).

A Secretaria Municipal de Turismo tem papéis fundamentais no desenvolvimento do turismo da cidade, através de:

- a) Manutenção, criação e consolidação de atrativos turísticos através da melhoria constante da infraestrutura receptiva;
- b) Promoção e rentabilização da oferta turística e comercial de Gramado;
- c) Administração e coordenação das Centrais de Atendimento ao turista: Praça Major Nicoletti, Pórticos de entrada por Nova Petrópolis e Várzea Grande, Rótula das Bandeiras, atendimento e informações ao turista, distribuição de mapas e folhetos com informações sobre os atrativos turísticos, levantamentos de dados turísticos através de pesquisa, oferta de hotéis, pousadas, restaurantes e comércio da cidade, acompanhamento de convidados e autoridades em visita a Gramado e;
- d) Difusão da imagem da cidade de Gramado como destino turístico, sua oferta e os serviços disponíveis, como entidade promotora da cidade, seguindo as políticas e estratégias de marketing previamente estabelecidas (GRAMADO, 2015).

### **3.2.3 O município de Gramado como destino turístico**

Em 2015, Gramado foi eleita pelo sétimo ano consecutivo, a cidade turística do Rio Grande do Sul preferida e mais lembrada por empresários gaúchos, através da votação realizada pelo evento Marcas de Quem Decide, organizado pelo Jornal do Comércio e pela Qualidata. O projeto Marcas de Quem Decide avalia simultaneamente os índices de lembrança e preferência de marcas em mais de 100 setores da economia. A pesquisa é feita no Rio Grande do Sul, junto a gestores

empresariais e profissionais liberais, distribuídos nos principais municípios com participação igual ou superior a 0,5% do PIB Gaúcho. (GRAMADO, 2015).

Recebeu também o prêmio 2015 Travelers Choice do site de avaliações Tripadvisor, que classifica a cidade como o melhor destino turístico do Brasil em 2015 (TRIPADVISOR, 2015).

O município de Gramado/RS possui um índice geral de competitividade acima da média dos demais destinos turísticos, conforme demonstrado no gráfico 2:

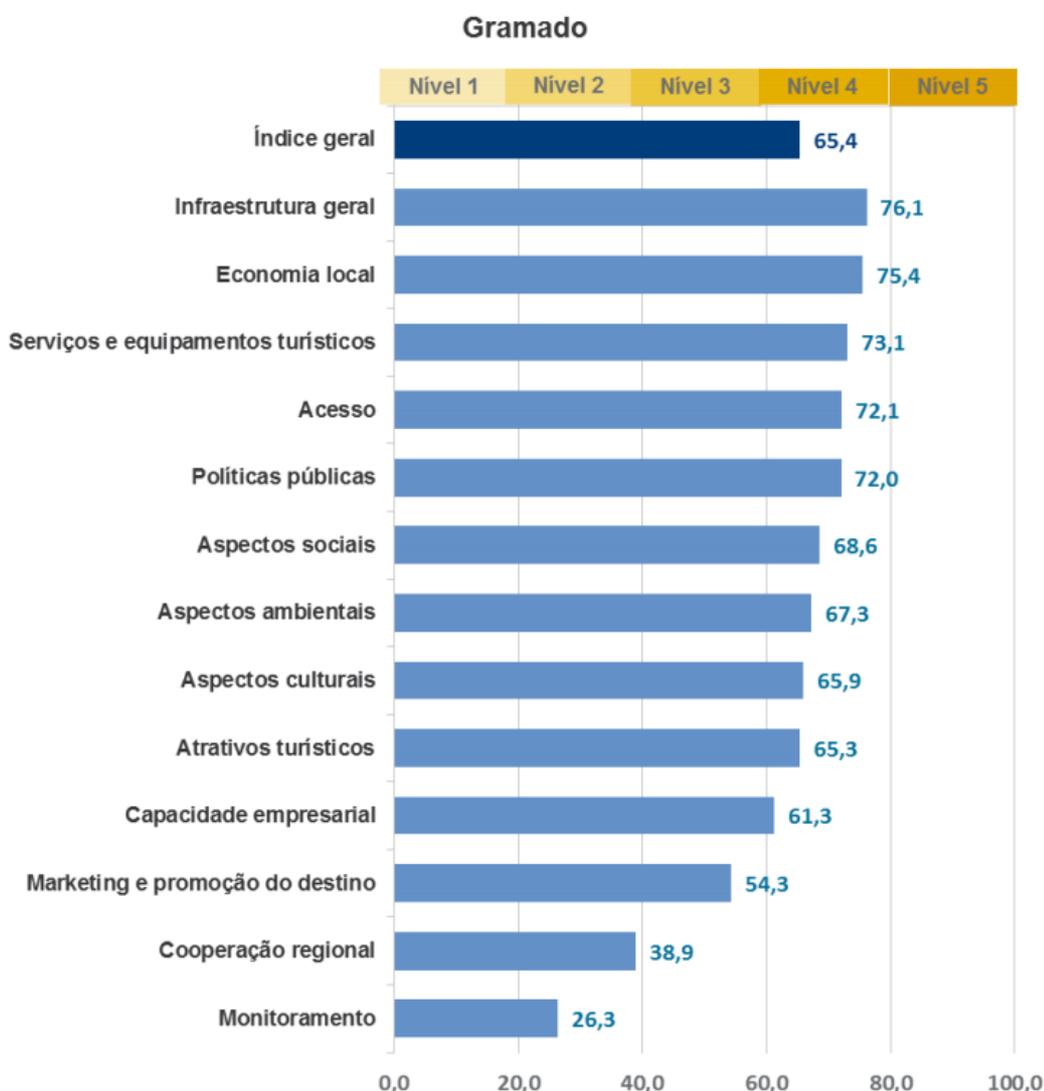
Gráfico 2 - Índices gerais de competitividade – destino x Brasil x não capitais: 2008-2015



Fonte: BRASIL. Ministério do Turismo; SERVIÇO BRASILEIRO DE APOIO ÀS MICRO E PEQUENAS EMPRESAS; FUNDAÇÃO GETULIO VARGAS, 2015.

Este índice foi influenciado pelos resultados de cada uma das 13 dimensões avaliadas. Os resultados são apresentados no gráfico 3:

Gráfico 3 - Índices do destino por dimensão, em ordem decrescente de desempenho



Fonte: BRASIL. Ministério do Turismo; SERVIÇO BRASILEIRO DE APOIO ÀS MICRO E PEQUENAS EMPRESAS; FUNDAÇÃO GETULIO VARGAS, 2015.

Na próxima seção são apresentadas fases e atividades planejadas para operacionalização deste estudo.

### 3.3 MÉTODO DE TRABALHO

A operacionalização deste estudo se deu através das seguintes fases e atividades:

**Fase 1 - Aprovação:** fase destinada a elaboração deste projeto de pesquisa e aprovação junto aos stakeholders envolvidos (universidade, avaliadores,

orientador e SMT). No Apêndice A pode ser visualizada a carta-convite enviada e retornada aprovada pela SMT.

**Fase 2 – Revisão Bibliográfica:** fase destinada à criação do referencial teórico envolvendo Turismo, Planejamento Turístico, Mineração de Opinião entre outras áreas de estudo abordadas na metodologia proposta para este estudo.

**Fase 3 – Definição:** fase destinada ao entendimento dos dados onde será realizada a coleta dos dados e definição das ferramentas utilizadas para a aplicação da técnica de Mineração de Opinião.

**Fase 4 – Execução:** fase destinada à execução da Mineração de Opinião, envolvendo a configuração das ferramentas selecionadas, pré-análise, reconfiguração e documentação dos resultados.

**Fase 5 - Apresentação dos resultados preliminares:** fase destinada à elaboração, aplicação e validação dos resultados junto aos gestores da prefeitura municipal de Gramado/RS.

**Fase 6 - Redação do relatório final:** fase destinada à avaliação e interpretação dos resultados e redação do relatório final juntamente com os gestores da SMT.

Os artefatos gerados ao final de cada atividade são apresentados no quadro 5.

Quadro 5 - Fases, Artefatos e Objetivos

<b>Fases da Pesquisa</b>	<b>Artefatos</b>	<b>Relação com os Objetivos da Pesquisa</b>
Fase 1 - Aprovação	Carta-convite assinada, apresentada nos anexos deste estudo	Objetivo Geral
Fase 2 – Revisão Bibliográfica	Referencial teórico (item 2 deste estudo)	Objetivo específico 1 e 4
Fase 3 – Definição	Documentação registrada no item 3.3.1 Etapa da Coleta de Dados	Objetivo específico 1
Fase 4 – Execução	Documentação registrada no item 3.3.2 A Identificação de Opinião ou Sentimento	Objetivo específico 2
Fase 5 - Apresentação dos resultados preliminares	Resultados preliminares consolidados, apresentados e submetidos à avaliação dos gestores da prefeitura municipal de Gramado/RS, com modelo apresentado nos anexos deste estudo	Objetivo específico 3
Fase 6 - Redação do relatório final	Resultados apresentados e discutidos no item 4 deste estudo	<b>Objetivo Geral</b>

Fonte: Elaborado pelo autor.

A seguir, é apresentada a técnica de coleta de dados que envolvem a Mineração de Opiniões sobre o município de Gramado/RS em dados de redes sociais.

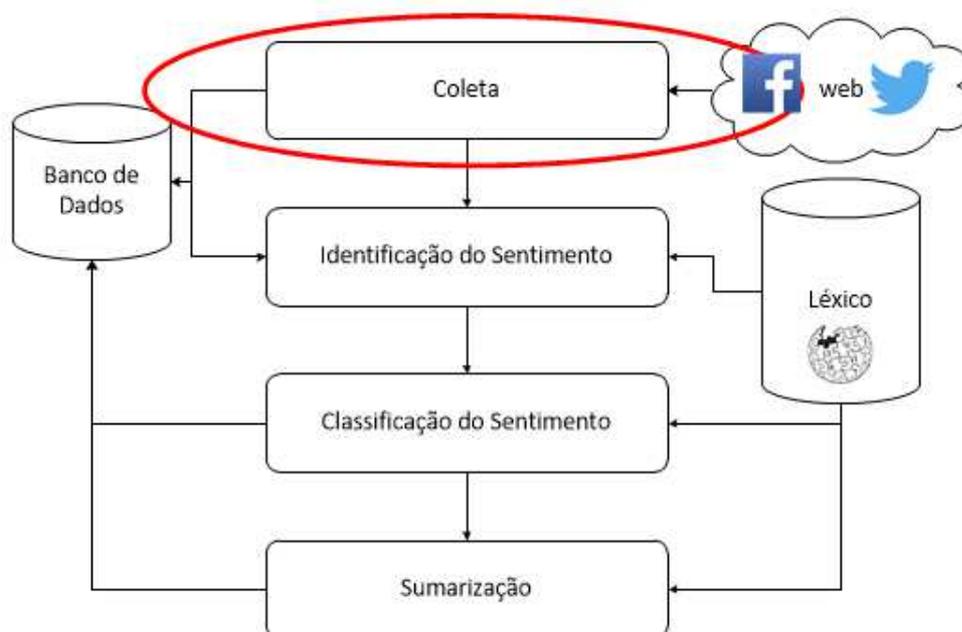
### 3.4 PROCEDIMENTOS PARA MINERAÇÃO DE OPINIÃO

O estudo considera duas fontes principais de informações, as redes sociais Twitter e Facebook. Uma pesquisa inicial foi realizada a fim de verificar quais as redes sociais mais usadas pelos turistas para registrar comentários sobre a cidade de Gramado/RS. Outras 5 fontes foram analisadas, Google, Google Blogs Tripadvisor RSS, Reclameaqui RSS e Yahoo Answers RRS. Porém, a quantidade de comentários extraídos é pouco significativa se comparada com as fontes principais.

### 3.4.1 Etapa da Coleta de Dados

A etapa de coleta de dados deste estudo está relacionada com as etapas necessárias para a Mineração de Opinião, conforme ilustra a figura 10:

Figura 10 - Etapa da Coleta de Dados



Fonte: Adaptado de TSYTSARAU; PALPANAS, 2012.

Para a extração dos dados foi utilizado o NodeXL versão PRO com licença acadêmica. Trata-se de um software de código aberto desenvolvido pela Social Media Research Foundation \*1, uma fundação de apoio à pesquisa em redes sociais sediada na Califórnia, US. O NodeXL funciona como um *add-on* do Microsoft Excel que inclui, entre outras funcionalidades, o acesso à rede de mídia social e à importadores de dados. O NodeXL apoia na exploração dos meios de comunicação social, com recursos de importação que extraem dados de redes como twitter, flickr, youtube, facebook, wikis e hiperlinks (NETWORK GRAPHS, 2015).

Segundo Hansen, Schneiderman e Smith (2010), o NodeXL possibilita uma análise e visualização de dados da rede mais fácil através da combinação de funções comuns de análise, visualização e facilidade de manipulação dos dados. A ferramenta implementa tarefas essenciais de análise com um grande número de usuários da rede ou mídia social. Um exemplo da interface do NodeXL pode ser visualizado na figura 11:

Figura 11 - Extração de Comentários do Twitter com o NodeXL

The screenshot displays the NodeXL Pro interface. On the left, a menu is open with 'From Twitter Search Network...' selected. The main window shows a table of extracted data with columns for user names, tweets, and timestamps. A dialog box titled 'Import from Twitter Search Network' is overlaid on the right, showing search options and a network diagram.

Relationship Date (UTC)	URLs in Tweet	Domains in Tweet	Hashtags in Tweet	Tweet Date (UTC)	Twitter Page for Twe	Latitude	Longitude
12/12/2015 11:35	https://www.swarma swarmapp.com				https://twitter.com/#!/guareis26/status/67564013361		
12/12/2015 12:23					65229905		
12/12/2015 12:31	http				67573888		
12/12/2015 17:39	http				676232975		
12/12/2015 18:07	http				676400188		
12/12/2015 19:41	http				67577781		
12/12/2015 19:47	http				682456404		
13/12/2015 01:08	http				508.787,4		
12/12/2015 19:27					589170205		
13/12/2015 12:00	http				67600867		
12/12/2015 22:24	https://www.swarma swarmapp.com				-2.937.989.641	-5.087.273,9	
13/12/2015 22:54	https://www.instagram.com/fontedoamore				-293.685.493	-508.787,4	
13/12/2015 23:02					https://twitter.com/#!/hurtzpcy/status/6761753512828		
13/12/2015 23:12		parisparade			https://twitter.com/#!/henriquesilvio/status/6761779		
13/12/2015 23:57					https://twitter.com/#!/housecrets/status/6761891298		

Fonte: Elaborado pelo autor.

As extrações foram realizadas durante o ano de 2015 e o número de comentários retornados foram limitados pela Application Program Interface (API) do Twitter, pelo tipo de licenciamento de uso do NodeXL e pela quantidade de comentários registrados e disponíveis na mídia social.

Tarefas de pré-processamento tais a identificação de *hashtags*, eliminação de duplicidades e tratamento de ambiguidades foram tratadas no NodeXL. Devido ao problema de ambiguidade quanto ao termo “Gramado” – comentários no contexto do esporte foram eliminados - a extração do Twitter teve sua configuração no NodeXL configurada para ignorar comentários com as palavras do quadro 6:

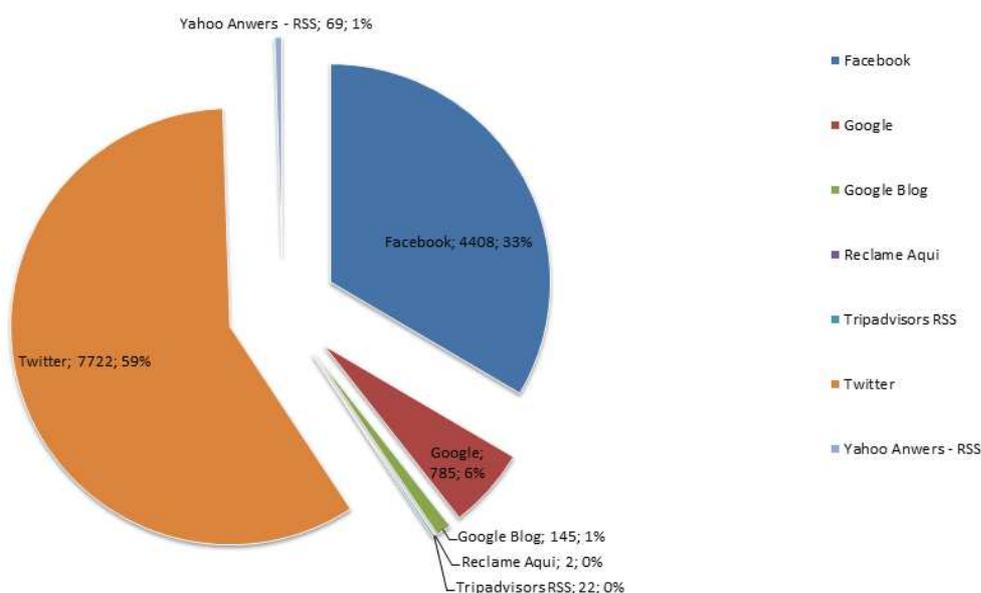
Quadro 6 - Configuração para Extração do Twitter

"gramado" -futebol -estadio -cruzeiro -espn -sporttv -esporte -brasileirão -bola -flamengo -time -campeonato -partida -brasileiro -gol -resultado -libertadores -g4 -rebaixamento -zona -corinthians -derrota -vitória -fluminense -coritiba -avai -figueirense -internacional -segundona -palmeiras -botafogo -mineirão -arena -vasco -maracanã -chupa -palmeiras -jogar -rivals -olimpica -river -santos -barcelona -barca -jogo -congresso -manifestantes -"o gramado" -"gramado do" -"gramado de" -"no gramado" -"do gramado" -"de gramado"

Fonte: Elaborado pelo autor.

A amostra final considerada para as próximas etapas da Mineração de Opinião foi 13.153 comentários e ficou distribuída conforme o gráfico 4:

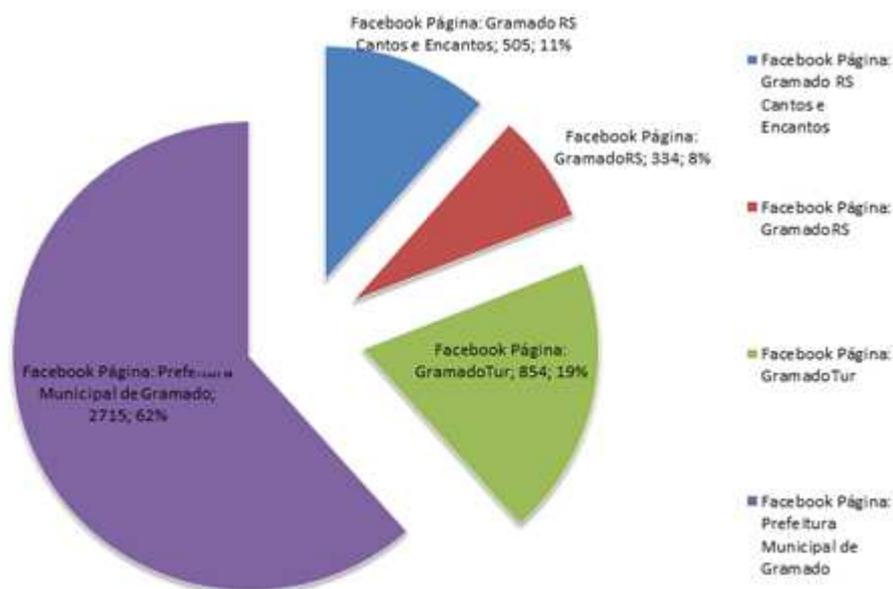
Gráfico 4 - Distribuição da Amostra Total



Fonte: Elaborado pelo autor.

Os 4.408 comentários extraídos da mídia Facebook foram realizadas em quatro páginas abertas e teve a distribuição conforme o gráfico 5:

Gráfico 5 - Distribuição da Amostra do Facebook



Fonte: Elaborado pelo autor.

### 3.4.2 A identificação da opinião ou sentimento

Esta etapa se refere ao Processamento da Linguagem Natural (PNL) e destinou-se à identificação da opinião ou sentimento de cada comentário coletado na etapa anterior, conforme figura 12:

Figura 12 - Etapa da Identificação do Sentimento



Fonte: Elaborado pelo autor.

Os termos sentimento e opinião frequentemente são usados como sinônimo neste contexto. A polaridade de um sentimento pode ser classificada em classes discretas (e.g. positiva, negativa ou neutra), ou como um intervalo que representa a intensidade deste sentimento, tipicamente [-1, 1].

Uma das abordagens para a extração automática de opinião ou sentimentos de textos considera o uso de léxicos de sentimentos, que podem ser gerais, como o SentiWordNet (ESULI; SEBASTIANI, 2006) ou específicos para determinadas tarefas (RILOFF; WIEBE, 2003; POIRIER et al., 2011). Outra abordagem utilizada por alguns autores usa o dicionário Wikipedia para determinar a polaridade de adjetivos vistos ao longo dos comentários (CHESLEY et al., 2006; WANG; DOMENICONI, 2008 ; WANG et al., 2013; CAMBRIA; HOWARD, 2014).

Várias ferramentas estão disponíveis para a execução da tarefa de identificação da opinião ou sentimento. A escolha depende da análise de vários critérios tais como a dimensão do problema que será analisado, o custo, o licenciamento, o sistema operacional e as técnicas disponibilizadas pela ferramenta.

Georgiou, MacFarlane e Russel-Rose (2015), em sua análise comparativa de ferramentas comerciais para análise de sentimentos, concluiu que Semantria e WEKA são amplamente utilizadas para análise de sentimento (CIELIEBAK; DURR; UZDILLI, 2013; FRANK et al., 2004; TURDAKOV et. al., 2014; WHITELAW; GARG; ARGAMON, 2005) e classificação. Outras alternativas foram analisadas como o SCUP e a biblioteca Natural Language Toolkit (NLTK) para Python. Este estudo levou em consideração requisitos como a facilidade de uso, usabilidade da ferramenta, custo, limite de buscas e idioma. Para a execução desta etapa e demais foi utilizada a versão do Semantria para o Microsoft Excel.

O Semantria é uma ferramenta on-line para análise de sentimento que utiliza o motor Saliency Lexalytics. Usado em companhias como Microsoft, Salesforce, Cisco e Thomson Reuters, o Semantria permite a extração de entidades, conceitos, resumos e sentimentos de dados não estruturados. O Semantria também faz uso de ligações entre artigos Wikipedia™ para desenvolver uma matriz de associações semânticas entre palavras-chave e frases, formando uma rede abrangente de grandes conceitos, temas e entidades que se ramificam de cada conceito, e as ligações entre todos estes artigos.

Desde 2011 um número de empresas e pesquisadores usaram Semantria para realizar análise de sentimento (LWARENCE, 2014; ASTON; LIDDLE; HU, 2014; ABEYWARDENA, 2014).

Segundo Georgiou, MacFarlane e Russel-Rose (2015), o Semantria pode melhorar o processo de tomada de decisões e garantir que um melhor serviço ao cliente através da análise dos consumidores.

Os 13.153 comentários coletados na etapa anterior foram submetidos à análise do Semantria. De acordo com Semantria, a extração de sentimentos em um documento segue as seguintes etapas:

- O documento dividido em partes do discurso (POS Tag);
- O algoritmo identifica e pontua o sentimento detectado em cada sentença e;
- Os escores ou pontuações são combinados para determinar a pontuação do sentimento geral no nível de documento.

Após o envio dos comentários para análise, uma tabela é retornada contendo a pontuação e polaridade para cada comentário analisado no nível de entidades, sentença e documento, conforme exemplo apresentado na tabela 2:

Tabela 2 - Modelo de Apresentação do Semantria

ID	Comentário	Sentimento detectado no document	Polaridade detectada no document	Intensificador	Negação	Sentenças	Sentimento detectado na sentença	Polaridade detectada na sentença	Entidade detectada	Tipo de Entidade	Sentimento detectado na entidade	Polaridade detectada na entidade	Sumarização
982	@Dropadex Eu amo Gramado. Os carros param pra ti, quase nada de lixo no chão, não vi nenhum cachorro de rua e nem mendigo. Acho lindo.	0,699999988	positive			amo	0,899999976	positive	@Dropadex	Pessoa	4,050000191	positive	Aspectos Ambientais
					nada	lixo	0,600000024	positive					Atrativos Turísticos
						lindo	0,600000024	positive					Desconforto Exclusão Social

Fonte: Elaborado pelo autor.

A pontuação final no nível de documento se dá através da combinação das pontuações ou scores parciais, conforme exemplo apresentado na tabela 3:

Tabela 3 - Pontuação final no nível do documento

ID	Comentário	Sentimento detectado no documento	Polaridade detectada no documento	Intensificador	Negação	Sentenças	Sentimento detectado na sentença
982	@Dropadex Eu amo Gramado. Os carros param pra ti, quase nada de lixo no chão, não vi nenhum cachorro de rua e nem mendigo. Acho lindo.	0,699999988	positive			amo	0,899999976
					nada	lixo	0,600000024
						lindo	0,600000024

Fonte: Elaborado pelo autor.

A definição do nível em que a opinião será classificada é de extrema importância, a este respeito Liu (2012) discorre que se um documento opinativo avalia mais de uma entidade então os sentimentos sobre as entidades podem ser diferentes, ou seja, o titular da opinião pode ter opiniões negativas em relação a algumas entidades e positivas sobre outras. O autor afirma ainda que, quando isto ocorre, não é recomendado que se faça uma classificação à nível de documento. (LIU, 2012).

### 3.4.3 A classificação da opinião ou sentimento

A próxima etapa da Mineração de Opinião, ilustrada na figura 13, destina-se a classificar cada comentário como Positivo, Neutro ou Negativo. Segundo Pang e Lee (2008), é tarefa de identificar se qualquer opinião expressa, sob a forma de um documento, sentença ou frase, é positiva ou negativa.

Figura 13 - Etapa de Classificação na Mineração de Opinião



Fonte: Elaborado pelo autor.

Os comentários foram classificados conforme uma escala de -2 a 2. A figura 14 ilustra o resultado apresentado pelo Semantria:

Figura 14 – Polaridade final no nível do documento

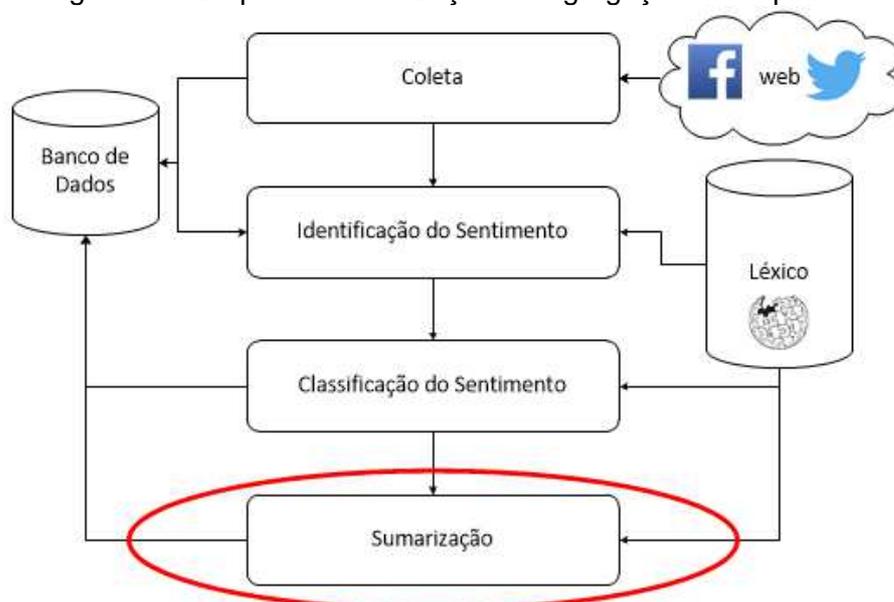
ID	Comentário	Sentimento detectado no documento	Polaridade detectada no documento	Intensificador	Negação	Sentenças	Sentimento detectado na sentença
982	@Dropadex Eu amo Gramado. Os carros param pra ti, quase nada de lixo no chão, não vi nenhum cachorro de rua e nem mendigo. Acho lindo.	0,699999983	positive			amo	0,899999976
					nada	lixo	0,600000024
						lindo	0,600000024

Fonte: Elaborado pelo autor.

### 3.4.4 A sumarização ou agregação

A última etapa da Mineração de Opinião, ilustrada na figura 15, trata da sumarização dos comentários de acordo com o assunto ou tema sobre o qual se falou. Sobre esta etapa, Liu (2012) destaca que a opinião de uma única pessoa não é suficiente para realizar a mineração de opinião. Na maioria das aplicações é preciso coletar um número expressivo de opiniões das pessoas. Este fato demonstra que é necessário sumarizar as informações (LIU, 2012), formando "o consenso dos sentimentos que os consumidores têm sobre um produto" (OSTROWSKI, 2010, p.1).

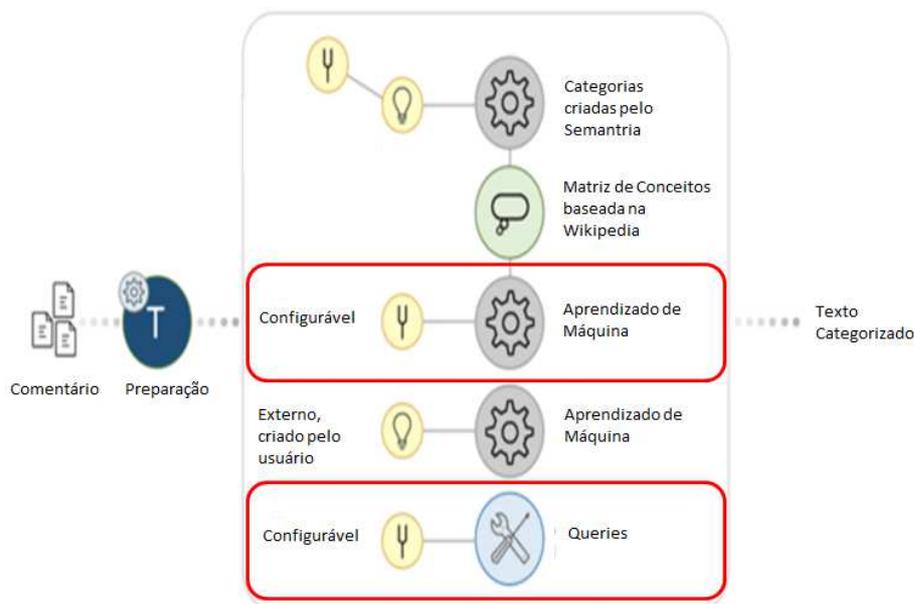
Figura 15 - Etapa de Sumarização ou Agregação das Opiniões



Fonte: Elaborado pelo autor.

O Semantria possui quatro alternativas de categorização, conforme a figura 16:

Figura 16 - Categorização no Semantria



Fonte: Adaptado de LEXALYTICS, 2015.

Este estudo utilizou dois métodos: a categorização com base na Matriz de Conceitos criada pelo Semantria e, adicionalmente, a categorização simples através de queries. A Matriz de Conceitos do Semantria baseia-se no dicionário Wikipedia (nota de rodapé) para encontrar uma relação com os termos configurados pelo usuário

Com a finalidade de aumentar a precisão, uma amostra de 10 comentários foi categorizada manualmente quanto às dimensões do ICT e, após, configurada e submetida à análise do Semantria até que os resultados fossem satisfatórios. Os termos configurados no Semantria foram baseados nos aspectos avaliados em cada destino turístico para cada dimensão obtidas no referencial sobre a criação e metodologia do ICT. Conforme ICT2015, os aspectos avaliados para cada dimensão são, respectivamente:

**a) Aspectos ambientais**

- Estrutura e legislação municipal de meio ambiente
- Atividades em curso potencialmente poluidoras
- Rede pública de distribuição de água
- Rede pública de coleta e tratamento de esgoto
- Coleta e destinação pública de resíduos
- Patrimônio natural e unidades de conservação no território municipal

**b) Infraestrutura geral**

- Capacidade de atendimento médico para o turista no destino

- Fornecimento de energia
- Serviço de proteção ao turista
- Estrutura urbana nas áreas turísticas

**c) Economia local**

- Aspectos da economia local
- Infraestrutura de comunicação
- Infraestrutura e facilidades para negócios
- Empreendimentos ou eventos alavancadores

**d) Aspectos culturais**

- Produção cultural associada ao turismo
- Patrimônio histórico-cultural
- Estrutura municipal para apoio à cultura

**e) Atrativos turísticos**

- Atrativos naturais
- Atrativos culturais
- Eventos programados
- Realizações técnicas, científicas ou artísticas
- Diversidade de atrativos, opções e equipamentos de lazer

**f) Capacidade empresarial**

- Capacidade de qualificação e aproveitamento do pessoal local
- Presença de grupos nacionais ou internacionais do setor de turismo
- Concorrência e barreiras de entrada
- Geração de negócios e empreendedorismo

**g) Acesso**

- Acesso aéreo
- Acesso rodoviário
- Acesso aquaviário
- Acesso ferroviário
- Sistema de transporte no destino
- Proximidade de grandes centros emissores de turistas

**h) Aspectos sociais**

- Acesso à educação
- Empregos gerados pelo turismo
- Uso de atrativos e equipamentos turísticos pela população
- Cidadania, sensibilização e participação na atividade turística
- Política de enfrentamento e prevenção à exploração de crianças e adolescentes

- i) Serviços e equipamentos turísticos**
  - Sinalização turística
  - Centro de atendimento ao turista
  - Espaços para eventos
  - Capacidade dos meios de hospedagem
  - Capacidade do turismo receptivo
  - Estrutura de qualificação para o turismo
  - Capacidade dos restaurantes
- j) Políticas públicas**
  - Estrutura municipal para apoio ao turismo
  - Grau de cooperação com o governo estadual
  - Grau de cooperação com o governo federal
  - Planejamento para a cidade e para a atividade turística
  - Grau de cooperação público-privada
- k) Cooperação regional**
  - Governança
  - Projetos de cooperação regional
  - Planejamento turístico regional
  - Roteirização
  - Promoção e apoio à comercialização de forma integrada
- l) Marketing e promoção do destino,**
  - Plano de marketing;
  - Participação em feiras e eventos;
  - Promoção do destino;
  - Estratégias de promoção digital;
- m) Monitoramento,**
  - Pesquisas de demanda;
  - Pesquisas de oferta;
  - Sistema de estatísticas do turismo;
  - Medição dos impactos da atividade turística;
  - Setor específico de estudos e pesquisas.

Outras nove categorias foram utilizadas conforme às dimensões criadas por Oliveira (2013):

- a) N1 - Hospitalidade dos residentes
- b) N2 - (In)segurança
- c) N3 - Preços
- d) N4 - (In)satisfação com a prestação dos serviços
- e) N5 - Assédio dos ambulantes

- f) N6 - Desconforto com a exclusão social
- g) N7 - Excesso de utilização do atrativo
- h) N8 - Locais para compras
- i) N9 - Poluição

Por fim, uma nova categoria foi criada pelo autor para cobrir agregar os comentários dos turistas sobre as condições climáticas. Segundo Thompson e Perry (1997), as condições climáticas, ainda que não seja um impedimento à realização da viagem, terão certamente um importante peso na satisfação do turista.

Após a configuração dos dicionários e posterior categorização dos comentários, o Semantria cria um índice de relevância ou pontuação, de 0 a 1, que representa o grau de confiança na categoria encontrada. Quanto mais alto o grau, maior a confiança. Os 13.153 comentários foram categorizados de forma automática, resultando em 21.037. Todas as categorizações foram consideradas, implicando em mais uma categorização do mesmo comentário, conforme exemplo do quadro 7.

Quadro 7 - Categorização no Semantria

@Dropadex Eu amo Gramado. Os carros param pra ti, quase nada de lixo no chão, não vi nenhum cachorro de rua e nem mendigo. Acho lindo.	Aspectos Ambientais	1,2402215
	Poluição	1,17342758
	Desconforto Exclusão Social	0,78387749
	Locais para compras	0,45379543

Fonte: Elaborado pelo autor

### 3.5 DELIMITAÇÃO DO ESTUDO

Este estudo está restrito à análise de duas redes sociais com uma ferramenta de Mineração de Opinião. Outras redes sociais foram analisadas preliminarmente. Esse procedimento foi necessário para a identificação das redes sociais sobre as quais o estudo de dedicaria. As configurações estarão limitadas ao escopo das dimensões do ICT e dimensões abordadas no referencial.

O estudo também esta delimitado às opiniões mineradas referentes exclusivamente ao município de Gramado/RS, apesar de que os resultados podem ser aplicados ao contexto de outros destinos turísticos nacionais.

A análise dos resultados de forma detalhada está restrita a algumas dimensões selecionadas. São elas: Condições Climáticas, Atrativos Turísticos e Segurança. A escolha foi baseada na importância destas dimensões ou variáveis no

contexto de Gramado/RS e pela importância das mesmas identificadas em outros estudos sobre fatores de motivação na escolha do destino turístico.

Rodrigues e Mallou (2014), em seu estudo sobre a influência da motivação na intenção de escolha de um destino turístico em tempo de crise econômica, demonstram que a segurança, hospitalidade, clima e os atrativos turísticos estão entre os atributos do destino que mais influenciam na decisão de escolha do destino, como mostra a tabela 4:

Tabela 4 - Motivações – Fatores Pull

<b>Indicador</b>	<b>Média</b>	<b>Desvio padrão</b>
Paisagens	4,39	0,72
Ambiente Natural	4,36	0,65
Atrações Culturais	4,33	0,76
Diferentes Culturas	4,16	0,88
Segurança	4,09	0,80
Hospitalidade	4,04	0,74
Clima	3,95	1,01
Atmosfera Relaxante	3,89	0,89
Alojamento	3,86	0,90
Gastronomia	3,82	0,86
Acessibilidades	3,78	0,83
Ambiente Social	3,75	0,98
Praias	3,59	1,27
Padrão de Vida	3,57	0,85
Transportes	3,43	0,92
Distância	3,29	0,88
Vida Noturna	3,24	1,09
Equipamentos Desportivos	2,69	1,09
Estabelecimentos Comerciais	2,62	1,03

Fonte: RODRIGUES; MALLOU, 2014.

Sobre a dimensão “Atrativos Turísticos”, a análise em particular desta dimensão também se deve a importância do turismo de eventos, em especial o Natal Luz, na consolidação do município de Gramado/RS como destino turístico. Dalonso

(2012) conclui que o forte potencial existente na cidade juntamente com um trabalho em rede dos diferentes atores locais contribuíram para a sua consolidação como destino turístico de excelência a partir do evento natalino. O Natal Luz se caracterizou como o grande responsável pela consolidação de Gramado como destino turístico de referência no Brasil. Para a SETUR, 2010, sob o ponto de vista da competitividade, portanto, o Natal luz impactou de forma efetiva a dimensão “atrativos turísticos” – especificamente a variável “eventos programados”.

A dimensão (In)segurança está relacionada à relacionada a dimensão Hospitalidade do Destino, esta considerada um fenômeno social, que está presente em todos os contextos e interações dos seres humanos. Aproxima as pessoas e possibilita trocas de conhecimento e afetos, apresentando-se como fator fundamental para o turismo (GENEROSI, 2011). O turista se desloca na busca de conhecer o outro lugar, o novo. Porém, é por meio da hospitalidade, pelo processo de interação social, pelas trocas realizadas entre anfitrião e hóspede, que a prática turística se efetiva e se potencializa (OLIVEIRA; GENEROSI; TOMAZZONI, 2010).

Por fim, a dimensão Condições Climáticas, categoria criada pelo autor para consolidar as opiniões sobre clima e tempo e sua influência na vivência da experiência turística ou na decisão de escolha do destino.

Também foi arbitrado pelo autor desconsiderar da análise os comentários de comentários da rede social Facebook, para fins de equalização à estrutura extraída do Twitter, bem como os comentários que não estivessem no idioma Português, assim como não houve distinção entre comentários envolvendo turismo de laser e negócios.

Sobre a rede social Facebook em especial, os comentários não estão disponíveis para toda a rede. Para a coleta de comentários do Facebook, este estudo arbitrou quatro páginas abertas criadas especificamente para a discussão sobre o município de Gramado, em geral. São elas: Gramado RS Cantos e Encantos, GramadoRS, GramadoTur e Prefeitura Municipal de Gramado.

## 4 APRESENTAÇÃO DO ESTUDO

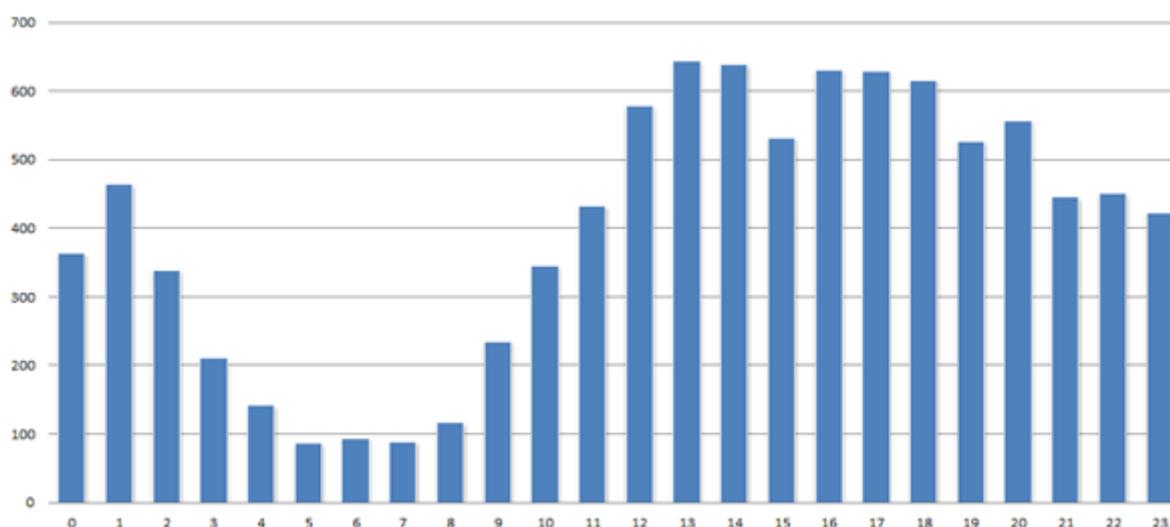
Este capítulo é destinado à apresentação, análise e discussão sobre os resultados encontrados neste estudo. Inicialmente, será abordado o contexto de aplicação da Mineração de Opinião. Após, serão abordados os resultados encontrados seguidos da análise e discussão.

Conforme abordado na Metodologia deste estudo, duas redes sociais foram escolhidas para a aplicação da Mineração de Opinião: o Twitter e o Facebook. O Twitter permite se extrair opiniões sobre determinado assunto, já que se trata de uma rede social com foco em opiniões dos usuários propriamente, além de estarem disponíveis a todos os usuários da rede. O Facebook possui uma grande quantidade de usuários que escrevem suas opiniões de forma textual ou na forma de “likes”.

### 4.1 RESULTADOS DA ETAPA DA COLETA

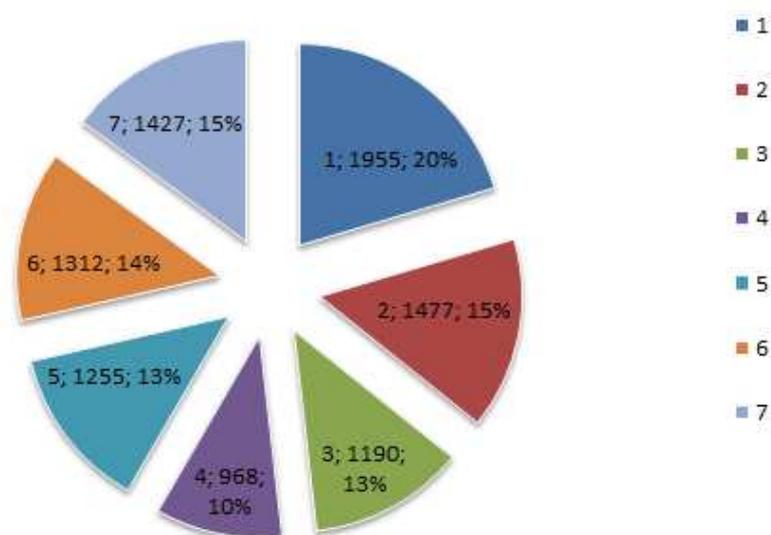
Conforme abordado no capítulo anterior, ao final da primeira etapa da etapa da Mineração de Opinião, a amostra disponível para análise de sentimento era de 13.153 comentários. Atributos como data da postagem do comentário, horário da postagem do comentário, local de origem, idioma, hashtags e número de likes foram identificados e extraídos pela ferramenta utilizada para a extração dos comentários. Estes atributos são apresentados nos gráficos 6 e 7:

Gráfico 6 - Distribuição da amostra conforme horário do post comentário



Fonte: Elaborado pelo autor.

Gráfico 7 - Distribuição da amostra conforme dia do post comentário



Fonte: Elaborado pelo autor.

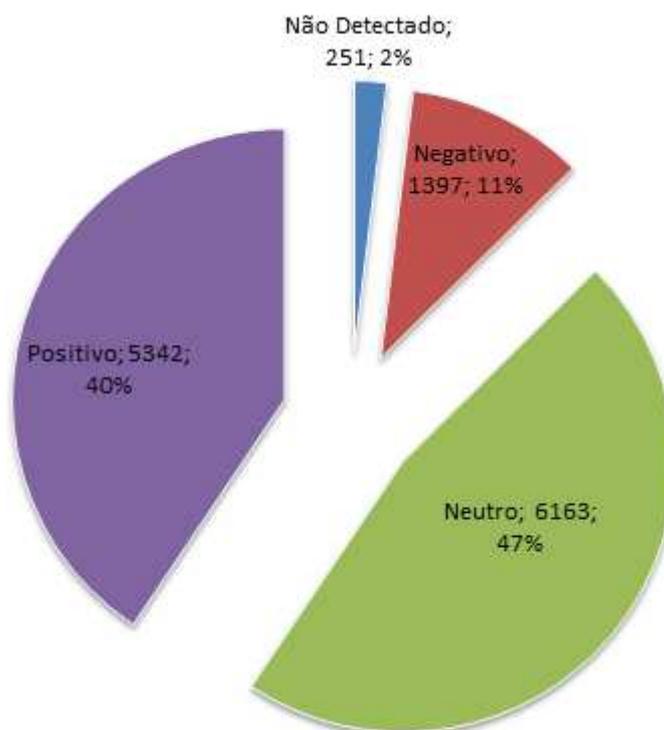
A etapa de coleta, através do NodeXL, identificou a origem ou localização do autor do comentário para 100% da amostra coletada da rede social Twitter. A distribuição dos comentários conforme a localização pode ser observada na nuvem de palavras apresentada na figura 17:

Figura 17 – Localizações Extraídas

alvorada (12) aracaju (10) argentina (17) bahia (11) bangu (7) barranquilla (13) belem (16) belo-horizonte-  
 --minas-gerais (43) belo-horizonte (40) bh (11) blumenau (8) bogota (19) br (12)  
**brasil** (456) brazil (39) buenos-aires (9) ca (24) cachoeirinha (11) campinas (14) canela (22) canoas (13)  
 caxias-do-sul (29) ceara (23) colombia (34) **curitiba** (59) ecuador (14) espirito-santo (18)  
 florianopolis (38) floripa (19) fortaleza (34) gaucha (16) goiania (28) goias (14) gramado---rs---  
 brasil (7) gramado---rs (18) gramado-rs (20) **gramado** (97) gravatai (36) guayaquil (9) hogwarts (7)  
 joao-pessoa (8) joinville (9) lages (9) los-angeles (12) maceio (11) manaus (12) medellin (14) mg (22) minas-gerais (36)  
 montenegro (7) morumbi (7) natal (14) niteroi (14) nova-petropolis (10) novo-hamburgo (21) osasco (11) paraiba (8)  
 parana (36) passo-fundo (14) pelotas (32) pernambuco (17) peru (8) poa (20) porto-alegre---  
 rs (42) porto-alegre-rs (15) **porto-alegre** (343) portugal (7) pr (10) recife (25) rio-de-janeiro---  
 brasil (8) **rio-de-janeiro** (297) rio-grande-do-su (12) **rio-grande-**  
**do-sul** (422) rio-grande (16) rj (71) rs---brasil (19) **rs** (216) salvador (17) santa-  
 catarina (71) santa-maria (18) santo-andre (8) sao-januario (10) **sao-leopoldo** (26) sao-luis (10) sao-paulo-  
 --brasil (16) sao-paulo---sp (14) **sao-paulo** (320) sc (34) **snap** (27) snapchat (13) sorocaba (8) sp---  
 brasil (8) **sp** (63) taquara (14) usa (15) ut (14) venezuela (9) viamao (16) vitoria (7) zona-leste (7)

Fonte: Elaborado pelo autor.





O número de comentários categorizados, cujos resultados são apresentados na próxima seção, é de 12.092.

#### 4.3 RESULTADOS DA ETAPA DE SUMARIZAÇÃO OU AGREGAÇÃO

De acordo com os objetivos específicos deste estudo, um dicionário de termos foi criado com base nas dimensões de análise de competitividade do destino. Dois dicionários foram criados e configurados conforme especificado na metodologia deste estudo. Estes dicionários foram configurados na ferramenta de análise para agregação automática dos comentários. O dicionário criado e configurado para a categorização baseada na Matriz de Conceitos é apresentado no quadro 8 e o dicionário criado e configurado para a categorização baseada em queries com algoritmos booleanos é apresentado no quadro 9:

Quadro 8 - Dicionário criado e configurado para a categorização baseada na Matriz de Conceitos

Índice	Dimensão	Dicionário
1	Aspectos Ambientais	poluição, esgoto, lixo, reciclagem, conservação, meio_ambiente, coleta_seletiva
2	Infraestrutura Geral	atendimento_médico, policia, delegacia, água, energia, militar, civil, bombeiro, vacinação, infraestrutura
3	Economia Local	negócios, empreendimento, investimento, expansão, salário, juros, pib
4	Aspectos Culturais	cultura, patrimônio, folclore, imigração, costume, artístico, arqueológico
5	Atrativos turísticos	neve, chocolate, natal_luz, festival, cinema, lago_negro, fondue, esqui, atrações, decoração, igreja, rua_coberta, ingresso
6	Capacidade Empresarial	Qualificação, mão_de_obra, barreiras, empresas, salário, concorrência, regulamentação
7	Acesso	acesso, ônibus, rodoviária, aeroporto, estrada, taxi, transfer, deslocamento, estacionamento_público
8	Aspectos Sociais	emprego, desemprego, educação, exploração_sexual, cidadania, longevidade, analfabetismo
9	Serviços e equipamentos turísticos	sinalização_turística, centro_de_informações, centro_de_eventos, gramadotur, feira, agência_de_turismo, guia_de_turismo, serviços_públicos, atendimento
10	Políticas públicas	governo, federal, estadual, secretaria, planejamento, política, políticas, autarquia, departamento
11	Cooperação Regional	Governança, cooperação, roteirização, integração, canela, conselho, associação, serra_gaúcha

(Continuação)

Índice	Dimensão	Dicionário
12	Marketing	propaganda, website, marketing, promoção, divulgação, pesquisa_de_opinião, folder
13	Monitoramento	pesquisa_de_demanda, pesquisa_de_oferta, estatística, pesquisa
N1	Hospitalidade dos Residentes	receptividade, hospitalidade, cortesia, cordialidade
N2	Assédio dos Ambulantes	ambulantes, pirataria, muamba
N3	Desconforto Exclusão Social	exclusão, mendigo, miséria, fome, esmola, indigente
N4	Insegurança	assalto, roubo, crime, estupro, morte, acidente, agressão, delito, droga, sinistro
N5	Qualidade dos Serviços	taxi, ônibus, estacionamento, pedágio, hospital, informações, hotel, restaurante, gasolina, rodoviária, aeroporto, aluguel, internet, pousada
N6	Preços	caro, barato, crédito, pagamento, preço, inflação
N7	Locais para compras	loja, shopping, comercial, horário, abertura, fechamento
N8	Excesso Utilização Atrativo	engarrafamento, lotado, esgotado
N9	Poluição	poluição, fumaça, cigarro, lixo, sujeira
X1	Tempo	frio, calor, chuva, nublado, úmido, tempestade, meteorologia

Fonte: Elaborado pelo autor.

Quadro 9 - Dicionário criado e configurado para a categorização baseada em queries com algoritmos booleanos

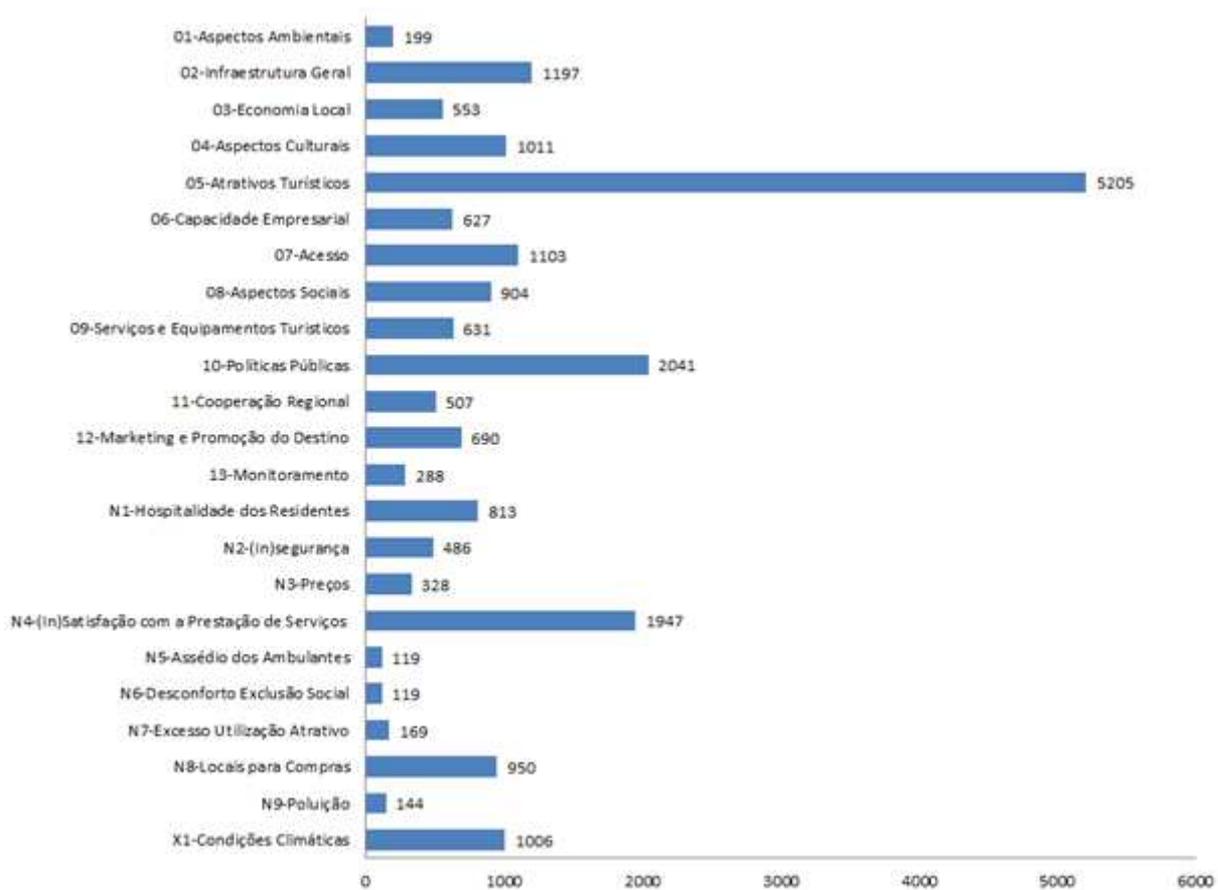
Índice	Dimensão	Dicionário
1	Aspectos Ambientais	poluição OR esgoto OR lixo OR reciclagem OR conservação OR "meio ambiente" OR "coleta seletiva"
2	Infraestrutura Geral	"atendimento médico" OR policia OR delegacia OR água OR energia OR militar OR civil OR bombeiro OR vacinação OR infraestrutura
3	Economia Local	negócios OR empreendimento OR investimento OR expansão OR salário OR juros OR pib
4	Aspectos Culturais	cultura OR patrimônio OR folclore OR imigração OR costume OR artístico OR arqueológico
5	Atrativos turísticos	neve OR chocolate OR "natal luz" OR festival OR cinema OR "lago negro" OR fondue OR esqui OR atrações OR decoração OR igreja OR "rua coberta" OR ingresso OR FRIO OR FRIA OR "LAGO NEGRO" OR CARACOL OR NEVE OR JARDIM OR JARDINS OR NEBLINA OR CHUVA OR GRANIZO OR GRAUS OR TEMPO OR TEMPERATURA OR INSTABILIDADE OR "SERRA GAÚCHA" OR "VALE DO QUILOMBO" OR "IGREJA MATRIZ" OR "SÃO PEDRO" OR CHOCOLATE OR CHOCOFEST OR "FESTIVAL DE CINEMA" OR "RUA COBERTA" OR COMER OR COMIDA OR "GRAMADO ZOO" OR FONDUE OR "PALÁCIO DOS FESTIVAIS" OR FESTIVAL OR "SERRA AZUL" OR KIKITO OR TROFÉU OR PRÊMIO OR FICÇÃO OR COMÉDIA OR DOCUMENTÁRIO OR DRAMA OR CURTA OR FILME OR "LONGA METRAGEM" OR ALPEN OR PARQUE OR "NATAL LUZ" OR MINIMUNDO OR "MINI MUNDO" OR FLORYBAL OR "SUPER CARROS" OR INGRESSO OR TICKET OR "HOLLYWOOD DREAM CARS" OR CULTURA OR "MUNDO ENCANTADO" OR NATIVITATEN OR SNOW OR SNOWBOARDING OR "MUSEU DE CERA" OR "MARIA FUMAÇA" OR "CASA SEGANFREDO" OR EVENTO OR CONGRESSO OR ENCONTRO OR AMOSTRA OR "PRÉ TEMPORADA" OR ARTESANATO OR RÚSTICA OR SICC OR FESTURIS OR TORNEIO OR REAVIVA
6	Capacidade Empresarial	Qualificação OR "mão de obra" OR barreiras OR empresas OR salário OR concorrência OR regulamentação
7	Acesso	acesso OR ônibus OR rodoviária OR aeroporto OR estrada OR taxi OR transfer OR deslocamento OR "estacionamento público" OR pedágio OR RS-235 OR "RS 235" OR RS235 OR RS-235 OR "RS 235" OR RS235 OR RODOVIÁRIA OR LINHA OR AEROPORTO OR TRASLADO OR TRANSLADO OR VAN OR RS-115 OR RS 115 OR RS115 OR RS-235 OR "RS 235" OR RS235 OR RODOVIÁRIO OR RODOVIÁRIA OR AÉREO OR PEDÁGIO OR MULTA OR LONGE OR ESTRADA OR RUA OR AVENIDA OR FAIXA OR PERTO OR ÔNIBUS OR VAN OR TAXI OR CARRO OR PASSAGEIRO OR MOTORISTA OR ASSENTO OR INFRA OR DESEMBARQUE OR TRANSLADO OR TRASLADO OR AVIÃO OR DISTÂNCIA OR ENGARRAFAMENTO OR CONGESTIONAMENTO OR ESTACIONAMENTO OR EXECUTIVO OR CONVENCIONAL
Índice	Dimensão	Dicionário
8	Aspectos Sociais	emprego OR desemprego OR educação OR "exploração sexual" OR cidadania OR longevidade OR analfabetismo
9	Serviços e equipamentos turísticos	"sinalização turística" OR "centro de informações" OR "centro de eventos" OR gramadotur OR feira OR "agência de turismo" OR "guia de turismo" OR "serviços públicos" OR atendimento OR GRAMADOTUR
10	Políticas públicas	governo OR federal OR estadual OR secretaria OR planejamento OR política OR políticas OR autarquia OR departamento
11	Cooperação Regional	Governança OR cooperação OR roteirização OR integração OR canela OR conselho OR associação OR "serra gaúcha"
12	Marketing	propaganda OR website OR marketing OR promoção OR divulgação OR "pesquisa de opinião" OR folder
13	Monitoramento	"pesquisa de demanda" OR "pesquisa de oferta" OR estatística OR pesquisa

Índice	Dimensão	Dicionário
N1	Hospitalidade dos Residentes	receptividade OR hospitalidade OR cortesia OR cordialidade
N2	Assédio dos Ambulantes	ambulantes OR pirataria OR muamba OR PIRATA OR PIRATARIA
N3	Desconforto Exclusão Social	exclusão OR mendigo OR miséria OR fome OR esmola OR indigente
N4	Insegurança	assalto OR roubo OR crime OR estupro OR morte OR acidente OR agressão OR delito OR droga OR sinistro
N5	Qualidade dos Serviços	taxi OR ônibus OR estacionamento OR pedágio OR hospital OR informações OR hotel OR restaurante OR gasolina OR rodoviária OR aeroporto OR aluguel OR internet OR pousada
N6	Preços	caro OR barato OR crédito OR pagamento OR preço OR inflação
N7	Locais para compras	loja OR shopping OR comercial OR horário OR abertura OR fechamento
N8	Excesso Utilização Atrativo	engarrafamento OR lotado OR esgotado
N9	Poluição	poluição OR fumaça OR cigarro OR lixo OR sujeira
X1	Condições Climáticas	frio OR calor OR chuva OR nublado OR úmido OR tempestade OR meteorologia

Fonte: Elaborado pelo autor.

Após a configuração dos dicionários e posterior categorização dos comentários, o Semantria retornou um índice de relevância para cada categoria versus comentário analisado e uma categoria de associado em uma análise. Conforme abordado na metodologia deste estudo, este índice de relevância ou pontuação, de 0 a 1, representa o grau de confiança do Semantria para a categoria encontrada. Quanto mais alto o grau, maior a confiança. Nos resultados apresentados a seguir, todas as categorizações foram consideradas, implicando em mais uma categorização do mesmo comentário, conforme demonstrado no gráfico 8:

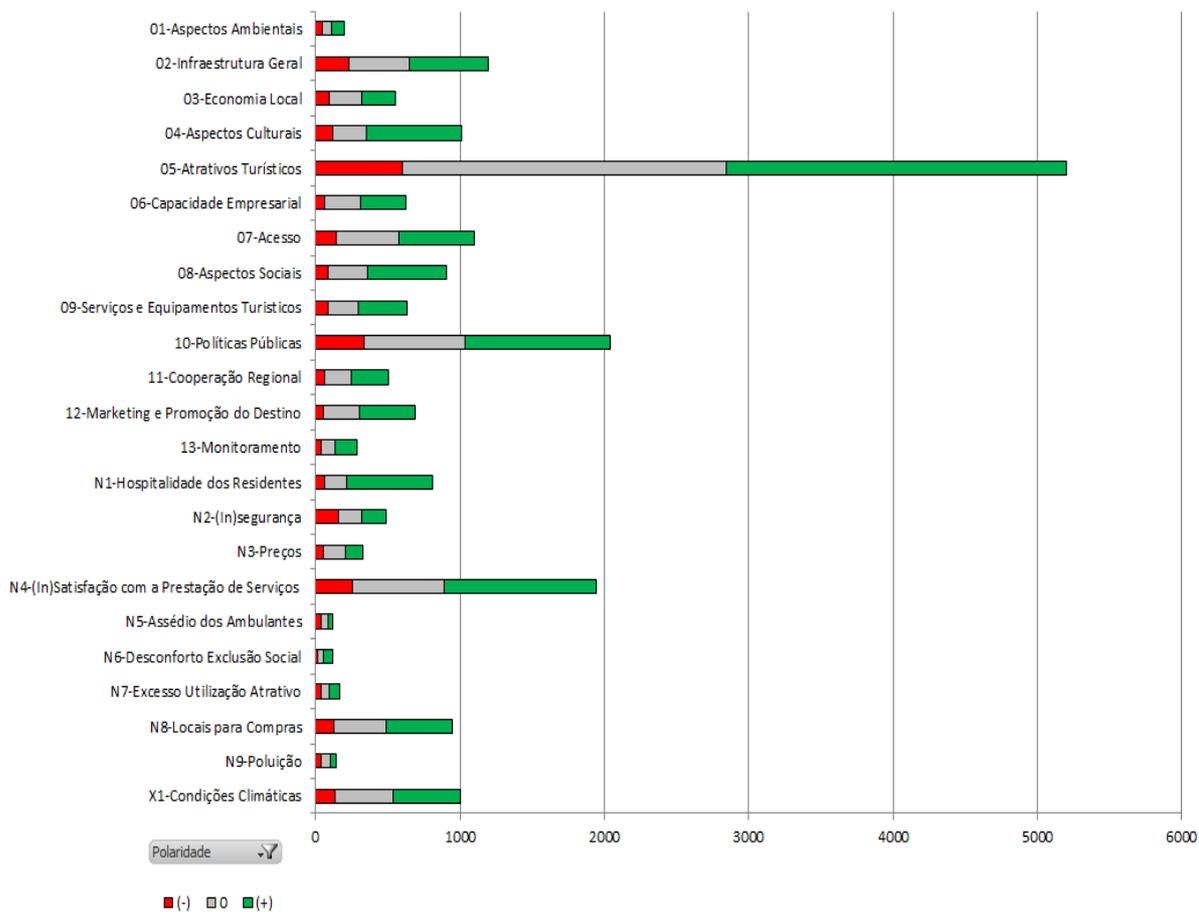
Gráfico 8 - Frequências por Categoria ou Dimensão de Análise



Fonte: Elaborado pelo autor.

As frequências das opiniões ou sentimentos, de acordo com cada categoria encontradas, são demonstradas no gráfico 9 e na tabela 5:

Gráfico 9 - Frequências por Sentimento x Categoria ou Dimensão de Análise



Fonte: Elaborado pelo autor.

Tabela 5 - Frequências por Sentimento x Categoria ou Dimensão de Análise

Dimensão de Análise	Polaridade (-)	Neutro	Positivo (+)	Total na Dimensão
01-Aspectos Ambientais	46	67	86	199
02-Infraestrutura Geral	228	422	547	1197
03-Economia Local	96	220	237	553
04-Aspectos Culturais	121	233	657	1011
05-Atrativos Turísticos	602	2240	2363	5205
06-Capacidade Empresarial	61	251	315	627
07-Acesso	145	431	527	1103
08-Aspectos Sociais	86	270	548	904
09-Serviços e Equipamentos Turísticos	85	211	335	631
10-Políticas Públicas	337	697	1007	2041
11-Cooperação Regional	60	187	260	507
12-Marketing e Promoção do Destino	54	249	387	690
13-Monitoramento	41	96	151	288
N1-Hospitalidade dos Residentes	59	152	602	813
N2-(In)segurança	159	157	170	486
N3-Preços	54	154	120	328
N4-(In)Satisfação com a Prestação de Serviços	257	633	1057	1947
N5-Assédio dos Ambulantes	36	47	36	119
N6-Desconforto Exclusão Social	14	42	63	119
N7-Excesso Utilização Atrativo	34	57	78	169
N8-Locais para Compras	130	360	460	950
N9-Poluição	36	66	42	144
X1-Condições Climáticas	137	399	470	1006
<b>Total Geral</b>	<b>2878</b>	<b>7641</b>	<b>10518</b>	<b>21037</b>

Fonte: Elaborado pelo autor.

#### 4.4 ANÁLISE DE RESULTADOS

A análise dos resultados é realizada segundo a categorização criada pela ferramenta de mineração. O conteúdo da amostra de 21.037 comentários permitiu explorar algumas dimensões e, conseqüentemente, melhor compreender algumas características do comportamento do consumidor em turismo nas redes sociais. Merece destaque as dimensões com maior. Apesar de 100% dos resultados demonstrarem a predominância do número de comentários com polaridade positiva sobre o número de comentários com polaridade negativa, a ênfase é dada às dimensões com grande número de comentários em geral, uma vez em que na amostra maior pode conter um número maior de comentários com informação relevante para o aprimoramento do produto turístico, objetivo principal deste estudo. Para uma análise em particular, foram selecionadas 3 dimensões, Atrativos Turísticos, (In)segurança e Condições Climáticas, dimensões relacionadas ao modelo de Barbosa (2012), Oliveira (2013) e proposta neste estudo, respectivamente.

#### 4.4.1 Dimensão de Análise: Atrativos Turísticos

O ICT, baseado no modelo de Barbosa (2012), inclui na avaliação da dimensão Atrativos Turísticos as variáveis Atrativos naturais, Atrativos culturais; Eventos programados, Realizações técnicas, científicas ou artísticas e Diversidade de atrativos e equipamentos de lazer. Nos últimos sete anos, o índice nesta dimensão tem diminuído de forma discreta, conforme demonstrado no gráfico 10:

Gráfico 10 - Índices Atrativos turísticos – destino x Brasil x não capitais: 2008-2015



Fonte: FGV/MTUR/SEBRAE, 2015

Entre os pontos fracos apontados pelo ICT estão:

- Inexistência de estudo de capacidade de carga e de monitoramento da quantidade de visitantes no principal atrativo natural do destino, o Lago Negro;
- Carência de recursos que viabilizem a circulação de pessoas com deficiência ou mobilidade reduzida no Lago Negro;
- Inexistência de estudo de capacidade de carga e de controle do número de visitantes para o Roteiro Raízes Coloniais;

- Ausência de estruturas de apoio no Tour Raízes Coloniais, como sinalização descritiva, restaurantes, balcão ou centro de visitantes, outros e;
- Carência de recursos que confirmam plena acessibilidade para pessoas com deficiência nos locais onde ocorre o Natal Luz.

Conforme pode ser verificado na tabela 5, sobre a amostra total de 21.037 comentários, 5.205 comentários estão relacionados à Atrativos Turísticos. Destes, 602 comentários foram classificados com polaridade negativa. Uma amostra selecionada de comentários categorizados na dimensão em questão demonstra descontentamento com fatores que, direta ou indiretamente, estão ligados a outras dimensões, mas que podem influenciar na decisão do turista em viajar. Na tabela 6 o comentário "...tem algum evento hoje a noite apesar dessa chuva.." é colocado com o intuito de receber recomendações de atrações para dias com chuva. De fato, o comentário está associado à existência de atrativos culturais, mas também poderia estar categorizado como um assunto ou opinião relacionado à "Condições Climáticas" ou à "Marketing do destino".

Tabela 6 - Amostra da Dimensão de Análise Atrativos Turísticos

Comentário	Polaridade do Documento	Dimensão de Análise
raíssa falou de gramado, bem que eu queria mas os pacotes são caríssimos para nosso orçamento de estudante pobre	Negativa	Atrativos turísticos
as pessoas de gramado ficaram aterrorizadas, com os pelotense louqueando na frente de um colégio	Negativa	Atrativos turísticos
Ontem fui pra gramado só pra comprar chocolate, pq tava uma neblina horrível e não dava pra enxergar nada	Negativa	Atrativos turísticos
Oi tem algum evento hoje, a noite a pesar dessa chuva obrigada...	Negativa	Atrativos turísticos

Fonte: Elaborado pelo autor.

#### 4.4.2 Dimensão de Análise: (In)segurança

Oliveira (2013) discutiu um método para captar a experiência turística em destinos que foi testado com diferentes tipos de turistas por meio de uma pesquisa qualitativa. Os dados foram submetidos à análise de conteúdo e à luz do modelo brasileiro de competitividade. Nos resultados apresentados, a sensação de insegurança foi relatada por 29% dos entrevistados como sendo um evento crítico na experiência de viagem. Conforme pode ser observado na tabela 7, o sentimento de insegurança foi manifestado devido a quatro causas: assaltos de rua; mendicância; consumo de drogas e assédio com o propósito de venda de produtos ou serviços (OLIVEIRA, 2013). Após a etapa de Sumarização ou Agregação, 486 comentários estão relacionados à (In)segurança. Destes, 159 comentários foram classificados com polaridade negativa.

Tabela 7 - Amostra da Dimensão de Análise (In)segurança

Comentário	Polaridade do Documento	Dimensão de Análise
Briga de vizinhos acaba em morte em Gramado <a href="http://t.co/laDNFLhk6L">http://t.co/laDNFLhk6L</a>	Negativa	(in)segurança
Gerente de lancheria é preso por tráfico de drogas em Gramado. <a href="http://t.co/wJFIS4Gv7S">http://t.co/wJFIS4Gv7S</a> <a href="http://t.co/jszBlzMvc6">http://t.co/jszBlzMvc6</a>	Negativa	(in)segurança
Traficante foge de cerco no Centro de Gramado. <a href="http://t.co/97JoQIDmZI">http://t.co/97JoQIDmZI</a>	Negativa	(in)segurança
Cara prefeitura peço que vocês averiguem os motoristas que contratam, agora de tardinha vi o carro 036 fazendo uma ultrapassagem um tanto quando perigosa quase ocasionando um acidente. Acho que o exemplo deve partir de cima 😊	Negativa	(in)segurança

Fonte: Elaborado pelo autor.

Segundo Cavaco e Simões (2009), o turismo é bastante influenciado por condições estruturais, nomeadamente pelos recursos naturais dos destinos, tendências dentro do próprio setor, mudanças de maior profundidade e de longo prazo, assim como por fatores conjunturais, como crises económicas, problemas de segurança ou flutuações do preço dos combustíveis.

#### 4.4.3 Dimensão de Análise: Condições Climáticas

A motivação para configuração de uma dimensão específica para receber os resultados relacionados às condições climáticas foi a importância dada pelo referencial teórico sobre a variável e a representação da frequência de comentários em relação à amostra total e a falta de uma categoria que pudesse receber esses resultados no modelo adotado pelo ICT. A frequência desta dimensão pode ser visualizada na tabela 8:

Tabela 8 - Frequências para a Dimensão Condições Climáticas

Dimensões de Análise	Negativo (-)	Neutro	Positivo (+)	Total	Freq Relativa
05-Atrativos Turísticos	602	2.240	2.363	5.205	24,74%
10-Políticas Públicas	337	697	1.007	2.041	9,70%
N4-(In)Satisfação com a Prestação de Serviços	257	633	1.057	1.947	9,26%
02-Infraestrutura Geral	228	422	547	1.197	5,69%
07-Acesso	145	431	527	1.103	5,24%
04-Aspectos Culturais	121	233	657	1.011	4,81%
X1-Condições Climáticas	137	399	470	1.006	4,78%
N8-Locais para Compras	130	360	460	950	4,52%

Fonte: Elaborado pelo autor.

Conforme abordado na metodologia deste estudo, a nova categoria foi criada para receber comentários relacionados às condições ou fatores climáticos, sejam eles tratados também como atrativos, como o frio e a neve, sejam eles tratados como fatores de diminuição da demanda turística. Mudanças climáticas fora do esperado para um determinado destino podem vir a provocar a diminuição da demanda turística. Daí a importância de se diversificarem as atividades e atrações existentes num destino, a fim de que este se torne menos vulnerável (IGNARRA, 1999. LOHMANN, 2004).

Matzarakis (2006) enfatiza a influência do clima e do estado do tempo nas três fases de uma viagem: antes, durante e depois. Isto é válido quer para os turistas, quer para a indústria turística. Na perspectiva do turista, o clima revela-se particularmente importante no momento anterior à viagem uma vez que, como se viu atrás, o clima expectável influi na decisão de viajar e na escolha do destino.

Fernandes (2009) analisa os fundamentos teóricos e práticos da relação clima-turismo para um melhor planejamento turístico e salienta:

A relação clima-turismo exige planejamento e a estratégia é adequar sistematicamente as diversas informações climatológicas e meteorológicas ao fator turismo. O conceito de turismo por si só já aborda e facilita a importância da união com Climatologia e a Meteorologia, pois devido ao deslocamento que o turista tem que fazer, ou seja, visitar uma região diferente e com condições meteorológicas e climatológicas divergentes das quais ele está acostumados presenciar, o planejamento climático-meteorológico torna-se imprescindível. Vale ressaltar que os danos causados pelos fenômenos da natureza, quando não anunciados de forma antecipada, são bastante prejudiciais às cidades e às pessoas, e ainda mais quando se trata de uma localidade turística. (FERNANDES, 2009, p.16).

Tabela 9 - Amostra da Dimensão de Análise Condições Climáticas

Comentário	Polaridade do Documento	Dimensão de Análise
Eu só queria ta em Gramado agora, essa merda dessa chuva fode tudo..	Negativa	Condições Climáticas
Ontem fui pra gramado só pra comprar chocolate, pq tava uma neblina horrível e não dava pra enxergar nada	Negativa	Condições Climáticas
@ingridzomk frio tipo Gramado um pouco pior. Mcksnsksk	Negativa	Condições Climáticas
Bem que poderia fazer tempo seco em gramado amanhã né :(	Negativa	Condições Climáticas
fiz a compra de quatro ingressos para o show nativiaten que foi cancelado em razão das fortes chuvas. fz posteriormente o pedido de devolução dos valores pagos, porém até o momento não houve manifestação da empresa.	Negativa	Condições Climáticas

Fonte: Elaborado pelo autor.

Na tabela 9, o comentário “...fiz a compra de quatro ingressos para o show nativiaten que foi cancelado em razão de fortes chuvas...” , apesar de poder influenciar na decisão de compra, uma vez em que os turistas que estão programados para viajar para o evento podem ter cancelado suas viagens devido às chuvas, também faz uma crítica ao atendimento por parte da organização do evento, o que justificaria a categorização do comentário na dimensão “(In)satisfação com a prestação dos serviços”.

Todos os casos acima expõe uma necessidade de melhoria do índice relacionado aos Atrativos Turísticos, através da criação de opções alternativas para dias chuvosos, com neblina ou até mesmo o frio, para os turistas que não veem nele um atrativo natural.

A Mineração de Opinião permitiu identificar inúmeros comentários que, apesar de participarem da amostra e estarem de acordo com os critérios considerados válidos pela ferramenta de mineração, apresentam conteúdos irrelevantes para a análise e posterior utilização como informação estratégica pelas organizações. A tabela 10 apresenta exemplos de comentários que poderiam ser identificados e descartados (ou não) através de análises qualitativas realizadas pelo trabalho humano ou através de refinamentos na configuração da ferramenta de mineração.

Tabela 10 - Comentários sem informação relevante

<a href="https://t.co/iASecg2fvO">https://t.co/iASecg2fvO</a> <a href="https://t.co/GIHFz3ZdDg">https://t.co/GIHFz3ZdDg</a>
Foii Gramado??????????
#Riograndedosul
saudade de gramado tbm... qro viajar...
É em Gramado? <a href="https://t.co/fpIhb1nBCC">https://t.co/fpIhb1nBCC</a>
BORA P GRAMADO SOZINHA ENTÃO NÉ

Fonte: Elaborado pelo autor.

Paine (2011) ressalta que do total existente de conteúdo coletado, aproximadamente 70% são irrelevantes, dando ênfase para o tamanho da amostra a ser selecionada.

#### 4.5 DISCUSSÃO

A análise do referencial envolvendo a criação do ICT foi necessária para, inicialmente, validar a conexão com a Mineração de Opinião, conexão esta sugerida e apresentada no framework proposto pelo autor, em acordo com os objetivos secundários estabelecidos neste trabalho. O framework apóia posteriormente na realização dos demais objetivos secundários propostos neste trabalho, uma vez em que as opiniões mineradas precisariam ser categorizadas de acordo com um critério

ou assunto, no caso, as variáveis ou dimensões do Índice de Competitividade do Turismo.

Dada a característica multidisciplinar das informações resultantes da Mineração de Opinião, uma vez categorizadas conforme as dimensões do ICT, podem ser distribuídas para as diversas secretarias de acordo com o assunto em questão ou mesmo compartilhadas com outras organizações da sociedade civil envolvidas no desenvolvimento das atividades do município de Gramado/RS, tais como a Agência de Desenvolvimento da Região das Hortências (VISÃO), a Autarquia de Turismo e Cultura Gramadotur, a Câmara de Dirigentes Logistas, o Sindicatos de Hotéis, Bares e Similares e o Convention & Visitors Bureau da região das Hortências. Segundo Gandara et al. (2003), é fundamental que todos os atores que participam da atividade turística assumam seus papéis, considerando sempre que devem ter benefícios e responsabilidades. Todos os atores devem participar e estar comprometidos com o planejamento, a implementação, o desenvolvimento, a gestão e o controle da sustentabilidade do desenvolvimento da atividade turística de forma estratégica.

A característica multidisciplinar dos diversos assuntos envolvendo as dimensões de análise ficou evidente quanto, com a finalidade da verificação de credibilidade dos resultados encontrados neste estudo, os mesmos foram apresentados aos gestores da SMT através de um relatório, conforme modelo apresentado no Apêndice 2.

Quando questionada a SMT sobre “Qual a sua avaliação sobre o método aplicado para a classificação dos comentários de acordo com as dimensões do índice de Competitividade do Turismo”, a SMT informou que as informações apresentadas “ficaram muito bem colocadas facilitando inclusive a interação entre as diversas secretarias da Administração que têm atuação nos assuntos que tratam as dimensões do Índice”.

Os resultados da análise particular das dimensões (In)segurança, Condições Climáticas e Atrativos Turísticos reforçam a necessidade de análise das informações por diferentes atores. Na dimensão “Atrativos Turísticos”, a amostra analisada indica que os comentários podem ser úteis para endossar a necessidade da geração e promoção de atrativos para dias em condições climáticas ruins. Já o comentário “...os pacotes são caros para o nosso orçamento...”, apesar de estar relacionado a

dimensão em questão, também está relacionado à dimensão “Preços” e que pode ser analisada por atores envolvidos com a prestação de serviços.

A mesma conclusão se aplica ao comentário “...fiz a compra de quatro ingressos para o show...” posteriormente cancelado devido às condições climáticas. O comentário sugere melhorias no sistema de atendimento do prestador de serviços e a análise de menções e retweets relacionados ao comentário podem indicar a necessidade de ações pontuais no atendimento ao cliente.

Além da análise de sentimento de cada comentário, a descoberta de conhecimento em textos permite a identificação de entidades tais como marcas, produtos, atributos sobre o autor do comentário e hashtags, conforme demonstrado anteriormente. As nuvens de palavras apresentadas com base na localização do autor e nas entidades extraídas do comentário fornecem informações que podem ser úteis para direcionar campanhas de marketing ou auxiliar na análise de sentimentos baseada em hashtags, técnica já implementada por Rodrigues Barbosa et al (2012) no trabalho onde foram coletados dados das eleições presidenciais do Brasil em 2010 utilizando hashtags para traçar o sentimento da população a respeito das eleições daquele ano.

Alinhado com os objetivos específicos II e III estabelecidos neste trabalho, a ferramenta escolhida para a Mineração de Opiniões requer a configuração de um dicionário particular de termos relacionados com as dimensões. Este dicionário de termos, documentados na seção 4.3, acelera o processo de implantação de uma ferramenta de mineração e é replicável para quaisquer dos 65 destinos indutores avaliados pela metodologia do ICT.

Num cenário de evolução das técnicas de Mineração de Opinião, com o uso frequente e variado de canais online e observada dependência de Investimentos em tecnologia e recursos humanos, a Mineração de Opinião, como ferramenta para análise preliminar de grandes volumes de conteúdo, fornece visibilidade nas conversações e temas de uma forma compreensível e útil, aumenta a eficácia do trabalho realizado na busca de informações atualizadas sobre a demanda e o mercado.

O uso do conhecimento gerado com o apoio de ferramentas de Mineração de Opinião pode ajudar na criação de um modelo de gestão competitivo e inovador que tome decisões mais alinhadas com as demandas dos turistas, garantindo assim que o que os clientes estão recebendo o que eles realmente desejam receber.

Através da análise sistemática das informações geradas pela Mineração de Opinião é possível identificar as necessidades dos clientes e se antecipar à concorrência, melhorando sua competitividade. Segundo Tomaél et al. (2006), o diferencial da organização competitiva está na sua capacidade de processar dados, informações e conhecimentos de forma rápida, clara, concisa e precisa, atributos fundamentais para a tomada de decisão. Segundo Xu et al. (2011), opiniões de clientes obtidas em redes sociais podem ser consideradas fontes naturais de informação para a Inteligência Competitiva.

Por fim, analisar as mudanças no ambiente externo e adaptar suas capacidades é primordial para que o destino turístico se mantenha competitivo e esta análise pode ser suportada pela Mineração de Opinião como ferramenta para obter o feedback da demanda. Nesse sentido, Paiva e Manfredini (2010) afirmam que a competitividade das destinações turísticas passa, necessariamente pelo processo de melhoria nas políticas de gestão e ferramentas de avaliação para as tomadas de decisão.

## 5 CONSIDERAÇÕES FINAIS

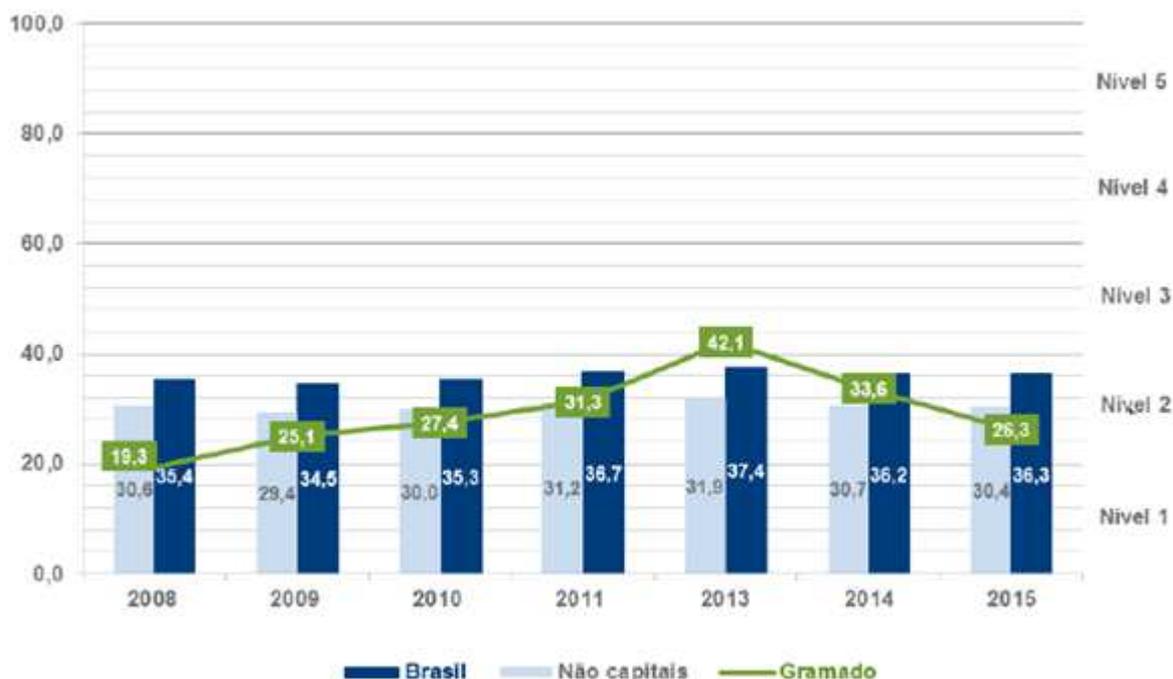
O propósito principal deste estudo foi analisar as oportunidades geradas pela Mineração de Opinião como processo de apoio no planejamento turístico do município de Gramado/RS. Para tanto, o autor deste estudo analisou o ICT e quais as bases teóricas que o sustentam. Em seguida, foram revisados os principais conceitos relacionados à KDT, mais especificamente à Mineração de Opinião. Por fim, foram analisados estudos sobre competitividade e turismo que abordaram a dimensão da demanda. Na sequência, aplicou a Mineração de Opinião em duas redes sociais utilizando como base para a categorização dos comentários um *framework* proposto pelo autor. Através deste estudo foi possível verificar alguns benefícios e oportunidades que as redes sociais podem oferecer às organizações. É possível afirmar que a Mineração de Opinião desempenha um papel importante no planejamento e gestão dos destinos turísticos.

O *framework* apresentado neste estudo destaca a importância do feedback sobre a experiência turística no apoio às ações visando a melhoria destas mesmas experiências e insere o turista como elemento fundamental no sistema turístico. O estudo piloto elaborado demonstra que a Mineração de Opinião pode ser um instrumento fundamental de apoio aos tomadores de decisão fornecendo, de forma ágil e constante, informações consistentes sobre a demanda e sobre a concorrência. As opiniões, polarizadas e categorizadas de acordo com as dimensões avaliadas em termos de competitividade, permitem a identificação rápida das áreas que necessitam maior atenção e que demandam ações de curto prazo, sejam elas pontos fortes ou fracos.

Foram 21.037 comentários agregados em 23 categorias de forma automática, como base em um dicionário de ontologias configurado em uma ferramenta de Mineração de Opiniões. De todas as categorias ou dimensões, uma merece destaque, a categoria “Monitoramento”.

Este estudo por si só colabora com a melhoria do índice de competitividade do turismo na dimensão Monitoramento na variável Pesquisa de Demanda. O baixo desempenho do município de Gramado/RS nesta dimensão observado no gráfico 11 se deve, entre outros fatores, pela ausência de pesquisa de demanda turística periódica, que poderia gerar dados relevantes para a gestão, o planejamento e a divulgação de informações sobre a atividade turística no destino (ICT2015).

Gráfico 11 - Índices Monitoramento – destino x Brasil x não capitais: 2008-2015



Fonte: Elaborado pelo autor.

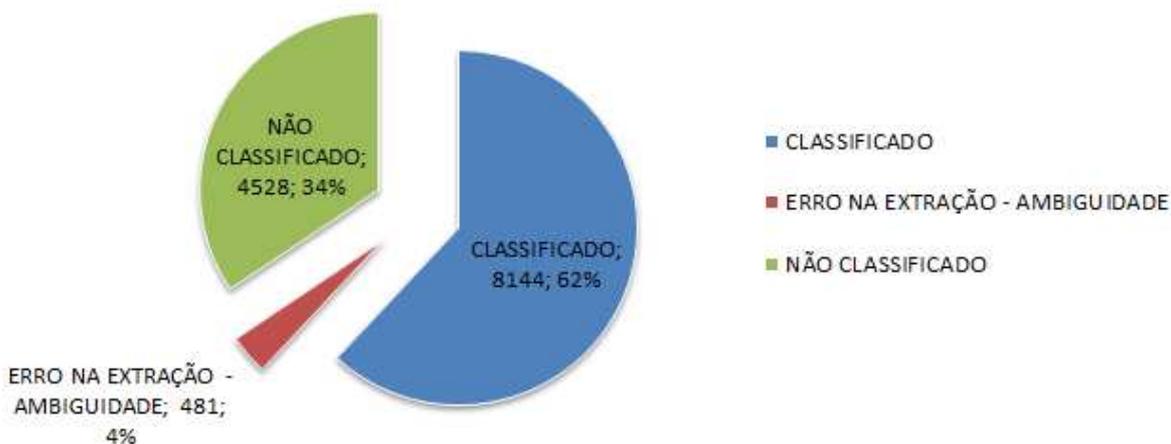
Por fim, cabe ressaltar que os resultados da Mineração de Opinião nas redes sociais não se restringem às opiniões coletadas e categorizadas. A sua utilização na prática pelas organizações responsáveis pelo turístico, sejam atores públicos ou privados, requer investimento em tecnologia e recursos humanos implementar, coletar e analisar os resultados para que, somente depois, sejam transformados em informação útil, dinâmica e em tempo real para a tomada de decisão. Este processo como um todo pode ser acelerado pelo framework criado e testado neste trabalho.

## 5.1 LIMITAÇÕES DA PESQUISA

Como fatores limitantes, ressalta-se que este estudo depende das técnicas de Mineração de Opinião implementadas pelo software utilizado e, portanto, os resultados são dependentes de eventuais problemas técnicos na sua construção ou devido à utilização de técnicas não usuais. Vale salientar que o fabricante do software Semantria não disponibiliza informações detalhadas sobre o método implementado para todas as atividades a que ele se propõe. Decidiu-se por um software amplamente utilizado no meio comercial e acadêmico, mas entre as sugestões para melhoria e continuidade deste estudo, foi citada a aplicação do

mesmo modelo ou framework em outra(s) ferramenta(s), com posterior comparação dos resultados quanto a polaridade e categoria encontradas. O gráfico 12 apresenta a frequência de comentários desconsiderados da pesquisa devido à problemas técnicos da ferramenta utilizada ou devido à falhas no dicionário configurado:

Gráfico 12 - Frequência de comentários não classificados



Fonte: Elaborado pelo autor.

Relativo à amostra extraída na etapa de coleta ou extração, a quantidade de comentários foi limitada a 18.000 extrações e disponíveis a uma semana da data da extração. Esta limitação foi imposta pela API utilizada para o Twitter. Quanto ao Facebook, o versão acadêmica do NodeXL limitava o número de comentários extraídos era limitado à 1000 para cada página ou grupo pesquisado. Já nas etapas de identificação, classificação e sumarização, realizadas pelo Semantria, a versão trial permitia a análise de até 10.000 comentários por registro de usuário.

## 5.2 SUGESTÕES PARA TRABALHOS FUTUROS

Os resultados encontrados neste trabalho não encerram a totalidade de benefícios da Mineração de Opinião no apoio ao processo de planejamento turístico do município de Gramado/RS, ao contrário, fomentam o desenvolvimento de novos e amplos estudos sobre o assunto. Como sugestões para a continuidade deste estudo pode-se citar:

- a. Segregar ou classificar os comentários relativos ao turismo de lazer e turismo de negócios com a finalidade de permitir analisar as

percepções dos dois atores, compará-las e analisar as oportunidades de forma distinta;

- b. Analisar outras informações adjacentes extraídas como hastags, retweets e emoticons e recursos peculiares da rede social Facebook tais como a funcionalidade “Curtir”, com o objetivo de aumentar a confiabilidade dos resultados na análise de sentimento;
- c. Explorar as oportunidades de informações em outras redes sociais tais como Instagram, Flickr, Myspace, Pinterest, Foursquare, Gogobot e Wayn. Isso permitirá melhorar o framework proposto, tornando mais amplo e exigindo um refinamento dos dicionários criados, especialmente em redes onde os comentários são baseados em imagens publicadas;
- d. Aprofundar a análise de cada dimensão ao nível de variáveis, para auxiliar no refinamento dos dicionários criados;
- e. Analisar e incorporar as dimensões de outros autores no framework. O trabalho se limitou a analisar as dimensões do índice juntamente com as novas dimensões propostas por Oliveira (2013). A análise detalhada de outros modelos de planejamento poderá apresentar novas dimensões ou variáveis não abordadas.
- f. Criar uma métrica de classificação das polaridades em cada dimensão analisada a fim de permitir a mineração e comparação com destinos concorrentes. O mesmo se aplica ao mesmo destino. Os resultados podem ser comparados com períodos passados ou com os índices apresentados no relatório do MTur;
- g. Testar os mesmos resultados extraídos com diferentes ferramentas de Mineração de Opiniões, com a finalidade de melhorar a confiabilidade dos resultados;

## REFERÊNCIAS

- ABEYWARDENA, Ishan Sudeera. Public Opinion on OER and MOOC: A Sentiment Analysis of Twitter Data. In: INTERNATIONAL CONFERENCE ON OPEN AND FLEXIBLE EDUCATION, 2014. **Proceedings...** Hong Kong, 2014.
- AGRA, Ronaldo et al. Análise de redes sociais pode aproximar governo e cidadão?. **Revista Brasileira de Administração Científica**, v. 5, n. 2, p. 172-185, 2014.
- ALVES, Simone; NOGUEIRA, Antônio Roberto Ramos. Rumo a um modelo de mensuração da competitividade turística sustentável de municípios: evidências empíricas do Brasil. In: SEMINÁRIO DA ASSOCIAÇÃO NACIONAL PESQUISA E PÓS-GRADUAÇÃO EM TURISMO, 11. **Anais...** 2014.
- ANDRADE, Maria Margarida. Pesquisa científica: noções introdutórias. In: \_\_\_\_\_. **Introdução à metodologia do trabalho científico**: elaboração de trabalhos na graduação. 6. ed. São Paulo: Atlas, 2003. p. 121-127.
- ANGELONI, Maria Terezinha. Elementos intervenientes na tomada de decisão tomada de decisão. **Ciência da Informação**, Brasília, v. 32, n. 1, p. 17-22, 2003.
- ASTON, Nathan; LIDDLE, Jacob; HU, Wei. Twitter Sentiment in Data Streams with Perceptron. **Journal of Computer and Communications**, [s. l.], v. 2, n. 3, p.11-16, 2014.
- BARBOSA, Luiz Gustavo Medeiros. The competitiveness of brazilian tourist destinations. Tese (Doutorado em filosofia) - Nottingham University Business School, Nottingham, 2012.
- BARCELOS, Yussif Tadeu et al. Ferramenta de visualização de dados e text mining: análise de reviews no tripadvisor. **Revista Eletrônica do Alto Vale do Itajaí**, Ibirama, v. 3, n. 4, p. 25-39, 2014.
- BARNEY, Jay B. Looking inside for competitive advantage. *The Academy of Management Executive*, v. 9, n. 4, p. 49-61, 1995.
- BARRETTO, Margarita. **Turismo e legado cultural**: as possibilidades do planejamento. Campinas: Papirus, 2000.
- BECKER, Karin; TUMITAN, Diego. Introdução à Mineração de Opiniões: Conceitos, Aplicações e Desafios. In: FERREIRA, João Eduardo. **Lectures of the 28th Brazilian Symposium on Databases**. Pernambuco, CIN-UFPE, 2013. P. 27-52.
- BENI, Mario Carlos. **Análise estrutural do turismo**. São Paulo: SENAC, 2001.
- BENI, Mário Carlos. Política e estratégia do desenvolvimento regional: planejamento integrado e sustentável do turismo. **Revista Turismo em Análise**, São Paulo, v. 10, n. 1, p. 7-17, 1999.

BENI, Mário Carlos. **Política e planejamento de turismo no Brasil**. São Paulo: Aleph, 2006.

BRAGA, Luis Paulo Vieira. **Introdução à Mineração de Dados**. 2. ed. Rio de Janeiro: E-papers, 2005.

BRASIL. Ministério do Turismo, **Índice de Competitividade do Turismo Nacional: Relatório Brasil 2014**. Brasília: SEBRAE, 2014. Disponível em: <[http://www.turismo.gov.br/sites/default/turismo/o\\_ministerio/publicacoes/downloads\\_publicacoes/indice\\_Brasil\\_2014\\_2.pdf](http://www.turismo.gov.br/sites/default/turismo/o_ministerio/publicacoes/downloads_publicacoes/indice_Brasil_2014_2.pdf)>. Acesso em: 30 abr. 2015.

BRASIL. Ministério do Turismo, **Índice de Competitividade do Turismo Nacional: Relatório Brasil 2015**. Brasília: MT, 2015.

BRASIL. Ministério do Turismo. **Plano Nacional de Turismo: o turismo fazendo muito mais pelo Brasil: 2013-2016**. Brasília: MT, 2013. Disponível em: <[http://www.turismo.gov.br/images/pdf/plano\\_nacional\\_2013.pdf](http://www.turismo.gov.br/images/pdf/plano_nacional_2013.pdf)>. Acesso em: 30 abr. 2015.

BRASIL. Ministério do Turismo. **Turismo já representa 3,7% do PIB**. 2012. Disponível em: <[http://www.turismo.gov.br/turismo/noticias/todas\\_noticias/20121010-2.html](http://www.turismo.gov.br/turismo/noticias/todas_noticias/20121010-2.html)>. Acesso em: 30 abr. 2015.

BRASIL. Ministério do Turismo; SERVIÇO BRASILEIRO DE APOIO ÀS MICRO E PEQUENAS EMPRESAS; FUNDAÇÃO GETULIO VARGAS. **Índice de Competitividade do Turismo Nacional**. Gramado: MTur/SEBRAE/FGV, 2015. Disponível em: <[http://www.dadosefatos.turismo.gov.br/export/sites/default/dadosefatos/espaco\\_academico/downloads\\_espaco/65\\_destinos/Gramado\\_RA\\_2015.pdf](http://www.dadosefatos.turismo.gov.br/export/sites/default/dadosefatos/espaco_academico/downloads_espaco/65_destinos/Gramado_RA_2015.pdf)>. Acesso em: 20 abr. 2015.

BROSS, Jürgen. **Aspect-Oriented Sentiment Analysis of Customer Reviews Using Distant Supervision Techniques**. 2013. Tese (Doutorado em Ciências Naturais) - Freie Universität Berlin. Berlin, 2013.

CAMBRIA, Erik; HOWARD, Newton. Common and Common-Sense Knowledge Integration for Concept-Level Sentiment Analysis. In: INTERNATIONAL FLORIDA ARTIFICIAL INTELLIGENCE RESEARCH SOCIETY CONFERENCE, 27. **Proceedings....** Menlo Park, 2014. P.170-173.

CAVACO, C.; SIMÕES, J. M. Turismos de nicho: uma introdução. **José Manuel Simões & Carlos Cardoso Ferreira Turismos de nicho: motivações, produtos, territórios**. Centro de estudos Geográficos, Lisboa, p. 15-39, 2009.

CHAPMAN, Pete et al. **CRISP-DM 1.0: Step-by-step data mining guide**. [s. l.]: SPSS, 2000.

CHESLEY, Paula et al. Using verbs and adjectives to automatically classify blog sentiment. In: AAAI SYMPOSIUM ON COMPUTATIONAL APPROACHES TO ANALYSING WEBLOGS. **Proceedings...** Menlo Park, 2006. p 27-29.

CIELIEBAK, Mark; DURR, Oliver; UZDILLI, Fatih. Potential and limitations of commercial sentiment detection tools. In: INTERNATIONAL WORKSHOP ON EMOTION AND SENTIMENT IN SOCIAL AND EXPRESSIVE MEDIA, 1. **Proceedings...** Torino, 2013.

CONWAY, Drew. **The Data Science Venn Diagram**. 2010. Disponível em: <<http://drewconway.com/zia/2013/3/26/the-data-science-venn-diagram>>. Acesso em: 30 abr. 2015.

DALONSO, Yoná da Silva et al. A influência dos eventos na construção das políticas públicas das cidades: o caso do Natal Luz em Gramado (Brasil). 2012.

DIAS, Maria Madalena. Parâmetros na escolha de técnicas e ferramentas de mineração de dados. **Acta Scientiarum. Technology**, Maringá, v. 24, p. 1715-1725, 2008.

DWYER, Larry; KIM, Chulwon. Destination competitiveness: determinants and indicators. **Current issues in tourism**, [s. l.], v. 6, n. 5, p. 369-414, 2003.

ESULI, Andrea; SEBASTIANI, Fabrizio. SentiWordNet: A Publicly Available Lexical Resource for Opinion Mining. In: INTERNATIONAL CONFERENCE ON LANGUAGE RESOURCES AND EVALUATION, 5. **Proceedings...** Genoa, 2006. p. 417-422. Disponível em: <[http://www.lrec-conf.org/proceedings/lrec2006/pdf/384\\_pdf.pdf](http://www.lrec-conf.org/proceedings/lrec2006/pdf/384_pdf.pdf)>. Acesso em: 27 jul. 2013.

FAYYAD, Usama M. et al. **Advances in Knowledge, Discovery and Data Mining**. Menlo Park: AAAI, 1996.

FERNANDES, Alisson Silva. A relação clima-turismo: fundamentos teóricos e práticos. 2009. Monografia (Especialização em Formação de Professores e Pesquisadores em Turismo)-Universidade de Brasília, Brasília, 2009.

FRANK, Eibe et al. Data mining in bioinformatics using Weka. **Bioinformatics**, Oxford, v.20, n.15, p. 2479-2481, 2004. Disponível em: <<http://bioinformatics.oxfordjournals.org/content/20/15/2479.short>>. Acesso em: 10 jan. 2015.

FREITAS, Cláudia. Sobre a construção de um léxico da afetividade para o processamento computacional do português. **Revista Brasileira de Linguística Aplicada**, Belo Horizonte, v. 13, n. 4, p. 1013-1059, 2013.

GANDARA, J. M. G.; TORRES, E.; LEFROU, D. A participação de todos os atores no processo turístico. **I Seminário de Pesquisa em Turismo do Mercosul**, 2003.

GAZZANA, Pedro Pires; SILVEIRA, Sidnei Renato. Sistema de recomendação para a área de turismo. In: SEMINÁRIO DE INFORMÁTICA, 8. **Anais...** Torres: ULBRA, 2009.

GENEROSI, Andressa. **Natal Luz em Gramado (RS), tradição e inovação: um olhar sobre os fatores que determinam o retorno dos turistas ao evento**. 2011. Tese de Doutorado. Universidade de Caxias do Sul.

GEORGIU, Despo; MACFARLANE, Andrew; RUSSELL-ROSE, Tony. Extracting sentiment from healthcare survey data: An evaluation of sentiment analysis tools. In: SCIENCE AND INFORMATION CONFERENCE, 2015. **Proceedings...** IEEE, 2015. p. 352-361.

GIL, Antonio Carlos. **Como elaborar projetos de pesquisa**. 4. ed. São Paulo: Atlas, 2002.

GOEBEL, Michael; GRUENWALD, Le. A survey of data mining and knowledge discovery software tools. **ACM SIGKDD Explorations Newsletter**, [s. l.], v. 1, n. 1, p. 20-33, 1999.

GRAMADO. Secretaria de Turismo. **Secretaria de Turismo**. Gramado, [2015]. Disponível em: <<http://www.gramado.rs.gov.br/pt-BR>>. Acesso em: 30 abr. 2015.

HALL, Michael. **Planejamento Turístico**: políticas, processos e relacionamentos. São Paulo: Contexto, 2004.

HANSEN, Derek; SHNEIDERMAN, Ben; SMITH, Marc A. **Analyzing social media networks with NodeXL**: Insights from a connected world. Philadelphia: Elsevier-MK, 2010.

HARRISON, Thomas H. **Intranet data warehouse**: ferramentas e técnicas para a utilização do data warehouse na intranet. Berkeley: ABDR, 1998.

HE, Wu; ZHA, Shenghua; LI, Ling. Social media competitive analysis and text mining: A case study in the pizza industry. **International Journal of Information Management**, Guildford, v. 33, n. 3, p. 464-472, 2013.

KAO, Anne; POTEET, Steve R. (Ed.). **Natural language processing and text mining**. London: Springer Science & Business Media, 2007.

KNUPP, Marcos Eduardo Carvalho Gonçalves et al. Competitividade das nações: uma análise do plano nacional de turismo do Brasil. 2012.

LANGIE, Leonardo Cavalheiro; LIMA, VLS. Classificação hierárquica de documentos textuais digitais usando o algoritmo kNN. In: WORKSHOP EM TECNOLOGIA DA INFORMAÇÃO E DA LINGUAGEM HUMANA.1. **Anais...** 2003.

LAROSE, Daniel T. **Discovering knowledge in data**: an introduction to data mining. New Jersey: John Wiley & Sons, 2014.

LAWRENCE, Linus. **Reliability of sentiment mining tools**: a comparison of semantia and social mention. Enschede, 2014.

LEIPER, N. The framework of tourism: towards a definition of tourism, tourist, and the tourist industry. **Annals of Tourism Research**, [s. l.], v. 6, p. 390-407, 1979.

LEXALYTICS. Categorization. **Technology**. [20--]. Disponível em: <<https://www.lexalytics.com/technology/categorization>>. Acesso em: 20 abr. 2015.

LIU, Bing. Sentiment analysis and opinion mining. **Synthesis lectures on human language technologies**, [s. l.], v. 5, n. 1, p. 1-167, 2012.

LOHMANN, Guilherme; NETTO, Alexandre Panosso. **Teoria do turismo**. São Paulo: Aleph, 2008.

MACEDO, Dayana Carla et al. Contribuições da Mineração de Dados para a Mensuração da Qualidade de Serviços. In: CONGRESSO BRASILEIRO DE ENGENHARIA DE PRODUÇÃO. 1. **Anais...** Ponta Grossa, 2011. Disponível em: <[http://www.cleibsonalmeida.blog.br/website/downloads/qualidade\\_estrategica/data\\_mining\\_servicos.pdf](http://www.cleibsonalmeida.blog.br/website/downloads/qualidade_estrategica/data_mining_servicos.pdf)>. Acesso em: 20 abr. 2015.

MARQUES, Maria Angela; BISSOLI, Ambrizi. **Planejamento turístico municipal com suporte em sistemas de informação**. São Paulo: Futura, 2000.

MATZARAKIS, Andreas. Weather-and climate-related information for tourism. **Tourism and Hospitality Planning & Development**, v. 3, n. 2, p. 99-115, 2006.

NASCIMENTO, Paula et al. Análise de sentimento de tweets com foco em notícias. In: **XXXII Congresso da Sociedade Brasileira de Computação, Curitiba, Brazil**. 2012. p. 16-19.

NETWORK GRAPHS. **Nodexl**. [2015]. Disponível em: <<https://nodexl.codeplex.com/>>. Acesso em: 20 abr. 2015.

NEVES, Augusto José Waszczynskyj Antunes das; MARCHIORI, Patricia Zeni. Qualidade de produtos e serviços turísticos em eventos: uma proposta para análise de conteúdos do twitter. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v. 12, n. 1, p. 61-77, 2014.

OLIVEIRA, Carlyle Tadeu Falcão de. **Competitividade de destinos turísticos: fatores de demanda e desempenho**. 194 f. Tese (Doutorado em Administração) – Escola Brasileira de Administração Pública e de Empresas. Fundação Getúlio Vargas. Rio de Janeiro, 2013.

OLIVEIRA, Carlyle Tadeu Falcão; ZOUAIN, Deborah Moraes; BARBOSA, Luiz Gustavo Medeiros. Management and competitiveness of 65 brazilian tourist destinations: a systemic approach. **Tourism & Management Studies**, Faro, v.4, p. 1085-1098, 2013.

OLIVEIRA, Luan S.; CAMPOS, Guilherme O.; DA SILVA, Rodrigo S. Mineração de dados e análise de opinião em redes sociais—Um estudo de caso sobre as Olimpíadas 2012 utilizando o Twitter.

OLIVER, Richard L. **Satisfaction: A behavioral perspective on the consumer**. Abingdon: Routledge, 2014.

OLSON, David L.; DELEN, Dursun. **Advanced data mining techniques**. Berlin: Springer, 2008.

ORGANIZAÇÃO MUNDIAL DO TURISMO. **Introducción al turismo**. Madrid: OMT, 1998.

OSTROWSKI, David Alfred. Sentiment mining within social media for topic identification. In: International Conference on Semantic Computing, 4. **Proceedings...** Pittsburgh: IEEE, 2010. p. 394-401.

PAIVA, Maria Lucia; MANFREDINI, Cíntia. Os stakeholders na administração pública do turismo. **ENANPAD**, v. 34, p. 1-17, 2010.

PANG, Bo; LEE, Lillian. Opinion mining and sentiment analysis. **Foundations and Trends in Information Retrieval**, [s. l.], v.2, n.1-2, p.1-90, 2008.

PATRUCCO, Luis Gustavo; BARRETTO, Margarita. O terceiro setor no Turismo: o caso da Câmara de Turismo de Rio Grande do Sul (Brasil). **Revista Turismo em Análise**, São Paulo, v. 22, n. 2, p. 294-321, 2011.

PEKAR, Viktor; OU, Shiyan. Discovery of subjective evaluations of product features in hotel reviews. **Journal of Vacation Marketing**, [s.l.], v. 14, n. 2, p. 145-155, 2008.

POIRIER, Damien et al. Automating opinion analysis in film reviews: the case of statistic versus linguistic approach. In: AHMAD, Khurshid. **Affective Computing and Sentiment Analysis**. Amsterdam: Springer, 2011. p. 125-140.

PORTER, Michael. **Estrategia competitiva**. Rio de Janeiro: Elsevier, 2004.

REATEGUI, Eliseo Berni; CAZELLA, Sílvio César. Sistemas de recomendação. In: CONGRESSO DA SOCIEDADE BRASILEIRA DE COMPUTAÇÃO, 25. **Anais...** São Leopoldo, 2005.

RILOFF, Ellen; WIEBE, Janyce. Learning extraction patterns for subjective expressions. In: CONFERENCE ON EMPIRICAL METHODS IN NATURAL LANGUAGE PROCESSING. ASSOCIATION FOR COMPUTATIONAL LINGUISTICS. **Proceedings...** Stroudsburg, 2003. p. 105-112.

RIO GRANDE DO SUL. Secretaria do Turismo, Esporte e Lazer. **Observatório de Turismo do RS: análises setoriais**. Porto Alegre, 2014.

RITCHIE, J. R.; CROUCH, Geoffrey I. A model of destination competitiveness/sustainability: Brazilian perspectives. **Revista de Administracao Publica**, Rio de Janeiro, v. 44, n. 5, p. 1049-1066, 2010.

RODRIGUES, Adriana Silva; MALLOU, Jesús Varela. A influência da motivação na intenção de escolha de um destino turístico em tempo de crise económica. **International Journal of Marketing, Communication and New Media**, v. 2, n. 2, 2014.

ROSE, Alexandre Turatti. **Turismo: planejamento e marketing**. Baurueri: Manole, 2002.

RUSCHMANN, Doris Van de Meene. **Turismo e planejamento sustentável: a proteção do meio ambiente**. Rio de Janeiro: Papirus, 1997.

RUSCHMANN, Doris Van de Meene. **Turismo no Brasil: análise e tendências**. Barueri: Manole, 2002.

SANTOS, Danilo Domingos dos; FERNANDES, Anita Maria da Rocha; KLEIN, Antonio Henrique da Fontoura. Sistema de apoio à decisão para o projeto segurança nas praias baseado em data mining. In: CONGRESO ARGENTINO DE CIENCIAS DE LA COMPUTACIÓN, 6. **Anais...** [s. l.], 2000.

SCHUTT, Rachel; O'NEIL, Cathy. **Doing data science: Straight talk from the frontline**. [s. l.]: O'Reilly, 2013.

THOMAZ, Guilherme Mendes. **Processo de mineração de conteúdos em redes sociais para auxílio na gestão de destinos turísticos**. Dissertação (Mestrado em Turismo) – Universidade Federal do Paraná. Curitiba, 2014.

THOMPSON, Russell D.; PERRY, Allen Howard. **Applied climatology: principles and practice**. Abingdon: Routledge, 1997.

TOMAÉL, M.I. et al. A contribuição do processo de inteligência competitiva para a inovação nas organizações. In: VALENTIM, M.L.P. (Org.). Informação, conhecimento e inteligência organizacional. 2. ed. Marília: FUNDEPE Editora, 2006. Cap. 9, p.155-175.

TRIGO, Luiz Gonzaga Godoi. **Turismo básico**. 6. ed. São Paulo: Senac, 2002.

TRIPADVISOR. Os 10 melhores destinos: Brasil. **TripAdvisor Brasil**. [2015]. Disponível em: <<http://www.tripadvisor.com.br/TravelersChoice-Destinations-cTop-g294280>> Acesso em: 30 abr. 2015.

TSYTSARAU, Mikalai; PALPANAS, Themis. Survey on mining subjective data on the web. **Data Mining and Knowledge Discovery**, [s. l.], v. 24, n. 3, p. 478-514, 2012.

TURDAKOV, D.Yu. et. al. Texterra: a framework for text analysis. **Programming and Computer Software**, [s. l.], v. 40, n.5, p. 288-295, 2014. Disponível em: <<http://link.springer.com/article/10.1134/S0361768814050090>>. Acesso em: 10 jan. 2015.

WANG, Pu; DOMENICONI, Carlotta. Building semantic kernels for text classification using wikipedia. In: ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING, 14. **Proceedings...** New York: ACM, 2008. p. 713-721

WANG, Xiang et al. Improving text categorization with semantic knowledge in Wikipedia. **IEICE TRANSACTIONS on Information and Systems**, [s. l.], v. 96, n. 12, p. 2786-2794, 2013.

WHITELAW, Casey; GARG, Navendu ; ARGAMON, Shlomo. Using appraisal groups for sentiment analysis. In: ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING, 14. **Proceedings...** New York: ACM, 2008. p. 625-631.

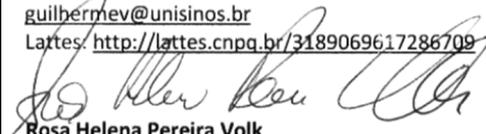
WORLD TOURISM ORGANIZATION. **Glossary of tourism terms**. 2014. Disponível em: < <https://s3-eu-west-1.amazonaws.com/staticunwto/Statistics/Glossary+of+terms.pdf> > . Acesso em: 30 abr. 2015.

XIANG, Zheng; GRETZEL, Ulrike. Role of social media in online travel information search. **Tourism management**, Gulildford, v. 31, n. 2, p. 179-188, 2010.

XU, Kaiquan et al. Mining comparative opinions from customer reviews for Competitive Intelligence. **Decision support systems**, v. 50, n. 4, p. 743-754, 2011.

YOON, Yooshik; UYSAL, Muzaffer. An examination of the effects of motivation and satisfaction on destination loyalty: a structural model. **Tourism management**, Gulildford, v. 26, n. 1, p. 45-56, 2005.

## APÊNDICE A – CARTA CONVITE

	<p>UNIVERSIDADE DO VALE DO RIO DOS SINOS Programa de Pós-Graduação em Gestão e Negócios Mestrado Profissional</p>	
<p>Porto Alegre, Abril de 2015.</p>		
<p>CARTA CONVITE</p>		
<p><b>Ilma. Sra . Rosa Helena Volk</b> <b>SECRETARIA DE TURISMO DE GRAMADO</b></p>		
<p>Vimos respeitosamente convidar a vossa Secretaria a participar do meu projeto de dissertação de mestrado do Programa de Pós Graduação em Gestão e Negócios da UNISINOS, orientado pelo Prof. Dr. Guilherme Luís Roehe Vaccaro.</p>		
<p>O projeto tem como objetivo avaliar as oportunidades de descoberta de informações através da aplicação de técnicas de mineração de dados para apoiar o aprimoramento do turismo da cidade de Gramado, RS. A participação da Secretaria Municipal de Turismo de Gramado é essencial no suporte à pesquisa, para quem o relatório final do projeto é destinado.</p>		
<p>Para que esse objetivo possa ser alcançado, necessitamos:</p>		
<ul style="list-style-type: none"> <li>• Permissão para coletar dados ao longo do desenvolvimento do projeto;</li> <li>• Permissão para que os dados sejam utilizados em publicações acadêmicas e científicas, como dados setoriais, resguardando-se os aspectos éticos e os de confidencialidade.</li> </ul>		
<p>Em termos de retorno à Secretaria de Turismo de Gramado e ao segmento, pela participação no projeto, oferecemos:</p>		
<ul style="list-style-type: none"> <li>• Um relatório das análises e resultados obtidos;</li> <li>• Divulgação da cidade e região em âmbito internacional, uma vez em que a dissertação também será apresentada para uma banca de análise da Universidade de Poitiers na França ( <a href="http://www.univ-poitiers.fr/">http://www.univ-poitiers.fr/</a> ), em 2016;</li> <li>• Menção de agradecimento à Secretaria pela contribuição.</li> </ul>		
<p>Atenciosamente,</p>		
<p><b>Marco Antonio Trois Endres</b> <a href="mailto:Marco.endres@globo.com">Marco.endres@globo.com</a> Lattes: <a href="http://lattes.cnpq.br/8596774459068062">http://lattes.cnpq.br/8596774459068062</a></p>		
<p><b>Guilherme Luís Roehe Vaccaro, Dr.</b> <a href="mailto:guilhermev@unisinovs.br">guilhermev@unisinovs.br</a> Lattes: <a href="http://lattes.cnpq.br/3189069617286709">http://lattes.cnpq.br/3189069617286709</a></p>  <p><b>Rosa Helena Pereira Volk</b> Secretária de Turismo do Município de Gramado/RS</p>		

## APÊNDICE B – RELATÓRIO PRELIMINAR PARA A SMT

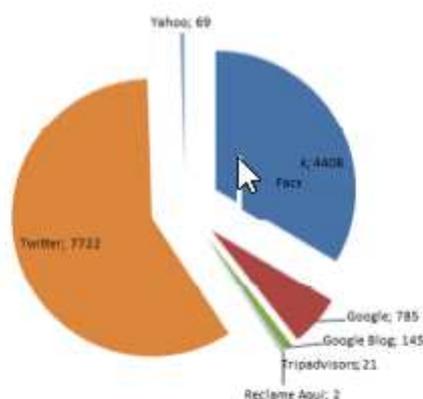
Ilma. Sra . Rosa Helena Volk  
SECRETARIA DE TURISMO DE GRAMADO

De acordo com o termo de parceria celebrado entre a Universidade do Vale do Rio dos Sinos e a Prefeitura Municipal de Gramado/RS, viemos através deste relatório apresentar os resultados obtidos pela pesquisa até o momento e solicitar a vossa avaliação.

A pesquisa, cujo objetivo principal é de avaliar as oportunidades da Mineração de Textos em mídias sociais para o planejamento turístico do município, utilizou uma amostra com as seguintes características:

1. Foram coletados em torno de 18 mil comentários sobre Gramado em períodos variados durante o ano de 2015 e eliminados os comentários que não se aplicavam ao município (gramado em esportes, futebol, política, entre outros).

### Redes sociais analisadas e respectiva amostra

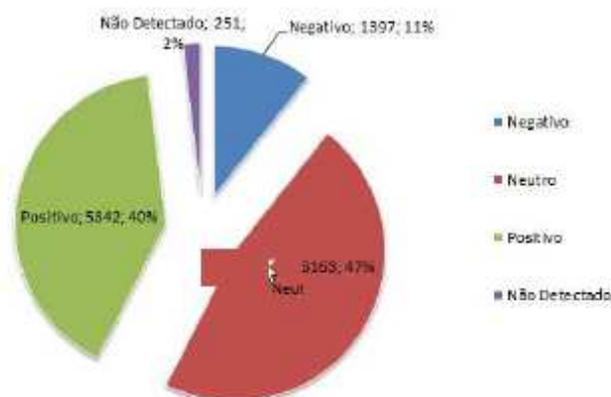


2. Os comentários ou opiniões sobre Gramado foram submetidos a uma técnica específica para avaliação do conteúdo (Análise de Sentimento) para classificação do comentário em "positivo", "negativo" ou "neutro".
3. Após a classificação quanto ao sentimento ou opinião, os comentários foram submetidos à uma técnica de análise de conteúdo para que cada comentário fosse classificado de acordo com 1 ou mais dimensões do Índice de Competitividade do Turismo, resultando em 33 mil comentários classificados de

acordo com a opinião do usuário e a que dimensão ou dimensões do ICT ele se refere.

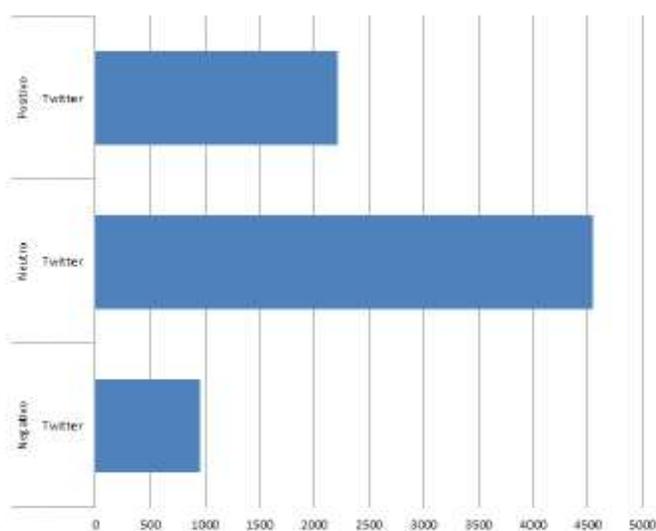
Após a aplicação de uma técnica para análise de sentimento ou opinião, os resultados encontrados foram os seguintes:

1. A opinião ou sentimento detectado nos comentários sobre Gramado, a seguinte distribuição foi encontrada, classificada em opiniões positivas, negativas, neutras ou não detectadas:

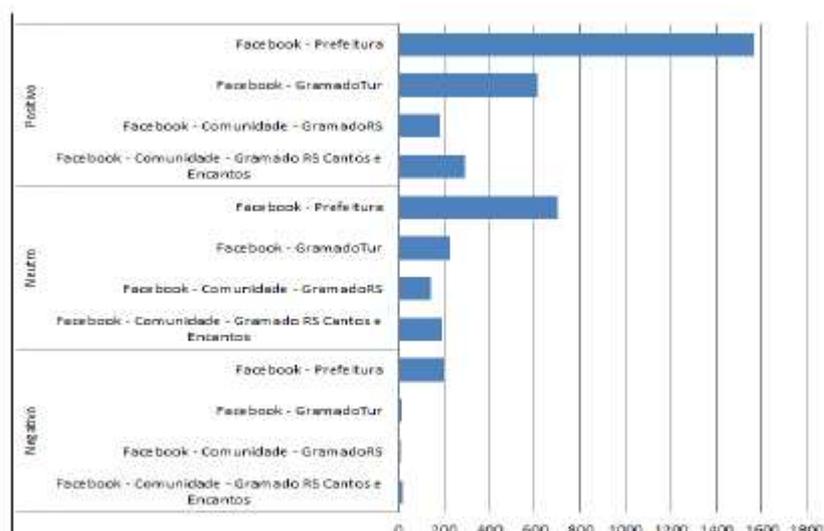


2. No Twitter, a distribuição das opiniões se mantém proporcional. O mesmo não ocorre no Facebook, onde o número de opiniões positivas é maior.

**Distribuição das opiniões no Twitter**

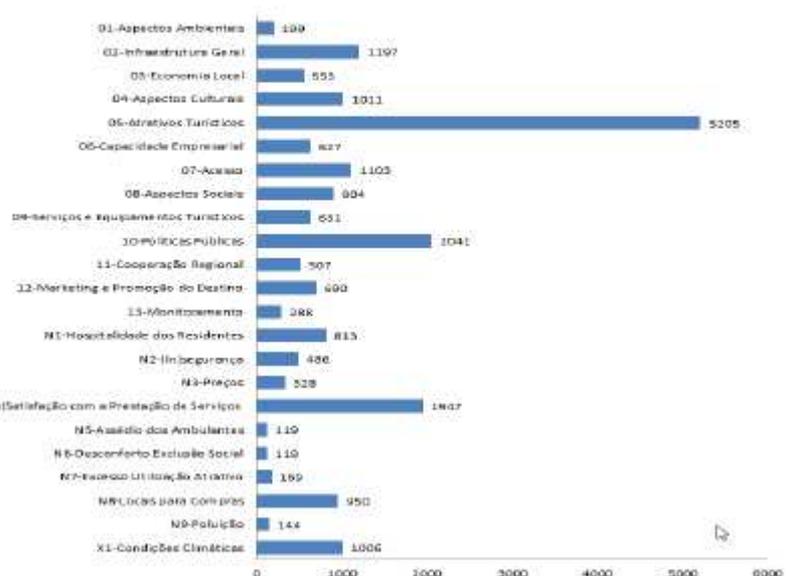


### Distribuição das opiniões no Facebook



3. Após a aplicação de uma técnica para classificação dos comentários de acordo com as dimensões do Índice de Competitividade do Turismo, a seguinte distribuição foi encontrada:

### Número de comentários por dimensão do ICT

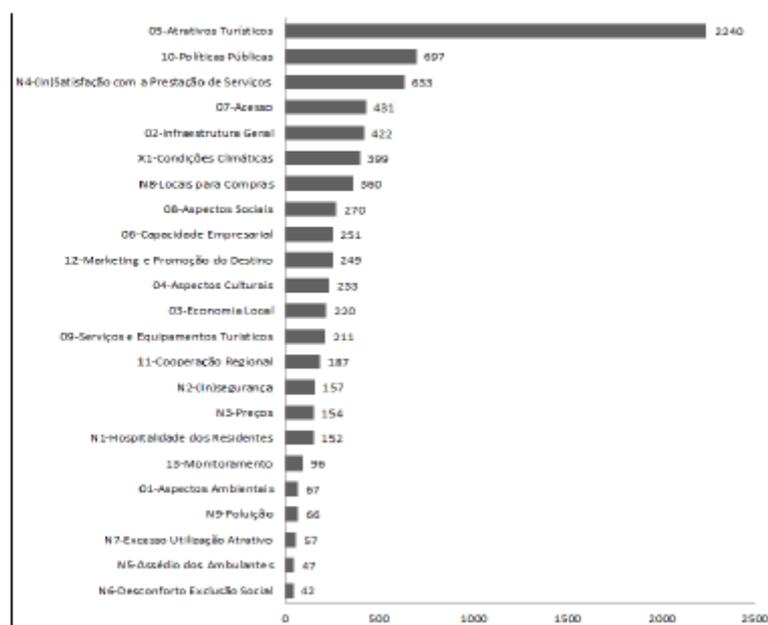


## 4. Distribuição do sentimento ou opiniões sobre Gramado de acordo com cada dimensão:

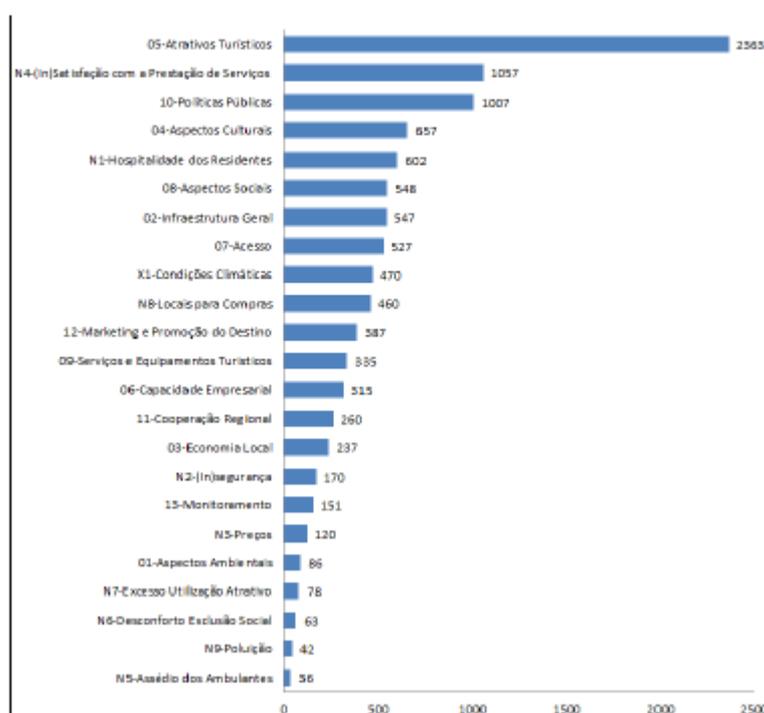
## Opiniões Negativas



## Opiniões Neutras

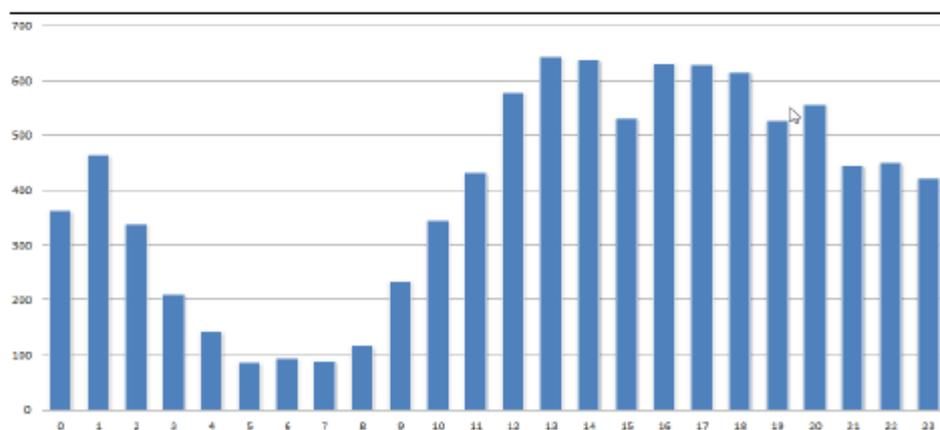


## Opiniões Positivas

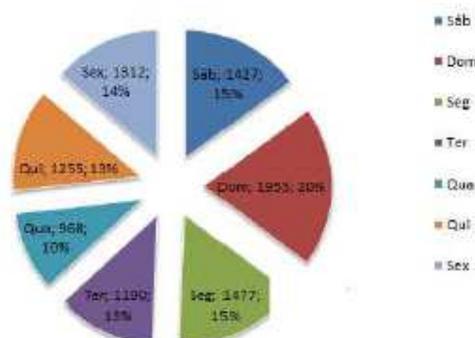


## 5. Outras constatações sobre a amostra analisada:

## Número de comentários x faixas de horários



Número de comentários x dias da semana



## 6. Avaliação manual de comentários classificados

Fonte	Likes	Comentário	Sentimento Detectado	Avaliação ON/NOK	Dimensão do ICT	Avaliação ON/NOK
Twitter		Não tem uma vez que eu venha nesse @McDonalds de Gramado RS que eu não me direito! Que merda gente, não abre 1 franquias pra fazer merda!!	Negativo		N2-(In)segurança	
Twitter		Gramado vota projeto que cria 49 CCs e FGs. Mais R\$ 1,7 milhão/ano dos cofres públicos. Ano que vem tem eleição. #BandRS	Negativo		10-Políticas Públicas	
Twitter		Dirigir em Gramado é a pior coisa, tenho dito.	Negativo		02-Infraestrutura Geral	
Twitter		raissa falou de gramado, bem que eu queria mas os pacotes são caríssimos para nosso orçamento de estudante pobre	Negativo		10-Políticas Públicas	
Twitter		TEMPORAL   Temporal em Igrejinha por Carine Godinho. Gramado com falta de luz por vento forte. <a href="https://t.co/WPhLVjgwtN">https://t.co/WPhLVjgwtN</a>	Negativo		X1-Condições Climáticas	
Twitter		eu não vi nenhum papai noel, EM GRAMADO, NO NATAL, to muito indignada ??	Negativo		05-Atrativos Turísticos	
Twitter		Eu achava caro os rodízios na grande Porto Alegre, aqui em gramado sai aquela lágrima involuntária no olho esquerdo na hora de pagar	Negativo		05-Atrativos Turísticos	
Twitter		Projeto de lei deve gerar aumento nos impostos de imóveis em Gramado <a href="http://t.co/v2lpeF2aNS">http://t.co/v2lpeF2aNS</a>	Negativo		04-Aspectos Culturais	
Twitter		SICC afasta sentimento de crise e promove negócios em Gramado <a href="http://t.co/DG93BZgL0C">http://t.co/DG93BZgL0C</a> (via @Jornaldegramado)	Negativo		03-Economia Local	
Twitter		Chegou a hora de dar tchau aos personagens do Mini Mundo! Última intervenção na Rua Coberta! A gente adorou ??... <a href="http://t.co/Wdh6rS3VIH">http://t.co/Wdh6rS3VIH</a>	Negativo		N6-Desconforto Exclusão Social	
Twitter		#ChuvasnoRS   RS-115 permanece bloqueada e prejudica o turismo <a href="http://t.co/81slacmLe5">http://t.co/81slacmLe5</a> <a href="http://t.co/nR9H8mrTey">http://t.co/nR9H8mrTey</a>	Negativo		10-Políticas Públicas	
Twitter		#ChuvasnoRS   Em Nova Petrópolis: queda de barranco interrompe parcialmente a BR-116 <a href="http://t.co/BmMqgxCxxR">http://t.co/BmMqgxCxxR</a> <a href="http://t.co/1teZMwn9mN">http://t.co/1teZMwn9mN</a>	Negativo		07-Acesso	

8. De forma geral, as informações apresentadas já eram conhecidas pela Secretaria? (S/N)

9. Se sim, através de fonte?

---

---

---

10. Qual a sua avaliação sobre o método aplicado para analisar o sentimento de cada comentário?

---

---

---

11. Qual a sua avaliação sobre o método aplicado para a classificação dos comentários de acordo com as dimensões do índice de Competitividade do Turismo?

---

---

---

12. A Secretaria possui interesse em investir em sistemas informatizados que apliquem os métodos apresentados?

---

---

---