



Universidade Federal de Uberlândia
Faculdade de Matemática
Bacharelado em Estatística

Reclamações de Consumidores: uma
abordagem multivariada

RAFAEL AMARAL DOS SANTOS

Uberlândia-MG
2020

RAFAEL AMARAL DOS SANTOS

Reclamações de Consumidores: uma abordagem multivariada

Trabalho de conclusão de curso apresentado
à Coordenação do Curso de Bacharelado em
Estatística como requisito parcial para
obtenção do grau de Bacharel em Estatística.

Orientador: Prof. Dr. Lúcio Borges de Araújo

Uberlândia-MG
2020



Universidade Federal de Uberlândia
Faculdade de Matemática

Coordenação do Curso de Bacharelado em Estatística

A banca examinadora, conforme abaixo assinado, certifica a adequação deste trabalho de conclusão de curso para obtenção do grau de Bacharel em Estatística.

Uberlândia, 14 de dezembro 2020

BANCA EXAMINADORA

Prof. Dr. Lucio Borges de Araújo

Prof.^a Dra. Mirian Fernandes Carvalho Araújo

Prof. Dr. Janser Moura Pereira

Uberlândia-MG
2020

AGRADECIMENTOS

A Deus que me iluminou e me deu fé para acreditar que era possível.

Á minha esposa Márcia que foi o pilar de sustentação de nossa família durante minha dedicação para conseguir esta conquista.

A minha filha Raquel que é o grande motivo da minha luta.

Aos meus pais, José Adalberto e Ubaldilei que me deram ensinamentos que possibilitaram esta caminhada.

Aos meus professores por terem me proporcionado um desenvolvimento e crescimento intelectual e profissional, em especial ao meu orientador Lúcio.

RESUMO

No trabalho utilizou-se a base de dados públicos do site consumidor do governo federal (consumidor.gov.br), que é um serviço público para solução alternativa de conflitos de consumo (conflitos entre consumidores e empresas). Para cada reclamação cadastrada são geradas 18 variáveis e, dessas, 11 foram consideradas nesse estudo. Foi aplicada a técnica de estatística multivariada de análise de correspondência com a finalidade de analisar a associação dessas 11 variáveis. Os resultados mostraram que houve associação entre todas as 11 variáveis e que as áreas de serviços financeiros e telecomunicações estão no topo em número de reclamações. As reclamações do setor de serviços financeiros estão associadas a faixa etária acima de 60 anos e consumidores acima de 60 anos está associado aos tipos de reclamação com forma de contratação (não comprei – não contratei), essa faixa etária também está associada a reclamações provenientes a região SUL do país. Reclamações provenientes na região NE (Nordeste) está associado aos setores de EDUCAÇÃO e SERVIÇOS FINANCEIROS. O setor de telecomunicação são os que mais recebem nota total na tratativa da reclamação por parte dos consumidores. Esse estudo e seus resultados podem contribuir para os órgãos do governo responsáveis pelo site fornecedor dos dados de pesquisa deste trabalho, para as empresas e ao público consumidor.

Palavras-chave: consumidor.gov.br. Variáveis. Análise de Correspondência.

ABSTRACT

The work used the public database of the federal government's consumer website (consumer.gov.br), which is a public service for the alternative solution of consumer conflicts (conflicts between consumers and companies). For each registered complaint, 18 variables are generated and, of these, 11 were considered in this study. The multivariate statistical technique of correspondence analysis was applied in order to analyze the association of these 11 variables. The results showed that there was an association between all 11 variables and that the areas of financial services and telecommunications are at the top in number of complaints. Complaints in the financial services sector are associated with the age group over 60 and consumers over 60 are associated with the types of complaints with the form of contracting (I did not buy - I did not contract), this age group is also associated with complaints from SOUTH region of the country. Complaints from the NE (Northeast) region are associated with the EDUCATION and FINANCIAL SERVICES sectors. The telecommunications sector are those that receive the highest score in the handling of complaints by consumers. This study and its results can contribute to the government agencies responsible for the website providing the research data for this work, for companies and the consumer public.

Keywords: consumidor.gov.br. Variables. Correspondence Analysis.

LISTA DE FIGURAS

Figura 1 - Biplot associado as variáveis X e Y	18
Figura 2 - Associação entre a Área e a Avaliação da reclamação	24
Figura 3 - Associação entre a Área e a Nota do Consumidor.....	26
Figura 4 - Associação entre a Área e a Faixa etária do consumidor	27
Figura 5 - Associação entre Faixa etária do consumidor e como comprou/Contratou.....	27
Figura 6 - Associação entre Faixa etária do consumidor e Grupo de Problema.....	28
Figura 7 - Associação entre Faixa etária do consumidor e Nota do Consumidor	29
Figura 8 - Associação entre Grupo de Problema e Nota do Consumidor.....	30
Figura 9 - Associação entre Região e Área do Consumidor.....	31
Figura 10 - Associação entre Região e Avaliação da Reclamação.....	32
Figura 11 - Associação entre Região e Como Comprou/ Contratou	32
Figura 12 - Associação entre Região e Faixa Etária.....	33
Figura 13 - Associação entre Região e Nota do Consumidor.....	34
Figura 14 - Associação entre Região e Tipo de Problema	34
Figura 15 - Associação entre Região e Grupo de Problema.....	35
Figura 16 - Associação entre Região e Tempo de Resposta.....	36
Figura 17 - Associação entre Grupo de Problema e Tempo de Resposta.....	37
Figura 18 - Associação entre UF e Área	37
Figura 19 - Associação entre UF e Avaliação da Reclamação	38
Figura 20 - Associação entre UF e Como Comprou/ Contratou	39
Figura 21 - Associação entre UF e Faixa Etária	39
Figura 22 - Associação entre UF e Grupo de Problema	40
Figura 23 - Associação entre UF e Nota do Consumidor.....	41
Figura 24 - Associação entre UF e Tempo de Resposta.....	41

LISTA DE TABELAS

Tabela 1– Estrutura geral de uma tabela de contingência utilizada na AC simples.....	15
Tabela 2 - Reclamações por Região	20
Tabela 3 - Reclamações por UF	21
Tabela 4 –Reclamações por Sexo.....	21
Tabela 5 –Reclamações por Faixa Etária	22
Tabela 6 –Reclamações por Tempo de Resposta	22
Tabela 7 –Reclamações por Área.....	23
Tabela 8 –Reclamações por Grupo de Problema	23
Tabela 9 –Reclamações por Como Comprou/Contratou.....	23
Tabela 10 –Reclamações por nota do consumidor.....	24
Tabela 11 - Reclamações por Avaliação da Reclamação.....	24

SUMÁRIO

1. Introdução.....	11
2. Metodologia.....	13
2.1. Descrição dos dados.....	13
2.2. Análise de Correspondência.....	14
2.2.1. Matriz de Correspondência.....	16
2.2.2. Matriz de perfil das linhas.....	16
2.2.3. Matriz de perfil das Colunas.....	16
2.2.4. Decomposição em coordenadas principais.....	16
2.2.5. Análise Gráfica (Biplot).....	18
3. Resultados e Discussão.....	20
4. Conclusões.....	43
5. Referências Bibliográficas.....	45

1. Introdução

Com o crescimento do setor comercial, houve a necessidade de melhoria na qualidade de produtos e serviços oferecidos aos clientes. Devido a isso, os consumidores se tornaram cada vez mais exigentes e cientes de seus direitos e também observa-se que os meios sociais de proteção ao consumidor (Instituto de Defesa do Consumidor – PROCON) e leis específicas (Código de defesa do Consumidor) estão presentes na relação de empresa e consumidor para assegurar a qualidade esperada dos serviços e produtos fornecidos. Além desses canais, quando ocorrem conflito na prestação de serviços, muitos clientes e consumidores tem buscado formas de manifestar sua insatisfação e frustração com a esperança de obter uma resposta e até solução por parte das empresas fornecedoras de forma mais ágil, sendo empresas públicas ou privadas. Cabe destacar que na internet existem vários sites de reclamação com o intuito de obter uma interatividade entre consumidores e fornecedores para redução de conflitos de consumo.

Nesse sentido, foi realizado um estudo estatístico, usando técnicas de estatística multivariada, das reclamações realizadas no site consumidor.gov.br. Esse site é um serviço público provido e mantido pelo governo federal e monitorado pela Secretaria Nacional do Consumidor (Senacon), Ministério da Justiça, PROCONS, Ministérios Públicos e também por toda sociedade, permite a interlocução direta entre consumidores e empresas para a solução de conflitos de consumo pela internet, de forma eficaz.

No site consumidor.gov.br, o consumidor (cliente) deve verificar se a empresa da qual quer reclamar está cadastrada no site e, posteriormente, o consumidor registra sua reclamação e, após isso, é iniciada a contagem do prazo para manifestação da empresa, que tem até 10 dias para analisar e responder a reclamação. Após esse trâmite, o consumidor tem até 20 dias para comentar a resposta recebida e classificá-la como resolvida ou não resolvida. O consumidor ainda registra seu nível de satisfação com o atendimento prestado pela empresa que é uma nota de 1 a 5. Os dados das reclamações vão para uma base de dados que é pública.

Atualmente, existem três estudos publicados voltados a análise de reclamações entre consumidor e fornecedor (prestador de serviços). No primeiro deles [2], busca demonstrar a evolução dos assuntos das reclamações pertinentes às ouvidorias dos bancos, considerando o porte e o tipo de controle no Brasil. Nesse artigo, realizou-se pesquisa exploratória numa amostra não probabilística de 26 bancos, entre públicos, privados nacionais e estrangeiros. Coletaram-se dados de 1.377 reclamações pertinentes às ouvidorias desses bancos, encerradas no 1º semestre de 2008 até o 1º semestre de 2011 e divulgadas pelo Banco Central do Brasil. No

tratamento dos dados empregou-se estatística descritiva, estatística inferencial e Análise de Correspondência Múltipla. O autor, conclui que os resultados podem contribuir para aperfeiçoar as atribuições regulamentares das ouvidorias dos bancos e, também, podem contribuir para entender a implementação das ações socialmente responsáveis de bancos, para atender seus stakeholders (são partes interessadas contemplando clientes, usuários, acionistas, empregados, fornecedores e a comunidade), clientes e usuários no Brasil.

No segundo trabalho [3], ajustou-se modelos de regressão logística para estudar a variável avaliação da reclamação, variável composta pelos níveis "Resolvida" e "Não Resolvida". O autor observou que na prática há muitas reclamações que não são avaliadas pelo consumidor e o governo as classifica como "Resolvidas", dificultando assim a obtenção do real índice de solução da reclamação calculado pelo site. O autor se empenhou em ajustar um modelo Binomial conjugado com critérios para obtenção de pontos de corte para classificação das reclamações não avaliadas como: "Resolvidas" e "Não Resolvidas"; e desse modo calcular o índice de solução. Nesse trabalho o autor utilizou o modelo de Regressão logística para classificação das reclamações não avaliadas em "Resolvida" e "Não Resolvida".

O terceiro trabalho [9], buscou analisar as relações entre reclamação, avaliação do consumidor quanto ao atendimento e a recuperação de serviços feita pela CEB (Companhia Energética de Brasília) no Reclame AQUI. A amostra foi composta por todas as reclamações registradas no Reclame AQUI no período de novembro 2011 a julho 2015. Os dados foram analisados utilizando-se Análise de Regressão e teste ANOVA. Os resultados obtidos indicam que o atendimento prestado pela CEB através do Reclame AQUI é considerado insatisfatório para os consumidores, não transparece preocupação com o consumidor, o que gera um aumento nos registros de insatisfação presentes no Reclame AQUI.

O objetivo deste trabalho é avaliar a associação entre as variáveis categóricas obtidas através das reclamações dos consumidores no site consumidor.gov.br com o status de finalizadas compreendidas entre os meses de janeiro à junho do ano 2020. Por meio da análise de correspondência. Espera-se que os resultados possam apresentar informações importantes para os PROCONs, para os consumidores e para as empresas cadastradas.

2. Metodologia

2.1. Descrição dos dados

Ao efetuar uma reclamação no site consumidor.gov.br, geram-se 18 variáveis associadas a essa reclamação, dessas 11 serão estudadas neste trabalho, que são: a região geográfica do consumidor reclamante (Região), a unidade federativa do consumidor (UF), o sexo do consumidor (Sexo), a faixa etária do consumidor (Faixa Etária), o número de dias para a resposta da reclamação pela empresa (Tempo de Resposta), área à qual pertence o assunto objeto da reclamação (Área), agrupamento do qual faz parte o problema classificado na reclamação (Grupo de Problema), descrição do problema objeto da reclamação (Problema), descrição do meio utilizado para contratação /aquisição do produto ou serviço reclamado (Como Comprou/Contratou), classificação atribuída pelo consumidor sobre o desfecho da reclamação (Avaliação da Reclamação) e nota atribuída pelo consumidor ao atendimento da empresa (Nota do Consumidor).

Para compreensão dos resultados, necessitou-se codificar os níveis das variáveis, especificamente para interpretação gráfica, assim sendo:

Regiões

CO: Centro-Oeste / N: Norte / NE: Nordeste / S: Sul / SE: Sudeste

Faixa Etária

A – Até 20 anos / B – Entre 21 e 30 anos / C – Entre 31 e 40 anos / D – Entre 41 e 50 anos / E – Entre 51 e 60 anos / F – Entre 61 e 70 anos / G – Mais de 70 anos

Em algumas associações que envolvem a variável faixa etária, a codificação foi realizada com outros códigos, conforme abaixo.

Faixa etária

AA - Até 20 anos / BB - Entre 21 e 30 anos / CC - Entre 31 e 40 anos / DD - Entre 41 e 50 anos / EE - Entre 51 e 60 anos / FF - Entre 61 e 70 anos / GG - Mais de 70 anos

Tempo de Resposta

A: 0 dia – B: 1 dia – C: 2 dias – D: 3 dias – E: 4 dias – F: 5 dias – G: 6 dias – H: 7 dias – I: 8 dias – J: 9 dias – K: 10 dias – L: 11 dias – M: 12 dias – N: 13 dias – O: 14 dias – P: 15 dias – Q: 16 dias - Vazias

Área

A – Água, Energia, Gás / B – Alimentos / C – Demais Produtos / D – Demais Serviços / E – Educação / F – Habitação / G – Produtos de Telefonia e Informática / H – Produtos Eletrodomésticos e Eletrônicos / I – Saúde / J – Serviços Financeiros / K – Telecomunicações / L – Transportes / M – Turismo e Viagens.

Em algumas associações que envolvem a variável Área, a codificação foi realizada com outros códigos, conforme abaixo:

Área

AA - Água, Energia, Gás / BB – Alimentos / CC - Demais Produtos / DD - Demais Serviços / EE – Educação / FF – Habitação / GG - Produtos de Telefonia e Informática / HH - Produtos Eletrodomésticos e Eletrônicos / II – Saúde / JJ - Serviços Financeiros / KK – Telecomunicações / LL – Transportes / MM – Turismo e Viagens.

Grupo Problema

A: Atendimento / SAC - B: Cobrança / Contestação - C: Contrato / Oferta – D: Entrega do Produto - E: Informação - F: Saúde e Segurança - G: Vício de Qualidade

Como Comprou / Contratou

A – Catálogo / B – Domicílio / C - Ganhei de presente / D – Internet / E - Loja física / F - (Não comprei_contratei) / G - Mensagem de texto SMS / H - (Stand, feiras e eventos) / I - telefone

Problema

p157 - SAC - Dificuldade em obter protocolo, gravação ou histórico / p92 - Funcionamento inadequado do serviço (má qualidade do sinal, instabilidade, queda) / p71 - Dificuldade para obter boleto de quitação ou informações acerca de cálculos, pagamentos, saldo devedor / p7 - Atraso de voo

Nota

A - Nota 1 / B - Nota 2 / C - Nota 3 / D - Nota 4 / E - Nota 5

2.2. Análise de Correspondência

A análise de correspondência (AC) é uma técnica estatística multivariada que permite realizar a análise simultânea de diferentes variáveis categóricas. Essas variáveis categorizadas podem possuir níveis de mensuração nominal e ordinal, ou ainda podem ser provenientes de variáveis quantitativas discretas ou contínuas. A AC é essencialmente uma análise exploratória e descritiva dos dados que permite uma análise gráfica das inter-relações ou correspondências entre as variáveis.

Segundo [4], denomina-se análise de correspondência simples quando AC é aplicada para duas variáveis. Quando AC é utilizada para três ou mais variáveis, a técnica é uma generalização da AC simples e recebe a denominação de análise de correspondência múltipla (ACM).

Quando se fala em análise de correspondência, estamos falando diretamente de uma tabela de dupla entrada (tabelas de contingência), conforme a Tabela 1.

Tabela 1– Estrutura geral de uma tabela de contingência utilizada na AC simples

		Variável Y						Total
		1	2	...	j	...	J	
Variável X	1	n_{11}	n_{12}	...	n_{1j}	...	n_{1J}	$n_{1.}$
	2	n_{21}	n_{22}	...	n_{2j}	...	n_{2J}	$n_{2.}$

	i	n_{i1}	n_{i2}	...	n_{ij}	...	n_{iJ}	$n_{i.}$

	I	n_{I1}	n_{I2}	...	n_{Ij}	...	n_{IJ}	$n_{I.}$
Total		$n_{.1}$	$n_{.2}$...	$n_{.j}$...	$n_{.J}$	$n_{..}$

As variáveis X e Y na Tabela 1, são variáveis categóricas (qualitativas) e ambas tem determinados níveis, a variável X tem I níveis e a variável Y tem J níveis. Se faz a contagem do número de elementos da combinação de cada nível n_{ij} , onde i se refere a linha da tabela e j se refere a coluna da tabela, ou seja, número de elementos em cada casela da tabela (frequências observadas). Utiliza-se o teste Qui-quadrado χ^2 para verificar se existe associação entre as variáveis X e Y. O teste Qui-quadrado é utilizado para avaliar se a informação contida nas linhas da tabela são independentes ou não da informação contida nas colunas. As hipóteses a serem testadas são:

H_0 : Independência entre as variáveis X e Y

H_1 : As variáveis estão associadas [7]

e a estatística de teste é

$$\chi^2 = \sum_{i=1}^{I*J} \frac{(O_i - E_i)^2}{E_i} \quad (1)$$

em que O_i é a frequência observada e E_i é a frequência esperada. A estatística tem distribuição Qui-quadrado com $k = (I - 1) * (J - 1)$ graus de liberdade. I é o número de linhas da tabela e J é o número de colunas da tabela.

Se a hipótese H_0 for rejeitada, aceita H_1 , conclui-se que as variáveis estão associadas. No entanto, confirmada a associação, o teste χ^2 não fornece informações a mais a respeito da associação entre as variáveis X e Y. Assim, a análise de correspondência vem com o objetivo de trabalhar com esses tipos de tabelas e então extrair conclusões detalhadas dos dados.

Para realizar a análise de correspondência é preciso obter as matrizes de correspondência, perfil de linhas e perfil de colunas. Estas matrizes são definidas abaixo:

2.2.1. Matriz de Correspondência

É a matriz \mathbf{P} de dimensão $I \times J$ constituída das proporções $p_{ij} = \frac{n_{ij}}{n}$, em que n é o total geral da matriz (Tabela 1), ou seja, $n_{..} = \sum_{i=1}^I \sum_{j=1}^J n_{ij}$ e n_{ij} representa o número de elementos que pertencem à categoria i da variável X e a categoria j da variável Y . [6]

2.2.2. Matriz de perfil das linhas

É a matriz diagonal, denotada por \mathbf{D}_r , cujos elementos diagonais são iguais a $\mathbf{r}' = (\frac{n_1}{n} \frac{n_2}{n} \dots \frac{n_p}{n})$. em que n_1 é o somatório dos elementos da primeira linha da matriz, ou seja, $n_1 = \sum_{j=1}^J n_{1j}$, n_2 é o somatório dos elementos da segunda linha, ou seja, $n_2 = \sum_{j=1}^J n_{2j}$ e, assim, prosseguindo até a última linha da matriz; n é o total geral da matriz que está na (Tabela 1) $n_{..} = \sum_{i=1}^I \sum_{j=1}^J n_{ij}$.

2.2.3. Matriz de perfil das Colunas

É a matriz diagonal, denotada por \mathbf{D}_c , cujos elementos diagonais são iguais a $\mathbf{c}' = (\frac{n_1}{n} \frac{n_2}{n} \dots \frac{n_q}{n})$. em que n_1 é o somatório dos elementos da primeira coluna da matriz, ou seja, $n_1 = \sum_{i=1}^I n_{i1}$, n_2 é o somatório dos elementos da segunda coluna, ou seja, $n_2 = \sum_{i=1}^I n_{i2}$ e, assim, prosseguindo até a última coluna da matriz; n é o total geral da matriz que está na (Tabela 1) $n_{..} = \sum_{i=1}^I \sum_{j=1}^J n_{ij}$. [6]

2.2.4. Decomposição em coordenadas principais

Considere a matriz: $\tilde{\mathbf{P}} = \mathbf{P} - \mathbf{r} \mathbf{c}'$. Essa matriz tem entradas do tipo: $(p_{ij} - \frac{n_i}{n} \frac{n_j}{n})$, ou seja, os termos representam uma comparação da proporção observada no interior de cada casela, com aquela esperada sob um modelo no qual as variáveis X e Y são independentes. Se as linhas e as colunas da matriz (Tabela 1) forem independentes, a matriz \mathbf{P} coincidirá com a matriz $\mathbf{r} \mathbf{c}'$. Dessa forma, não faz sentido aplicar a análise de correspondência. O posto (rank) da matriz $\tilde{\mathbf{P}}_{I \times J}$, ou seja, o número de linhas ou colunas linearmente independentes que $\tilde{\mathbf{P}}_{I \times J}$ possui é dado por $k = \min(I-1, J-1)$, em que k também pode ser definido como o número de autovalores não nulos que se irá obter da matriz $\tilde{\mathbf{P}}_{I \times J}$, obtidos pela Decomposição em Valor Singular (DVS).

$$\tilde{\mathbf{P}}_{I \times J} = \mathbf{A} \mathbf{\Lambda} \mathbf{B}' \quad (2)$$

em que $\mathbf{A} = \mathbf{D}_r^{1/2} \mathbf{U}_{I \times k}$ é uma matriz de dimensão $I \times k$, $\mathbf{B} = \mathbf{D}_c^{1/2} \mathbf{V}_{J \times k}$ é uma matriz de dimensão $J \times k$. \mathbf{U} e \mathbf{V} são matrizes ortogonais e $\mathbf{\Lambda}$ é a matriz de dimensão $k \times k$ contendo os autovalores não nulos provenientes da DVS da matriz $\tilde{\mathbf{P}}$ ordenados em ordem decrescente. A matriz \mathbf{U} contém os autovetores associados aos autovalores da matriz $\tilde{\mathbf{P}}\tilde{\mathbf{P}}'$, e a matriz \mathbf{V} contém os autovetores associados aos autovalores da matriz $\tilde{\mathbf{P}}'\tilde{\mathbf{P}}$.

Como consequência da equação (2), tem-se que as linhas da matriz $\tilde{\mathbf{P}}$ podem ser escritas como combinações lineares das linhas da matriz \mathbf{B}' , e as colunas da matriz $\tilde{\mathbf{P}}$ podem ser escritas como combinações lineares das colunas da matriz \mathbf{A} . Assim, são definidas as coordenadas principais das linhas (Variável X) e das colunas (Variável Y) como segue.

As coordenadas principais das linhas da matriz $\tilde{\mathbf{P}}$ são definidas como:

$$\mathbf{Y}_{I \times k} = \mathbf{D}_r^{-1} \mathbf{A}_{I \times k} \mathbf{\Lambda}_{k \times k} \quad (3)$$

em que I é o número de linhas (número de variáveis nas linhas) da tabela de contingência (Tabela 1) e k é posto (rank) da matriz $\tilde{\mathbf{P}}_{I \times J}$

As coordenadas principais das colunas da matriz $\tilde{\mathbf{P}}$ são definidas como:

$$\mathbf{Z}_{J \times k} = \mathbf{D}_c^{-1} \mathbf{B}_{J \times k} \mathbf{\Lambda}_{k \times k} \quad (4)$$

em que J é o número de colunas (número de variáveis nas colunas) da tabela de contingência (Tabela 1) e k é posto (rank) da matriz $\tilde{\mathbf{P}}_{I \times J}$

As coordenadas principais podem ser representadas em um gráfico (Biplot). O gráfico denomina-se como Biplot pelo fato de representar duas variáveis X e Y.

As duas primeiras coordenadas principais das linhas e das colunas (colunas 1 e 2 de \mathbf{Y} e \mathbf{Z}) são as mais representativas em termos da associação total que existe entre as variáveis X e Y, pois estão relacionadas aos maiores autovalores da matriz $\tilde{\mathbf{P}}$.

A variação total (variabilidade total) existente nesse sistema, chamada também de inércia total, é dada por:

$$\sum_{i=1}^k \lambda_i^2 \quad (5)$$

em que λ_i (obtidos na DVS) são os autovalores não-nulos da diagonal da matriz $\mathbf{\Lambda}$, $i = 1, 2, \dots, k$. A proporção de explicação da i -ésima coordenada principal em relação a inércia total é dada por:

$$\frac{\lambda_i^2}{\sum_{i=1}^k \lambda_i^2} \quad (6)$$

A inércia total definida em (6) está relacionada com a estatística Qui-quadrado da seguinte forma:

$$\sum_{i=1}^k \lambda_i^2 = \frac{\chi^2}{n} = \frac{1}{n} \sum_{(i,j)} \frac{(n_{ij} - E_{ij})^2}{E_{ij}} \quad (7)$$

em que n_{ij} representa o número de observações na casela (i,j) da tabela de contingência e $E_{ij} = n P_{ij}$ representa o número esperado de observações na casela (i,j) calculado sob a hipótese de que as duas variáveis X e Y são independentes, isto é, as linhas e colunas da tabela de contingência são independentes. Sob a hipótese de independência, E_{ij} é dado por:

$$E_{ij} = n p_i.p_j = n \frac{n_i}{n} \frac{n_j}{n} = \frac{n_i.n_j}{n} \quad (8)$$

2.2.5. Análise Gráfica (Biplot)

A interpretação gráfica será realizada utilizando o gráfico representado na Figura 1. Na Figura 1, o eixo X está representando a **coordenada principal 1** e o eixo Y está representando a **coordenada principal 2**, tem-se a origem do gráfico (0,0) e os pontos;

- C1, C2, C3 e C4 são as codificações dos níveis de uma variável Y.
- Linha 1, linha 2 e linha 3 são as codificações dos níveis de uma variável X.

A análise é realizada construindo vetores para esses pontos. Por exemplo, olhando para os pontos C1, C2 e linha 2, na Figura 1, constrói-se vetores partindo da origem do gráfico (0,0) até o ponto analisado. [6]

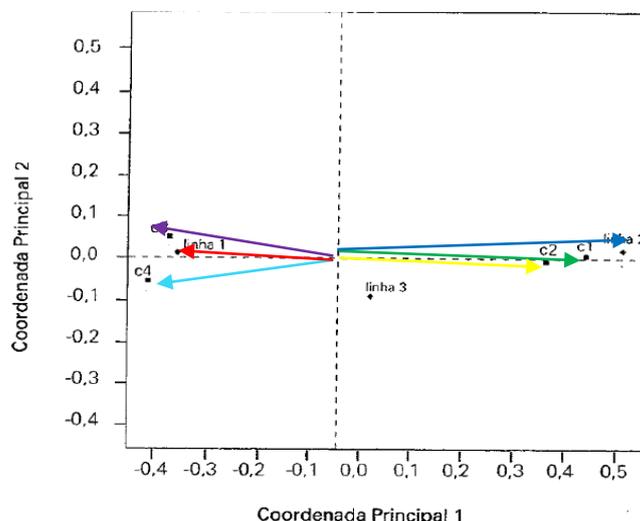


Figura 1 - Biplot associado as variáveis X e Y
Fonte: [7]

Veja que os vetores na cor, verde, azul e amarelo estão todos no mesmo sentido e praticamente na mesma direção (o ângulo entre os vetores é agudo). Isso quer dizer que esses três pontos estão associados entre si, ou seja, a linha 2, C1 e C2 estão associados entre si. Se um desses pontos cresce os demais também irão crescer.

Igualmente, olhando para os vetores nas cores roxo, vermelho e azul turquesa que se direcionam aos pontos C3, linha 1 e C4, respectivamente, estão todos no mesmo sentido e praticamente na mesma direção, indicando que esses pontos também estão associados entre si.

Olhando para o ponto linha 1 em relação aos pontos C1 e C2, o vetor que representa a linha 1 e os vetores que representam os pontos C1 e C2 estão em sentidos opostos (o ângulo entre os vetores é obtuso). Isso quer dizer que a linha 1 e em relação aos pontos C1 e C2 são inversamente associados, se a linha 1 aumenta os pontos C1 e C2 diminuem. Em relação a linha 3, em comparação aos demais pontos, está muito próximo da origem do gráfico, indicando que não tem associação com os demais pontos. A medida que os vetores vão se aproximando do eixo vertical, maior será o indicativo de não associação entre as variáveis, ou seja, independência (o ângulo entre os vetores é próximo de 90°). A linha 3 está muito próxima da origem, assim, ela não tem associação com as demais.

3. Resultados e Discussão

Com a finalidade de obter a maior quantidade possível de informação dos dados, realizou-se uma análise exploratória das variáveis: Região, UF, Sexo, Faixa Etária, Tempo de Resposta, Área, Grupo de Problema, Problema, Como Comprou/Contratou, Avaliação da Reclamação e a Nota do Consumidor.

A variável região possui nível de mensuração nominal, observa-se na Tabela 2 que houve um destaque para as regiões Sudeste, Sul e Nordeste, respectivamente como regiões onde houve mais registro de reclamações. Nota-se que a região Sudeste com 46,86% do total de reclamações, a região Sul com 21,07%, a região Nordeste com 17,26%, a região Norte com 3,54% e a região Centro Oeste com 11,27%. De acordo com as estimativas da população divulgadas no site IBGE com data de referência 1º de julho de 2020, período abordado para o estudo desse trabalho, na região Sudeste está concentrado 42% no número total de residentes no país, a região Sul 14,25%, região Norte 8,8%, região Nordeste 27,1% e a região Centro-oeste 7,8%. Comparando as porcentagens entre o número de pessoas e quantidade de reclamações em cada região, nota-se que para algumas existe uma grande diferença nessas porcentagens, em destaque as regiões Norte e Nordeste. De acordo com o site IBGE¹, PNAD contínua TIC correspondente ao ano de 2018, em todas as regiões aproximadamente 80% dos domicílios na área urbana tem acesso à internet. Em área rural, na região Norte aproximadamente 30% dos domicílios tem acesso à internet, região Nordeste aproximadamente 45%, região Sudeste aproximadamente 59%, região Sul aproximadamente 62% e a região Centro-oeste 56% dos domicílios tem acesso à internet. Essa diferença talvez esteja relacionada a falta de uma melhor divulgação do site consumidor.gov.br em algumas regiões do que em outras, o que sugere um investigação detalhada.

Tabela 2 - Reclamações por Região

Regiões	Frequência Absoluta	Frequência Relativa (%)
CO	57423	11,27%
N	18032	3,54%
NE	87933	17,26%
S	107318	21,07%
SE	238710	46,86%
Total Geral	509416	100,00%

A variável UF possui nível de mensuração nominal, observa-se na Tabela 3 que a unidade federativa SP é responsável por 25,84% do total de reclamações registradas no período, RS com 7,69%, RJ com 8,65%, PR com 9,41%, MG com 9,83% e BA com 5,11%.

¹ <https://educa.ibge.gov.br/jovens/materias-especiais/20787-uso-de-internet-televisao-e-celular-no-brasil.html>

Tabela 3 - Reclamações por UF

UF	Frequência Absoluta	Frequência Relativa (%)
AC	1547	0,30%
AL	3111	0,61%
AM	4357	0,86%
AP	632	0,12%
BA	26045	5,11%
CE	14015	2,75%
DF	22422	4,40%
ES	12940	2,54%
GO	15099	2,96%
MA	9407	1,85%
MG	50063	9,83%
MS	10021	1,97%
MT	9881	1,94%
PA	6339	1,24%
PB	5475	1,07%
PE	12832	2,52%
PI	10876	2,13%
PR	47912	9,41%
RJ	44054	8,65%
RN	3769	0,74%
RO	2177	0,43%
RR	651	0,13%
RS	39192	7,69%
SC	20214	3,97%
SE	2403	0,47%
SP	131653	25,84%
TO	2329	0,46%
Total Geral	509416	100,00%

Analisando a variável Sexo, cujo nível de mensuração é qualitativa nominal, nota-se na Tabela 4 que há mais reclamações de consumidores do Sexo masculino com percentual igual a 57,07%. Segundo os dados do site educa.ibge², a PNAD contínua correspondente ao ano de 2019, o número de mulheres no Brasil é superior ao de homens. A população brasileira é composta por 48,2% de homens e 51,8% de mulheres. A verificação das causas dessa diferença entre o número de reclamações por sexo requer um estudo mais aprofundado.

Tabela 4 –Reclamações por Sexo

Sexo	Frequência Absoluta	Frequência Relativa (%)
F	218671	42,93%
M	290742	57,07%
O	3	0,00%
Total Geral	509416	100,00%

Analisando a variável Faixa Etária, cujo nível de mensuração é qualitativa ordinal, nota-se na Tabela 5 que há mais reclamações na plataforma do site dos consumidores que estão nas faixas etárias entre 31 a 40 anos com percentual igual a 32,97% e entre 21 a 30 anos com percentual igual a 26,48%. Através de pesquisas do autor no site educa.ibge³, os dados da

² <https://educa.ibge.gov.br/jovens/conheca-o-brasil/populacao/18320-quantidade-de-homens-e-mulheres.html#:~:text=Segundo%20dados%20da%20PNAD%20Cont%C3%ADnu,51%2C8%25%20de%20mulheres.>

³ <https://educa.ibge.gov.br/jovens/materias-especiais/20787-uso-de-internet-televisao-e-celular-no-brasil.html#subtitulo-0>

PNAD contínua 2017-2018 informam que no ano de 2018, do total de pessoas compreendidas na faixa etária 20 a 24 anos, 91% utilizaram internet. Na faixa etária entre 25 a 29 anos 90,7% utilizaram a internet, entre 30 a 34 anos 89% e entre 35 e 39 anos 86,8%.

O comportamento da frequência relativa das reclamações por faixa etária, observado na tabela 5, é muito parecido com o comportamento das frequências relativas das faixas etárias de pessoas que utilizaram a internet no ano de 2018, apesar não haver coincidência de ano na comparação.

Tabela 5 –Reclamações por Faixa Etária

Faixa Etária	Frequência Absoluta	Frequência Relativa (%)
Até 20 anos	9511	1,87%
Entre 21 a 30 anos	134887	26,48%
Entre 31 a 40 anos	167969	32,97%
Entre 41 a 50 anos	85207	16,73%
Entre 51 a 60 anos	54128	10,63%
Entre 61 a 70 anos	39384	7,73%
Mais de 70 anos	18330	3,60%
Total Geral	509416	100,00%

A variável Tempo de Resposta é o tempo que empresas gastam para resolverem o conflito com o consumidor. O nível de mensuração dessa variável é quantitativa discreta. Conforme a análise da Tabela 6 de frequências, nota-se que a maior parte das empresas respondem as reclamações entre seis dias (incluso) e dez dias (incluso). 11,66% das empresas respondem em até 10 dias. A variável Tempo de resposta possui média 7,2, desvio padrão 4,21 e mediana 7.

Tabela 6 –Reclamações por Tempo de Resposta

Tempo de Resposta	Frequência Absoluta	Frequência Relativa (%)
0	24255	4,82%
1	38028	7,55%
2	28711	5,70%
3	27570	5,47%
4	26793	5,32%
5	27871	5,53%
6	34589	6,87%
7	45843	9,10%
8	50828	10,09%
9	49665	9,86%
10	58734	11,66%
11	9977	1,98%
12	11108	2,21%
13	15618	3,10%
14	25322	5,03%
15	28752	5,71%
16	5	0,00%
Total Geral	503669	100,00%

A Tabela 7 representa as frequências Absolutas e Relativas da variável área comercial que se refere a reclamação do consumidor. Nota-se que a área comercial que mais recebeu reclamações foi a de Serviços Financeiros com 35,84% do número total de reclamações registradas no período. Em segundo lugar está a área de Telecomunicações com 26,74%. As duas áreas juntas correspondem a 62,58% no número total de reclamações registradas no site.

Tabela 7 –Reclamações por Área

Áreas	Frequência Absoluta	Frequência Relativa (%)
Água, Energia, Gás	14761	2,90%
Alimentos	2657	0,52%
Demais Produtos	33649	6,61%
Demais Serviços	25117	4,93%
Educação	787	0,15%
Habitação	373	0,07%
Produtos de Telefonia e Informática	37140	7,29%
Produtos Eletrodomésticos e Eletrônicos	23068	4,53%
Saúde	4265	0,84%
Serviços Financeiros	182566	35,84%
Telecomunicações	136208	26,74%
Transportes	36431	7,15%
Turismo/Viagens	12394	2,43%
Total Geral	509416	100,00%

A Tabela 8 representa as frequências Absolutas e Relativas da variável grupo de problema, agrupamento do qual faz parte o problema classificado na reclamação. Nota-se que o agrupamento COBRANÇA/ CONTESTAÇÃO é o que mais se destaca com 43,45% do total de reclamações registradas e em segundo lugar o grupo de problema CONTRATO/ OFERTA com 18,06%.

Tabela 8 –Reclamações por Grupo de Problema

Grupo de Problema	Frequência Absoluta	Frequência Relativa (%)
Atendimento / SAC	73086	14,35%
Cobrança / Contestação	221352	43,45%
Contrato / Oferta	92013	18,06%
Entrega do Produto	27873	5,47%
Informação	27457	5,39%
Saúde e Segurança	3120	0,61%
Vício de Qualidade	64515	12,66%
Total Geral	509416	100,00%

A Tabela 9 representa as frequências Absolutas e Relativas da variável Como Comprou/ Contratou, meio utilizado pelo consumidor para contratar ou adquirir o produto ou serviço. Nota-se que os meios de aquisição que mais se destacaram foram através da INTERNET corresponde a 42,04% do total de reclamações, de LOJA FÍSICA com 18,64% e NÃO COMPREI/CONTRATEI com 17,74%. Juntas correspondem a 78,42% do total de reclamações.

Tabela 9 –Reclamações por Como Comprou/Contratou

Como comprou / Contratou	Frequência Absoluta	Frequência Relativa (%)
Catálogo	1833	0,36%
Domicílio	11877	2,33%
Ganhei de presente	1873	0,37%
Internet	214134	42,04%
Loja física	94951	18,64%
Não comprei / contratei	90387	17,74%
SMS / Mensagem de texto	2470	0,48%
Stand, feiras e eventos	1837	0,36%
Telefone	90054	17,68%
Total Geral	509416	100,00%

A Tabela 10 representa as frequências Absolutas e Relativas da variável Nota do Consumidor, nota atribuída pelo consumidor ao atendimento da empresa. As notas variam de 1 a 5, sendo 1 a menor e 5 a maior. As notas que se destacaram foram a nota 1 com 35,22% de todas as atribuições e a nota 5 com 38,38%.

Tabela 10 – Reclamações por nota do consumidor

Nota	Frequência Absoluta	Frequência Relativa (%)
1	100636	35,22%
2	14429	5,05%
3	28908	10,12%
4	32129	11,24%
5	109669	38,38%
Total Geral	285771	100,00%

A Tabela 11 representa as frequências Absolutas e relativas da variável avaliação da reclamação. Classificação atribuída pelo consumidor ao atendimento da empresa. 43,9% do total de reclamações registradas no site são não avaliadas pelo consumidor.

Tabela 11 - Reclamações por Avaliação da Reclamação

Avaliação da Reclamação	Frequência Absoluta	Frequência Relativa (%)
Não Avaliada	223645	43,90%
Não Resolvida	110123	21,62%
Resolvida	175648	34,48%
Total Geral	509416	100,00%

Todas as associações foram avaliadas pelo teste Qui-quadrado e em todas, tais associações foram significativas ao nível de 5%, p valor < 0,0001, devido a isso utilizou-se a técnica estatística multivariada que se chama análise de correspondência. Todas as análises foram realizadas utilizando o software SAS University Edition.[8]

Para a associação entre Área e Avaliação da Reclamação, a primeira coordenada principal explica 86,35% dos dados e a segunda coordenada principal explica 13,65% dos dados. O Biplot dessa associação é apresentado na Figura 2.

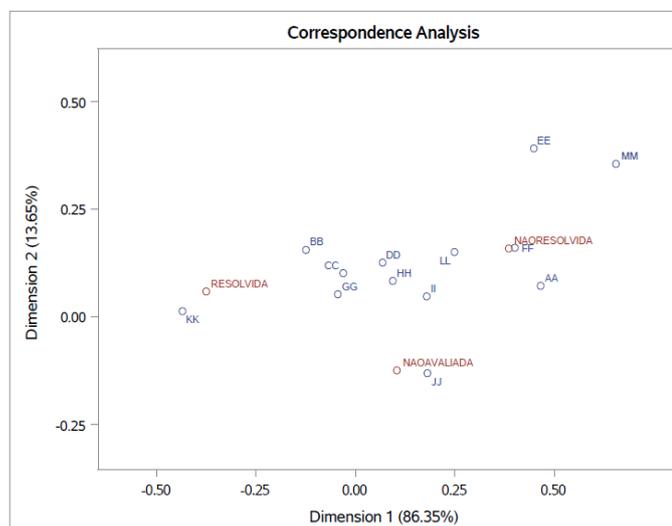


Figura 2 - Associação entre a Área e a Avaliação da reclamação

Conforme a análise da Figura 2, que se refere a associação entre a Área e a Avaliação da reclamação (Avaliação da reclamação é a classificação atribuída pelo consumidor sobre o desfecho da reclamação), nota-se que o nível NÃORESOLVIDO da variável avaliação da reclamação está mais associado aos níveis AA e FF da variável Área. AA se refere ao nível (ÁREA, ENERGIA, GÁS) e FF se refere ao nível (HABITAÇÃO). Isso significa que aumentando o número de reclamações nas áreas AA e FF a tendência é que ocorra um aumento na categoria de NÃO RESOLVIDA da variável avaliação da reclamação. O nível KK que se refere a área de TELECOMUNICAÇÕES, está associado ao tipo de avaliação RESOLVIDO, o que significa que aumentando a quantidade de reclamações a tendência por parte das empresas desse setor é RESOLVER o conflito.

Note que o nível JJ da variável Área (serviços financeiros) está moderadamente associado com a avaliação da reclamação NÃOAVALIADA. [3] em seu trabalho, menciona que as reclamações avaliadas pelo consumidor como NÃOAVALIADA são aquelas a qual o indivíduo não avaliou a tratativa da empresa na solução da reclamação.

Nesses dados há uma variável chamada de Avaliação Reclamação, que é a situação final da reclamação; e tem as seguintes classificações: "Resolvida", "Não Resolvida" e "Não Avaliada". As reclamações classificadas como não avaliadas, são aquelas a qual o indivíduo não avaliou; isso dificulta a obtenção do real valor do índice de solução calculado pelo site, que é a proporção de reclamações resolvidas em relação ao total de reclamações por empresa.[3]

Analisando a Tabela 7 de frequências correspondente a variável Área, nota-se que os níveis Telecomunicações e Serviços Financeiros possuem 26,74% e 35,84%, respectivamente, do número total de reclamações no período avaliado. Isso significa que essas duas áreas estão em primeiro e segundo colocados em quantidade de reclamações, coincidindo com os resultados de [3].

Para a associação entre Área e Nota do consumidor, a primeira coordenada principal explica 98,81% dos dados e a segunda coordenada principal explica 0,83% dos dados. O Biplot dessa associação é apresentado na Figura 3.

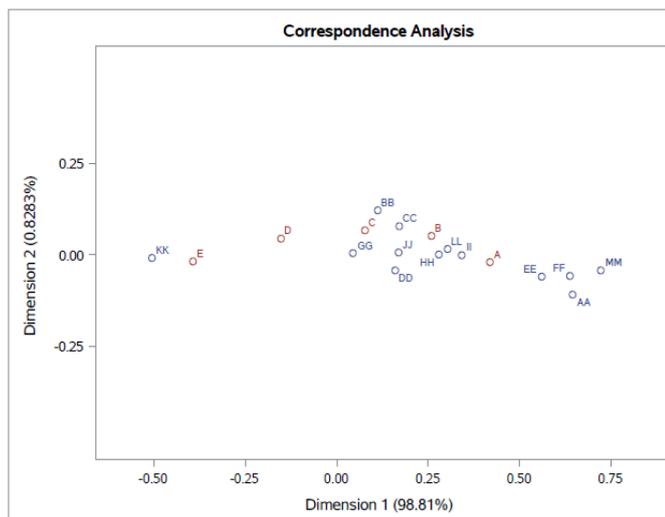


Figura 3 - Associação entre a Área e a Nota do Consumidor

Conforme a análise da Figura 3, que se refere a associação entre a Área e a Nota do Consumidor, nota-se que o nível E (nota 5) da variável Nota do Consumidor está associado ao nível KK variável Área. O código KK se refere ao setor de Telecomunicações. Interpretando o gráfico Biplot, nota-se que aumentando a quantidade de reclamações na área de Telecomunicações, a tendência é que a tratativa por parte das empresas desse setor em resolver o conflito receba nota 5 da parte dos consumidores.

O nível A (nota 1) da variável Nota está associados aos níveis EE, FF, MM e AA (educação, habitação, turismo/viagens e água/energia/gás). Aumentando a quantidade de reclamações nas áreas EE, FF, MM e AA a tendência é que a tratativa por parte das empresas desse setor em resolver o conflito receba nota 1 da parte dos consumidores. Conforme a análise da Tabela 10, a nota 5 representa 38,38% do número total de reclamações registradas no site no período considerado. Já a nota 1 representa 35,22%.

Para a associação entre Área e Faixa Etária do Consumidor, a primeira coordenada principal explica 83,30% dos dados e a segunda coordenada principal explica 13,36% dos dados. O Biplot dessa associação é apresentado na Figura 4.

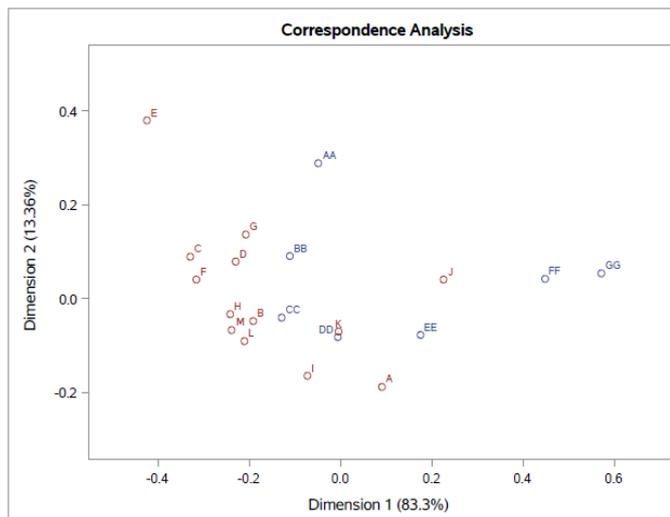


Figura 4 - Associação entre a Área e a Faixa etária do consumidor

Conforme a análise da Figura 4, que se refere a associação entre a Área e a Faixa etária do consumidor, nota-se que o nível J da variável Área está associado aos níveis FF e GG da variável Faixa etária. J se refere ao nível (Serviços Financeiros) e FF se refere a faixa etária entre 61 e 70 anos e GG se refere ao nível (mais de 70 anos).

Aumentando a quantidade de reclamações na área de Serviços Financeiros a tendência é que também aumente o número de reclamações envolvendo consumidores nas faixas etárias acima de 61 anos. Conforme a análise da Tabela 5, as faixas etárias entre 61 e 70 anos e mais de 70 anos correspondem a 7,73% e 3,6%, respectivamente, do número total de reclamações registradas no período considerado para análise desse trabalho.

Para a associação entre Faixa Etária do Consumidor e Como Comprou/Contratou, a primeira coordenada principal explica 98,06% dos dados e a segunda coordenada principal explica 1,25% dos dados. O Biplot dessa associação é apresentado na Figura 5.

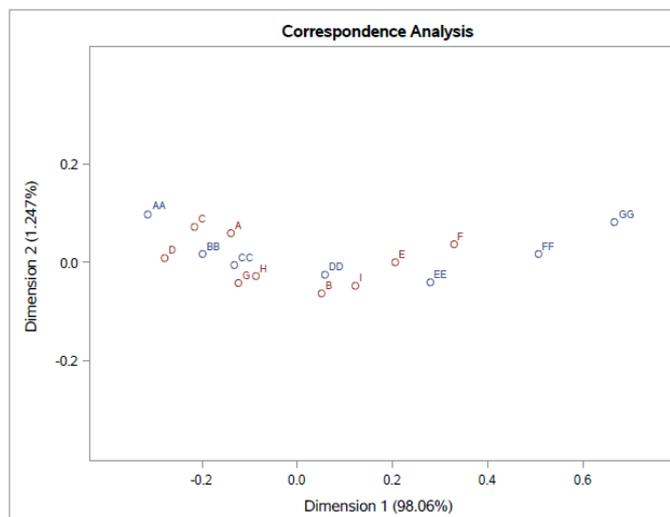


Figura 5 - Associação entre Faixa Etária do consumidor e como Comprou/Contratou

Conforme a análise da Figura 5, que se refere a associação entre a variável Faixa etária do consumidor e a variável (Como Comprou/Contratou), nota-se que o nível F da variável como Como Comprou/Contratou está associado aos níveis FF e GG da variável Faixa etária. F se refere ao nível (Não comprei_contratei) da variável (Como Comprou/Contratou) e FF se refere ao nível da faixa etária (entre 61 e 70 anos) e GG se refere ao nível (mais de 70 anos).

Aumentando a quantidade de reclamações com categorias Não comprei_contratei, a tendência é que ocorra um aumento de reclamações nas faixa etárias (entre 61 e 70 anos) e (mais de 70 anos) de consumidores. Conforme os dados da Tabela 9, a categoria Não comprei_contratei representa 17,74% do número total de reclamações registradas no período considerado. Note que essa faixa etária também está associada com a área de Serviços Financeiros.

O nível C da variável (Como Comprou/Contratou) está associado ao nível AA da variável (Faixa etária). O nível C é ganhei de presente e o nível AA se refere a faixa etária (até 20 anos). Aumentando a quantidade de reclamações com categorias Ganhei de presente, a tendência é que ocorra um aumento de reclamações nas faixa etárias (até 20 anos) de consumidores.

Para a associação entre Faixa Etária do Consumidor e Grupo de Problema, a primeira coordenada principal explica 84,02% dos dados e a segunda coordenada principal explica 9,29% dos dados. O Biplot dessa associação é apresentado na Figura 6.

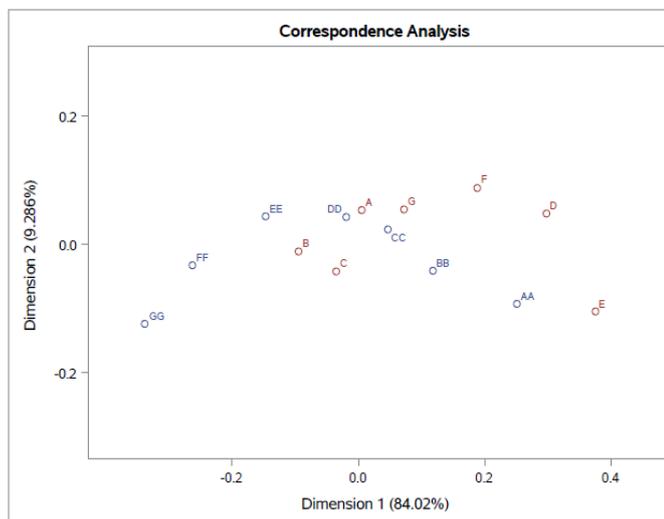


Figura 6 - Associação entre Faixa Etária do consumidor e Grupo de Problema

Conforme a análise da Figura 6, que se refere a associação entre a variável Faixa Etária do consumidor e a variável (Grupo de Problema), nota-se que o nível E da variável Grupo de Problema está associado aos níveis AA e BB da variável Faixa etária. E se refere ao nível

(Informação) da variável (Grupo de Problema) e AA se refere ao nível da faixa etária (até 20 anos) e BB se refere ao nível faixa etária (entre 21 e 30 anos).

Aumentando a quantidade de reclamações com categorias Informação da variável Grupo de Problema, a tendência é que ocorra um aumento de reclamações nas faixas etárias (até 20 anos) e (entre 21 e 30 anos) de consumidores. Conforme os dados da Tabela 8, a categoria Informação representa 5,39% dos dados do número total de reclamações registradas no período considerado.

Para a associação entre Faixa Etária do Consumidor e Nota do Consumidor, a primeira coordenada principal explica 89,13% dos dados e a segunda coordenada principal explica 7,51% dos dados. O Biplot dessa associação é apresentado na Figura 7.

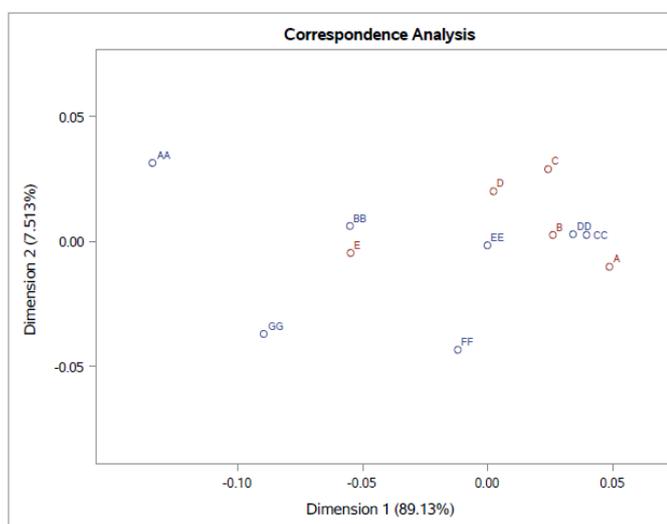


Figura 7 - Associação entre Faixa Etária do Consumidor e Nota do Consumidor

Conforme a análise da Figura 7, que se refere a associação entre a variável Faixa Etária do consumidor e a variável (Nota do consumidor), nota-se que o nível E da variável Nota do Consumidor está associado ao nível BB da variável faixa etária do consumidor. E se refere ao nível (nota 5) da variável (Nota do consumidor) e BB se refere ao nível da faixa etária (entre 21 e 30 anos).

Para a associação entre Grupo de Problema e Nota do Consumidor, a primeira coordenada principal explica 85,28% dos dados e a segunda coordenada principal explica 13,12% dos dados. O Biplot dessa associação é apresentado na Figura 8.

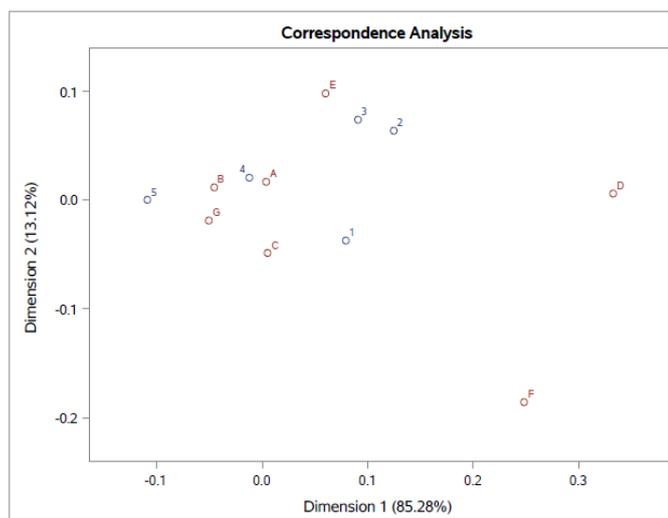


Figura 8 - Associação entre Grupo de Problema e Nota do Consumidor

Conforme a análise da Figura 8, que se refere a associação entre a variável Grupo de Problema e a variável (Nota do consumidor), nota-se que o nível E da variável grupo de problema (Informação) está associado as notas 2 e 3 da variável Nota do Consumidor. Aumentando o número de reclamações no Grupo de Problema (Informação), a tendência é ocorrer um aumento nas notas 2 e 3.

O nível F da variável Grupo de Problema (Saúde e Segurança) está associado a nota 1 do consumidor. Aumentando o número de reclamações no Grupo de Problema (Saúde e Segurança), tende-se a ocorrer um aumento na nota 1.

A Nota do Consumidor segue uma escala de 1 a 5, onde 1 é a pior nota e 5 é a melhor nota. É a nota de satisfação do consumidor dada pelo tratamento da reclamação por parte da empresa.

Para a associação entre a Região e Área do Consumidor, a primeira coordenada principal explica 61,07% dos dados e a segunda coordenada principal explica 30,48% dos dados. O Biplot dessa associação é apresentado na Figura 9.

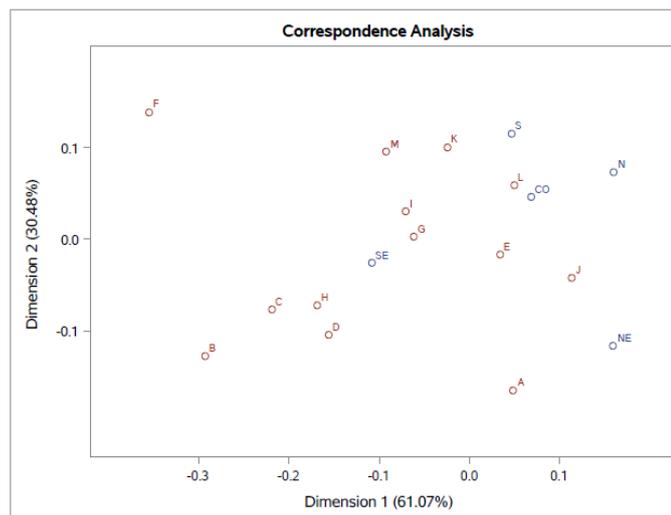


Figura 9 - Associação entre Região e Área do Consumidor

Conforme a análise da Figura 9, que se refere a associação entre a Área e a Região do Consumidor, nota-se que as reclamações provenientes da região NE (Nordeste) está associado aos níveis E e J da variável Área. O código E se refere ao nível (Educação) e J se refere ao nível (Serviços Financeiros). Se aumentar a quantidade de reclamações na região Nordeste, a tendência é que também ocorra aumento de reclamações nas áreas da Educação e Serviços Financeiros.

A região SE (sudeste) está associada as áreas (H) que se refere a produtos eletrodomésticos e eletrônicos, (C) demais produtos, (D) demais serviços e (B) alimentos. Aumentando o número de reclamações na região Sudeste, a tendência é também ocorra um aumento nas áreas B, C, D e H.

Conforme a análise da Tabela 2, a região SE possui 46,86% do número total de reclamações registradas no site consumidor.gov.br no período analisado. A região Nordeste 17,26%.

Para a associação entre a Região e Avaliação da Reclamação, a primeira coordenada principal explica 66,70% dos dados e a segunda coordenada principal explica 33,30% dos dados. O Biplot dessa associação é apresentado na Figura 10.

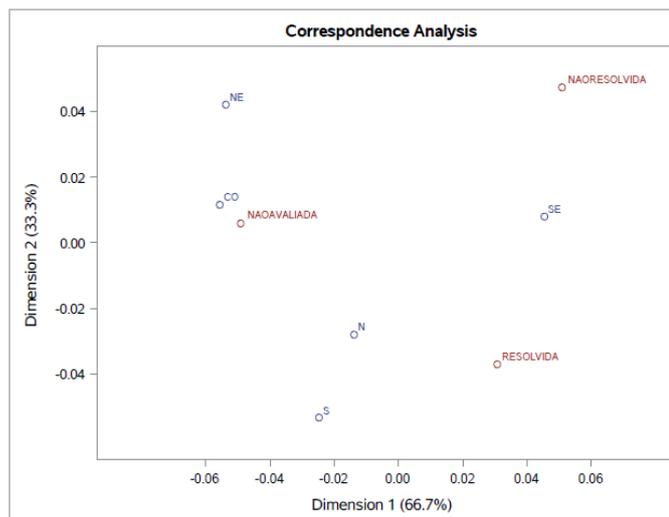


Figura 10 - Associação entre Região e Avaliação da Reclamação

Conforme a análise da Figura 10, que se refere a associação entre a avaliação da reclamação e a Região do consumidor, nota-se que a região CO (Centro-oeste) está associado ao nível NÃO AVALIADO da variável avaliação da reclamação. Se ocorrer aumento na quantidade de reclamações na região Centro-oeste, a tendência é que aumente o número de não avaliação por parte do cliente da tratativa da empresa na solução da reclamação.

Conforme a análise da Tabela 11, a avaliação da reclamação NÃO AVALIADO corresponde a 43,90% da quantidade total de avaliações realizadas pelos consumidores. Quase 50% dos consumidores não avaliaram a tratativa das empresas na solução das reclamações. Analisando o Biplot (Figura 10), NÃO AVALIADO está associado a área de Serviços Financeiros.

Para a associação entre a Região e Como Comprou/Contratou, a primeira coordenada principal explica 60,52% dos dados e a segunda coordenada principal explica 35,87% dos dados. O Biplot dessa associação é apresentado na Figura 11.

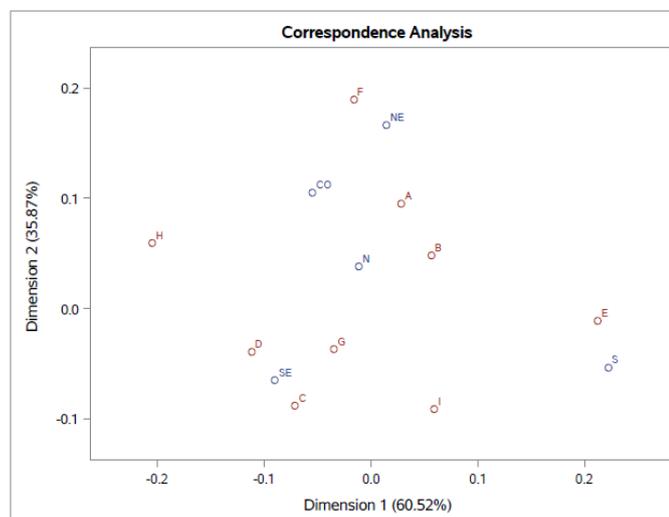


Figura 11 - Associação entre Região e Como Comprou/Contratou

Conforme a análise do Biplot (Figura 11), nota-se que existe associação entre a variável como comprou contratou e a variável Região do Consumidor, nota-se que a região S (Sul) está associado ao nível LOJA FÍSICA, codificado com a letra E, da variável Como Comprou/Contratou. Aumentando o número de reclamações na região SUL, a tendência é que ocorra aumento no número de reclamações relacionado a compras através da LOJA FÍSICA.

Como Comprou/Contratou é a descrição do meio utilizado para contratação/aquisição do produto ou serviço reclamado. Conforme a análise da Tabela 9, o meio de compra LOJA FÍSICA corresponde a 18,64% do número total de reclamações registradas no período estudado.

Para a associação entre a Região e Faixa Etária, a primeira coordenada principal explica 71,59% dos dados e a segunda coordenada principal explica 23,73% dos dados. O Biplot dessa associação é apresentado na Figura 12.

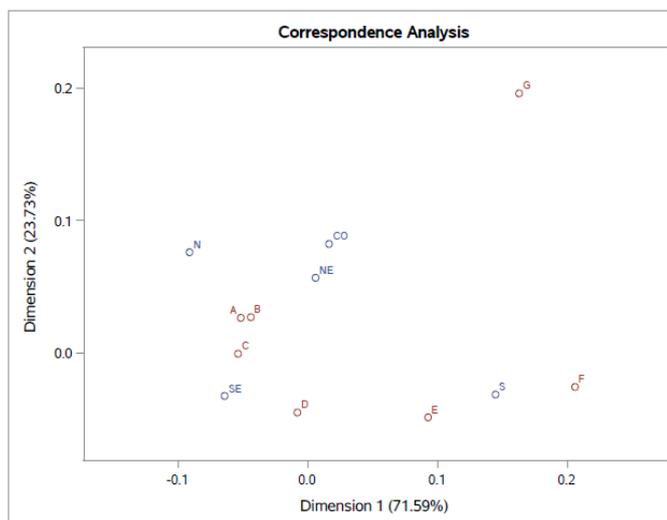


Figura 12 - Associação entre Região e Faixa Etária

Conforme a análise do Biplot (Figura 12), que se refere a associação entre a variável Faixa Etária e a variável Região do consumidor, nota-se que a região S (Sul) está associado ao nível F da variável da variável faixa etária (ENTRE 61 E 70 ANOS). Aumentando o número de reclamações na região SUL, a tendência é haver um aumento no número de reclamações de consumidores nessa faixa etária.

Analisando a Tabela 2, a região Sul representa 21,07% de todas as reclamações registradas no período em estudo. Já a faixa etária (ENTRE 61 E 70 ANOS), corresponde a 7,73% do total de reclamações registradas, vide a Tabela 5.

Para a associação entre a Região e Nota do Consumidor, a primeira coordenada principal explica 93,01% dos dados e a segunda coordenada principal explica 6,52% dos dados. O Biplot dessa associação é apresentado na Figura 13.

Conforme a análise do Biplot (Figura 13), que se refere a associação entre a variável Nota do Consumidor e a variável Região do consumidor, nota-se que a região S (Sul) está associado ao nível E da variável Nota do Consumidor (Nota 5). Aumentando o número de reclamações na região Sul, a tendência é haver um aumento no número de reclamações que receberão nota 5 de satisfação do consumidor devido a empresa ter resolvido o conflito.

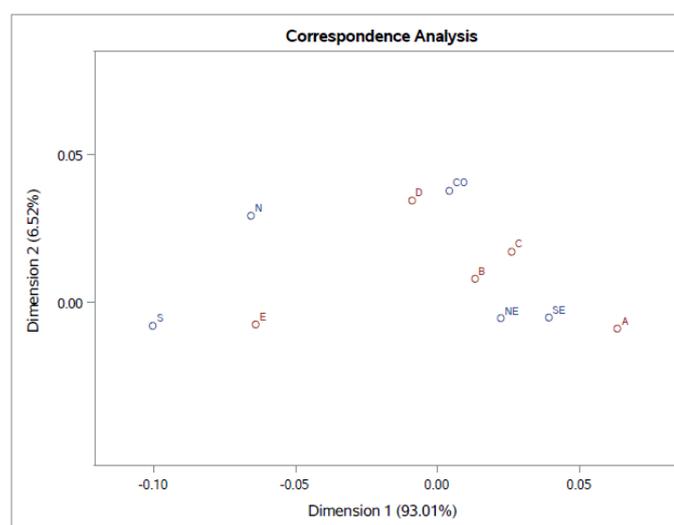


Figura 13 - Associação entre Região e Nota do Consumidor

Para a associação entre a Região e Tipo de Problema, a primeira coordenada principal explica 50,68% dos dados e a segunda coordenada principal explica 26,64% dos dados. O Biplot dessa associação é apresentado na Figura 14.

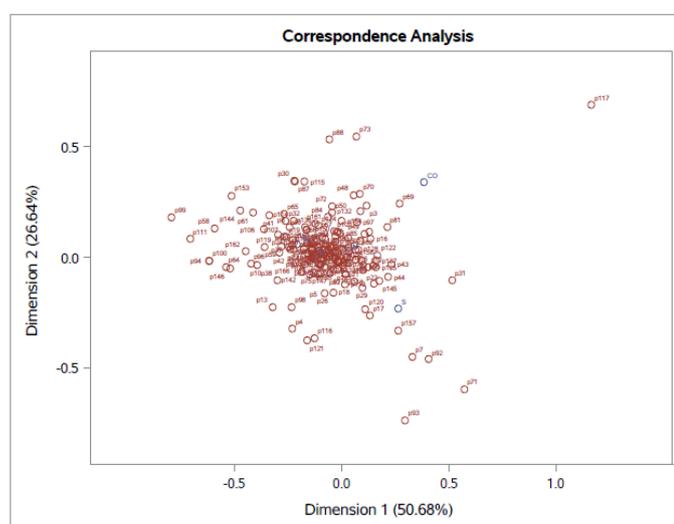


Figura 14 - Associação entre Região e Tipo de Problema

Conforme a análise do Biplot (Figura 14), que se refere a associação entre a variável Tipo de Problema e a variável Região do consumidor, nota-se que a região CO (Centro-Oeste)

está associado ao nível P117 da variável tipo de problema que é (Não entrega do contrato ou documentação relacionada ao serviço). Aumentando o número de reclamações na região CO (Centro-Oeste), a tendência é haver um aumento no número de reclamações com tipo de problema (Não entrega do contrato ou documentação relacionada ao serviço).

A região S (Sul) está associado aos níveis p157, p7, p92 e p71 da variável tipo de problema que é (Não entrega do contrato ou documentação relacionada ao serviço). Existe 166 tipos de problemas, ou seja, a variável Problema possui 166 níveis.

Para a associação entre a Região e Grupo de Problema, a primeira coordenada principal explica 42,28% dos dados e a segunda coordenada principal explica 33,91% dos dados. O Biplot dessa associação é apresentado na Figura 15.

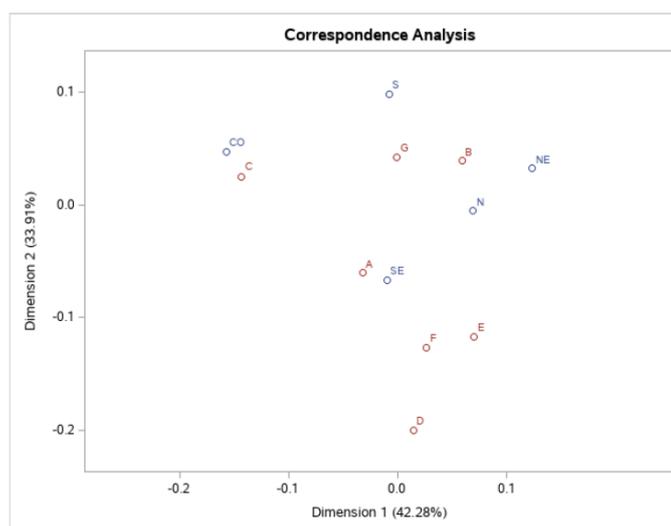


Figura 15 - Associação entre Região e Grupo de Problema

Conforme a análise da Figura 15, que se refere a associação entre a variável Grupo de problema e a variável Região do consumidor, nota-se que a região CO (Centro-Oeste) está associado ao nível C da variável Grupo de Problema (Contrato/Oferta). Aumentando o número de reclamações na região CO (Centro-Oeste), a tendência é haver um aumento no número de reclamações com grupo de problema (Contrato/Oferta).

Grupo de Problema é o agrupamento do qual faz parte o problema classificado na reclamação. Conforme a análise da Tabela 8, Contrato/Oferta corresponde 18,06% do total de reclamações registradas no período de estudo desse trabalho.

Para a associação entre a Região e Tempo de Resposta, a primeira coordenada principal explica 56,66% dos dados e a segunda coordenada principal explica 28,30% dos dados. O Biplot dessa associação é apresentado na Figura 16.

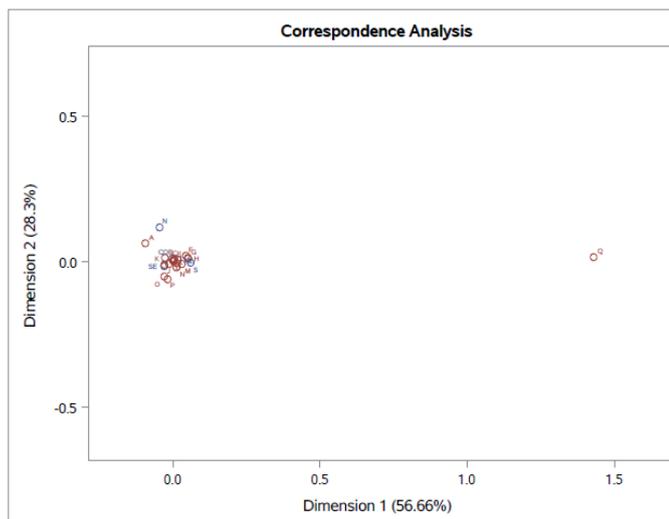


Figura 16 - Associação entre Região e Tempo de Resposta

Conforme a análise da Figura 16, que se refere a associação entre as variáveis Região e Tempo de Resposta, nota-se que a região N (norte) está associada ao tempo de resposta A (Menos de um dia). Aumentando o número de reclamações na região Norte, a tendência é haver um aumento no número de reclamações com tempo de tratativa, por parte das empresas, com menos de um dia. O tempo de resposta é o número de dias para a resposta da reclamação pela empresa, e a diferença, em dias, entre a data da resposta e a data de abertura. O ponto isolado Q que se refere 16 dias não tem associação com as regiões do país.

Para a associação entre as variáveis Grupo de Problema e Tempo de Resposta, a primeira coordenada principal explica 88,05% dos dados e a segunda coordenada principal explica 11,02% dos dados. O Biplot dessa associação é apresentado na Figura 17.

Conforme a análise do Biplot Figura 17, que se refere a associação entre a variável Grupo de Problema e a variável Tempo de Resposta, nota-se que o nível Tempo de Resposta igual a zero (0 dia) está associado ao nível E do Grupo de Problema. O nível E se refere a Informação. Aumentando o número de reclamações no Grupo de Problema Informação, a tendência é haver um aumento no número de reclamações com tempo de tratativa, por parte das empresas, com menos de um dia.

unidades federativas MS e PI, a tendência é haver um aumento no número de reclamações na área de Serviços Financeiros.

Para a associação entre a UF e Avaliação da Reclamação, a primeira coordenada principal explica 82,77% dos dados e a segunda coordenada principal explica 17,23% dos dados. O Biplot dessa associação é apresentado na Figura 19.

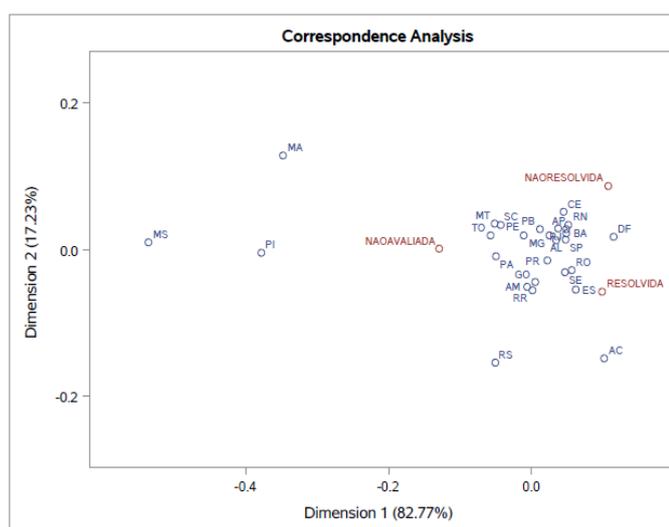


Figura 19 - Associação entre UF e Avaliação da Reclamação

Conforme a análise da Figura 19, que se refere a associação entre a avaliação da reclamação e UF, nota-se que os níveis SE, ES e RO da variável UF estão associados ao nível RESOLVIDA da variável avaliação da reclamação. Aumentando o número de reclamações nas unidades federativas SE, ES e RO, a tendência é haver um aumento no número de reclamações atribuídas com RESOLVIDA por parte dos consumidores.

Para a associação entre a UF e Como Comprou/Contratou, a primeira coordenada principal explica 61,93% dos dados e a segunda coordenada principal explica 31,53% dos dados. O Biplot dessa associação é apresentado na Figura 20.

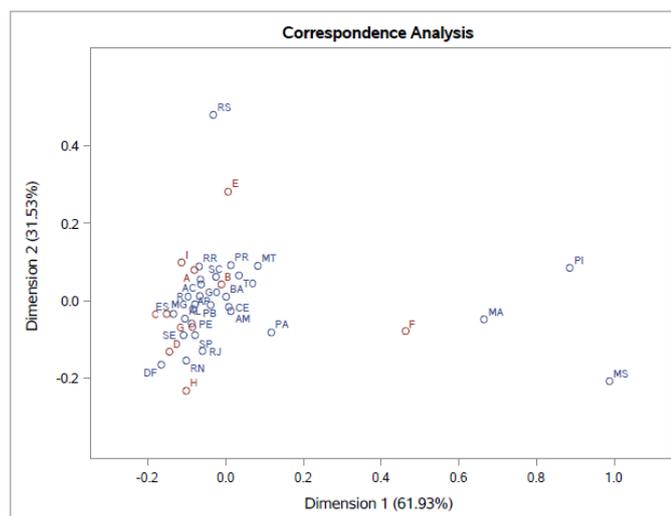


Figura 20 - Associação entre UF e Como Comprou/Contratou

Conforme a análise da Figura 20, que se refere a associação entre a variável como comprou contratou e a variável UF, nota-se que os níveis PI, MA e MS da variável UF estão associados ao nível F da variável Como Comprou/Contratou. O nível F é (Não_comprei_contratei). Aumentando o número de reclamações nas unidades federativas PI, MA e MS, a tendência é haver um aumento no número de reclamações envolvendo a categoria (Não_comprei_contratei).

Para a associação entre a UF e Faixa Etária, a primeira coordenada principal explica 79,08% dos dados e a segunda coordenada principal explica 13,35% dos dados. O Biplot dessa associação é apresentado na Figura 21.

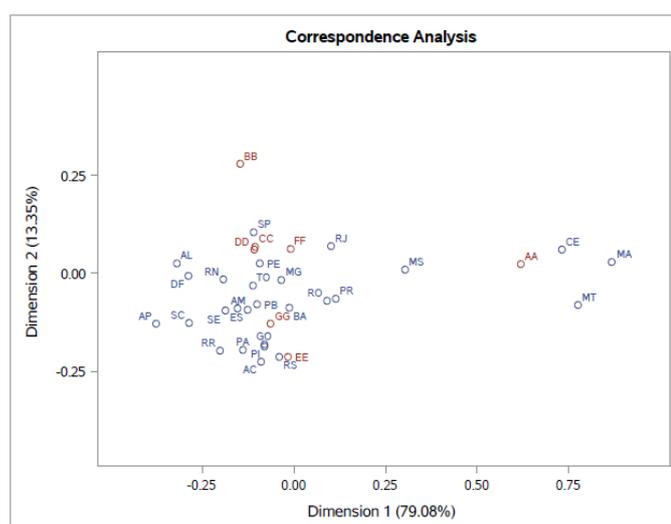


Figura 21 - Associação entre UF e Faixa Etária

é que a tratativa, por parte das empresas, para resolver conflitos nesse Estado receba nota 1 pelos consumidores.

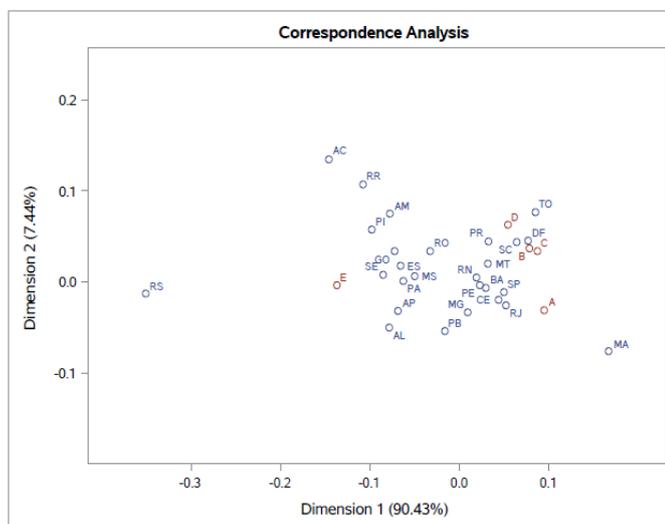


Figura 23 - Associação entre UF e Nota do Consumidor

Aumentando o número de reclamações na unidade federativa RS, a tendência é que a tratativa por parte das empresas para resolver conflitos nesse Estado receba nota 5 pelos consumidores. Assim, nota-se que no estado do Rio Grande do Sul há uma melhor satisfação na tratativa das reclamações, por parte dos consumidores, em relação ao estado do Maranhão.

Para a associação entre a UF e Tempo de Resposta, a primeira coordenada principal explica 39,35% dos dados e a segunda coordenada principal explica 25,96% dos dados. O Biplot dessa associação é apresentado na Figura 24.

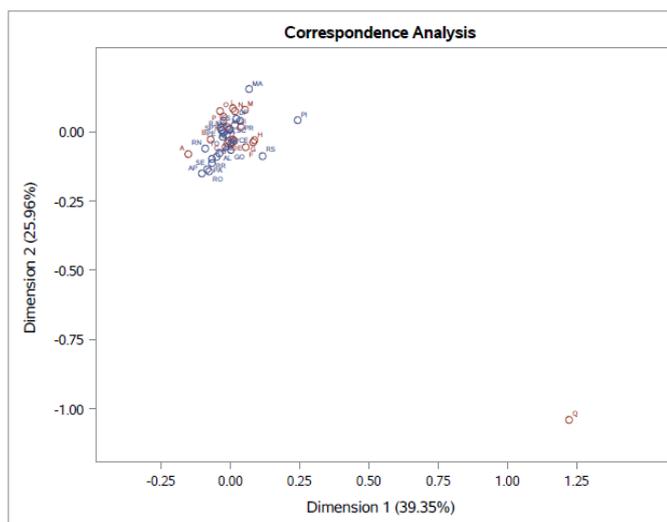


Figura 24 - Associação entre UF e Tempo de Resposta

Conforme a análise da Figura 24, que se refere a associação entre a variável Tempo de Resposta e a variável UF, nota-se que o nível RN da variável UF está associado ao nível A da variável Tempo de Resposta. O nível A é (0 dia). Aumentando o número de reclamações na unidade federativa RN, a tendência é que o tempo de tratativa, por parte das empresas, para resolver conflitos nesse Estado é menos de um dia. O ponto isolado Q que se refere 16 dias não tem associação com as unidades federativas do país.

4. Conclusões

Com base nos objetivos propostos neste trabalho de se analisar a associação entre as variáveis categóricas obtidas através das reclamações dos consumidores, todas variáveis consideradas nesse estudo tiveram associação significativa. Constatou-se que as áreas de Serviços Financeiros e Telecomunicações foram as que mais receberam reclamações na análise realizada. O setor de telecomunicações, apesar de estar no topo das reclamações, está associado a maior nota de satisfação dos reclamantes na solução dos conflitos. A maior nota de satisfação está associada à faixa etária entre 21 e 30 anos e também associada a unidade federativa RS.

O setor de serviços financeiros recebem mais reclamações da faixa etária acima de 60 anos, que por sua vez está associado a não aquisição e não contratação. As reclamações associadas a faixa etária entre 61 a 70 anos tiveram associação com a região SUL do país. A região Sul está associada a forma de contratação LOJA FÍSICA. O setor de serviços financeiros está associado as unidades federativas MS e PI. A região NE está associado a conflitos relacionados as áreas de educação e Serviços Financeiros. Conforme pesquisa publicada em março de 2020, no site Valorinveste⁴, Bancos e telefonia são líderes em reclamações de consumidores com fonte da SENACON.

Os setores que receberam as piores notas de avaliação da solução dos conflitos foram o de educação, habitação, Turismo e Viagens e Água/Energia/Gás. O setor de Transportes, Turismo e Viagens estão associados a reclamações da unidade federativa DF.

O grupo de problema Saúde e Segurança está associado a pior nota dos consumidores na solução do conflito, recebeu nota 1. Já o grupo de problema Informação está associado a faixa etária até 30 anos e com tempo de resposta por parte da empresa para resolver o conflito em menos de um dia.

A forma de compra Ganhei de presente está associada a faixa etária até 20 anos que por sua vez está associada as unidades federativas MT, MA e CE.

A partir dos resultados obtidos pode-se concluir que a análise das reclamações são essenciais para as empresas e organizações a verificarem os tipos de seguimentos que merecem mais atenção no pós-venda de serviços e produtos e a fim de ajudá-los nas tomadas de decisão. Para os consumidores, esse trabalho apresenta uma visão geral dos segmentos de mercado que tem mais conflitos na resolução dos problemas em relação aos produtos e serviços adquiridos.

⁴ <https://valorinveste.globo.com/objetivo/gastar-bem/noticia/2020/03/10/bancos-e-telefonia-sao-lideres-em-reclamacoes-de-consumidores-diz-senacon.ghtml>

Quanto ao aspecto metodológico, este trabalho mostrou uma aplicação original da técnica AC, usada para auxiliar a interpretação de dados qualitativos a respeito das reclamações registradas no site consumidor.gov.br, isso considerando todos os seguimentos de mercado.

Como recomendação de estudos futuros, pode-se estudar as causas que levaram a essas associações.

5. Referências Bibliográficas

- [1]. ALMEIDA, T. N. V; RAMOS, A. S. M. **Os impactos das reclamações on line na lealdade dos consumidores: um estudo experimental.** Revista de Administração Contemporânea, v.16, n. 5, p.664–683, 2012. Disponível em < https://www.scielo.br/scielo.php?script=sci_arttext&pid=S1415-65552012000500003&lng=pt&tlng=pt >. Acesso em: 25 agosto 2020.
- [2]. ALVES, C.A.M. **Estudo sobre a Responsabilidade Social corporativa e as reclamações pertinentes às ouvidorias dos bancos, considerando o porte e o tipo de controle no Brasil.** Art. Rev. Adm. UFSM, Santa Maria, v. 10, número 3, p.402-419, JUL.-SET. 2017. Disponível em < <https://periodicos.ufsm.br/reaufsm/issue/view/1183> >. Acesso em: 25 agosto 2020.
- [3]. CAMPOS, M. H. C. **Aplicação do Modelo Binomial aos Dados do Sítio Consumidor.gov.br.** Trabalho de conclusão de Curso. Universidade Federal de Juiz de Fora. Departamento de Estatística. 2016. Disponível em < https://www.ufjf.br/cursoestatistica/files/2016/08/monografia_mario_halfeld_clark_campos.pdf >. Acesso em: 18 novembro 2020.
- [4]. FAVERO, L.P. et.al. **Análise de dados:** modelagem multivariada para tomada de decisões. Rio de Janeiro: Elsevier, 2009.
- [5]. IBGE. Instituto Brasileiro de Geografia e Estatística. **Pesquisa Nacional por Amostra de Domicílios Contínua - PNAD contínua.** Disponível em: < <https://agenciadenoticias.ibge.gov.br/agencia-noticias/2012-agencia-de-noticias/noticias/27522-rendimento-impacta-meio-de-acesso-da-populacao-a-bens-tecnicos-e-internet> >. Acesso em: 30 novembro 2020.
- [6]. MINGOTI, S. A. **Análise de dados através de métodos de estatística multivariada:** uma abordagem aplicada. Belo Horizonte: Ed. UFMG, 2005. 257 p.
- [7]. MORETTIN, Pedro A.; BUSSAB, Wilton de O. **Estatística Básica.** Ed. Saraiva. 2010. 76 p.
- [8]. ANALYTICS SOFTWARE & SOLUTIONS. University Edition. Disponível em: < <https://welcome.oda.sas.com/login> >. Acesso em: 04 de outubro 2020.
- [9]. SILVA, L.F.F. **Análise da avaliação do consumidor, sua reclamação e da recuperação de serviços da CEB no Reclame Aqui.** Trabalho de conclusão de curso. UnB/DF. Sistema de Bibliotecas, 2015. Disponível em < <https://bdm.unb.br/handle/10483/17576> >. Acesso em: 25 agosto 2020.