

# UNIVERSIDADE FEDERAL DO PARANÁ CAMPUS PONTAL DO PARANÁ CENTRO DE ESTUDOS DO MAR CURSO DE LICENCIATURA EM CIÊNCIAS EXATAS

César Eduardo Maciel

# ANÁLISE DE COMPONENTES PRINCIPAIS APLICADO AO PROBLEMA DE EVASÃO UNIVERSITÁRIA: UM ESTUDO SOBRE O QUESTIONÁRIO DO VESTIBULAR DA UFPR

Pontal do Paraná 2021

#### César Eduardo Maciel

# ANÁLISE DE COMPONENTES PRINCIPAIS APLICADO AO PROBLEMA DE EVASÃO UNIVERSITÁRIA: UM ESTUDO SOBRE O QUESTIONÁRIO DO VESTIBULAR DA UFPR

Trabalho de Conclusão de Curso apresentado à disciplina Trabalho de Conclusão de Curso II do Curso de Licenciatura em Ciências Exatas do Campus Pontal do Paraná, Centro de Estudos do Mar, Universidade Federal do Paraná, como exigência parcial para obtenção do grau de Licenciado em Ciências Exatas - Habilitação em Matemática.

Orientador: Eduardo Tadeu Bacalhau

Pontal do Paraná 2021

# DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP) UNIVERSIDADE FEDERAL DO PARANÁ SISTEMA DE BIBLIOTECAS – BIBLIOTECA DO CENTRO DE ESTUDOS DO MAR

#### Maciel, César Eduardo

Análise de componentes principais aplicado ao problema de evasão universitária: um estudo sobre o questionário do vestibular da UFPR / César Eduardo Maciel. – Curitiba, 2021. 1 recurso on-line: PDF.

Monografia (Graduação) – Universidade Federal do Paraná, Campus Pontal do Paraná, Centro de Estudos do Mar, Curso de Licenciatura em Ciência Exatas – Habilitação em Matemática. Orientador: Prof. Dr. Eduardo Tadeu Bacalhau.

1. Evasão universitária. 2. Análise de componentes principais - ACP. 3. Questionário do estudante. I. Bacalhau, Eduardo Tadeu. II. Universidade Federal do Paraná. III. Título.

Bibliotecária: Liliam Maria Orquiza CRB-9/712



# UNIVERSIDADE FEDERAL DO PARANÁ CENTRO DE ESTUDOS DO MAR Campus Pontal do Paraná



### TERMO DE APROVAÇÃO

# CÉSAR EDUARDO MACIEL

Análise de Componentes Principais Aplicado ao Problema de Evasão Universitária: Um estudo sobre o questionário do vestibular da UFPR.

Trabalho de Conclusão de Curso aprovado como requisito parcial para a obtenção do grau de Licenciado em Ciências Exatas - Matemática, da Universidade Federal do Paraná, pela Comissão formada pelos professores:

duar de Fadm Bacallau Dr. Eduardo Tadeu Bacalhau Orientador e Presidente

Dr. Fernando Araújo Borges Membro Examinador

> Dr<sup>a</sup>. Luciana Casacio Membro Examinador

Pontal do Paraná, 23/12/2021.

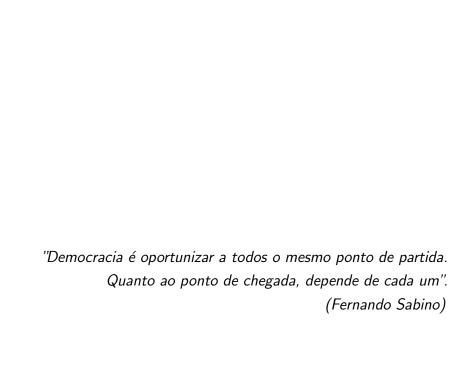
# Agradecimentos

Agradeço primeiro a Deus por ter me mantido na trilha certa durante este projeto de pesquisa com saúde e forças para chegar até o final.

Agradeço ao meu orientador, Prof. Dr. Eduardo Tadeu Bacalhau, pelo acompanhamento, orientação e paciência.

Aos meus pais José Carlos Maciel e Rita de Cássia Marinho Maciel que sempre estiveram ao meu lado me apoiando ao longo de toda a minha trajetória.

Aos professores, que estiveram presente em algum momento de minha graduação, por todo apoio e repasse de conhecimento. Aos colegas de curso, por todos os momentos vividos. Ao Curso de Licenciatura em Ciências Exatas, da Universidade Federal do Paraná. Aos funcionários do Campus Mirassol, por toda ajuda e apoio.



# Resumo

A evasão universitária é um problema árduo e que nos últimos anos, vem se acentuando nas universidade públicas brasileiras. Este problema afeta o sistema educacional, dado que o custo operacional do serviço público prestado para a manutenção de um curso de graduação é muito alto, e essencialmente, o próprio aluno, simbolizando uma perda de autoestima e também social. Neste contexto, este trabalho tem como objetivo principal identificar e diagnosticar os perfis de estudantes, classificado como evadido ou matriculado, utilizando os dados referentes ao questionário do estudante promovido pela Universidade Federal do Paraná, para o Curso de Licenciatura em Ciências Exatas, do Campus Pontal do Paraná. Este diagnostico é produzido a partir de um método matemático, denominado Análise de Componentes Principais, do inglês Principal Components Analysis (PCA), relacionando as principais questões da base de dados com os indicadores de abandono disponibilizados pelo curso. Um estudo computacional é elaborado, compreendendo quatro experimentos acerca da aplicação do PCA na linguagem R, um software livre de análise de dados. Estes experimentos destacam as variáveis com maior representatividade quanto da base de dados e uma base de dados reduzida, ilustrando os resultados via gráficos, mapas e tabelas. Finalmente as conclusões apontam para um diagnóstico acurado sobre os perfis dos estudantes e para o desenvolvimento de futuros desdobramentos do trabalho.

**Palavras-chave**: Evasão universitária, análise de componentes principais, questionário do estudante.

# **Abstract**

University dropout is a hard problem, and in recent years, it has been increasing in Brazilian public universities. It affects not only the educational system, once the public service operational cost to maintain an undergraduate degree is very expensive, but also the student, representing a loss of self-esteem and social as well. In this context, the main objective of this work is to identify and diagnose the student profiles, classifying as evaded or not, using data from a student questionnaire promoted by the Federal University of Paraná for the Lecture Degree in Exact Sciences, at Pontal do Paraná Campus. This diagnosis is elaborated through a mathematical method, called Principal Components Analysis (PCA), listing the main questions in the database as regards the dropout indicators provided by the course. A computational study is produced, comprising four experiments considering the PCA application in R language, a free data analysis software. These experiments show the variables with high representation in relation to the database and the reduced database, demonstrating the results via graphs, maps and tables. Finally, conclusions are presented, pointing to an accurate diagnosis of student profiles and the future developments of this work.

Key-words: University dropout, principal components analysis, student questionnaire.

# Sumário

| 1       | INTRODUÇÃO   | 1  |
|---------|--|----|
| 2       | REVISÃO BIBLIOGRÁFICA                                | 3  |
| 2.1     | A Evasão escolar no Ensino Superior                  | 3  |
| 2.2     | Métodos Inteligentes Aplicados ao Problema da Evasão | 3  |
| 2.3     | Método PCA aplicado ao Problema de Evasão            |    |
| 3       | ANÁLISE DE COMPONENTES PRINCIPAIS                    | 8  |
| 3.1     | Padronização dos Dados                               | 8  |
| 3.2     | Matriz de Covariância                                | g  |
| 3.3     | Componentes Principais (CPs)                         | g  |
| 3.4     | Exemplo de Cálculo de Componentes Principais         | 10 |
| 3.4.1   | Autovetores e Autovalores                            | 12 |
| 3.5     | O Algoritmo PCA                                      | 12 |
| 3.6     | Descrição Matemática                                 | 13 |
| 4       | IMPLEMENTAÇÃO DO PCA                                 | 16 |
| 4.1     | PCA do conjunto de dados do vinho                    | 16 |
| 4.2     | PCA com pcaMethods                                   | 18 |
| 5       | ESTUDOS COMPUTACIONAIS                               | 21 |
| 5.1     | Experimento I  | 21 |
| 5.1.1   | Implementação e Resultados                           | 21 |
| 5.2     | Experimento II                                       | 24 |
| 5.2.1   | Implementação e Resultados                           | 25 |
| 5.3     | Experimento III                                      | 28 |
| 5.3.1   | Implementação e Resultados                           | 29 |
| 5.4     | Experimento IV                                       | 33 |
| 5.4.1   | Implementação e Resultados                           | 33 |
| 5.4.2   | Questões Relacionadas às Matriculas Ativas           | 34 |
| 5.4.3   | Questões Relacionadas à Evasão                       | 36 |
| 5.4.3.1 | Comentários Finais                                   | 39 |
| 6       | CONCLUSÕES   | 41 |
|         | REFERÊNCIAS  | 42 |

| ÊNDICE 2                         | 7   | APÊNDICE 1                         |
|----------------------------------|-----|------------------------------------|
| dos Estatísticos Base de Dados R | 7.1 | Código R                           |
| dos Estatísticos Base de Dados R | 8   | APÊNDICE 2                         |
| relograma                        | 8.1 | Base de dados                      |
|                                  | 8.2 | Dados Estatísticos Base de Dados R |
| ESTIONÁRIO SOCIOECONÔMICO UFPR   | 8.3 | Correlograma                       |
|                                  | 9   |                                    |

# 1 Introdução

Um dos principais problemas enfrentados pelas universidades públicas brasileiras é o elevado índice de evasão de escolar. Este problema vem se agravando nas últimas décadas, dado o grande número de cursos e universidades disponíveis após a elaboração de dois processos de expansão do ensino superior, ambos classificados como Programa de Apoio a Planos de Reestruturação e Expansão das Universidades Federais (Reuni). O primeiro passo do programa, instituído no ano de 2008, buscava a ampliação e inclusão de novos estudantes para as instituições federais de ensino superior. Já em 2010, o programa buscou novos objetivos, destacando o processo de interiorização das instituições federais de ensino. Contudo, é intuito notar que os problemas sociais enfrentados pela maioria da população brasileira, tais como a desigualdade regional, a má distribuição de renda, recursos limitados para o ensino básico, médio e superior, têm uma grande parcela de responsabilidade pelo aumento desses índices de desistência.

A evasão, além de representar perda social e de recursos públicos, também representa uma perda de autoestima para o estudante e uma perda enorme de produção de conhecimento científico para o curso envolvido, significando uma ociosidade e um desperdício intelectual. Neste contexto, destacam-se as particularidades enfrentadas pelo Campi alocados em regiões com grande índice de vulnerabilidade social, o qual é representado pelos dados obtidos no estudo de casos proposto neste trabalho.

A região do litoral paranaense, especificamente o município de Pontal do Paraná, sofre com os problemas sociais já mencionados, além de uma já demonstrada falta de professores e profissionais qualificados (IBGE..., 2022). Sendo assim, a partir de 2014, caracterizado pelo segundo processo do REUNI, novos cursos de graduação foram criados e implementados no município, pela Universidade Federal do Paraná. Os novos cursos deram origem ao Campus Pontal do Paraná, mais do que dobrando o número de vagas ofertadas para o processo seletivo. Sendo assim, considerando todas as particularidades da região litorânea, as dificuldades de infraestrutura presentes instituição, o grande número de alunos com alguma vulnerabilidade social, incentivou este estudo, que busca estudar o processo de evasão enfrentado pelo Campus Pontal do Paraná, mais especificamente, o Curso de Licenciatura em Ciências Exatas.

Cabe ressaltar que, alguns dos principais problemas relacionados à evasão no Curso está relacionado à necessidade de se atuar em mais de uma atividade, além do próprio estudo, associada sobretudo ao sustento familiar, onde o indivíduo prioriza a questão financeira frente a formação proporcionada. Outro problema intuitivamente relacionado é o baixo nível escolar médio da população, demonstrado pela defasagem oriunda de um ensino básico e médio muito enfraquecido. Isto mostra a importância social explicitamente justificada na criação dos novos cursos de graduação na região.

A partir deste panorama, este trabalho buscar realizar um estudo sobre o questionário

do estudante promovido pela Universidade Federal do Paraná, dentro do processo seletivo para as vagas. Esta estratégia consiste em identificar e diagnosticar os principais fatores que indicam um perfil de estudante com a probabilidade de sofrer evasão, dada as observações mencionadas. O principal objetivo é realizar um diagnóstico através de um método matemático, Análise dos Componentes Principais, do inglês *Principal Components Analisys* (PCA), o qual promove uma análise das principais perguntas do questionário, relacionando-as com os indicadores de abandono dos cursos disponibilizados pela própria instituição. A identificação dos alunos, que correm o risco de abandonar a universidade, tem como finalidade fornecer assistência através de apoio estudantil promovido pela Universidade Federal do Paraná. Busca-se um plano de intervenção, prevenção e recuperação do indivíduo identificado, incentivando a permanência e a redução dos índices de abandono.

A abordagem é validada por um estudo computacional, o qual promove uma análise gráfica dos componentes socioeconômicos do questionário, uma investigação sobre as principais questões, demonstrando suas representatividades nas componentes escolhidas pelo método, e um experimento para delimitação dos dados do questionário, descrevendo-os e relacionando-os com os índices de evasão.

Por fim, na conclusão do trabalho, considerando uma análise geral apresentada no estudo computacional proposto, o desempenho do PCA é discutido para a base de dados levantada e novos apontamentos são apresentados para continuidade da pesquisa iniciada, enfatizando o desenvolvimento de um novo questionário voltado unicamente para as características da região do litoral do Paraná.

# 2 Revisão Bibliográfica

A capacidade de prever o desempenho dos alunos em cursos cria oportunidades para melhorar os resultados educacionais de uma instituição de ensino, seja da rede pública de ensino ou particular. Por meio de métodos computacionais eficazes de previsão de desempenho, os professores podem direcionar esforços e utilizar diferentes metodologias de ensino com mais precisão. Assim, esta revisão bibliográfica visa identificar recursos que podem ser usados para fazer previsões, identificar algoritmos que podem melhorar as previsões e quantificar vários aspectos do desempenho do aluno. Além disso, também são explorados trabalhos que descrevem as consequências da evasão escolar e alguns estudos sobre o tema onde os recursos de previsão não são utilizados. Por fim, é apresentada uma análise que ilustra o número de estudos nesta área e seu aumento significativo, além das tecnologias utilizadas, as quais têm se tornado cada vez mais diversificadas.

# 2.1 A Evasão escolar no Ensino Superior

Vários estudos mostram o impacto da identificação no desempenho acadêmico, bemestar e aprendizagem, por exemplo, Filho et al. (2007) apresentam o trabalho intitulado "Evasão no ensino superior no Brasil", onde a evasão foi considerada como um problema de gestão. O artigo fez um resumo da evolução histórica dos estudos internacionais mais importantes sobre as causas da evasão do aluno, indicando as ações necessárias para combatê-la.

Mais recentemente, o foco de estudo passou da organização para o estudo de identificação do tipo de estudante, como mostra o trabalho de Shirasu, Albuquerque et al. (2015), intitulado "Determinantes da Evasão Escolar e Repetência Escolar no Ensino Médio do Ceará", onde os autores mostraram que a defasagem idade-série, que identifica os repetentes, foram apenas 23% nas séries escolares. Assim, os autores concluíram que o fator pontuação média da escola, no SPAECE 2008 (SPAECE, 2021) (nos exames de português e matemática), mostrou-se nulo, ou seja, o índice repetências foi irrelevante, pois as escolas possuem praticamente as mesmas estruturas administrativas e quase todas estão localizadas em áreas urbanas. Já o trabalho de Bayer et al. (2012) investigou a identificação com grupos de comportamento social. Este artigo se concentrou em prever o abandono e o fracasso escolar relacionados a dados de alunos enriquecidos com dados derivados do comportamento social destes.

# 2.2 Métodos Inteligentes Aplicados ao Problema da Evasão

Ao longo dos anos métodos numéricos de Álgebra linear e da estatística, têm se tornado fundamental para o sucesso das mais variadas tarefas computacionais, pela robustez com que esses resolvem problemas computacionais comuns a uma série de campos, bem como pela elegante formalização matemática aplicável. Assim, métodos numéricos computacionais podem ser ferramentas essenciais para auxiliar no diagnóstico da evasão, sendo utilizados no processamento dos dados para identificação de alunos em risco e permitir um planejamento oportuno para intervenções educacionais.

No trabalho intitulado "Predição do Grupo de Risco de Evasão Discente em Cursos Superiores Presenciais Utilizando uma Rede Neural ArtMap- Fuzzy", Martinho, Nunes e Minussi (2013), desenvolveram técnicas da inteligência artificial, uma Rede Neural ARTMAP-Fuzzy, que possibilitou o aprendizado continuado do sistema. Dados socioeconômicos e acadêmicos dos estudantes matriculados nos cursos superiores de tecnologia do Instituto Federal de Mato Grosso - IFMT foram utilizados para o treinamento da rede neural. Foi apresentado um método inovador para identificar os estudantes considerados do grupo de risco de evasão escolar. A partir da análise dos resultados obtidos, os autores puderam deduzir que o sistema foi robusto, eficaz e com significativo grau de confiabilidade, atingindo uma taxa de sucesso global de 76% e 94% na identificação precoce do grupo de risco evasivo. O estudo apresenta a possibilidade de elaboração de uma rede neural artificial explorando os resultados apresentados.

Lykourentzou et al. (2009) propõe um sistema de predição de alunos em situação de risco de evasão, combinando os resultados da aplicação de três algoritmos diferentes: Redes Neurais, Máquinas de suporte vetorial e sequência mínima de otimização e conjunto probabilístico simplificado ARTMAP-Fuzzy. Utiliza-se dados demográficos constantes no decorrer do curso como sexo e residência, performance e nível acadêmicos, e dados dinâmicos, como notas e a data da entrega dos trabalhos por exemplo. Usando esses esquemas, o método proposto atingiu uma taxa de classificação geral dos alunos de 75-85% da primeira seção dos dois cursos, para chegar a uma taxa de 97-100% nas seções finais.

Em (BAYER et al., 2012), o objetivo foi desenvolver um método de mineração de dados, do inglês *Data Mining*, educacional com o intuito de elaborar um classificador para prever o sucesso de um aluno, com validação a partir de dados reais. Os dados estudados descrevem dependências sociais coletadas principalmente por conversas de e-mail e fóruns de discussão entre os alunos, além de outras fontes. Estes dados de comportamento social resultaram em aumento significativo da precisão da previsão, sendo possível estimar o número de semestres necessários para que o modelo pudesse prever uma desistência com alta probabilidade de acerto. Neste contexto, os autores ainda afirmaram que informações das redes sociais poderiam ser futuramente consideradas, com a finalidade de melhorar a previsão.

Em (TIMBAL, 2019), foi destacado que a detecção da evasão precoce pode reduzir os riscos de abandono escolar por meio de ações oportunas e medidas proativas, e ainda, o avanço do país depende excepcionalmente da educação de seus cidadãos, uma verdade incontestável e de perspectiva mundial. Assim, os autores afirmam que a educação é o elemento principal no desenvolvimento de um indivíduo para se tornar um cidadão produtivo e é o fator mais importante que contribui para o progresso do país. Além destes aspectos educacionais muito

bem redigidos, os quais foram muito importantes na construção dos ideais apresentados neste trabalho, este artigo teve como objetivo analisar o conjunto de dados, utilizando a técnica de mineração de dados. As técnicas de mineração de dados também são usadas para prever os resultados, a identificação de indicadores para uma melhor previsão também foi feita para determinar quais características dos alunos são os preditores com mais impacto.

Segundo Assis (2020), a evasão nos cursos de Licenciatura é fenômeno recorrente, por isso a necessidade de se prever os potenciais a evasão torna-se necessário. Foi avaliada a capacidade de uma rede neural artificial, do tipo Perceptron Multicamadas, buscando identificar indivíduos de cursos de Licenciatura em Matemática com possibilidade de evadir-se do curso. Após o treinamento da rede, foi obtida uma taxa de assertividade de 81.97% na identificação de alunos com potencial à desistência, com um erro de apenas 18.03%, existindo a possibilidade de aumento da acurácia desta rede, com a inserção de mais dados para o treinamento.

Braga, Carmo e Bogutchi (2003) apresentam uma análise da evasão e repetência no sistema de ensino brasileiro. Trata-se do problema de reprovação e evasão escolar principalmente na UFMG (Universidade Federal de Minas Gerais), problema que ocorria nas décadas de 60 e 70. A análise realizada resultou de uma pesquisa com base na análise do perfil socioeconômico dos estudantes e do seu desempenho escolar, visando uma compreensão da complexidade das relações que se desenvolvem no ambiente escolar e que podem vir a ser causadoras desse fenômeno. Verificou-se a fraca correlação entre a evasão e a classificação obtida no concurso vestibular, renda familiar e o grau de instrução dos pais destes discentes, ao mesmo tempo em que a decepção com expectativas não correspondidas pelo curso foi apontada pelos evadidos como o motivo que mais contribuiu para sua decisão de abandonar o curso. Também verificou a existência de uma forte correlação entre a evasão e os altos índices de repetência registrados no primeiro período do curso. Outro aspecto importante relaciona-se à cultura escolar e ao papel do professor, sendo este diagnóstico foi realizado com base na análise do perfil socioeconômico dos estudantes e do seu desempenho escolar.

No experimento de (JAYAPRAKASH et al., 2014) é criado um algoritmo para alertar o risco relacionado ao desempenho do aluno, com a intenção da diminuição das taxas de evasão, gerando ao aluno um relatório a respeito do seu rendimento escolar. São utilizados dados como sexo e idade, índice do desempenho acadêmico, tempo de vida na universidade, tempo online no ambiente virtual (AVA), entre outros, e foram analisados dados de 9938 alunos utilizando o algoritmo das Redes Bayesianas <sup>1</sup>. Testes de predição do algoritmo classificador mostraram resultados de 94,20% acurácia geral e 66,70% na predição do risco de evasão.

Em Cambruzzi, Rigo e Barbosa (2015) propõe-se um sistema de aprendizagem educacional combinando ferramentas em uma camada da rede neural chamada de *MultiTrail*, que é responsável pela interpretação dos dados, processamento e estratificação dos dados para análise. Para predição são utilizadas Redes Neurais e são utilizados dados dos alunos da Universidade

<sup>&</sup>lt;sup>1</sup>Matematicamente, uma rede bayesiana é uma representação compacta de uma tabela de conjunção de probabilidades do universo do problema.

do Vale do Rio dos Sinos (UNISINOS). Uma precisão de 87% na predição de alunos em risco de evasão foi atingida.

Baker, Isotani e Carvalho (2011) mostram um estudo sobre a Mineração de Dados Educacionais, do inglês *Educational Data Mining* (EDM). O termo, também conhecido como *Knowledge Discovery in Databases*, refere-se à descoberta de "novas" informações através da análise de grande quantidade de dados. Os autores utilizaram variáveis preditivas para desenvolver modelos que consigam descrever características e informações sobre os dados obtidos para melhorar a qualidade do ensino. Este estudo proposto tem como objetivo validar as pesquisas realizadas fora do país, com a finalidade de implementá-las no país, principalmente em relação a cursos na modalidade educação a distância.

Yadav e Pal (2012) utilizaram árvores de decisão para a predição final da situação de alunos quanto à evasão. Apresentaram dados do curso de engenharia da *Veer Bahadur Singh Purvanchal University de Jaunpur*, Índia (VBS). Foram consideradas as décadas de 90 até 2010, são utilizadas variáveis demográficas, tais como sexo, renda familiar, escolaridade e ocupação dos pais, também os rendimentos acadêmicos no primeiro ano do curso. É analisado a classificação dos alunos em risco de evasão, através de alguns algoritmos de testes desenvolvidos, resultando em valores de acurácia acima de 60%.

# 2.3 Método PCA aplicado ao Problema de Evasão

Os estudos sobre a evasão acabaram se tornando essenciais, determinando muitas pesquisas na busca por métodos matemáticos e metodologia que representem um processo de resolução para o problema. Em (HUNT et al., 2002), intitulado "A comprehensive needs assessment to facilitate prevention of school drop out and violence", os autores abordaram a violência escolar e o abandono escolar, os dados foram coletados em cinco escolas em um distrito de uma pequena cidade no norte da Geórgia, EUA. Entrevistas com grupos e individuais serviram como base para a construção de um relatório contendo 78 itens, considerando 304 funcionários de escolas. É promovida a análise de componentes principais, a qual revelou cinco fatores distintos: vínculos escolares associados ao clima escolar positivo, causas da violência, causas do abandono escolar, intervenções para abandono escolar, e intervenções para violência. Um dos focos desta pesquisa foi descobrir se há coincidência entre as percepções dos entrevistados sobre as causas e as intervenções para abandono escolar e violência. Os resultados mostraram que, embora haja percepções desses problemas, também podem o abandono e a violência serem vistos como problemas distintos. Este trabalho foi intensamente explorado, buscando associar as ideias elaboradas na pesquisa com as contribuições deste trabalho. A análise de componentes principais é objeto para o desenvolvimento dos estudos computacionais propostos.

Em (HEGDE, 2016), intitulado "Dimensionality Reduction Technique for Developing Undergraduate Student Dropout Model using Principal Component Analysis through R Package", os autores descreveram que, os alunos do ensino superior, em sua maioria, evadem por atritos,

descritos como problemas comportamentais, oriundos da rigidez institucional, ou mesmo da condição da família, apesar de não ser tão considerável no problema. Essa retenção do acadêmico, segundo os autores, reflete no *ranking* global da universidade. Assim, o estudo proposto se concentrou em identificar o motivo do abandono escolar do curso superior, utilizando também a análise de componentes principais, o PCA. Os dados utilizados são oriundos de questionários separados por período da graduação, com base no desempenho acadêmico, perfil socioeconômico, motivação, relacionamento familiar, aceitação e outros. O estudo também destacou que é muito importante analisar e prever a evasão inicial de alunos em uma organização acadêmica. Da mesma forma que fora relatado no estudo anterior apresentado, os autores destacam o PCA usado para descrever atributos não correlacionados (componentes principais) extraídos do enorme conjunto de dados. Por fim, os autores demonstram que o conjunto de dados transformado pode ser usado para classificação e previsão, resultado muito importante na justificativa da aplicação do PCA neste trabalho.

Antes de encerrar esta seção, mais quatro artigos serão mencionados. Esta menção é considerada como um breve histórico para a parte teórica do método PCA. O primeiro deles, de Anderson (1963), é o mais teórico dos quatro. Ele discutiu as distribuições de amostragem assintótica dos coeficientes e variâncias dos PCs de amostra, com base no trabalho anterior de Girshick (1939), e foi frequentemente citado em desenvolvimentos teóricos subsequentes. O artigo PCA de variáveis instrumentais de Rao (1964) é notável pelo grande número de novas ideias sobre os usos, interpretações e extensões do PCA. Gower (1967) discutiu as ligações entre o PCA e várias outras técnicas estatísticas, e também forneceu uma série de *insights* geométricos importantes. Finalmente, Jeffers (1967) deu um impulso ao lado realmente prático do assunto, discutindo dois estudos de caso em que o uso de PCA vai além de uma simples ferramenta de redução de dimensão.

# 3 Análise de Componentes Principais

Análise de Componente Principal, do inglês *Principal Components Analysis* (PCA), é um método de redução de dimensionalidade usado em grandes conjuntos de dados, transformando um grande conjunto de variáveis em um menor, que ainda represente o conjunto em sua maior parte, com as variáveis mais relevantes (JOLLIFFE, 2002).

A técnica do componente principal foi descrita pela primeira vez por Pearson (1901) com o uso do ajuste no subespaço de nuvem de pontos. Posteriormente, Hotelling (1933) consolidou essa técnica em 1933 e em 1936 com o propósito específico de analisar estruturas relacionadas (XAVIER, 2000). No entanto, o uso da análise só foi amplamente utilizada após o desenvolvimento dos computadores eletrônicos. Atualmente, devido à grande disponibilidade de complexos recursos computacionais e softwares, essa tecnologia tem sido amplamente utilizada em diversos campos científicos (TYSKA, 2017).

Basicamente, a redução dá origem a um novo conjunto de variáveis, os componentes principais, os quais não estão correlacionados e são ordenados de modo que os primeiros retenham a maior parte da variabilidade. O cálculo dos componentes principais é dado por um problema de autovalor-autovetor para uma matriz simétrica definida semi-positiva (JOLLIFFE, 2002). Isto é, o PCA realiza uma mudança da base do espaço vetorial, utilizando a construção de uma matriz de covariância entre os dados estudados.

Sendo assim, este capítulo apresenta a descrição da padronização dos dados, a Matriz de covariância, as componentes principais, um exemplo de cálculo das componentes principais, o algoritmo do PCA e uma descrição matemática para todos os itens apresentados.

# 3.1 Padronização dos Dados

Antes de iniciar o método de PCA aplicado a alguma base de dados, é necessário observar se os dados estão padronizados, ou seja, não é possível que se tenha dados com grandes valores e dados com valores muito pequenos, pois a matriz covariância, definida nas próximas seções, poderá encontrar informações errôneas. Assim, a equação a seguir, mostra como normalizar esses dados a partir do cálculo do desvio padrão das variáveis.

$$z = \frac{valor - media}{desviopadr\~ao}. (3.1)$$

Uma vez feita a normalização, todas as variáveis estão transformadas na mesma escala.

### 3.2 Matriz de Covariância

O objetivo desta seção é entender como as variáveis do conjunto de dados de entrada estão variando da média entre si, ou seja, verificar se existe alguma relação entre elas. Se as variáveis são altamente correlacionadas, elas podem conter informações redundantes. Para identificar essas correlações, calcula-se a matriz de covariâncias.

A matriz de covariância X é uma matriz simétrica  $p \times p$ , onde p é a dimensão da matriz, que tem como entradas as covariâncias associadas com todos os possíveis pares de variáveis iniciais. Por exemplo, para um conjunto de dados tridimensionais com 3 variáveis x, y e z, a matriz de covariância é uma matriz  $3 \times 3$  (JOLLIFFE, 2002):

$$X = \begin{bmatrix} Cov_{xx} & Cov_{xy} & Cov_{xz} \\ Cov_{yx} & Cov_{yy} & Cov_{yz} \\ Cov_{zx} & Cov_{zy} & Cov_{zz} \end{bmatrix}$$

Uma vez que a covariância de uma variável consigo mesma, é a sua variância (Cov(a,a) = Var(a)), na diagonal principal (superior esquerdo para inferior direito) tem-se as variâncias de cada variável inicial. Ainda, considerando que a covariância é comutativa (Cov(a,b) = Cov(b,a)), as entradas da matriz de covariância são simétricas em relação à diagonal principal, o que significa que as matrizes triangulares superior e inferior são iguais (JOLLIFFE, 2002).

Observando a matriz o sinal da covariância, quando positivo, as duas variáveis aumentam ou diminuem juntas (correlacionadas). Caso contrário, um aumenta quando o outro diminui (inversamente correlacionado). Ainda, como a matriz da covariância é simétrica, sabe-se que sempre possível encontrar um conjunto de p autovalores e autovetores ortonormais. Portanto, vê-se a matriz de covariância como uma tabela que mostra todas as correlações entre todos os pares de variáveis possíveis.

# 3.3 Componentes Principais (CPs)

As componentes principais são novas variáveis construídas como combinações lineares ou "misturas das variáveis iniciais". Essas combinações são feitas de tal forma que as novas variáveis (ou seja, CP) não estão correlacionadas e a maioria das informações, dentro das variáveis iniciais, é espremida ou comprimida nos primeiros componentes. Por exemplo, se a base de dados possui dimensão igual a 10, ou seja, 10 variáveis de entrada, o cálculo da matriz de covariância vai fornecer 10 componentes principais nas primeiras 10 linhas da matriz, mas o PCA tenta colocar o máximo de informações possíveis nos primeiros componentes, como ilustra a Figura 1, um exemplo de porcentagem de variação para cada componente em um caso fictício.

A redução de informações em componentes principais permitirá a redução da dimensionalidade sem perder informações relevantes. Desta forma, descartam-se os componentes com pouca informação, determinando novas variáveis para a base de dados (SANTO et al., 2021).

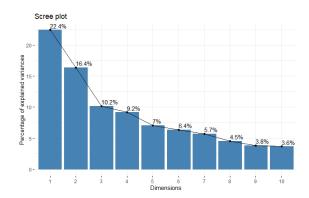


Figura 1 – Porcentagem de variação para cada CP - Fonte do autor.

Como existem tantos componentes principais quanto variáveis nos dados, os componentes principais são construídos de tal maneira que o primeiro componente principal é responsável pela maior variância possível no conjunto de dados.

Geometricamente, as principais componentes, ordenadas em uma matriz de covariância, representam eixos de um plano, ou do espaço, sendo que todas as variáveis originais serão apresentadas como linhas nesta representação geométrica, demonstrando a dispersão dos pontos em relação às componentes principais. Dada uma variável, quanto maior a dispersão da linha em direção a uma componente principal (eixo), mais informações esta componente principal carrega da variável dada.

# 3.4 Exemplo de Cálculo de Componentes Principais

Considere este exemplo descrito em (JOLLIFFE, 2002): Suponha um vetor  $\mathbf{x}$ , que contenha p variáveis aleatórias e que as variâncias das  $\mathbf{p}$  variáveis aleatórias e a estrutura das covariâncias ou correlações entre as  $\mathbf{p}$  variáveis são de interesse. Alternativamente, procuram-se algumas variáveis derivadas que preservam a maioria das informações fornecidas por essas variâncias e correlações ou covariâncias.

A PCA se concentra nas variâncias, não ignorando as covariâncias e correlações. A primeira etapa consiste em uma função linear  $\alpha_1^T \mathbf{x}$  de elementos de  $\mathbf{x}$  com variância máxima, dada por:

$$\alpha_1^T = \alpha_{11} \mathbf{x}_1 + \alpha_{12} \mathbf{x}_2 + \dots + \alpha_{1p} \mathbf{x}_p = \sum_{j=1}^p \alpha_{1j} \mathbf{x}_j$$
 (3.2)

onde  $\alpha_1$  é um vetor de tamanho p e constantes  $\alpha_{11}, \alpha_{12}, ..., \alpha_{1p}$ , T denota transposição (JOLLIFFE, 2002).

Na segunda etapa, procura-se uma função linear  $\alpha_2^T\mathbf{x}$ , não correlacionada com  $\alpha_1^T\mathbf{x}$  mas com variância máxima. Assim, segue-se analogamente, de modo que a k-ésima função linear  $\alpha_k^T\mathbf{x}$  seja encontrada, também com variância máxima sujeita a não estar correlacionada com  $\alpha_1\mathbf{x},\ \alpha_2\mathbf{x},\ \ldots,\ \alpha_{k-1}\mathbf{x}$ . A k-ésima variável encontrada,  $\alpha_k\mathbf{x}$  é a k-ésima CP. Contudo,

espera-se que a maior parte da variação em  ${\bf x}$  seja contabilizada por m CPs, onde  $m \ll p$ . Sendo assim, as Figuras 2 e 3 apresentam os gráficos gerados para 50 variáveis aleatórias, considerando as duas primeiras variáveis altamente correlacionadas e as duas componentes principais, respectivamente.

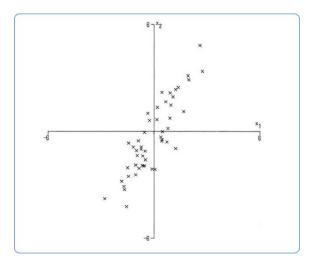


Figura 2 – Gráfico com 50 observações em relação as duas variáveis  $x_1$  e  $x_2$  (JOLLIFFE, 2002)

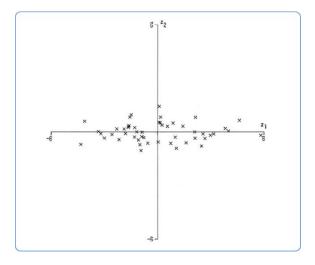


Figura 3 – Gráfico com 50 observações em relação as duas variáveis  $\mathbf{z}_1$  e  $\mathbf{z}_2$  (JOLLIFFE, 2002)

Nota-se pelas figuras apresentadas que, existe uma maior amplitude na direção de  $\mathbf{z}_1$  do que em qualquer uma das variáveis originais, mas pouquíssima variação na direção de  $\mathbf{z}_2$ . Isto mostra que, se um conjunto de dados tem correlações entre elas, então os primeiros CPs serão responsáveis pela maior parte da variação nas variáveis originais (JOLLIFFE, 2002).

Tendo os CPs definidos, precisa-se saber como encontrá-los. Considere, por enquanto, o caso em que o vetor de variáveis aleatórias  ${\bf x}$  tem uma matriz de covariância conhecida  $\Sigma$ . Esta é a matriz cujo (i,j)ésimo elemento é a covariância (conhecida) entre os i-ésimo elementos de  ${\bf x}$  quando i=j, e a variância do j-ésimo elemento de  ${\bf x}$  quando i=j. O caso mais realista, onde  $\Sigma$  é desconhecida, segue substituindo está por uma matriz de covariância

de amostra **S** . Acontece que para  $k=1,2,\ldots,p$ , o k-ésimo CP é dado por  $z_k=\alpha_k^Tx$ , onde  $\alpha_k$  é um autovetor de  $\Sigma$  correspondente ao seu k-ésimo maior autovalor  $\lambda_k$ . Além disso, se  $\alpha_k$  for escolhido para ter comprimento unitário ( $\alpha_k^T\alpha_k=1$ ), então  $var(z_k)=\lambda_k$ , onde  $var(z_k)$  denota a variância de  $z_k$ .

Antes de apresentar o resultado, dois tópicos devem ser explicados. Primeiramente, os autovalores e o autovetor são descritos juntamente com um exemplo, e então o método dos Multiplicadores de Lagrange será apresentado.

#### 3.4.1 Autovetores e Autovalores

Autovetores e autovalores são conceitos de álgebra linear, os quais são calculados a partir da matriz de covariância, a fim de determinar a **componente principal** para os dados originais apresentados. Estes serão em pares, de modo que cada autovetor tem um autovalor correspondente. E seu número é igual à dimensão dos dados.

Os autovetores e autovalores representam as direções dos eixos onde há mais variância, ou seja, mais informações, as quais define-se como Componentes Principais. E os autovalores são os coeficientes relacionados aos autovetores, que fornecem a quantidade de variância correspondente a cada componente principal (JOLLIFFE, 2002).

Classificando os autovetores em ordem de seus autovalores, do mais alto para o mais baixo, obtêm-se os componentes principais em ordem de significância, obtendo-se  $\lambda_1 \geq \lambda_2$ . A componente principal (CP1) é o autovetor associado ao autovalor  $\lambda_1$  e o que corresponde ao segundo componente (CP2) é o autovetor associado ao autovalor  $\lambda_2$ . Importante também considerar a propriedade que se refere à direção do autovetor, ou seja, os autovetores são perpendiculares (ortogonais) entre si. Essa propriedade é importante dada a construção de gráficos com as componentes principais.

Portanto, são encontrados dois pares de autovalores e autovetores. Os autovalores são calculados por algoritmos melhores do que encontrar as raízes das equações algébricas.

# 3.5 O Algoritmo PCA

O algoritmo para aplicação resume-se nos seguintes passos:

- Organização dos dados das medições em uma matriz  $n \times m$ , onde m é número de variáveis medidas, ou dimensões, e n é o número de elementos;
- Caso necessário, será dividido cada valor pelo seu desvio padrão para normalização e evitar a sensibilidade da PCA às diferenças entre estas variáveis;
- Calculo a matriz de covariância da matriz resultante dos itens anteriores, no caso, se realizado o passo 2, cálculo do desvio padrão, esta será a matriz de correlação;

- Cálculo dos Autovetores e Autovalores associados à matriz de covariância;
- Ordenação dos Autovetores de acordo com os Autovalores associados. Deste modo, o primeiro autovetor é a primeira componente principal, o segundo é a segunda componente principal, e assim, sucessivamente;
- Serão eliminadas as componentes principais de menor relevância. Para isso, serão definidos os percentuais da variância original que serão mantidos. Assim, serão escolhidas as componentes principais de modo que soma dos autovalores associados seja maior ou igual a este percentual.

### 3.6 Descrição Matemática

Os dados são armazenados em uma matriz, em que cada amostra é uma observação, cada coluna é uma variável  $[x_1,...,x_m]$ , sendo a última coluna a variável resposta, y:

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & y_{1n} \\ x_{21} & x_{22} & \dots & y_{2n} \\ \vdots & \vdots & & \vdots \\ x_{m1} & x_{m2} & \dots & y_{mn} \end{bmatrix}_{m \times n}$$

Seja uma matriz de dados  $X_{(m \times n)}$ , onde cada uma das n linhas representa uma observação e cada uma das m colunas representa uma variável, com vetor de valores médios  $\mu$  e matriz de variâncias covariâncias  $\Sigma$ . Pretende-se encontrar um novo conjunto de variáveis  $Y_1, Y_2, ..., Y_m$ , não correlacionadas entre si e sendo estas ordenadas pela máxima variância:

Figura 4 – ACP - Fonte: (VICINI, 2005)

Os vetores que representam a máxima variância possuem uma dependência linear (ortogonais), observa-se a nuvem dos dados, representado pela elipsoide a seguir.

O primeiro CP, possui maior variância, sendo o eixo maior. O segundo maior, de menor variância, CP2 é ortogonal ao eixo maior.

As componentes principais são os Autovetores associados aos autovalores. Afim de facilitar as comparações entres os CPs, usualmente, normaliza-se os Autovetores.

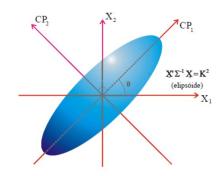


Figura 5 - ACP Fonte: (VICINI, 2005)

A componentes principais são obtidas de uma forma geral, é um conjunto de p variáveis  $X_1, X_2, \ldots, X_p$  com médias  $\mu_1, \mu_2, \ldots, \mu_p$  e variância  $\sigma_1, \sigma_2, \ldots, \sigma_p$ . Estas variáveis não são linearmente independentes, portanto possuem covariância entre a i-ésima e k-ésima variável definida por  $\sigma_{ik}$ , para  $i \neq k = 1, 2, \ldots, p$ . As variáveis podem ser expressas na forma vetorial por:  $X = [X_1, X_2, \ldots X_p]$ , com vetor das médias  $\mu = [\mu_1, \mu_2, \ldots, \mu_p]$  e matriz covariante  $\Sigma$  (HONGYU; SANDANIELO; JUNIOR, 2016).

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2p} \\ \vdots & \vdots & & \vdots \\ \sigma_{n1} & \sigma_{n2} & \dots & \sigma_{nn} \end{bmatrix}_{n \times r}$$

Os pares de autovalores e autovetores são encontrados,  $(\lambda_1,\alpha_1),(\lambda_2,\alpha_2),\ldots,(\lambda_p,\alpha_p)$  em que são ordenados do maior para o menor,  $\lambda_1 \geq \lambda_2 \geq \lambda_p$ , sendo  $\lambda_1$  o de máxima variância, associados a  $\Sigma$  e então o i-ésimo componente principal é definido (HONGYU; SANDANIELO; JUNIOR, 2016).

A redução dos dados pode ser feita por Singular Value Decomposition (SVD) (WALL; RECHTSTEINER; ROCHA, 2003) ou Empirical Orthogonal Functions (EOF) (SMITH et al., 1996). Busca-se planos ortogonais entre as variáveis que maximize a variança, neste estudo é utilizado o pacote do R PCA Methods SVD (STACKLIES et al., 2007).

Os vetores são representados em um novo plano ortogonal, saindo do ponto (0,0) até x máximo e y máximo.

Sendo a equação  $\Sigma \alpha_1 = \lambda_1 \alpha_1$ , o maior autovalor  $\alpha_2$  é escolhido como um autovetor de  $\Sigma$  que corresponde ao segundo maior auto valor de  $\alpha_2^T X$ .

Este processo para k=1...p, resultando diferentes autovetores de  $\Sigma$  junto aos autovalores  $\lambda_1, \lambda_2, ..., \lambda_p$ . A variação de cada CP é dada por:

$$Var(\alpha_k^T X) = \lambda_k.$$

A transformação linear ortogonal é dada por  $1 \leq q \leq p$ , para um inteiro q:  $y = B^T X$ , onde y é um vetor de q-elementos,  $B^T$  é matriz  $q \times p$  e  $\Sigma y = B^T \Sigma B$  é a matriz de variância-covariância para y (JOLLIFFE, 2002).

As projeções dos vetores não estão correlacionadas, o percentual de variância é retido nas primeiras q CPs, sendo:

$$\frac{\sum_{k=1}^{q} \lambda k}{\sum_{k=1}^{p} \lambda k} .100 \tag{3.3}$$

Sendo a matriz de covariância da amostra:

$$S = \frac{1}{n-1} X^T X,$$

onde X é uma matriz  $(n \times p)$  com a média zero.

sendo gerada uma matriz A, combinando p autovetores de S (ou autovetores de  $X^TX$ ), definida então a matriz de CP (scores) (JOLLIFFE, 2002),

$$Z = XA$$
.

• Obter a média e centralizar os dados:

$$\overline{\mu} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

• Calcular a matriz de covariância:

$$C = \frac{1}{n-1}A^T A$$

• Decompor a matriz de covariância em autovalores e autovetores

 Gerar uma matriz de projeção, com os k autovetores correspondentes ao k maiores autovalores

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}_{n \times n}$$

 Projeção dos dados originais no novo espaço com k dimensões

Neste capítulo foi descrição da padronização da matemática da PCA e o algorítimo aplicado ao estudo.

# 4 Implementação do PCA

A análise de PCA foi implementada no Software estatístico **R** (R Core Team, 2021). Há uma infinidade de algoritmos que simulam o método PCA, será utilizado um exemplo do funcionamento, sendo este nativo do R para estudos.

Existem diversos sites contendo informações para utilização do PCA no R, documentação da própria ferramenta, além de exemplos encontrados no google (KULKARNI, 2021).

Primeiramente a análise será utilizada a biblioteca "PRCOMP"4.1 nativa do R e a seguir a biblioteca "pcaMethods", um pacote que traz uma infinidade de recursos para nosso estudo (BIOCONDUCTOR, 21).

# 4.1 PCA do conjunto de dados do vinho

Para a PCA apresentada em dados reais, está disponibilizado um conjunto de dados de vinho no  $UCIMachineLearningRepository^1$  que pode servir como um bom exemplo inicial.

```
wine <- read.table("http://archive.ics.uci.edu/ml/machine-learning-
databases/wine/wine.data", sep=",")</pre>
```

Listing 4.1 – Código fonte em R

De acordo com a documentação<sup>2</sup>, esses dados consistem em 13 parâmetros físicoquímicos medidos em 178 amostras de vinho de três culturas distintas cultivadas na Itália. Verificam-se os padrões em pares de variáveis e, em seguida, ver o que um PCA faz sobre isso plotando PC1 contra PC2.

```
# Variaveis nomeadas
colnames(wine) <- c("Cvs","Alcohol","Malic acid","Ash","Alcalinity of ash",
    "Magnesium", "Total phenols", "Flavanoids", "Nonflavanoid phenols", "
    Proanthocyanins", "Color intensity", "Hue", "OD280/OD315 of diluted
    wines", "Proline")

# A primeira coluna corresponde as classes
wineClasses <- factor(wine$Cvs)

# Pareamento
pairs(wine[,-1], col = wineClasses, upper.panel = NULL, pch = 16, cex =
    0.5)</pre>
```

<sup>&</sup>lt;sup>1</sup>https://archive.ics.uci.edu/ml/index.php

<sup>&</sup>lt;sup>2</sup>https://archive.ics.uci.edu/ml/datasets/Wine

Listing 4.2 - Código fonte em R

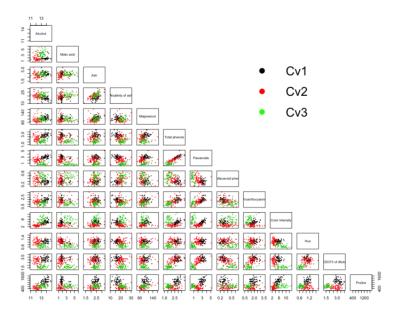


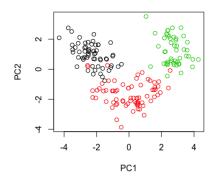
Figura 6 – Classes

Entre outras coisas, observam-se correlações entre variáveis (por exemplo, fenóis totais e flavonoides) e, ocasionalmente, a separação bidimensional dos três cultivares (por exemplo, usando % de álcool e a razão OD).

Se já é difícil olhar para todas as interações de pares em um conjunto de 13 variáveis, muito menos em conjuntos de centenas ou milhares de variáveis. Nestes casos, o PCA é de grande ajuda. Tenta-se neste conjunto de dados:

```
dev.off() # clear the format from the previous plot
winePCA <- prcomp(scale(wine[,-1]))
plot(winePCA$x[,1:2], col = wineClasses)</pre>
```

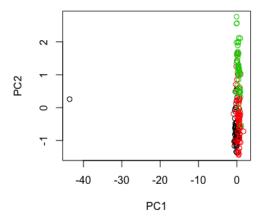
Listing 4.3 - Código fonte em R



Três linhas de código, e vê-se uma separação clara entre os cultivares de videira. Além disso, os pontos de dados estão espalhados uniformemente em intervalos relativamente estreitos em ambos os PCs.

```
wineOutlier <- wine
wineOutlier[10,] <- wineOutlier[10,]*10 # change the 10th obs. into an
    extreme one by multiplying its profile by 10
outlierPCA <- prcomp(scale(wineOutlier[,-1]))
plot(outlierPCA$x[,1:2], col = wineClasses)</pre>
```

Listing 4.4 – Código fonte em R



Como esperado, a grande variação decorrente da separação da décima observação do núcleo de todas as outras amostras são totalmente absorvidas pelo primeiro PC. A amostra periférica torna-se evidente.

# 4.2 PCA com pcaMethods

Agora pode-se voltar para pcaMethods, um pacote compacto de ferramentas PCA.

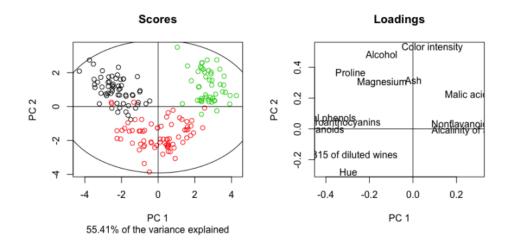
```
#instalacao do pacote a partir do repositorio
source("https://bioconductor.org/biocLite.R")
biocLite("pcaMethods")
library(pcaMethods)
```

Listing 4.5 - Código fonte em R

```
winePCAmethods \leftarrow pca(wine[,-1], scale = "uv", center = T, nPcs = 2, method = "svd")

slplot(winePCAmethods, scoresLoadings = c(T,T), scol = wineClasses)
```

Listing 4.6 - Código fonte em R



Portanto, tem-se uma reprodução fiel dos dados do PCA anterior. Então, tendo o painel de cargas do lado direito, pode-se verificar que:

- O vinho de Cv2 (tinto) tem uma intensidade de cor mais clara, menor% de álcool, uma maior proporção de DO e matiz, em comparação com o vinho de Cv1 e Cv3.
- O vinho de Cv3 (verde) tem um maior teor de ácido málico e fenóis não flavonoides, e uma alcalinidade de cinzas mais alta em comparação com o vinho de Cv1 (preto)

Finalmente, embora a variação explicada em conjunto pelos dois primeiros PCs seja impressa por padrão (55,41%), pode ser mais informativo consultar a variação explicada em PCs individuais. Pode-se chamar a estrutura de métodos winePCA, inspecionar os slots e imprimir os de interesse, já que há muita informação contida. A variação explicada por componente é armazenada em um slot denominado R2.

```
str(winePCAmethods) # slots are marked with @ winePCAmethods@R2
```

Listing 4.7 - Código fonte em R

Aparentemente, PC1 e PC2 explicam 36,2% e 19,2% da variância no conjunto de dados do vinho, respectivamente.

# 5 Estudos Computacionais

O PCA tem a capacidade de projetar as observações descritas por p variáveis em alguns componentes ortogonais definidos onde os dados 'se estendem' mais, renderizando uma visão geral simplificada. O PCA é particularmente poderoso para lidar com multicolinearidade e variáveis que superam as amostras (p > n). É um método não supervisionado, o que significa que sempre analisará as maiores fontes de variação, independentemente da estrutura de dados.

A partir do estudo bibliográfico realizado para este trabalho, uma base de dados para a uma análise por PCA, visando a predição ou identificação de estudantes propensos a evadir, é constituída por características que compreendem fatores demográficos, fatores internos e externos à universidade. Neste trabalho, os dados coletados para o estudo foram obtidos através de um questionário socioeconômico composto por 40 perguntas, as quais são aplicadas no ato da inscrição para os vestibulares. A base utilizada compreende os anos de 2014 a 2020, referentes ao CPP-CEM. Em março de 2021, a Coordenação do Curso de Licenciatura em Ciências Exatas fez uma primeira coleta de dados para identificar o status dos alunos (cancelado ou não), utilizando um sistema da UFPR para gerar um relatório com apenas as matrículas, sempre preservando o anonimato dos alunos.

A primeira tarefa a ser realizada no trabalho foi a categorização desses campos quanto ao tema, abrangência e qualidade informacional para a atividade de interesse na pesquisa, de forma a identificar informações redundantes, complementares ou potencialmente interagentes.

Antes da aplicação do PCA, foram realizadas algumas análises exploratórias de dados, como estatísticas descritivas, matriz de correlação e matriz de gráfico de dispersão.

Decide-se então utilizar o PCA para reduzir os elementos do questionário em indicadores ou variáveis latentes, passíveis de interpretação, capazes de sintetizar grande parte da diversidade das variáveis em si.

# 5.1 Experimento I

O principal objetivo deste experimento é uma análise completa e sintetizada do método PCA aplicado à base de dados estudada de forma completa. Ainda, o experimento busca mostrar que é possível utilizar o PCA para esse conjunto de dados estudados, reduzindo a dimensionalidade e mantendo aproximadamente 17,70% da variância original, ou seja, os 40 indicadores originais serão substituídos ou representados em duas novas dimensões, aqui denominadas CP1 e CP2.

# 5.1.1 Implementação e Resultados

Os procedimentos adotados neste experimento são divididos em duas etapas:

#### i- Coleta e tratamento dos dados:

Os dados foram sintetizados, reunidos e completados à medida que a falta de informações fora identificada. Foi elaborada apenas uma planilha com os dados sintetizados, e uma nova informação foi inserida com o auxílio de indicadores acadêmicos coletados pela Coordenação do Curso. Tais dados mostram se o aluno analisado está com matrícula ativa ou evadiu;

#### ii- Aplicação das técnicas estatísticas:

Dado o tratamento dos dados, o PCA foi aplicado, apresentando-se a matriz de covariância, os autovalores e autovetores calculados, a proporção destes autovalores e autovetores, destacando a porcentagem de variância dos autovalores e a contribuição de cada uma das questões quanto às duas principais componentes selecionadas.

A Tabela 1 apresenta cada uma das questões, denotadas como componentes principais (CPs), além dos autovalores, a proporção dos autovalores calculada, e a proporção acumulada. Pode-se notar na Tabela 1 42 CPs, isto é, existem duas componentes principais a mais do que o número de questões disponibilizadas na base de dados. Isto é devido ao fato descrito no Item (i), onde duas novas componentes foram inseridas, o Índice de Rendimento Acadêmico (IRA) e a condição do aluno (matriculado ou não).

Tabela 1 – Componentes principais (CPs), autovalores  $\lambda_i$  e porcentagem da variância explicada e proporção acumulada (%) pelos componentes

| Componente Principal | Autovalores | Proporção | Proporção Acumulada |
|----------------------|-------------|-----------|---------------------|
| CP 1                 | 4,245987    | 10,109492 | 10,109492           |
| CP 2                 | 3,185041    | 7,583431  | 17,692923           |
| CP 3                 | 2,750645    | 6,549155  | 24,242078           |
| CP 4                 | 2,374023    | 5,652436  | 29,894514           |
| CP 5                 | 2,302836    | 5,482944  | 35,377458           |
| CP 6                 | 2,000944    | 4,764153  | 40,141611           |
| CP 7                 | 1,917535    | 4,565559  | 44,707169           |
| CP 8                 | 0,18885     | 4,496421  | 49,20359            |
| CP 9                 | 1,672718    | 3,982661  | 53,186251           |
| CP 10                | 1,566667    | 3,730159  | 56,91641            |
| CP 11                | 0,14885     | 3,544056  | 60,460466           |
| CP 12                | 1,366717    | 3,254089  | 63,714555           |
| CP 13                | 1,280612    | 3,049077  | 66,763632           |
| CP 14                | 1,200993    | 2,859508  | 69,62314            |

Continua na próxima página

Tabela 1 – Continuação da tabela

| Componente Principal | Autovalores | Proporção | Proporção Acumulada |
|----------------------|-------------|-----------|---------------------|
| CP 15                | 1,129098    | 2,688329  | 72,311469           |
| CP 16                | 1,059571    | 2,522789  | 74,834258           |
| CP 17                | 1,031084    | 2,454961  | 77,289219           |
| CP 18                | 0,943618    | 2,24671   | 79,535929           |
| CP 19                | 0,835545    | 1,989393  | 81,525322           |
| CP 20                | 0,811697    | 1,932613  | 83,457935           |
| CP 21                | 0,735258    | 1,750615  | 85,20855            |
| CP 22                | 0,678944    | 1,616533  | 86,825083           |
| CP 23                | 0,610233    | 1,452936  | 88,278019           |
| CP 24                | 0,580724    | 1,382677  | 89,660696           |
| CP 25                | 0,519915    | 1,237892  | 90,898589           |
| CP 26                | 0,496159    | 1,181331  | 92,079919           |
| CP 27                | 0,468524    | 1,115534  | 93,195453           |
| CP 28                | 0,373267    | 0,888732  | 94,084185           |
| CP 29                | 0,349516    | 0,832182  | 94,916367           |
| CP 30                | 0,32743     | 0,779595  | 95,695961           |
| CP 31                | 0,279553    | 0,665603  | 96,361564           |
| CP 32                | 0,255521    | 0,608383  | 96,969948           |
| CP 33                | 0,226202    | 0,538576  | 97,508524           |
| CP 34                | 0,194489    | 0,463068  | 97,971592           |
| CP 35                | 0,191479    | 0,455903  | 98,427495           |
| CP 36                | 0,163615    | 0,389559  | 98,817053           |
| CP 37                | 0,129449    | 0,308211  | 99,125264           |
| CP 38                | 0,108171    | 0,257551  | 99,382815           |
| CP 39                | 0,09747     | 0,232071  | 99,614886           |
| CP 40                | 0,07705     | 0,183452  | 99,798337           |
| CP 41                | 0,043537    | 0,103659  | 99,901997           |
| CP 42                | 0,041161    | 0,098003  | 100,000000          |

Fim da tabela

Com base nos resultados apresentados na Tabela 1, os dois primeiros CPs foram responsáveis por 17,70% da variação total, sendo o CP1 foi responsável por 10,10% e o segundo, CP2, por 7,58% das variações dos dados. Neste primeiro momento, sem um ajuste dos dados, não se tem a garantia de toda representatividade das informações, ou seja, teria a necessidade de mais CPs.

Analisando a porcentagem de variância, é produzido um gráfico que ilustra em barras a contribuição de cada autovalor calculado. A Figura 7 ilustra esta análise, onde os valores próprios são grandes para o CP1 e pequenos para os CPs subsequentes.

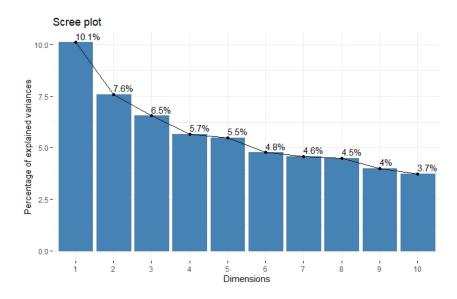


Figura 7 - Porcentagem de variância - Experimento I

Ainda analisando a Figura 7, 58% das informações (variações) contidas nos dados são retidas pelos dez primeiros componentes principais, isto mostra a base de dados é extremamente complexa para o estudo proposto.

Um CP com um autovalor calculado > 1 indica que o mesmo é responsável por uma maior variância do que ele mesmo representa nos dados ainda padronizados. Isso é comumente usado como um ponto de corte para determinar o número de CPs a serem retidos. Neste contexto, foi conduzido um novo experimento visando uma análise do PCA em relação a redução do número de questões, buscando uma maior representatividade para as duas principais componentes, dado que as variáveis que são correlacionadas com CP1 e CP2 são as mais importantes para explicar a variabilidade no conjunto de dados.

# 5.2 Experimento II

Neste novo experimento, é proposta a redução do número de variáveis no estudo inicial, filtrando a contribuição de cada uma com fator inferior à 3.5% de representatividade, gerando uma nova base de dados composta por 16 variáveis, conforme apresenta a Tabela 2. Essas são as perguntas que mais correspondem a quase 75% da representatividade acumulada.

Tabela 2 – Variáveis com maior representatividade

| Pergunta | Descrição  |
|----------|--|
| 2        | Quantos anos você completará em 2020?                              |
| 7        | Qual o nível de instrução do seu pai?                              |
| 8        | Qual o nível de instrução da sua mãe?                              |
| 10       | Qual a principal ocupação da sua mãe?                              |
| 11       | Você trabalha atualmente em atividade remunerada?                  |
| 12       | Qual é a sua participação na vida econômica de seu grupo familiar? |
| 14       | Quantas pessoas contribuem para a obtenção da renda familiar?      |
| 16       | Com que idade você começou a exercer atividade remunerada?         |
| 17       | Durante o curso, você terá obrigatoriamente que trabalhar?         |
| 18       | Como fez seus estudos do ensino fundamental e ensino médio?        |
| 19       | Em que ano você concluiu (ou concluirá) o curso de Ensino Médio?   |
| 23       | Você já fez o vestibular em outros anos?                           |
| 24       | Considerando sua capacidade de leitura em Língua Estrangeira,      |
|          | em que situação você melhor se enquadra?                           |
| 25       | Você já iniciou algum curso superior?                              |
| 26       | Se já iniciou um curso universitário, qual sua principal razão     |
|          | para participar deste Processo Seletivo?                           |
| 35       | A sua cor ou raça é  |

Fim da tabela

# 5.2.1 Implementação e Resultados

Os procedimentos adotados neste experimento são divididos em três etapas:

#### i- Execução do Experimento I:

A partir os resultados obtidos no Experimento I, uma análise dos dados é conduzida e uma redução na base de dados é identificada;

#### ii- Redução do número de questões:

Identificou-se as componentes com fator de contribuição menor que 3,5%, efetuou-se um corte na base de dados, dado um processo algorítmico desenvolvido na Linguagem R, em conjunto com a aplicação do PCA;

#### iii- Nova aplicação das técnicas estatísticas:

Dada a redução do conjunto de dados, o PCA foi novamente aplicado, apresentando-se a matriz de covariância, os autovalores e autovetores calculados, a proporção destes

autovalores e autovetores, destacando a porcentagem de variância dos autovalores e a contribuição de cada uma das questões quanto às duas principais componentes selecionadas.

Após o ajuste para às variáveis de maior representatividade, a Tabela 3) mostra que as duas primeiras componentes aumentaram sua representatividade, passando para 38,83% das informações retidas, determinando uma maior robustez para a análise dos dados.

Tabela 3 – Componentes principais (CPs), autovalores  $\lambda_i$  e porcentagem da variância explicada e proporção acumulada (%) pelos componentes

| Componente Principal | Autovalores | Proporção | Proporção Acumulada |
|----------------------|-------------|-----------|---------------------|
| CP 1                 | 3,591423    | 22,446397 | 22,446397           |
| CP 2                 | 2,621635    | 16,38522  | 38,831617           |
| CP 3                 | 1,625246    | 10,157787 | 48,989404           |
| CP 4                 | 1,465015    | 9,156345  | 58,145749           |
| CP 5                 | 0,112476    | 7,029752  | 65,175501           |
| CP 6                 | 1,018026    | 6,362663  | 71,538164           |
| CP 7                 | 0,908683    | 5,679267  | 77,217431           |
| CP 8                 | 0,721043    | 4,506522  | 81,723952           |
| CP 9                 | 0,606384    | 3,7899    | 85,513853           |
| CP 10                | 0,583857    | 3,649108  | 89,162961           |
| CP 11                | 0,461246    | 2,882786  | 92,045747           |
| CP 12                | 0,416767    | 2,604794  | 94,650541           |
| CP 13                | 0,331419    | 2,071367  | 96,721908           |
| CP 14                | 0,263976    | 1,64985   | 98,371759           |
| CP 15                | 0,146243    | 0,914021  | 99,285779           |
| CP 16                | 0,114275    | 0,714221  | 100                 |

Fim da tabela

Cabe ainda notar na Tabela 3 que a representatividade atinge os 75% já na sétima questão, viabilizando uma análise mais simples em relação às perguntas que podem estar ligadas ao processo de evasão universitária. Esta informação pode ser melhor visualizada pelo gráfico ilustrado na Figura 8.

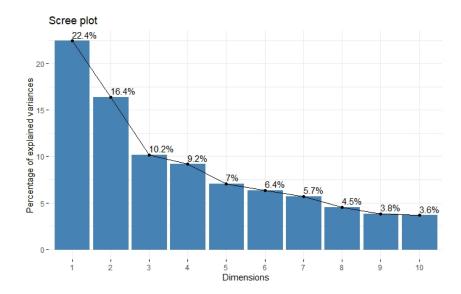


Figura 8 - Porcentagem de variância - Experimento II

A Figura 8 apresenta um decaimento bem menor a partir da quinta questão, mostrando que as quatro primeiras questões em nível de representatividade podem já garantir uma boa análise sobre a base de dados estudada.

Neste estudo, é perceptível o efeito que as variâncias dos caracteres originais exercem na determinação das CPs, especialmente neste caso em que o PCA foi aplicado a uma redução da base de dados. Pode-se notar que a CP1, vide Figura 9), possui grande representatividade pela variável 12 e 11 (Qual é a sua participação na vida econômica de seu grupo familiar? e Você trabalha atualmente em atividade remunerada?), cuja variância representa quase 30%. Em relação à CP2, é possível notar uma grande influência da variável 8 e 7 (Qual o nível de instrução da sua mãe? e Qual o nível de instrução de seu pai?), representando mais de 35%. Analisando em um aspecto temático, é possível notar a forte influência da variável participação na vida econômica de seu grupo familiar.

Ademais, pode-se analisar também a contribuição em conjunto das duas principais componentes, e assim produzir mais um gráfico ilustrado na Figura 11.

Torna-se explícita a influência de uma questão que não estava classificada com muita importância em uma análise separada, como é o caso da questão 19 (Em que ano você concluiu (ou concluirá) o curso de Ensino Médio?). Alocando em conjunto as variáveis mencionadas, questões 12, 11, 8, 7 e 19, pode-se assumir que estas cinco questões podem ser determinantes na análise produzida pelo PCA. Contudo, esta não é uma tarefa tão fácil, uma vez que não estão dispostas as inserções de dados, ou seja, não está demonstrada a relação das componentes principais com a finalidade do estudo, o fato do aluno estar matriculado ou não. Uma forma simples de examinar esta informação, e relacionar as questões com o aluno evadido, é dada por meio de seus "escores", representando graficamente os valores de saída, ordenados pelas primeiras CPs. Esta estratégia determina a elaboração do Experimento III.

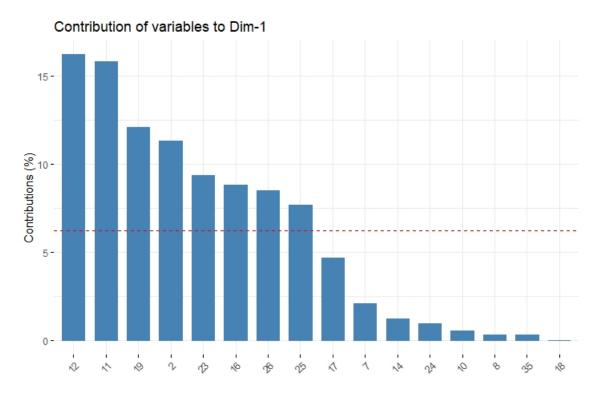


Figura 9 - Contribuições de variáveis no PC1

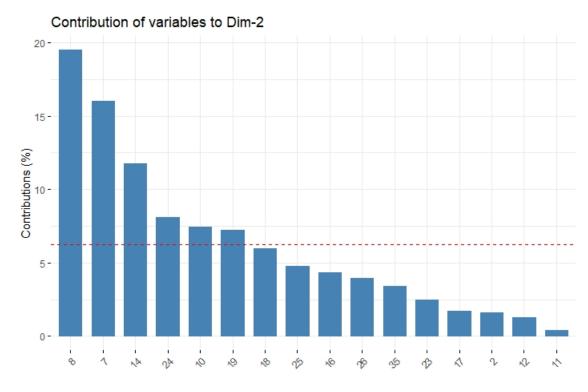


Figura 10 - Contribuições de variáveis no PC2

## 5.3 Experimento III

Este experimento busca ilustrar a relação das variáveis referentes às questões e as variáveis referentes aos alunos, representando-as em um plano cartesiano dado pelas compo-

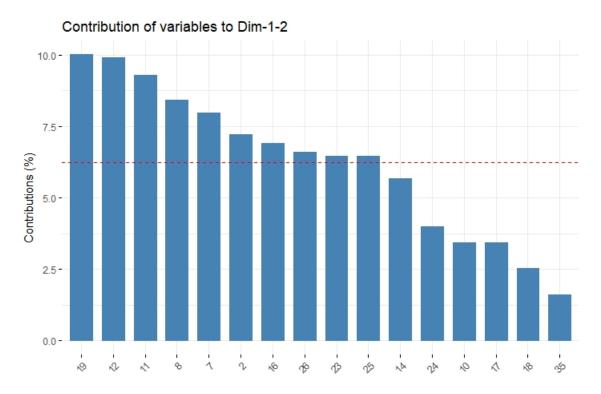


Figura 11 – Contribuição total nos PC1 e PC2

nentes CP1 e CP2 provenientes do Experimento II. O principal objetivo é conseguir, ao final desta análise, demonstrar quais questões do banco de dados reduzido estão relacionadas ao processo de evasão estudado.

### 5.3.1 Implementação e Resultados

Os procedimentos adotados neste experimento são divididos em três etapas:

#### i- Execução do Experimento II:

A partir a execução do PCA no Experimento II, os autovalores e autovetores calculados são guardados para a elaboração dos gráficos;

#### ii- Elaboração de gráficos das variáveis da base de dados:

São elaborados dois gráficos, buscando produzir uma análise separada dos pesos das respostas de cada aluno e das variáveis quanto às componentes principais.

#### iii- Análise conjunta das questões e condição do estudante:

Um gráfico é feito, determinando uma análise em conjunto dos dados apresentados na etapa anterior. Porém, é ilustrada a relação entre os pesos das questões e a condição do estudante.

Após os resultados colhidos no experimento anterior, um gráfico mostrando as distribuições dos estudantes em relação as componentes principais são ilustradas na Figura 12. Esta

figura mostra no plano cartesiano, o peso do questionário de cada um dos alunos em relação às duas componentes principais, CP1 e CP2.

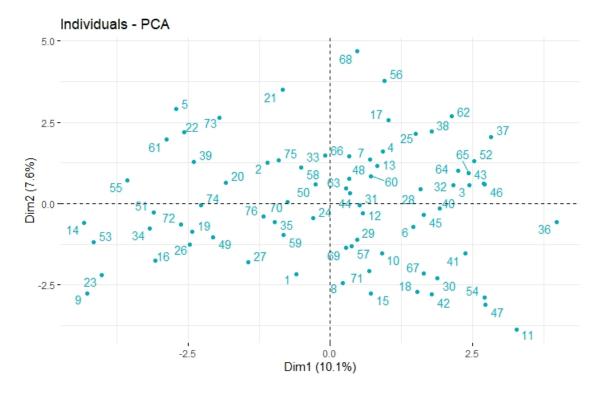


Figura 12 — Gráfico dos pesos de cada questionário da base de dados em relação às componentes principais

Partindo deste princípio, é elaborado um mapa para as variáveis, representado na Figura 13. As retas mostradas indicam os ângulos formados pela posição de cada variável em relação à origem dos eixos das componentes principais. Quanto menor for o ângulo formado pelas retas para um par de variáveis, maior é a semelhança entre o comportamento dessas variáveis, mostrando uma maior correlação linear. Em contraste, ângulos próximos de 180° demostram uma a correlação linear inversa, tendo as variáveis investigadas coeficientes de correlação negativos (BRAGA; ALBUQUERQUE, ).

Através do mapa, busca-se identificar melhor as correlações entre as variáveis, quanto mais próxima uma variável for do círculo de correlações, melhor sua representação no mapa (e mais importante é a variável para a interpretação desses componentes)(LEONI, 2021). As variáveis próximas ao centro do gráfico são menos importantes para os primeiros componentes. Portanto, a Figura 13 apresenta este mapa, destacando as variáveis com maior representatividade quanto às componentes principais CP1 e CP2, incluindo o rendimento acadêmico e a condição dos alunos.

A Figura 13 mostra que o rendimento acadêmico pouco influenciou nas componentes principais calculadas, mesmo após a redução da base de dados, e assim, foi excluído das análises feitas neste estudo computacional. Além disso, o fator de condição dos alunos, plotado como variável, também não teve grande influência quanto à representatividade, mas será

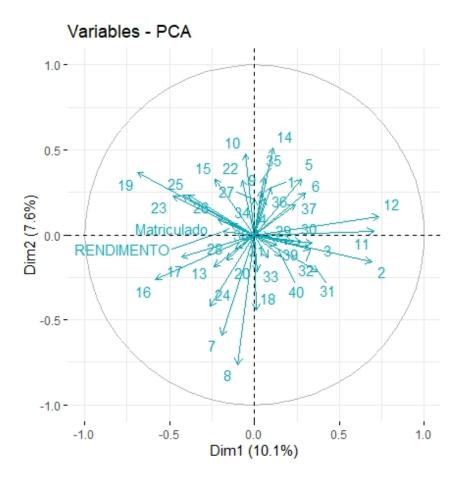


Figura 13 – Plotagem Variáveis PCA

essencial na condução da próxima análise, sendo representado apenas como um atributo para cada questionário estudado. Para isso, um gráfico de dispersão dos escores dos dois primeiros componentes baseados na matriz de correlações juntamente com os respectivos autovetores é elaborado, mostrando a forte influência das variáveis nas componentes principais CP1 e CP2. Este gráfico pode ser visto na Figura 14, sendo classificado como um *Biplot*, uma representação exploratória que generaliza um gráfico de dispersão de duas variáveis, ou seja, é um gráfico que mostra duas informações em uma única exibição.

Nesta Figura 14, pode-se observar que foi considerado a base de dados completa, proveniente do Experimento I, composta com 40 variáveis. Além disso, utilizou-se a variável referente à condição do aluno como um atributo de cada questionário apresentado pelos alunos, definida como 0, se não está matriculado e 1, caso contrário. Sendo assim, é visto que as questões 2, 11 e 12 estão relativamente próximas, refletindo sua correlação positiva relativamente grande, e visualmente relacionadas à evasão dos alunos. Já as questões 16 e 19 estão em quadrantes diferentes com correlação negativa pequena apenas em relação a segunda componente principal, mas com boa representatividade. Cabe também notar, que as questões 16 e 19 estão relacionadas aos alunos ainda matriculados.

Apesar da fácil visualização feita, nota-se que as questões 7 e 8, com grande repre-

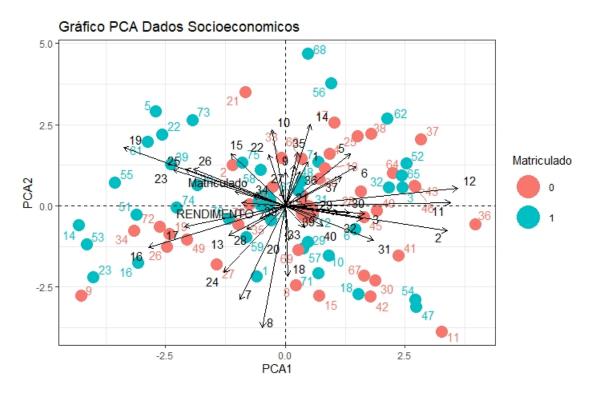


Figura 14 – Dispersão Componentes Principais de maior representatividade referente ao Experimento I

sentatividade no Experimento II, ficaram com correlação positiva alta, mas muito difíceis de visualizar sua relação com algum dos atributos. Portanto, um novo gráfico deve ser elaborado, explorando a redução da base de dados, ilustrado pela Figura 15.

A Figura 15 mostra claramente a redução das questões, sem reduzir o número de questionários, e ainda, mostra de forma mais clara quais questões estão relacionadas à evasão dos alunos, considerando que os evadidos ocupam as posições à direita do gráfico, e os matriculados estão à esquerda. As questões 16, 19, 23, 25, 26 estão visualmente relacionadas aos alunos ainda matriculados, da mesma forma que 2, 11 e 12 estão relacionadas aos alunos evadidos, mas esta foi uma análise possível de ser realizada sem a redução da base de dados. Porém, agora também é possível visualizar a relação das questões 7 e 8 com os alunos com matriculas ativas, determinada essencialmente pela primeira componente principal. Uma última observação mostra que a questão 14 pode ser considerada relevante para o estudo, pois tem uma representação considerável e está relacionada com a evasão.

A análise mostrou que existem correlações positivas entre as variáveis, que pode ser interpretado como uma significativa contribuição para não evasão do indivíduo ou não, e assim foi elaborado um Experimento IV, buscando explorar mais as particularidades.

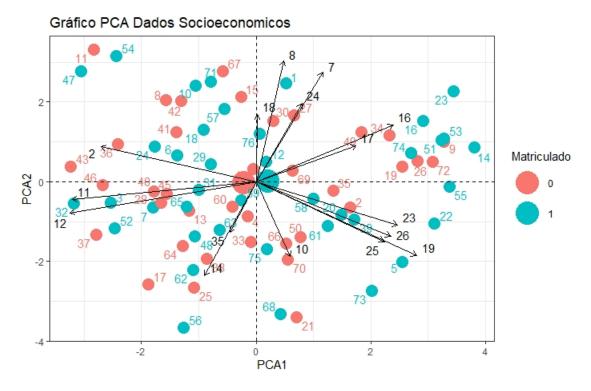


Figura 15 – Dispersão Componentes Principais de maior representatividade referente ao Experimento II

### 5.4 Experimento IV

Nesta seção, avalia-se a metodologia proposta que vem da aplicação do PCA no experimento III. Busca-se apresentar as questões relacionadas ao processo de evasão, e justificar a aplicação do PCA, mostrando que a análise produzirá uma base de estudos sólida para uma futura continuidade da pesquisa.

#### 5.4.1 Implementação e Resultados

Os procedimentos adotados neste estudo são divididos em duas etapas:

#### i- Execução do Experimento III:

A partir os resultados obtidos no Experimento III, uma análise dos dados é conduzida buscando a identificação;

#### ii- Nova aplicação das técnicas estatísticas:

Com o intuito de se entender a importância de cada variável na construção deste PCA reduzido em 16 CPs, busca-se compreender os comportamentos dos indivíduos na análise exploratória das questões abordadas no final do Experimento III.

Pretende-se através do Experimento IV, uma maior explicitação e familiaridade com as questões apontadas no Experimento III, que forneçam elementos para uma investigação de aprofundamento, apontando dificuldades e consequências para evasão.

A relação entre as variáveis elencadas nos experimentos anteriores e os indivíduos analisados conforme sua condição junto ao curso é sintetizada e apresentada nos gráficos abaixo. Nestes gráficos são adotadas a cor azul, para alunos com matrícula ativa, e a cor vermelha, para alunos que evadiram, os quais representam 48% dos investigados.

#### 5.4.2 Questões Relacionadas às Matriculas Ativas

No primeiro grupo, alunos não evadidos representam 51%, tendo correlação com as perguntas 19, 23, 25 e 26. Nota-se que, os que concluíram o ensino médio antes do ano de 2008 contemplam 13%, os que informaram também já haverem prestado vestibular em anos anteriores um percentual de 29%, e os que cursavam ou haviam iniciado algum curso superior 19%. A questão envolvendo o aluno estar decidido em relação ao vestibular é um pouco mais pulverizada: apenas 1% informaram estarem mais preparados; pouco mais de 2% alegam não terem gostado do curso anterior, sendo o motivo de um novo vestibular; os que alegam terem disponibilidade para o curso são 5%; e, finalmente, os que buscam uma nova formação ou complemento dos estudos são 6%.

No gráfico da Figura 16, a questão tem forte correlação com a escolha do curso. Conforme o tempo entre a conclusão do ensino médio e o ingresso no Curso aumenta, maior será a porcentagem de alunos matriculado. Por exemplo, os alunos matriculados ativos que concluíram o ensino médio até o ano de 2013 representam 19% do total.

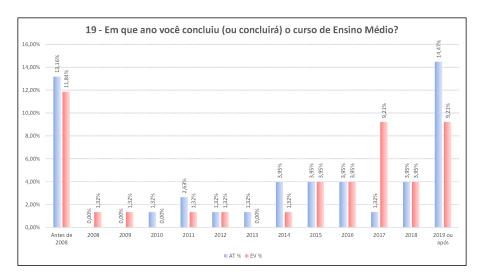


Figura 16 – Questão 19 - Em que ano você concluiu (ou concluirá) o curso de Ensino Médio?

Quanto a questão 23, Você já fez o vestibular em outros anos? (não leve em conta a possível situação de "treineiro"), ilustrada na Figura 17, os alunos que já prestaram vestibulares

em anos anteriores ficam próximo de 29% (somando-se todas as porcentagens), tendendo à um menor risco para evasão na base analisada.

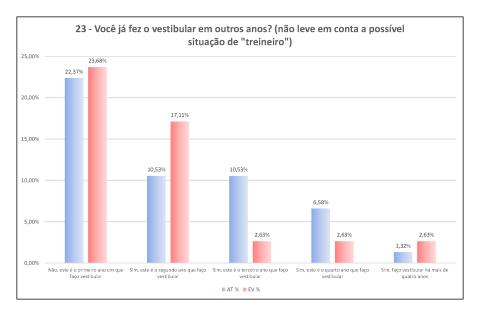


Figura 17 – Questão 23 - Você já fez o vestibular em outros anos? (não leve em conta a possível situação de "treineiro")

Verifica-se uma forte correlação entre as perguntas 25 e 26, alunos que responderam na 18, "Sim, estou cursando", "Sim, mas não concluí"ou "Sim, mas já concluí", totalizam 19,74% e 19, "Por estar mais decidido e preparado", "Por não ter gostado", "Por ter maior disponibilidade para fazer este curso", "Por desejar outra formação"ou "Porque este curso complementa os estudos já feitos"chegam a 15,79%, são pessoas que estão mais decididas em relação à escolha do curso aqui analisada.

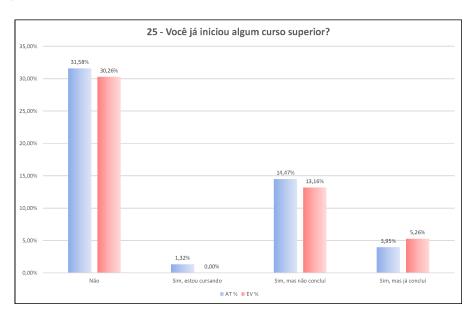


Figura 18 – Questão 25 - Você já iniciou algum curso superior?)

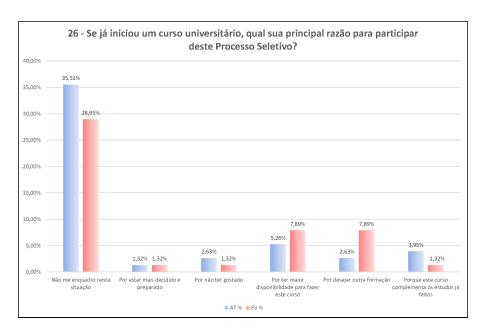


Figura 19 – Questão 26 - Se já iniciou um curso universitário, qual sua principal razão para participar deste Processo Seletivo?

#### 5.4.3 Questões Relacionadas à Evasão

No mesmo sentido de análise, busca-se entender um pouco melhor o perfil dos alunos que atualmente estão em situação de evasão, são informações sobre as perguntas, 2, 11, 12 e 14, esta última aponta como muito relevante ao estudo, e pode-se destacar que está fortemente relacionada a questão 35, que é a respeito da raça ou cor.

A 14, é relacionada ao número de pessoas que contribuem para a obtenção da renda familiar, alunos que informaram apenas uma, representam 19,47% dos evadidos.

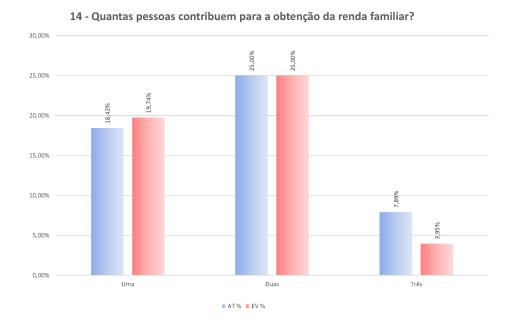


Figura 20 – Questão 14 - Quantas pessoas contribuem para a obtenção da renda familiar?

No gráfico, 21, vê-se que 30% dos evadidos trabalham em atividade remunerada, trazendo uma forte relação com a questão 14, renda familiar.

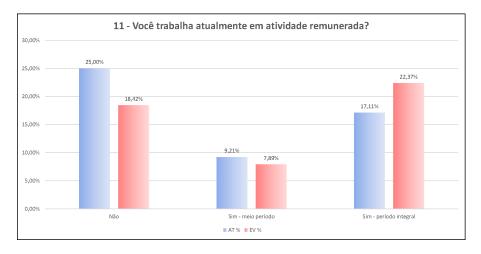


Figura 21 – Questão 11 - Você trabalha atualmente em atividade remunerada?

As informações da questão 12, alertam para os evadidos com participação significativa na renda da família, como principal responsável e/ou contribuí nesta, apontam 26% que contribuem de maneira substancial, ver 22

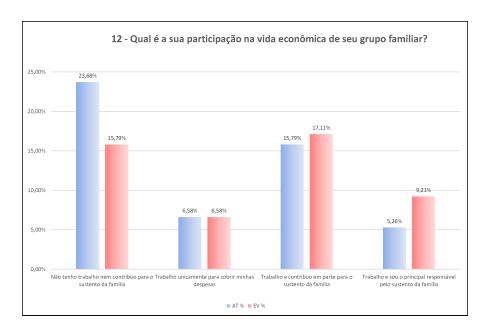


Figura 22 – Questão 12 - Qual é a sua participação na vida econômica de seu grupo familiar?

Verifica-se na questão 2, Quantos anos o aluno completará no ano do vestibular, um percentual de 26% evadidos maiores de 23 anos.

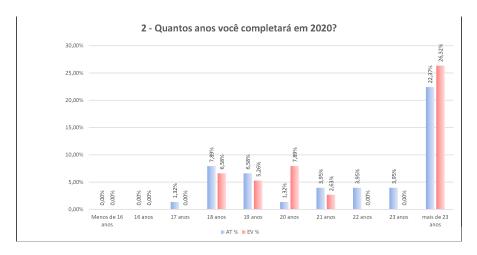


Figura 23 – Questão 2 - Quantos anos você completará em ...?

E por final, analisa-se a última questão, a 35, a qual aparece fortemente relacionada a pergunta 14 (ver Figura 15), o número de evadidos, autodeclarados como brancos, representam 34% dos evadidos.

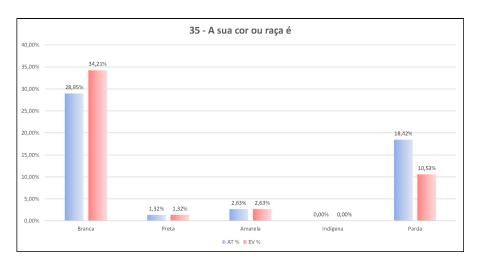


Figura 24 – Questão 35 - Raça ou cor

#### 5.4.3.1 Comentários Finais

Neste Experimento IV, verificou-se que subcategorias como "Realização Pessoal" e "Realização Profissional", apesar destas razões seres similares, apresentam um peso considerável para decisão da evasão do aluno. Chama também a atenção o número de autodeclarados brancos, que trabalham em atividade remunerada, que concluíram o ensino médio antes de 2013 e participam do sustento para obtenção da renda familiar, estes se apresentaram mais propensos a entrarem nos índices de abandono.

Por outro lado, vê-se que pardos, nativos da região, apresentam um índice de evasão de apenas 10%, apontando um fato representativo. Conclui-se que estes são mais interessados e mais dedicados ao curso, talvez por significar uma grande oportunidade de realizar um curso superior e garantir uma melhor condição de vida.

Em uma análise mais geral sobre o estudo computacional realizado, pode-se afirmar que o método de PCA se revelou consistente e importante neste estudo. Ele funcionou como um balizador dos perfis dos estudantes para o momento da inscrição no vestibular. Assim, o estudo pode ser estendido, visando ações educacionais programas para que o índice de evasão caia, ao mesmo tempo que se viabiliza um planejamento para melhorar o rendimento daqueles que permanecem com a matrícula ativa.

Contudo, um fator importante a ser relatado neste estudo é que, não existem informações de aluno evadido que reingressou no mesmo curso da universidade, por um novo vestibular. Esta estratégia do estudante tem como objetivo "limpar"seu histórico escolar, reingressando e aplicando todas equivalências das disciplinas já cursadas. Este caso não muito raro em cursos com altas taxas de reprovação, como é o caso do curso de Licenciatura em Ciências Exatas.

Dadas conclusões deste estudo computacional, a principal tarefa agora seria realizar pesquisas com os estudantes evadidos, buscando identificar as razões que os levaram a abandonar o curso. Além disso, fazer uma análise dos indicadores dos índices de integralização, para tentar encontrar o estudante com baixo rendimento, a fim de evitar uma possível evasão. Por

fim, promover ações e metas pedagógicas, reformulações no curso, e propor, finalmente, um questionário mais apropriado para a região e mais completo, para que o mapeamento dos perfis resulte em uma maior acurácia.

## 6 Conclusões

Este trabalho teve como objetivo principal descobrir quais as razões que levaram determinados grupos de alunos do curso de Licenciatura em Ciências Exatas ao processo de evasão universitária, através de um método matemático, a Análise de Componentes Principais, aplicado ao questionário do estudante promovido pela Universidade Federal do Paraná durante seu processo seletivo, o vestibular.

Esta base de dados representou apenas os questionários respondidos pelos alunos do curso de Licenciatura em Ciências Exatas, os quais prestaram vestibular a partir do ano de 2014. Após o tratamento e padronização dos dados, foi elaborado um estudo computacional com o objetivo de promover o conhecimento das particularidades da vida do estudante, sua situação socioeconômica, permitindo uma análise multivariada das possíveis questões relacionadas à evasão ocorrida no Curso. No início da pesquisa, através da revisão bibliográfica que o este trabalho apresenta, ficou evidente que este estudo é novo, pois tratou de um caso específico, dado a análise do questionário do estudante, e constatou-se também que o tema evasão é considerado um problema recorrente e em pleno crescimento.

As análises produzidas em cada um dos experimentos realizados nos estudos computacionais mostraram que a aplicação do PCA foi oportuna, permitindo uma análise do horizonte do curso mais completa. Os estudos também mostraram que o algoritmo para predição, previamente desenvolvido na Linguagem R, também é promissor, sendo possível ser aplicado às distintas bases de dados e diferentes problemas. As conclusões tomadas nos experimentos permitiram traçar um panorama inicial do perfil do estudante, apresentando oportunidade para discussões e planejamento de ações a fim de combater e minimizar a taxa de evasão do Curso. As informações apontadas devem ser vistas como uma precursão da formação do planejamento pedagógico do Curso.

Contudo, o trabalho evidenciou a dificuldade na obtenção de dados. Foi possível apenas analisar a base de dados do vestibular, não sendo disponibilizadas informações sobre alunos oriundos do Sistema de Seleção Unificada (SISU), além de outros programas de cotas. Desta forma, busca-se num futuro próximo, dar continuidade ao trabalho, estudando estas novas bases de dados, e aumentando a robustez da análise produzida.

Em síntese, os resultados indicaram a necessidade de um grande esforço na busca de uma base de dados maior e de qualidade, para continuidade das investigações e para futuras discussões, sempre com o objetivo final de reduzir o percentual de evasão. Deste modo, apesar das falhas e omissões diagnosticadas nessa base de dados do questionário, pôde-se estabelecer conceitos, indicar procedimentos e o critérios, culminando em um estudo instigante, que contribuiu para um melhor conhecimento e condução na troca de experiências educacionais. Isto favorece a avaliação do perfil de cada estudante ingressante e colabora para a melhoria da administração desta variedade de pessoas.

## Referências

- ANDERSON, T. W. The use of factor analysis in the statistical analysis of multiple time series. Psychometrika, Springer, v. 28, n. 1, p. 1–25, 1963.
- ASSIS, W. R. de. Rede neural artificial: Identificação de acadêmicos em curso de licenciatura em matemática com possibilidade de desistência. <u>Cadernos Cajuína</u>, v. 5, n. 3, p. 534–544, 2020.
- BAKER, R.; ISOTANI, S.; CARVALHO, A. Mineração de dados educacionais: Oportunidades para o brasil. Revista Brasileira de Informática na Educação, v. 19, n. 02, p. 03, 2011.
- BAYER, J. et al. Predicting drop-out from social behaviour of students. <u>International</u> Educational Data Mining Society, ERIC, 2012.
- BIOCONDUCTOR. <u>Bioconductor pcaMethods</u>. 21. <a href="https://www.bioconductor.org/packages/release/bioc/html/pcaMethods.html">https://www.bioconductor.org/packages/release/bioc/html/pcaMethods.html</a>. (Accessed on 12/12/2021).
- BRAGA, M. M.; CARMO, L. P. Maria do; BOGUTCHI, T. F. A evasão no ensino superior brasileiro: o caso da ufmg. <u>Avaliação: Revista da Avaliação da Educação Superior</u>, v. 8, n. 3, 2003.
- BRAGA, T. C.; ALBUQUERQUE, É. L. de. Avaliação estatística multivariada aplicada a qualidade do ar em salvador-ba: O caso da estação itaigara.
- CAMBRUZZI, W. L.; RIGO, S. J.; BARBOSA, J. L. Dropout prediction and reduction in distance education courses with the learning analytics multitrail approach. <u>J. UCS</u>, v. 21, n. 1, p. 23–47, 2015.
- FILHO, R. L. S. et al. A evasão no ensino superior brasileiro. <u>Cadernos de pesquisa</u>, SciELO Brasil, v. 37, p. 641–659, 2007.
- GIRSHICK, M. On the sampling theory of roots of determinantal equations. The Annals of Mathematical Statistics, JSTOR, v. 10, n. 3, p. 203–224, 1939.
- GOWER, J. C. Multivariate analysis and multidimensional geometry. <u>Journal of the Royal Statistical Society: Series D (The Statistician)</u>, Wiley Online Library, v. 17, n. 1, p. 13–28, 1967.
- HEGDE, V. Dimensionality reduction technique for developing undergraduate student dropout model using principal component analysis through r package. In: IEEE. <u>2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)</u>. [S.I.], 2016. p. 1–6.
- HONGYU, K.; SANDANIELO, V. L. M.; JUNIOR, G. J. de O. Análise de componentes principais: resumo teórico, aplicação e interpretação. <u>E&S Engineering and science</u>, v. 5, n. 1, p. 83–90, 2016.
- HOTELLING, H. Analysis of a complex of statistical variables into principal components. Journal of educational psychology, Warwick & York, v. 24, n. 6, p. 417, 1933.
- HUNT, M. H. et al. A comprehensive needs assessment to facilitate prevention of school drop out and violence. <u>Psychology in the Schools</u>, Wiley Online Library, v. 39, n. 4, p. 399–416, 2002.

- $\label{localization} IBGE \mid Cidades@ \mid Paraná \mid Pesquisa \mid Censo escolar sinopse \mid Ensino básico. 2022. \\ < https://cidades.ibge.gov.br/brasil/pr/pesquisa/13/78117?tipo=ranking&indicador=5929>. \\ (Accessed on 02/02/2022). \\$
- JAYAPRAKASH, S. M. et al. Early alert of academically at-risk students: An open source analytics initiative. Journal of Learning Analytics, v. 1, n. 1, p. 6–47, 2014.
- JEFFERS, J. N. Two case studies in the application of principal component analysis. <u>Journal of the Royal Statistical Society: Series C (Applied Statistics)</u>, Wiley Online Library, v. 16, n. 3, p. 225–236, 1967.
- JOLLIFFE, I. Principal component analysis for special types of data. <u>Principal component</u> analysis, Springer, p. 338–372, 2002.
- KULKARNI, A. R. RPubs Wine Data Principal Component Analysis (PCA) & Samp; Clustering. 2021. <a href="https://rpubs.com/amolrk7/289957">https://rpubs.com/amolrk7/289957</a>. (Accessed on 12/12/2021).
- LEONI, R. C. P. D. <u>Análise de Componentes Principais ACP</u>. 2021. <a href="https://rstudio-pubs-static.s3.amazonaws.com/251265\_9f65a6ca663f4eff814ae8b1f5a04ffe.html">https://rstudio-pubs-static.s3.amazonaws.com/251265\_9f65a6ca663f4eff814ae8b1f5a04ffe.html</a>. (Accessed on 12/09/2021).
- LYKOURENTZOU, I. et al. Dropout prediction in e-learning courses through the combination of machine learning techniques. <u>Computers & Education</u>, Elsevier, v. 53, n. 3, p. 950–965, 2009.
- MARTINHO, V. R.; NUNES, C.; MINUSSI, C. R. Predição do grupo de risco de evasão discente em cursos superiores presenciais utilizando uma rede neural artmap-fuzzy. In: CONGRESSO BRASILEIRO DE EDUCAÇÃO EM ENGENHARIA-COBENGE. [S.I.: s.n.], 2013. v. 41, p. 2013.
- PEARSON, K. Liii. on lines and planes of closest fit to systems of points in space. The London, Edinburgh, and Dublin philosophical magazine and journal of science, Taylor & Francis, v. 2, n. 11, p. 559–572, 1901.
- R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria, 2021. Disponível em: <https://www.R-project.org/>.
- RAO, C. R. The use and interpretation of principal component analysis in applied research. Sankhyā: The Indian Journal of Statistics, Series A (1961-2002), Springer, v. 26, n. 4, p. 329–358, 1964.
- SANTO, B. C. E. et al. Reciprocidade e compromisso: como o marketing bancário supera o paradoxo do banqueiro. Dissertação de Mestrado, Florianópolis, SC, 2021.
- SHIRASU, M. R.; ALBUQUERQUE, R. de et al. Determinantes da evasão e repetência escolar no ensino médio do ceará. Revista Econômica do Nordeste, v. 46, n. 4, p. 117–136, 2015.
- SMITH, T. M. et al. Reconstruction of historical sea surface temperatures using empirical orthogonal functions. <u>Journal of Climate</u>, American Meteorological Society, v. 9, n. 6, p. 1403–1420, 1996.
- SPAECE. SPAECE. 2021. <a href="https://spaece.caedufjf.net/">https://spaece.caedufjf.net/</a>. (Accessed on 12/02/2021).

- STACKLIES, W. et al. pcamethods a bioconductor package providing pca methods for incomplete data. Bioinformatics, v. 23, p. 1164–1167, 2007.
- TIMBAL, M. A. Analysis of student-at-risk of dropping out (sardo) using decision tree: An intelligent predictive model for reduction. <u>International Journal of Machine Learning and Computing</u>, v. 9, n. 3, p. 273–278, 2019.
- TYSKA, D. U. Avaliação das características produtivas de codornas de corte e qualidade de ovos por meio de análises multivariadas. <u>Dissertação de Mestrado</u>, UNIVERSIDADE FEDERAL DE PELOTAS, FACULDADE DE AGRONOMIA ELISEU MACIEL, DEPARTAMENTO DE ZOOTECNIA, 2017.
- VICINI, L. Análise multivariada: da teoria à prática. Universidade Federal de Santa Maria, p. 30, 2005.
- WALL, M. E.; RECHTSTEINER, A.; ROCHA, L. M. Singular value decomposition and principal component analysis. In: <u>A practical approach to microarray data analysis</u>. [S.I.]: Springer, 2003. p. 91–109.
- XAVIER, L. H. Modelos univariado e multivariado para análise de medidas repetidas e verificação da acurácia do modelo univariado por meio de simulação. Tese (Doutorado) Universidade de São Paulo, 2000.
- YADAV, S. K.; PAL, S. Data mining: A prediction for performance improvement of engineering students using classification. arXiv preprint arXiv:1203.3832, 2012.

# 7 Apêndice 1

### 7.1 Código R

```
2 #Pacotes do repositorio CRAN do R
3 #Importados e instalados
4 #Analise PCA PCAmethods
5 install.packages('factoextra')
6 install.packages("PerformanceAnalytics") #Utilizado PCA
7 #Importação arquivos de dados
8 install.packages("readxl")
9 #Grafico e plotagens
10 install.packages(corrplot)
install.packages("ggcorrplot")
12
13 #Comando utilizados para carregamentos dos pacotes
14 library ('factoextra')
15 library ("readxl")
16 library (corrplot)
17 library ("FactoMineR")
18 library (PerformanceAnalytics)
19 library (ggcorrplot)
20 library (ggplot2)
23 #Base de dados importada do repositorio local
24 mat <- read_excel("X:/caminho/Dados_socioeconomicos/PCA-R/varesocio1.xlsx")
26 #Usando a funcao summary verificamos algumas estatisticas.
27 summary (mat)
29 #Excluir algumas linhas e colunas suplementares
30 mat.active <- mat[5:46]
32 #calculando o desvio padrao
33 sapply (mat.active, sd)
34
35 #Matriz de correlação, antes de PCA
36 res <- cor(mat.active) # Corr matrix</pre>
37 round (res, 2) #arredonda os valores em seu primeiro argumento para o numero
       especificado
38 #exportar CSV podrao da Matriz de correlação para analise
39 write.csv(res, "X:\\caminho\\corr.csv", row.names = FALSE)
40
41
```

```
42 #Visualização dessa matriz de correlação com a função corrplot:
43
  ggcorrplot(res, hc.order = TRUE,
44
              type = "lower",
45
              lab = TRUE.
46
47
              lab\_size = 2,
              method="circle",
48
              colors = c("tomato2", "white", "springgreen3"),
49
              title="Correlograma",
              ggtheme=theme_bw)
51
52
53
54 #Gerar o PCA
| res.pca <- PCA(mat.active, graph = FALSE) #library ("FactoMineR")
56 print (res.pca)
57 summary (res.pca)
59 #Autovalores e suas representatividades
60 eigenvalues <- res.pca$eig
61 head (eigenvalues [, 1:2])
62
63 # Criando uma Classe, matriculado Sim ou Nao
64 matClasses <- factor(mat$Matriculado)
65
66 #extrair resultados das variaveis do PCA para plotar
67 var <-get_pca_var(res.pca)
68 ind <-get_pca_ind(res.pca)
69
70 #representa a qualidade de representacao das variaveis no mapa de fatores.
71|#e calculado como as coordenadas quadradas: var.cos2 = var.coord * var.
      coord.
72 var$cos2
73
75 #PCA Correlação entre as variaveis
76 corrplot (varscos2, is.corr = T)\#, addCoef.col = 'black', cl.cex = 0.8, tl.cex
     = 0.8, number.cex = 0.5) \#\cos 2 = a qualidade dos individuos no mapa de
77 corrplot (var$contrib, is.corr=FALSE)#,addCoef.col = 'black',cl.cex = 0.8,
      tl.cex = 0.8, number.cex = 0.5)
78
79 #
80 |#fviz_eig(res.pca,addlabels = T,ylim=c(0,25)) #escala do grafico eixo y
81 fviz_eig(res.pca, addlabels = TRUE) #sem escala do grafico eixo y
82
83
84 var <- get_pca_var(res.pca)
```

```
85 var
86
87 #Quanto cada variavel tem de representatividade
|\#fviz\_cos2(res.pca, choice = "var") \#cos2 = a qualidade dos individuos no
      mapa de fatores
89 # Contributions of variables to PC1
90 fviz_contrib(res.pca, choice = "var", axes = 1)
91 # Contributions of variables to PC2
92 fviz_contrib (res.pca, choice = "var", axes = 2)
93 # Contributions of variables toPC 1 e 2
94 fviz_contrib (res.pca, choice = "var", axes = c(1,2))
95
96 #---
97 # Autovalores
98 eig.val <- get_eigenvalue(res.pca)
99 eig.val
100 #
101
102 # Resultados para as variaveis
103 res.var <- get_pca_var(res.pca)</pre>
                          # Coordenadas
104 res.var$coord
                          # Contribuicoes para as PCs
105 res.var$contrib
106 res.var$cos2
                           # Qualidade da representacao
107
108
109 # Resultado das observações
| res.ind <- get_pca_ind(res.pca)
111 res.ind$coord
                           # Coordenadas
112 res.ind$contrib
                           # Contribuicoes para as PCs
113 res.ind$cos2
                           # Qualidade da representacao
114
#criando grupo cluster
116 grupo <-as.factor(mat$Matriculado)
117 grupo
118
119 #Plotagem das contribuicoes das caracteristicas originais
120
121
122 #
123 fviz_pca_biplot(res.pca,
                   \#axes = c(1, 2),
124
                    repel = TRUE,
125
                    col.var = "black", # cor das variaveis
126
                    col.ind = as.factor(mat$Matriculado),
127
                    #habillage=mat$RENDIMENTO, # Color individuals by groups
128
                    #geom.ind = "point",
129
                    pointshape=16,
130
```

```
pointsize = 5,
131
                    label = "all",
132
                    labelsize = 4,
133
                    addEllipses = FALSE,
134
                   \#select.var = list(contrib = 50), \# Selecionar top 3
135
                       contributing variables
                   \#select.var= list(name = c(1:10)),\# Select var by names
136
                   \#var = list(cos2 = 0.0), \#Select and visualize variables
137
                       with cos2 >= 0.96
                   \#select.ind = list(cos2 = 30), \# Select the top 20
138
                       according the cos2
                   \#select.ind = list(name = c("20", "33", "119")), \# Select
139
                       nome indiv
                   \#select.ind = list(contrib = 30), \# Select the top 30
140
                       contributing individuals
                   legend.title = "Matriculado"
141
142 )+
     theme_bw()+
143
     labs(title = "Grafico PCA Dados Socioeconomicos",
144
          fill = "Matricula",
145
          x="PCA1"
146
          y="PCA2")
147
148
149
150 #Remocao das questoes com menores significancias
  dim(mat.active)
151
152 tamanho<-dim(mat.active)
153
154 for (j in 1:tamanho [2]) {
     if (res.var\$contrib [(tamanho [2]+1) -j,1] < 3.5 &&
155
156
        res.var\contrib[(tamanho [2]+1) -j,2]<3.5)
     157
158
159 }
```

Listing 7.1 - Código fonte em R

# 8 Apêndice 2

# 8.1 Base de dados

| 10 | MATRICILLA | ID A   | CDD D    |   |    |   |   |   |   |   |   |    | F  | Pergu | ıntas |    |    |    |    |    |    |    |    |
|----|------------|--------|----------|---|----|---|---|---|---|---|---|----|----|-------|-------|----|----|----|----|----|----|----|----|
| ID | MATRICULA  | IRA    | GRP_Rend | 1 | 2  | 3 | 4 | 5 | 6 | 7 | 8 | 9  | 10 | 11    | 12    | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| 1  | GRR—-3334  | 0,9194 | OTIMO    | 2 | 10 | 1 | 1 | 3 | 1 | 3 | 7 | 12 | 1  | 1     | 1     | 2  | 1  | 2  | 4  | 5  | 1  | 0  | 2  |
| 2  | GRR—-3379  | 0      | EVADIDO  | 2 | 10 | 1 | 1 | 8 | 3 | 3 | 3 | 12 | 12 | 1     | 1     | 2  | 1  | 4  | 2  | 4  | 3  | 12 | 3  |
| 3  | GRR—-3330  | 0,4264 | BAIXO    | 2 | 10 | 1 | 2 | 8 | 6 | 2 | 5 | 11 | 12 | 3     | 3     | 2  | 2  | 2  | 3  | 3  | 1  | 3  | 3  |
| 4  | GRR—-3383  | 0      | EVADIDO  | 2 | 5  | 1 | 1 | 1 | 6 | 4 | 3 | 4  | 11 | 2     | 2     | 2  | 1  | 4  | 3  | 2  | 1  | 11 | 3  |
| 5  | GRR—-3338  | 0,3733 | BAIXO    | 2 | 4  | 1 | 1 | 8 | 8 | 2 | 4 | 4  | 12 | 1     | 1     | 2  | 2  | 4  | 5  | 3  | 1  | 12 | 3  |
| 6  | GRR—-3327  | 0,9002 | OTIMO    | 2 | 10 | 2 | 1 | 3 | 6 | 2 | 2 | 3  | 3  | 1     | 1     | 2  | 1  | 3  | 1  | 2  | 4  | 0  | 4  |
| 7  | GRR—-3365  | 0,6034 | REGULAR  | 2 | 7  | 1 | 1 | 8 | 1 | 2 | 4 | 10 | 10 | 3     | 3     | 1  | 2  | 4  | 3  | 3  | 1  | 9  | 3  |
| 8  | GRR—-4841  | 0      | EVADIDO  | 1 | 6  | 1 | 1 | 1 | 1 | 6 | 6 | 1  | 1  | 3     | 3     | 2  | 2  | 4  | 3  | 3  | 2  | 7  | 3  |
| 9  | GRR—-3372  | 0      | EVADIDO  | 1 | 5  | 1 | 1 | 2 | 1 | 3 | 6 | 1  | 10 | 1     | 1     | 3  | 1  | 3  | 5  | 4  | 3  | 12 | 3  |
| 10 | GRR—-3398  | 0,393  | BAIXO    | 2 | 10 | 1 | 1 | 2 | 3 | 7 | 7 | 11 | 11 | 3     | 2     | 1  | 1  | 1  | 4  | 3  | 1  | 0  | 2  |
| 11 | GRR—-4110  | 0      | EVADIDO  | 1 | 10 | 2 | 1 | 8 | 4 | 5 | 7 | 2  | 2  | 3     | 4     | 2  | 1  | 6  | 3  | 3  | 3  | 2  | 5  |
| 12 | GRR—-4113  | 0,3576 | BAIXO    | 2 | 10 | 1 | 1 | 8 | 4 | 2 | 3 | 2  | 11 | 1     | 1     | 1  | 1  | 3  | 2  | 4  | 1  | 5  | 3  |
| 13 | GRR—-4435  | 0      | EVADIDO  | 2 | 10 | 1 | 1 | 4 | 4 | 2 | 2 | 10 | 12 | 1     | 1     | 1  | 2  | 3  | 2  | 2  | 1  | 0  | 3  |
| 14 | GRR—-4502  | 0,6334 | REGULAR  | 2 | 3  | 1 | 1 | 3 | 1 | 7 | 5 | 1  | 11 | 1     | 1     | 3  | 1  | 3  | 5  | 4  | 1  | 12 | 3  |
| 15 | GRR—-4109  | 0      | EVADIDO  | 1 | 10 | 1 | 1 | 8 | 3 | 8 | 8 | 12 | 2  | 3     | 3     | 2  | 1  | 1  | 4  | 3  | 1  | 0  | 3  |
| 16 | GRR—-4120  | 0,6843 | REGULAR  | 2 | 5  | 1 | 1 | 1 | 1 | 5 | 7 | 2  | 4  | 1     | 1     | 3  | 1  | 3  | 5  | 3  | 4  | 12 | 1  |
| 17 | GRR—-5418  | 0      | EVADIDO  | 1 | 10 | 1 | 1 | 2 | 3 | 1 | 2 | 5  | 2  | 2     | 4     | 1  | 2  | 4  | 1  | 2  | 1  | 8  | 3  |
| 18 | GRR—-4117  | 0,7736 | ВОМ      | 1 | 10 | 1 | 1 | 8 | 3 | 5 | 5 | 2  | 10 | 3     | 2     | 4  | 1  | 1  | 3  | 3  | 3  | 0  | 5  |
| 19 | GRR—-5630  | 0      | EVADIDO  | 2 | 4  | 1 | 1 | 2 | 1 | 4 | 5 | 12 | 2  | 1     | 1     | 1  | 1  | 4  | 5  | 1  | 1  | 12 | 3  |

| 20 | GRR—-4107 | 0,4851 | BAIXO   | 1 | 8  | 1 | 1 | 2 | 1 | 4 | 2 | 12 | 12 | 1 | 1 | 1 | 2 | 4 | 4 | 4 | 1 | 8  | 3 |
|----|-----------|--------|---------|---|----|---|---|---|---|---|---|----|----|---|---|---|---|---|---|---|---|----|---|
| 21 | GRR—-4104 | 0      | EVADIDO | 1 | 4  | 1 | 1 | 8 | 2 | 3 | 2 | 12 | 12 | 3 | 2 | 2 | 3 | 4 | 2 | 4 | 1 | 12 | 3 |
| 22 | GRR—-4700 | 0,6325 | REGULAR | 1 | 4  | 1 | 1 | 8 | 8 | 5 | 2 | 12 | 10 | 1 | 1 | 2 | 2 | 4 | 5 | 4 | 1 | 12 | 3 |
| 23 | GRR—-4121 | 0,5654 | REGULAR | 1 | 5  | 1 | 1 | 1 | 2 | 8 | 7 | 12 | 1  | 1 | 1 | 3 | 1 | 2 | 5 | 4 | 1 | 11 | 3 |
| 24 | GRR—-4660 | 0,6311 | REGULAR | 1 | 7  | 1 | 1 | 2 | 2 | 5 | 4 | 1  | 1  | 2 | 3 | 2 | 2 | 5 | 4 | 3 | 1 | 8  | 3 |
| 25 | GRR—-4658 | 0      | EVADIDO | 2 | 6  | 1 | 4 | 8 | 1 | 2 | 3 | 4  | 11 | 3 | 3 | 2 | 2 | 4 | 2 | 2 | 1 | 11 | 3 |
| 26 | GRR—-4652 | 0      | EVADIDO | 1 | 5  | 1 | 6 | 3 | 1 | 5 | 5 | 4  | 10 | 1 | 1 | 2 | 1 | 5 | 3 | 4 | 1 | 10 | 2 |
| 27 | GRR—-4653 | 0      | EVADIDO | 2 | 6  | 1 | 1 | 2 | 1 | 3 | 5 | 4  | 2  | 1 | 1 | 2 | 2 | 4 | 4 | 5 | 1 | 9  | 3 |
| 28 | GRR—-4663 | 0      | EVADIDO | 1 | 7  | 1 | 1 | 8 | 4 | 4 | 4 | 12 | 5  | 3 | 3 | 1 | 3 | 3 | 3 | 3 | 1 | 8  | 2 |
| 29 | GRR—-4644 | 0,1019 | BAIXO   | 1 | 5  | 1 | 1 | 2 | 1 | 4 | 6 | 3  | 2  | 3 | 2 | 1 | 2 | 2 | 3 | 3 | 1 | 10 | 2 |
| 30 | GRR—-4670 | 0      | EVADIDO | 2 | 10 | 3 | 1 | 1 | 4 | 5 | 7 | 10 | 1  | 1 | 2 | 1 | 1 | 1 | 3 | 3 | 1 | 0  | 3 |
| 31 | GRR—-4667 | 0,5502 | REGULAR | 1 | 9  | 1 | 1 | 2 | 3 | 5 | 5 | 3  | 1  | 3 | 3 | 3 | 2 | 3 | 2 | 3 | 1 | 6  | 3 |
| 32 | GRR—-4671 | 0,2946 | BAIXO   | 2 | 10 | 1 | 1 | 8 | 1 | 3 | 3 | 12 | 2  | 3 | 3 | 3 | 3 | 4 | 3 | 3 | 1 | 4  | 3 |
| 33 | GRR—-4643 | 0      | EVADIDO | 2 | 6  | 1 | 1 | 8 | 1 | 2 | 2 | 2  | 10 | 3 | 3 | 2 | 2 | 3 | 3 | 5 | 1 | 9  | 3 |
| 34 | GRR—-4650 | 0      | EVADIDO | 1 | 5  | 1 | 1 | 8 | 2 | 2 | 7 | 11 | 1  | 1 | 1 | 4 | 2 | 5 | 5 | 5 | 1 | 10 | 3 |
| 35 | GRR—-4654 | 0      | EVADIDO | 1 | 6  | 1 | 1 | 2 | 1 | 5 | 5 | 12 | 12 | 1 | 1 | 2 | 2 | 5 | 1 | 3 | 3 | 10 | 5 |
| 36 | GRR—-4657 | 0      | EVADIDO | 2 | 10 | 3 | 1 | 8 | 3 | 7 | 5 | 12 | 11 | 3 | 4 | 2 | 1 | 2 | 1 | 3 | 1 | 0  | 1 |
| 37 | GRR—-4655 | 0      | EVADIDO | 2 | 10 | 1 | 1 | 8 | 1 | 2 | 2 | 12 | 11 | 3 | 3 | 2 | 2 | 3 | 1 | 3 | 1 | 5  | 1 |
| 38 | GRR—-4656 | 0      | EVADIDO | 2 | 10 | 1 | 1 | 8 | 3 | 2 | 2 | 12 | 11 | 3 | 3 | 2 | 2 | 3 | 3 | 3 | 1 | 4  | 1 |
| 39 | GRR—-4669 | 0,8154 | ВОМ     | 1 | 10 | 1 | 1 | 8 | 1 | 2 | 2 | 1  | 11 | 1 | 1 | 2 | 2 | 6 | 4 | 4 | 1 | 8  | 3 |
| 40 | GRR4648   | 0      | EVADIDO | 1 | 10 | 2 | 6 | 2 | 4 | 3 | 5 | 2  | 1  | 3 | 2 | 1 | 2 | 2 | 1 | 3 | 1 | 10 | 3 |
| 41 | GRR—-4647 | 0      | EVADIDO | 1 | 10 | 2 | 1 | 8 | 3 | 7 | 4 | 12 | 12 | 3 | 3 | 4 | 2 | 2 | 4 | 3 | 3 | 0  | 3 |
| 42 | GRR—-5965 | 0      | EVADIDO | 1 | 10 | 1 | 1 | 2 | 4 | 6 | 5 | 2  | 2  | 3 | 2 | 1 | 1 | 1 | 3 | 3 | 3 | 0  | 3 |
| 43 | GRR—-4834 | 0      | EVADIDO | 2 | 10 | 1 | 1 | 8 | 1 | 2 | 5 | 12 | 10 | 3 | 4 | 1 | 2 | 6 | 3 | 3 | 1 | 1  | 3 |
| 44 | GRR—-0540 | 0      | EVADIDO | 1 | 10 | 2 | 6 | 8 | 3 | 3 | 4 | 10 | 11 | 1 | 1 | 2 | 1 | 4 | 2 | 3 | 1 | 0  | 3 |
|    |           |        |         |   |    |   |   |   |   |   |   |    |    |   |   |   |   |   |   |   |   |    |   |

| 45 | GRR—-0980 | 0      | EVADIDO | 2 | 7  | 1 | 1 | 1 | 4 | 2 | 6 | 2  | 1  | 3 | 4 | 1 | 1 | 1 | 2 | 2 | 1 | 8  | 3             |
|----|-----------|--------|---------|---|----|---|---|---|---|---|---|----|----|---|---|---|---|---|---|---|---|----|---------------|
| 46 | GRR—-4467 | 0      | EVADIDO | 2 | 10 | 1 | 1 | 8 | 3 | 1 | 3 | 2  | 11 | 3 | 4 | 2 | 1 | 3 | 4 | 3 | 1 | 0  | 2             |
| 47 | GRR—-3928 | 0,4325 | BAIXO   | 1 | 10 | 1 | 1 | 8 | 3 | 6 | 7 | 1  | 1  | 3 | 4 | 1 | 1 | 1 | 2 | 2 | 1 | 0  | 3             |
| 48 | GRR—-1979 | 0,727  | ВОМ     | 1 | 10 | 1 | 1 | 8 | 1 | 3 | 2 | 12 | 11 | 2 | 3 | 6 | 2 | 3 | 3 | 2 | 1 | 4  | 1             |
| 49 | GRR—-1963 | 0      | EVADIDO | 1 | 10 | 1 | 1 | 8 | 2 | 5 | 5 | 12 | 11 | 1 | 1 | 2 | 1 | 3 | 4 | 2 | 4 | 10 | 3             |
| 50 | GRR—-3876 | 0      | EVADIDO | 2 | 4  | 2 | 4 | 2 | 3 | 5 | 6 | 4  | 12 | 2 | 3 | 1 | 2 | 2 | 2 | 2 | 1 | 11 | 3             |
| 51 | GRR—-5026 | 0      | BAIXO   | 2 | 6  | 1 | 1 | 8 | 1 | 8 | 5 | 12 | 11 | 1 | 1 | 1 | 1 | 3 | 4 | 4 | 1 | 12 | 3             |
| 52 | GRR—-4207 | 0,75   | ВОМ     | 2 | 10 | 1 | 1 | 8 | 4 | 2 | 5 | 12 | 10 | 3 | 4 | 2 | 3 | 4 | 2 | 3 | 1 | 0  | 3             |
| 53 | GRR—-4345 | 0,9067 | OTIMO   | 2 | 4  | 1 | 2 | 4 | 1 | 7 | 7 | 1  | 1  | 1 | 1 | 4 | 2 | 4 | 5 | 4 | 1 | 12 | 3             |
| 54 | GRR—-4343 | 0,56   | REGULAR | 1 | 10 | 2 | 1 | 8 | 6 | 6 | 7 | 6  | 1  | 3 | 4 | 2 | 1 | 4 | 3 | 3 | 2 | 0  | 4             |
| 55 | GRR—-4339 | 0      | BAIXO   | 1 | 4  | 1 | 1 | 8 | 2 | 5 | 4 | 1  | 11 | 1 | 1 | 2 | 1 | 4 | 5 | 4 | 1 | 12 | 3             |
| 56 | GRR—-4315 | 0      | BAIXO   | 2 | 10 | 1 | 2 | 8 | 6 | 1 | 3 | 12 | 12 | 2 | 4 | 1 | 3 | 3 | 1 | 2 | 1 | 11 | 5             |
| 57 | GRR—-4353 | 0,8638 | ОТІМО   | 2 | 10 | 1 | 4 | 8 | 4 | 5 | 5 | 2  | 12 | 1 | 2 | 1 | 2 | 2 | 4 | 2 | 4 | 0  | 3             |
| 58 | GRR—-4344 | 0,475  | BAIXO   | 2 | 8  | 1 | 1 | 8 | 4 | 8 | 2 | 12 | 12 | 1 | 1 | 2 | 1 | 2 | 2 | 3 | 1 | 9  | 1             |
| 59 | GRR—-5017 | 0,7094 | ВОМ     | 1 | 4  | 1 | 1 | 2 | 1 | 3 | 6 | 5  | 11 | 3 | 3 | 3 | 3 | 4 | 3 | 2 | 4 | 11 | 4             |
| 60 | GRR—-4696 | 0      | EVADIDO | 1 | 6  | 1 | 1 | 8 | 1 | 3 | 2 | 2  | 2  | 1 | 1 | 2 | 2 | 4 | 2 | 3 | 1 | 10 | 3             |
| 61 | GRR—-5109 | 0,4106 | BAIXO   | 2 | 4  | 1 | 1 | 8 | 1 | 4 | 4 | 4  | 1  | 2 | 2 | 3 | 2 | 5 | 1 | 4 | 1 | 12 | 3             |
| 62 | GRR—-5009 | 0,5006 | REGULAR | 2 | 10 | 3 | 1 | 8 | 8 | 1 | 2 | 12 | 12 | 3 | 3 | 2 | 2 | 4 | 4 | 3 | 1 | 0  | 3             |
| 63 | GRR—-4637 | 0,6221 | REGULAR | 2 | 9  | 1 | 6 | 2 | 1 | 3 | 5 | 2  | 2  | 3 | 3 | 2 | 3 | 3 | 4 | 3 | 1 | 7  | 3             |
| 64 | GRR—-4654 | 0      | EVADIDO | 1 | 10 | 3 | 1 | 8 | 3 | 2 | 2 | 4  | 2  | 3 | 4 | 1 | 1 | 4 | 1 | 3 | 1 | 9  | 3             |
| 65 | GRR—-4898 | 0,478  | BAIXO   | 2 | 9  | 2 | 1 | 8 | 3 | 5 | 5 | 5  | 2  | 2 | 3 | 2 | 2 | 3 | 2 | 2 | 1 | 7  | 3             |
| 66 | GRR—-4759 | 0      | EVADIDO | 1 | 10 | 1 | 1 | 8 | 6 | 4 | 5 | 11 | 11 | 1 | 3 | 1 | 3 | 5 | 3 | 3 | 1 | 10 | 5             |
| 67 | GRR—-4766 | 0      | EVADIDO | 2 | 10 | 2 | 1 | 8 | 6 | 7 | 7 | 1  | 1  | 2 | 3 | 3 | 2 | 2 | 4 | 2 | 4 | 0  | 3             |
| 68 | GRR—-4675 | 0,7377 | ВОМ     | 2 | 5  | 1 | 1 | 8 | 8 | 2 | 2 | 10 | 12 | 2 | 3 | 1 | 3 | 4 | 3 | 2 | 1 | 12 | 1             |
| 69 | GRR—-4679 | 0      | EVADIDO | 1 | 10 | 2 | 1 | 3 | 3 | 8 | 8 | 12 | 12 | 2 | 3 | 2 | 2 | 2 | 2 | 3 | 1 | 12 | 5             |
|    |           |        |         |   |    |   |   |   |   |   |   |    |    |   |   |   |   |   |   |   |   |    | $\overline{}$ |

| 70 | GRR—-4657 | 0      | EVADIDO | 1 | 4  | 1 | 1 | 2 | 2 | 2 | 5 | 2  | 2  | 2 | 3 | 4 | 2 | 5 | 3 | 2 | 1 | 12 | 2 |
|----|-----------|--------|---------|---|----|---|---|---|---|---|---|----|----|---|---|---|---|---|---|---|---|----|---|
| 71 | GRR—-4688 | 0,6464 | REGULAR | 1 | 10 | 1 | 1 | 8 | 1 | 6 | 5 | 12 | 1  | 1 | 1 | 2 | 2 | 4 | 4 | 3 | 2 | 0  | 5 |
| 72 | GRR—-4768 | 0      | EVADIDO | 2 | 4  | 1 | 1 | 3 | 1 | 5 | 7 | 12 | 1  | 1 | 1 | 2 | 2 | 3 | 5 | 4 | 1 | 12 | 3 |
| 73 | GRR—-4638 | 0,7374 | BOM     | 1 | 5  | 3 | 1 | 8 | 2 | 1 | 2 | 11 | 11 | 1 | 1 | 1 | 2 | 3 | 5 | 2 | 1 | 12 | 3 |
| 74 | GRR—-4712 | 0,3088 | BAIXO   | 2 | 10 | 2 | 1 | 2 | 7 | 8 | 5 | 12 | 11 | 1 | 1 | 2 | 1 | 3 | 5 | 4 | 1 | 12 | 3 |
| 75 | GRR—-4776 | 0,7211 | BOM     | 2 | 7  | 1 | 1 | 2 | 1 | 2 | 2 | 12 | 11 | 2 | 3 | 2 | 2 | 3 | 2 | 2 | 1 | 9  | 3 |
| 76 | GRR—-4685 | 0,5326 | REGULAR | 1 | 8  | 1 | 1 | 8 | 1 | 5 | 5 | 12 | 2  | 1 | 1 | 2 | 2 | 4 | 5 | 3 | 1 | 7  | 3 |

| ID |    |    |    |    |    |    |    |    |    | Perg | untas | ,  |    |    |    |    |    |    |    |    |
|----|----|----|----|----|----|----|----|----|----|------|-------|----|----|----|----|----|----|----|----|----|
| טו | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30   | 31    | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 |
| 1  | 2  | 5  | 3  | 2  | 1  | 6  | 1  | 6  | 3  | 8    | 8     | 5  | 1  | 4  | 1  | 1  | 2  | 9  | 2  | 13 |
| 2  | 2  | 5  | 5  | 1  | 4  | 6  | 3  | 6  | 2  | 4    | 3     | 7  | 5  | 8  | 1  | 1  | 2  | 5  | 2  | 13 |
| 3  | 1  | 2  | 3  | 1  | 1  | 3  | 1  | 6  | 5  | 4    | 8     | 1  | 1  | 4  | 4  | 1  | 2  | 7  | 2  | 13 |
| 4  | 1  | 5  | 1  | 2  | 4  | 3  | 2  | 1  | 6  | 3    | 8     | 6  | 3  | 4  | 4  | 32 | 2  | 5  | 2  | 13 |
| 5  | 2  | 5  | 5  | 1  | 4  | 6  | 3  | 6  | 2  | 1    | 3     | 2  | 3  | 8  | 4  | 1  | 2  | 5  | 2  | 13 |
| 6  | 1  | 5  | 2  | 1  | 1  | 6  | 2  | 6  | 6  | 3    | 8     | 2  | 6  | 10 | 1  | 1  | 2  | 3  | 2  | 13 |
| 7  | 1  | 5  | 2  | 1  | 1  | 3  | 2  | 6  | 1  | 4    | 1     | 2  | 5  | 1  | 1  | 1  | 2  | 5  | 2  | 13 |
| 8  | 1  | 5  | 1  | 4  | 3  | 1  | 2  | 6  | 2  | 5    | 2     | 2  | 4  | 2  | 1  | 1  | 2  | 8  | 2  | 13 |
| 9  | 1  | 5  | 5  | 4  | 4  | 6  | 1  | 6  | 6  | 3    | 6     | 2  | 3  | 4  | 1  | 1  | 1  | 4  | 2  | 13 |
| 10 | 1  | 5  | 1  | 5  | 3  | 4  | 1  | 6  | 6  | 1    | 7     | 2  | 6  | 10 | 4  | 32 | 2  | 4  | 2  | 13 |
| 11 | 4  | 1  | 1  | 4  | 1  | 3  | 1  | 6  | 5  | 4    | 7     | 7  | 5  | 4  | 1  | 1  | 2  | 5  | 2  | 13 |
| 12 | 2  | 5  | 1  | 5  | 3  | 5  | 1  | 6  | 6  | 3    | 8     | 7  | 5  | 1  | 1  | 1  | 2  | 5  | 2  | 13 |
| 13 | 2  | 3  | 4  | 4  | 1  | 4  | 2  | 6  | 2  | 2    | 3     | 2  | 3  | 4  | 4  | 1  | 2  | 4  | 2  | 13 |

| 14 | 1 | 5 | 5 | 4 | 4 | 6 | 3 | 6 | 2 | 3 | 3 | 2 | 1 | 4  | 1 | 1  | 2 | 1 | 2 | 13 |
|----|---|---|---|---|---|---|---|---|---|---|---|---|---|----|---|----|---|---|---|----|
| 15 | 2 | 2 | 5 | 1 | 4 | 6 | 1 | 6 | 4 | 1 | 6 | 2 | 3 | 4  | 1 | 1  | 2 | 6 | 2 | 13 |
| 16 | 3 | 5 | 5 | 1 | 4 | 6 | 1 | 6 | 5 | 3 | 4 | 6 | 1 | 4  | 1 | 1  | 2 | 5 | 2 | 13 |
| 17 | 2 | 5 | 5 | 1 | 4 | 5 | 3 | 3 | 1 | 1 | 1 | 2 | 1 | 4  | 4 | 1  | 2 | 5 | 2 | 13 |
| 18 | 1 | 5 | 2 | 2 | 4 | 6 | 1 | 6 | 6 | 8 | 8 | 8 | 6 | 4  | 1 | 1  | 2 | 9 | 2 | 13 |
| 19 | 1 | 2 | 5 | 4 | 4 | 6 | 1 | 6 | 3 | 3 | 8 | 2 | 3 | 8  | 1 | 1  | 2 | 5 | 2 | 13 |
| 20 | 1 | 5 | 2 | 2 | 4 | 6 | 1 | 6 | 3 | 3 | 6 | 7 | 3 | 10 | 1 | 1  | 2 | 8 | 2 | 13 |
| 21 | 4 | 5 | 5 | 1 | 4 | 6 | 3 | 6 | 6 | 3 | 3 | 2 | 1 | 8  | 4 | 1  | 2 | 1 | 2 | 13 |
| 22 | 1 | 5 | 5 | 2 | 4 | 6 | 3 | 6 | 6 | 3 | 8 | 2 | 5 | 10 | 1 | 32 | 2 | 8 | 2 | 13 |
| 23 | 1 | 5 | 5 | 4 | 4 | 6 | 3 | 6 | 3 | 2 | 3 | 2 | 3 | 4  | 1 | 1  | 2 | 3 | 2 | 13 |
| 24 | 1 | 5 | 1 | 2 | 1 | 1 | 1 | 6 | 3 | 2 | 2 | 1 | 2 | 10 | 4 | 6  | 2 | 3 | 2 | 13 |
| 25 | 4 | 5 | 1 | 1 | 4 | 6 | 1 | 6 | 6 | 3 | 8 | 6 | 1 | 1  | 4 | 6  | 2 | 5 | 2 | 13 |
| 26 | 1 | 1 | 5 | 6 | 4 | 6 | 1 | 6 | 5 | 4 | 8 | 2 | 6 | 4  | 3 | 6  | 2 | 3 | 2 | 13 |
| 27 | 1 | 1 | 1 | 4 | 1 | 3 | 1 | 6 | 4 | 4 | 3 | 3 | 1 | 4  | 1 | 6  | 2 | 7 | 2 | 13 |
| 28 | 1 | 3 | 1 | 2 | 4 | 2 | 1 | 6 | 1 | 5 | 6 | 8 | 1 | 9  | 1 | 6  | 2 | 9 | 2 | 13 |
| 29 | 1 | 5 | 1 | 1 | 1 | 5 | 1 | 6 | 4 | 2 | 8 | 8 | 1 | 4  | 1 | 6  | 2 | 5 | 2 | 13 |
| 30 | 4 | 1 | 4 | 1 | 4 | 6 | 3 | 6 | 6 | 6 | 8 | 8 | 3 | 4  | 1 | 6  | 2 | 3 | 2 | 13 |
| 31 | 1 | 5 | 3 | 1 | 4 | 6 | 4 | 4 | 6 | 7 | 2 | 1 | 5 | 4  | 1 | 6  | 2 | 3 | 2 | 13 |
| 32 | 5 | 2 | 3 | 1 | 1 | 1 | 3 | 6 | 4 | 1 | 8 | 2 | 6 | 4  | 4 | 6  | 2 | 9 | 2 | 13 |
| 33 | 1 | 5 | 1 | 2 | 4 | 6 | 3 | 6 | 3 | 3 | 6 | 2 | 3 | 4  | 1 | 6  | 2 | 5 | 2 | 13 |
| 34 | 1 | 5 | 1 | 4 | 4 | 6 | 2 | 6 | 1 | 3 | 1 | 2 | 6 | 1  | 1 | 6  | 2 | 7 | 2 | 13 |
| 35 | 1 | 5 | 3 | 4 | 4 | 6 | 2 | 6 | 2 | 4 | 8 | 8 | 6 | 10 | 3 | 5  | 2 | 5 | 3 | 13 |
| 36 | 1 | 5 | 1 | 2 | 3 | 5 | 1 | 6 | 6 | 8 | 7 | 6 | 2 | 4  | 1 | 6  | 2 | 1 | 2 | 13 |
| 37 | 1 | 5 | 1 | 1 | 1 | 5 | 3 | 6 | 5 | 7 | 2 | 7 | 2 | 4  | 1 | 6  | 2 | 3 | 2 | 13 |
| 38 | 1 | 5 | 5 | 1 | 4 | 6 | 1 | 6 | 5 | 7 | 6 | 7 | 2 | 4  | 1 | 6  | 2 | 1 | 2 | 13 |
|    |   |   |   |   |   |   |   |   |   |   |   |   |   |    |   |    |   |   |   |    |

| 39 | 1 | 5 | 5 | 4 | 4 | 6 | 1 | 6 | 4 | 5 | 1 | 2 | 6 | 4  | 1 | 6 | 2 | 3 | 2 | 13 |
|----|---|---|---|---|---|---|---|---|---|---|---|---|---|----|---|---|---|---|---|----|
| 40 | 2 | 5 | 5 | 1 | 1 | 3 | 2 | 6 | 2 | 7 | 3 | 2 | 5 | 4  | 2 | 6 | 2 | 1 | 2 | 13 |
| 41 | 1 | 5 | 3 | 2 | 1 | 6 | 1 | 6 | 6 | 7 | 8 | 8 | 3 | 10 | 1 | 6 | 2 | 3 | 2 | 13 |
| 42 | 3 | 5 | 2 | 1 | 3 | 5 | 2 | 6 | 6 | 1 | 8 | 7 | 6 | 10 | 1 | 6 | 2 | 9 | 2 | 13 |
| 43 | 2 | 5 | 1 | 2 | 1 | 3 | 1 | 6 | 1 | 3 | 7 | 5 | 1 | 4  | 1 | 6 | 2 | 3 | 2 | 13 |
| 44 | 3 | 5 | 5 | 4 | 1 | 6 | 2 | 6 | 4 | 1 | 1 | 6 | 5 | 4  | 4 | 1 | 2 | 4 | 2 | 13 |
| 45 | 1 | 5 | 1 | 1 | 4 | 6 | 3 | 6 | 6 | 3 | 3 | 6 | 1 | 5  | 1 | 1 | 2 | 7 | 2 | 13 |
| 46 | 2 | 5 | 1 | 1 | 1 | 5 | 3 | 6 | 4 | 5 | 7 | 2 | 1 | 4  | 1 | 1 | 2 | 3 | 2 | 13 |
| 47 | 4 | 5 | 2 | 6 | 2 | 2 | 1 | 6 | 3 | 3 | 8 | 2 | 6 | 4  | 3 | 1 | 2 | 8 | 2 | 13 |
| 48 | 1 | 5 | 2 | 2 | 4 | 6 | 1 | 6 | 5 | 2 | 7 | 4 | 6 | 4  | 1 | 1 | 2 | 3 | 2 | 13 |
| 49 | 3 | 2 | 5 | 4 | 4 | 6 | 3 | 6 | 1 | 1 | 1 | 2 | 1 | 8  | 1 | 1 | 2 | 5 | 2 | 13 |
| 50 | 1 | 5 | 5 | 1 | 4 | 6 | 1 | 6 | 3 | 3 | 6 | 2 | 5 | 9  | 1 | 1 | 2 | 3 | 2 | 13 |
| 51 | 3 | 5 | 5 | 4 | 4 | 6 | 1 | 6 | 6 | 3 | 3 | 2 | 1 | 8  | 1 | 1 | 2 | 8 | 2 | 13 |
| 52 | 2 | 5 | 1 | 2 | 4 | 6 | 1 | 6 | 6 | 8 | 8 | 2 | 3 | 4  | 1 | 1 | 2 | 4 | 2 | 13 |
| 53 | 1 | 5 | 5 | 2 | 4 | 6 | 2 | 6 | 1 | 3 | 6 | 2 | 3 | 4  | 1 | 1 | 2 | 5 | 2 | 13 |
| 54 | 1 | 3 | 3 | 6 | 1 | 3 | 2 | 6 | 6 | 3 | 7 | 8 | 6 | 10 | 4 | 1 | 2 | 1 | 2 | 13 |
| 55 | 5 | 5 | 5 | 4 | 4 | 6 | 4 | 5 | 2 | 4 | 3 | 2 | 3 | 8  | 4 | 1 | 2 | 5 | 2 | 13 |
| 56 | 2 | 5 | 5 | 1 | 4 | 6 | 3 | 6 | 1 | 4 | 1 | 2 | 5 | 4  | 3 | 1 | 2 | 8 | 2 | 13 |
| 57 | 1 | 5 | 2 | 4 | 3 | 4 | 1 | 6 | 6 | 1 | 8 | 2 | 1 | 10 | 1 | 1 | 2 | 5 | 2 | 13 |
| 58 | 3 | 5 | 5 | 1 | 1 | 6 | 2 | 6 | 6 | 3 | 8 | 2 | 6 | 10 | 4 | 1 | 2 | 8 | 2 | 13 |
| 59 | 2 | 2 | 1 | 4 | 4 | 6 | 3 | 6 | 6 | 5 | 3 | 2 | 5 | 4  | 1 | 1 | 2 | 8 | 2 | 13 |
| 60 | 1 | 5 | 2 | 2 | 1 | 3 | 4 | 2 | 6 | 7 | 8 | 2 | 5 | 1  | 4 | 1 | 2 | 5 | 2 | 13 |
| 61 | 1 | 5 | 5 | 4 | 4 | 6 | 1 | 6 | 6 | 5 | 3 | 2 | 3 | 8  | 4 | 1 | 2 | 5 | 1 | 2  |
| 62 | 2 | 5 | 5 | 1 | 4 | 6 | 3 | 6 | 5 | 1 | 7 | 2 | 5 | 1  | 4 | 1 | 2 | 5 | 3 | 13 |
| 63 | 4 | 5 | 4 | 2 | 4 | 6 | 2 | 6 | 2 | 3 | 7 | 7 | 5 | 4  | 4 | 1 | 2 | 5 | 2 | 13 |
|    |   |   |   |   |   |   |   |   |   |   |   |   |   |    |   |   |   |   |   |    |

| 64 | 2 | 5 | 5 | 1 | 4 | 6 | 2 | 6 | 6 | 1 | 8 | 6 | 5 | 10 | 1 | 1  | 2 | 5 | 2 | 13 |
|----|---|---|---|---|---|---|---|---|---|---|---|---|---|----|---|----|---|---|---|----|
| 65 | 1 | 5 | 1 | 1 | 4 | 6 | 4 | 1 | 6 | 8 | 7 | 7 | 5 | 4  | 4 | 1  | 2 | 5 | 2 | 13 |
| 66 | 2 | 5 | 5 | 1 | 4 | 6 | 1 | 6 | 6 | 8 | 8 | 2 | 3 | 10 | 1 | 1  | 2 | 5 | 2 | 13 |
| 67 | 1 | 4 | 5 | 4 | 3 | 5 | 4 | 1 | 1 | 7 | 7 | 2 | 1 | 4  | 1 | 1  | 2 | 5 | 2 | 13 |
| 68 | 1 | 5 | 5 | 1 | 4 | 6 | 2 | 4 | 6 | 3 | 8 | 7 | 1 | 8  | 1 | 32 | 2 | 3 | 2 | 13 |
| 69 | 4 | 5 | 5 | 2 | 4 | 6 | 1 | 6 | 6 | 3 | 7 | 2 | 6 | 5  | 1 | 1  | 2 | 3 | 2 | 13 |
| 70 | 2 | 1 | 5 | 1 | 4 | 6 | 1 | 6 | 6 | 2 | 8 | 5 | 5 | 4  | 4 | 1  | 2 | 5 | 2 | 13 |
| 71 | 4 | 5 | 2 | 2 | 1 | 3 | 3 | 6 | 5 | 4 | 8 | 1 | 3 | 4  | 1 | 1  | 2 | 3 | 3 | 13 |
| 72 | 1 | 5 | 5 | 1 | 4 | 6 | 3 | 6 | 3 | 5 | 8 | 5 | 3 | 4  | 1 | 1  | 2 | 7 | 2 | 13 |
| 73 | 4 | 5 | 5 | 1 | 4 | 6 | 1 | 6 | 5 | 2 | 3 | 2 | 1 | 10 | 4 | 1  | 2 | 9 | 2 | 13 |
| 74 | 1 | 5 | 5 | 1 | 4 | 6 | 1 | 6 | 6 | 2 | 1 | 3 | 5 | 8  | 2 | 1  | 2 | 5 | 2 | 13 |
| 75 | 2 | 5 | 5 | 6 | 4 | 6 | 1 | 6 | 1 | 4 | 1 | 7 | 5 | 4  | 1 | 1  | 2 | 5 | 2 | 13 |
| 76 | 1 | 5 | 1 | 2 | 1 | 4 | 1 | 6 | 4 | 1 | 3 | 4 | 1 | 7  | 4 | 1  | 2 | 5 | 2 | 13 |
|    |   |   |   |   |   |   |   |   |   |   |   |   |   |    |   |    |   |   |   |    |

#### 8.2 Dados Estatísticos Base de Dados R

```
> summary(mat)
                                                                     Matriculado
    sq_id
                MATRICULA
                                      IRA
                                                    GRP_Rend
                                                                                      RENDIMENTO
                                                                                                         1
Min. : 1.00
                Length:76
                                  Min. :0.0000
                                                  Length:76
                                                                    Min. :0.0000
                                                                                    Min.
                                                                                          :1.000
                                                                                                   Min.
                                                                                                         :1.000
                                                                                                                  Min. : 3.000
1st Qu.:19.75
                Class :character
                                  1st Qu.:0.0000
                                                  class :character
                                                                   1st Qu.:0.0000
                                                                                    1st Qu.:1.000
                                                                                                   1st Qu.:1.000
                                                                                                                  1st Qu.: 5.000
Median :38.50
                Mode :character
                                  Median :0.0000
                                                  Mode :character
                                                                    Median :1.0000
                                                                                    Median :2.000
                                                                                                   Median :2.000
                                                                                                                  Median : 9.000
Mean :38.50
                                  Mean :0.2797
                                                                    Mean :0.5132
                                                                                    Mean
                                                                                         :2.039
                                                                                                   Mean :1.526
3rd Qu.:57.25
                                  3rd Qu.: 0.5749
                                                                    3rd Qu.:1.0000
                                                                                    3rd Qu.:3.000
                                                                                                   3rd Qu.:2.000
Max. :76.00
                                  Max. :0.9194
                                                                    Max.
                                                                           :1.0000
                                                                                    Max.
                                                                                           :5.000
                                                                                                   мах.
                                                                                                          :2.000
                                                                                                                  мах.
                                                                                                 9
      3
                                    5
                                                    6
                                                                                  8
                                                                                                                  10
Min. :1.000
               Min. :1.000
                               Min. :1.000
                                              Min. :1.000
                                                             Min. :1.000
                                                                            Min. :2.000
                                                                                           Min. : 1.000
                                                                                                            Min. : 1.000
               1st Qu.:1.000
                               1st Qu.:2.000
                                              1st Qu.:1.000
                                                             1st Qu.:2.000
                                                                            1st Qu.:3.000
                                                                                           1st Qu.: 2.000
1st Ou.:1.000
                                                                                                            1st Ou.: 2.000
                                              Median :3.000
Median :1.000
               Median :1.000
                               Median :8.000
                                                             Median :4.000
                                                                            Median :5.000
                                                                                           Median :10.000
                                                                                                            Median :10.000
Mean :1.276
                               Mean :5.526
                                              Mean :2.868
                                                             Mean :4.039
                                                                            Mean :4.474
                                                                                           Mean : 7.237
                Mean :1.421
                                                                                                            Mean : 7.053
                                                                            3rd Qu.:6.000
3rd Qu.:1.000
                3rd Qu.:1.000
                               3rd Qu.:8.000
                                              3rd Qu.:4.000
                                                             3rd Qu.:5.000
                                                                                           3rd Qu.:12.000
                                                                                                            3rd Qu.:11.000
                                                             Max. :8.000
Max. :3.000
               мах.
                     :6.000
                               Max. :8.000
                                              Max. :8.000
                                                                            Max. :8.000
                                                                                           Max.
                                                                                                :12.000
                                                                                                            Max. :12.000
      11
                     12
                                    13
                                                14
                                                               15
                                                                              16
                                                                                             17
                                                                                                            18
                                                                                                       Min.
Min. :1.000
               Min.
                     :1.000
                               Min. :1
                                          Min.
                                               :1.000
                                                         Min. :1.000
                                                                        Min. :1.000
                                                                                        Min. :1.000
                                                                                                            :1.000
1st Qu.:1.000
               1st Qu.:1.000
                               1st Qu.:1
                                          1st Ou.:1.000
                                                         1st Qu.:2.750
                                                                        1st Qu.:2.000
                                                                                        1st Qu.:2.000
                                                                                                       1st Qu.:1.000
Median :2.000
               Median :2.000
                               Median :2
                                          Median :2.000
                                                         Median :3.000
                                                                        Median :3.000
                                                                                        Median :3.000
                                                                                                       Median :1.000
                Mean :2.224
                                          Mean :1.737
                                                         Mean :3.289
                                                                        Mean :3.132
Mean :1.961
                               Mean :2
                                                                                        Mean
                                                                                             :3.039
                                                                                                       Mean :1.461
                                                         3rd Qu.:4.000
3rd Qu.:3.000
                3rd Qu.:3.000
                               3rd Qu.:2
                                          3rd Qu.:2.000
                                                                        3rd Qu.:4.000
                                                                                        3rd Qu.:4.000
                                                                                                       3rd Qu.:1.000
Max. :3.000
               Max.
                      :4.000
                               Max.
                                     :6
                                          мах.
                                                :3.000
                                                         мах.
                                                                :6.000
                                                                        Max.
                                                                               :5.000
                                                                                        Max.
                                                                                              :5.000
                                                                                                       Max.
                                                                                                             :4.000
      19
                      20
                                     21
                                                     22
                                                                    23
                                                                                   24
                                                                                                  25
                                                                                                                  26
                                                              Min. :1.000
Min. : 0.000
                      :1.000
                               Min. :1.000
                                               Min. :1.000
                                                                             Min. :1.000
                                                                                            Min. :1.000
                                                                                                            Min. :1.000
                Min.
1st Qu.: 0.750
                1st Qu.:3.000
                                1st Qu.:1.000
                                               1st Qu.:5.000
                                                              1st Qu.:1.000
                                                                             1st Qu.:1.000
                                                                                            1st Qu.:1.000
                                                                                                            1st Qu.:4.750
Median : 8.500
                Median :3.000
                                Median :1.000
                                               Median:5.000
                                                              Median :3.500
                                                                             Median :2.000
                                                                                            Median :4.000
                                                                                                            Median :6.000
                Mean :2.947
                                Mean :1.855
                                               Mean :4.408
                                                                             Mean :2.368
Mean : 6.921
                                                              Mean :3.276
                                                                                            Mean :3.053
3rd Qu.:11.000
                 3rd Qu.:3.000
                                3rd Qu.:2.000
                                               3rd Qu.:5.000
                                                              3rd Qu.:5.000
                                                                             3rd Qu.:4.000
                                                                                             3rd Qu.:4.000
                                                                                                            3rd Qu.:6.000
     :12.000
                мах.
                     :5.000
                                Max. :5.000
                                               мах.
                                                    :5.000
                                                              мах.
                                                                   :5.000
                                                                             мах.
                                                                                  :6.000
                                                                                            Max. :4.000
                                                                                                            Max. :6.000
      27
                     28
                                    29
                                                    30
                                                                   31
                                                                                  32
                                                                                                 33
     :1.000
               Min.
                     :1.000
                               Min. :1.000
                                              Min. :1.000
                                                             Min. :1.000
                                                                            Min. :1.000
                                                                                           Min. :1.000
                                                                                                           Min. : 1.000
1st Qu.:1.000
               1st Qu.:6.000
                               1st Qu.:2.750
                                              1st Qu.:2.000
                                                             1st Qu.:3.000
                                                                            1st Qu.:2.000
                                                                                           1st Qu.:1.000
                                                                                                           1st Ou.: 4.000
Median :2.000
               Median :6.000
                               Median :5.000
                                              Median :3.000
                                                             Median :7.000
                                                                            Median :2.000
                                                                                           Median :3.000
                                                                                                           Median: 4.000
Mean :1.908
               Mean :5.645
                               Mean :4.171
                                              Mean :3.711
                                                             Mean :5.447
                                                                            Mean :3.829
                                                                                            Mean :3.487
                                                                                                           Mean : 5.566
                3rd Qu.:6.000
                               3rd Qu.:6.000
                                              3rd Qu.:5.000
                                                             3rd Qu.:8.000
                                                                             3rd Qu.:6.000
                                                                                            3rd Qu.:5.000
3rd Qu.:3.000
                                                                                                           3rd Qu.: 8.000
     :4.000
               Max. :6.000
                               Max. :6.000
                                              Max. :8.000
                                                             Max. :8.000
                                                                            Max. :8.000
                                                                                           Max. :6.000
                                                                                                           Max. :10.000
мах.
                                  37
                                                                               40
      35
                 36
                                                 38
                                                                39
      :1
                 : 1.000
                            Min. :1.000
                                           Min. :1.000
                                                          Min. :1.000
                                                                          Min. : 2.00
                                           1st Qu.:3.000
1st Ou.:1
            1st Qu.: 1.000
                            1st Qu.:2.000
                                                          1st Qu.:2.000
Median :1
            Median : 1.000
                            Median :2.000
                                           Median :5.000
                                                          Median :2.000
                                                                          Median :13.00
            Mean : 3.934
                            Mean :1.987
                                           Mean
                                                :4.974
                                                          Mean
                                                                :2.026
                                                                              :12.86
                                                                          Mean
            3rd ou.: 6.000
                            3rd Qu.:2.000
                                           3rd Qu.:6.250
                                                          3rd Ou.:2.000
                                                                          3rd ou.:13.00
Max. :4
          Max. :32.000
                            Max. :2.000
                                           Max.
                                                 :9.000
                                                          Max.
                                                                 :3.000
```

Figura 25 - Dados Estatísticos

### 8.3 Correlograma

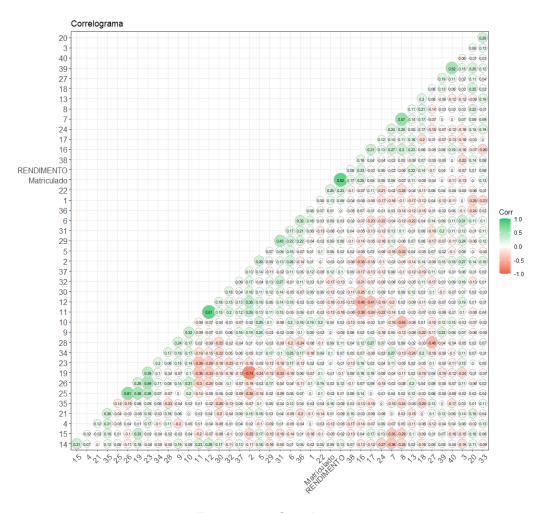


Figura 26 - Correlograma

# 9 Questionário Socioeconômico UFPR

| 1 | 0 l 2 /D: l / : )                     |    | DESCRIÇÃO          |
|---|---------------------------------------|----|--------------------|
|   | Qual o seu sexo? (Biológico)          | 1  | Masculino          |
| 1 | Qual o seu sexo? (Biológico)          | 2  | Feminino           |
| 2 | Quantos anos você completará em 2020? | 1  | Menos de 16 anos   |
| 2 | Quantos anos você completará em 2020? | 2  | 16 anos            |
| 2 | Quantos anos você completará em 2020? | 3  | 17 anos            |
| 2 | Quantos anos você completará em 2020? | 4  | 18 anos            |
| 2 | Quantos anos você completará em 2020? | 5  | 19 anos            |
| 2 | Quantos anos você completará em 2020? | 6  | 20 anos            |
| 2 | Quantos anos você completará em 2020? | 7  | 21 anos            |
| 2 | Quantos anos você completará em 2020? | 8  | 22 anos            |
| 2 | Quantos anos você completará em 2020? | 9  | 23 anos            |
| 2 | Quantos anos você completará em 2020? | 10 | mais de 23 anos    |
| 3 | Qual o seu estado civil?              | 1  | Solteiro(a)        |
| 3 | Qual o seu estado civil?              | 2  | Casado(a)          |
| 3 | Qual o seu estado civil?              | 3  | Outro              |
| 4 | Qual o Estado em que nasceu?          | 1  | Paraná             |
| 4 | Qual o Estado em que nasceu?          | 2  | Santa Catarina     |
| 4 | Qual o Estado em que nasceu?          | 3  | Rio Grande do Sul  |
| 4 | Qual o Estado em que nasceu?          | 4  | São Paulo          |
| 4 | Qual o Estado em que nasceu?          | 5  | Mato Grosso do Sul |
| 4 | Qual o Estado em que nasceu?          | 6  | Outro              |

| 5 | Qual o local de sua residência?       | 1 | Curitiba                                     |
|---|---------------------------------------|---|--|
| 5 | Qual o local de sua residência?       | 2 | Demais municípios da Região Metropolitana    |
| 5 | Qual o local de sua residência?       | 3 | Interior do Paraná                           |
| 5 | Qual o local de sua residência?       | 4 | Santa Catarina                               |
| 5 | Qual o local de sua residência?       | 5 | Rio Grande do Sul                            |
| 5 | Qual o local de sua residência?       | 6 | São Paulo                                    |
| 5 | Qual o local de sua residência?       | 7 | Mato Grosso do Sul                           |
| 5 | Qual o local de sua residência?       | 8 | Outro  |
| 6 | Qual a sua situação quanto à moradia? | 1 | Mora em casa dos pais, quitada ou financiada |
| 6 | Qual a sua situação quanto à moradia? | 2 | Mora em casa dos pais, alugada               |
| 6 | Qual a sua situação quanto à moradia? | 3 | Mora em casa própria, quitada ou financiada  |
| 6 | Qual a sua situação quanto à moradia? | 4 | Mora em casa alugada, paga por você          |
| 6 | Qual a sua situação quanto à moradia? | 5 | Mora em república, casa de estudante,        |
| U | Quai a sua situação quanto a moradia: | J | pensão ou pensionato                         |
| 6 | Qual a sua situação quanto à moradia? | 6 | Mora em casa de parentes ou amigos           |
| 6 | Qual a sua situação quanto à moradia? | 7 | Mora em casa alugada para você,              |
| U | Quai a sua situação quanto a moradia: | ' | paga por seus pais                           |
| 6 | Qual a sua situação quanto à moradia? | 8 | Outros                                       |
| 7 | Qual o nível de instrução do seu pai? | 1 | Sem escolaridade                             |
| 7 | Qual o nível de instrução do seu pai? | 2 | Ensino Fundamental incompleto                |
| 7 | Qual o nível de instrução do seu pai? | 3 | Ensino Fundamental completo                  |
| 7 | Qual o nível de instrução do seu pai? | 4 | Ensino Médio incompleto                      |
| 7 | Qual o nível de instrução do seu pai? | 5 | Ensino Médio completo                        |
| 7 | Qual o nível de instrução do seu pai? | 6 | Superior incompleto                          |
| 7 | Qual o nível de instrução do seu pai? | 7 | Superior completo                            |
|   |                                       |   |  |

| 7 | Qual o nível de instrução do seu pai? | 8 | Não sei informar   |
|---|---------------------------------------|---|--|
| 8 | Qual o nível de instrução da sua mãe? | 1 | Sem escolaridade   |
| 8 | Qual o nível de instrução da sua mãe? | 2 | Ensino Fundamental incompleto                            |
| 8 | Qual o nível de instrução da sua mãe? | 3 | Ensino Fundamental completo                              |
| 8 | Qual o nível de instrução da sua mãe? | 4 | Ensino Médio incompleto                                  |
| 8 | Qual o nível de instrução da sua mãe? | 5 | Ensino Médio completo                                    |
| 8 | Qual o nível de instrução da sua mãe? | 6 | Superior incompleto                                      |
| 8 | Qual o nível de instrução da sua mãe? | 7 | Superior completo  |
| 8 | Qual o nível de instrução da sua mãe? | 8 | Não sei informar   |
| 0 | Qual a principal ocupação do seu pai? | 1 | Funcionário público do governo Federal,                  |
| 9 | Quai a principai ocupação do seu par: | 1 | Estadual ou Municipal                                    |
| 9 | Qual a principal ocupação do seu pai? | 2 | Empregado de empresa                                     |
| 9 | Qual a principal ocupação do seu pai? | 3 | Sócio ou proprietário de empresa                         |
| 9 | Qual a principal ocupação do seu pai? | 4 | Trabalho remunerado por conta própria,                   |
| 9 | Quai a principai ocupação do seu par: | 4 | com auxílio de parentes e/ou de familiares               |
| 9 | Qual a principal ocupação do seu pai? | 5 | Trabalho remunerado por conta própria,                   |
| 9 | Quai a principal ocupação do seu par: | 5 | com empregados   |
| 9 | Qual a principal ocupação do seu pai? | 6 | Artista (pintor, escritor, músico, cantor,               |
| 9 | Quai a principal ocupação do seu par: | U | ator, etc.)  |
|   |                                       |   | Trabalha em entidade, organização ou                     |
| 9 | Qual a principal ocupação do seu pai? | 7 | instituição não governamental de cunho                   |
|   |                                       |   | filantrópico, assistencial, religioso, de lazer ou outro |
| 9 | Qual a principal ocupação do seu pai? | 8 | Parlamentar ou cargo eleitoral, diplomata, militar       |
| 9 | Qual a principal ocupação do seu pai? | 9 | Atleta profissional                                      |
|   |                                       |   |  |

| 9  | Qual a principal ocupação do seu pai? | 11 | Não trabalha   |
|----|---------------------------------------|----|--|
| 9  | Qual a principal ocupação do seu pai? | 12 | Outros   |
| 10 | Qual a principal ocupação da sua mãe? | 1  | Funcionária pública do governo Federal,                    |
| 10 | Quai a principal ocupação da sua mae: | 1  | Estadual ou Municipal                                      |
| 10 | Qual a principal ocupação da sua mãe? | 2  | Empregada de empresa                                       |
| 10 | Qual a principal ocupação da sua mãe? | 3  | Sócia ou proprietária de empresa                           |
| 10 | Qual a principal ocupação da sua mãe? | 4  | Trabalho remunerado por conta própria, com                 |
| 10 | Quai a principal ocupação da sua mae: | 4  | auxílio de parentes e/ou de familiares                     |
| 10 | Qual a principal ocupação da sua mãe? | 5  | Trabalho remunerado por conta própria, com                 |
| 10 | Quai a principal ocupação da sua mae: | 5  | empregados   |
| 10 | Qual a principal ocupação da sua mãe? | 6  | Artista (pintora, escritora, música, cantora, atriz, etc.) |
|    |                                       |    | Trabalha em entidade, organização ou instituição           |
| 10 | Qual a principal ocupação da sua mãe? | 7  | não governamental de cunho filantrópico,                   |
|    |                                       |    | assistencial, religioso, de lazer ou outro                 |
| 10 | Qual a principal ocupação da sua mãe? | 8  | Parlamentar ou cargo eleitoral, diplomata, militar         |
| 10 | Qual a principal ocupação da sua mãe? | 9  | Atleta profissional  |
| 10 | Qual a principal ocupação da sua mãe? | 10 | Trabalha em casa e/ou não tem atividade remunerada         |
| 10 | Qual a principal ocupação da sua mãe? | 11 | Não trabalha   |
| 10 | Qual a principal ocupação da sua mãe? | 12 | Outros   |
| 11 | Você trabalha atualmente em atividade | 1  | Não  |
| 11 | remunerada?                           | 1  | IVaO   |
| 11 | Você trabalha atualmente em atividade | 2  | Sim - meio período   |
| 11 | remunerada?                           | 2  | Siii - Meio periodo  |
| 11 | Você trabalha atualmente em atividade | 3  | Sim - período integral                                     |
| 11 | remunerada?                           | 3  | Siiii - periodo iiitegiai                                  |

| 12 | Qual é a sua participação na vida   | 1 | Não tenho trabalho nem contribuo para o                  |
|----|-------------------------------------|---|--|
|    | econômica de seu grupo familiar?    | 1 | sustento da família                                      |
| 12 | Qual é a sua participação na vida   | 2 | Trabalho unicamente para cobrir minhas despesas          |
|    | econômica de seu grupo familiar?    | 2 |  |
| 12 | Qual é a sua participação na vida   | 3 | Trabalho e contribuo em parte para o sustento da família |
|    | econômica de seu grupo familiar?    | 3 |  |
| 12 | Qual é a sua participação na vida   | 4 | Trabalho e sou o principal responsável pelo              |
|    | econômica de seu grupo familiar?    | 4 | sustento da família                                      |
| 13 | A renda total mensal de sua família | 1 | Até R\$ 1.448,00   |
|    | se situa na faixa:                  | 1 |  |
| 13 | A renda total mensal de sua família | 2 | De R1.449,00aR 2.896,00                                  |
|    | se situa na faixa:                  | 2 |  |
| 13 | A renda total mensal de sua família | 3 | De R2.897,00aR 4.344,00                                  |
|    | se situa na faixa:                  | 3 |  |
| 13 | A renda total mensal de sua família | 4 | De R4.345,00aR 5.792,00                                  |
|    | se situa na faixa:                  | 4 |  |
| 13 | A renda total mensal de sua família | 5 | De R5.793,00aR 8.688,00                                  |
|    | se situa na faixa:                  | 5 |  |
| 13 | A renda total mensal de sua família | 6 | De R8.689,00aR 11.584,00                                 |
|    | se situa na faixa:                  | Ü |  |
| 13 | A renda total mensal de sua família | 7 | De R11.585,00aR 14.480,00                                |
|    | se situa na faixa:                  | , |  |
| 13 | A renda total mensal de sua família | 8 | Acima de R\$ 14.481,00                                   |
|    | se situa na faixa:                  | 0 |  |
|    |                                     |   |  |

| 14 | Quantas pessoas contribuem para a | 1 | Uma          |
|----|-----------------------------------|---|--------------|
|    | obtenção da renda familiar?       | 1 |              |
| 14 | Quantas pessoas contribuem para a | 2 | Duas         |
| 14 | obtenção da renda familiar?       | 2 | Duds         |
| 14 | Quantas pessoas contribuem para a | 3 | Três         |
| 14 | obtenção da renda familiar?       | 3 | ites         |
| 14 | Quantas pessoas contribuem para a | 4 | Quatro       |
| 14 | obtenção da renda familiar?       | 4 | Quatro       |
| 14 | Quantas pessoas contribuem para a | 5 | Cinco        |
| 14 | obtenção da renda familiar?       | 5 | Cinco        |
| 14 | Quantas pessoas contribuem para a | 6 | Seis ou mais |
| 14 | obtenção da renda familiar?       |   |              |
| 15 | Quantas pessoas são sustentadas   | 1 | Uma          |
| 15 | com a renda familiar?             |   |              |
| 15 | Quantas pessoas são sustentadas   | 2 | Duas         |
|    | com a renda familiar?             |   |              |
| 15 | Quantas pessoas são sustentadas   | 3 | Três         |
| 15 | com a renda familiar?             | 3 |              |
| 15 | Quantas pessoas são sustentadas   | 4 | Quatro       |
| 15 | com a renda familiar?             | 7 | Quatro       |
| 15 | Quantas pessoas são sustentadas   | 5 | Cinco        |
| 10 | com a renda familiar?             | 3 | Cirico       |
| 15 | Quantas pessoas são sustentadas   | 6 | Seis ou mais |
| 15 | com a renda familiar?             | О | Seis ou mais |
|    |                                   |   |              |

| 16 | Com que idade você começou a    | 1 | Antes dos 14 anos                            |
|----|---------------------------------|---|--|
| 10 | exercer atividade remunerada?   | 1 |  |
| 16 | Com que idade você começou a    | 2 | Entre 14 e 16 anos                           |
| 10 | exercer atividade remunerada?   | 2 | Little 14 e 10 anos                          |
| 16 | Com que idade você começou a    | 3 | Entre 16 e 18 anos                           |
| 10 | exercer atividade remunerada?   | 3 | Little 10 e 10 anos                          |
| 16 | Com que idade você começou a    | 4 | Após 18 anos                                 |
| 10 | exercer atividade remunerada?   | 4 | Apos 10 anos                                 |
| 16 | Com que idade você começou a    | 5 | Nunca trabalhei                              |
| 10 | exercer atividade remunerada?   | 5 | ivunca tradainei                             |
| 17 | Durante o curso, você terá      | 1 | Sim, mas apenas nos últimos anos             |
| 11 | obrigatoriamente que trabalhar? | 1 |  |
| 17 | Durante o curso, você terá      | 2 | Sim, desde o primeiro ano, em tempo parcial  |
| 11 | obrigatoriamente que trabalhar? | ۷ |  |
| 17 | Durante o curso, você terá      | 3 | Sim, desde o primeiro ano, em tempo integral |
| 11 | obrigatoriamente que trabalhar? | 3 |  |
| 17 | Durante o curso, você terá      | 4 | Não sei                                      |
| 11 | obrigatoriamente que trabalhar? | 4 |  |
| 17 | Durante o curso, você terá      | 5 | Não  |
| 17 | obrigatoriamente que trabalhar? | 5 | INAO   |
| 18 | Como fez seus estudos do ensino | 1 | Integralmente em escela pública              |
| 10 | fundamental e ensino médio?     | 1 | Integralmente em escola pública              |
| 10 | Como fez seus estudos do ensino | 2 | Integralmente em escala particular           |
| 18 | fundamental e ensino médio?     | 2 | Integralmente em escola particular           |
|    |                                 |   |  |

| 18 | Como fez seus estudos do ensino         | 3        | Maior parte em escola pública             |
|----|---|----------|---|
|    | fundamental e ensino médio?             | 3        |   |
| 18 | Como fez seus estudos do ensino         | 4        | Maior parte em escola particular          |
| 10 | fundamental e ensino médio?             | 4        | ivialor parte em escola particular        |
| 18 | Como fez seus estudos do ensino         | 5        | Em escolas comunitárias/CNEC ou outro     |
| 10 | fundamental e ensino médio?             | 5        | LITT escolas comunicarias/ CIVEC ou outro |
| 19 | Em que ano você concluiu (ou concluirá) | 0        | Antes de 2008                             |
| 19 | o curso de Ensino Médio?                | U        | Antes de 2006                             |
| 19 | Em que ano você concluiu (ou concluirá) | 1        | 2008                                      |
| 19 | o curso de Ensino Médio?                | 1        |   |
| 19 | Em que ano você concluiu (ou concluirá) | 2        | 2009                                      |
| 19 | o curso de Ensino Médio?                | 2        |   |
| 19 | Em que ano você concluiu (ou concluirá) | 3        | 2010                                      |
| 19 | o curso de Ensino Médio?                | 3        |   |
| 19 | Em que ano você concluiu (ou concluirá) | 4        | 2011                                      |
| 19 | o curso de Ensino Médio?                | 4        |   |
| 19 | Em que ano você concluiu (ou concluirá) | 5        | 2012                                      |
| 19 | o curso de Ensino Médio?                | 5        |   |
| 19 | Em que ano você concluiu (ou concluirá) | 6        | 2013                                      |
| 19 | o curso de Ensino Médio?                | O        | 2013                                      |
| 19 | Em que ano você concluiu (ou concluirá) | 7        | 2014                                      |
| 19 | o curso de Ensino Médio?                |          | 2014                                      |
| 19 | Em que ano você concluiu (ou concluirá) | 8        | 2015                                      |
| 19 | o curso de Ensino Médio?                | <b>O</b> | 2013                                      |
|    |   |          |   |

| 19 | Em que ano você concluiu (ou concluirá) | 9  | 2016                                   |
|----|---|----|--|
|    | o curso de Ensino Médio?                |    | 2010                                   |
| 19 | Em que ano você concluiu (ou concluirá) | 10 | 2017                                   |
|    | o curso de Ensino Médio?                |    | 2017                                   |
| 19 | Em que ano você concluiu (ou concluirá) | 11 | 2018                                   |
|    | o curso de Ensino Médio?                | 11 | 2010                                   |
| 19 | Em que ano você concluiu (ou concluirá) | 12 | 2019 ou após                           |
|    | o curso de Ensino Médio?                | 12 | 2019 Oil apos                          |
| 20 | No ensino médio, você fez (ou faz):     | 1  | Curso de magistério                    |
| 20 | No ensino médio, você fez (ou faz):     | 2  | Curso profissionalizante               |
| 20 | No ensino médio, você fez (ou faz):     | 3  | Ensino médio regular                   |
| 20 | No ensino médio, você fez (ou faz):     | 4  | Ensino médio com terceirão ou cursinho |
| 20 | No ensino médio, você fez (ou faz):     | 5  | Outro                                  |
| 21 | Em que turno você fez (ou faz) o        | 1  | Todo diurno                            |
| 21 | curso de ensino médio?                  |    | rodo diarno                            |
| 21 | Em que turno você fez (ou faz) o        | 2  | Todo noturno                           |
| 21 | curso de ensino médio?                  | ۷  | 1000 HOLUFHO                           |
| 21 | Em que turno você fez (ou faz) o        | 3  | Maior parte diurno                     |
| 21 | curso de ensino médio?                  | 3  | Maior parte diumo                      |
| 21 | Em que turno você fez (ou faz) o        | 4  | Maior parte noturno                    |
| 21 | curso de ensino médio?                  | 4  | Maior parte noturno                    |
| 21 | Em que turno você fez (ou faz) o        | 5  | Outro                                  |
| 21 | curso de ensino médio?                  | 3  | Outio                                  |
| 22 | Por quanto tempo você fez cursinho?     | 1  | Por menos de um semestre               |
| 22 | Por quanto tempo você fez cursinho?     | 2  | Por um semestre                        |
|    |   |    |  |

| 22 | Por quanto tempo você fez cursinho?                    | 3 | Por um ano  |
|----|--|---|---|
| 22 | Por quanto tempo você fez cursinho?                    | 4 | Por mais de um ano                                |
| 22 | Por quanto tempo você fez cursinho?                    | 5 | Não fiz cursinho                                  |
| 23 | Você já fez o vestibular em outros anos?               | 1 | Sim, este é o segundo ano que faço vestibular     |
| 23 | (não leve em conta a possível situação de "treineiro") | 1 | Sim, este e o segundo ano que raço vestibular     |
| 23 | Você já fez o vestibular em outros anos?               | 2 | Since and a financian and more form conditions.   |
| 23 | (não leve em conta a possível situação de "treineiro") | 2 | Sim, este é o terceiro ano que faço vestibular    |
| 23 | Você já fez o vestibular em outros anos?               | 2 | Simo anto financiato and must force went holders  |
| 23 | (não leve em conta a possível situação de "treineiro") | 3 | Sim, este é o quarto ano que faço vestibular      |
|    | Você já fez o vestibular em outros anos?               | 4 | C: C ::   |
| 23 | (não leve em conta a possível situação de "treineiro") |   | Sim, faço vestibular há mais de quatro anos       |
| 23 | Você já fez o vestibular em outros anos?               | 5 | Não, este é o primeiro ano em que faço vestibular |
| 23 | (não leve em conta a possível situação de "treineiro") |   |   |
|    | Considerando sua capacidade de leitura                 |   |   |
| 24 | em Língua Estrangeira, em que situação                 | 1 | Não leio em nenhuma Língua Estrangeira            |
|    | você melhor se enquadra?                               |   |   |
|    | Considerando sua capacidade de leitura                 |   |   |
| 24 | em Língua Estrangeira, em que situação                 | 2 | Leio apenas em Espanhol                           |
|    | você melhor se enquadra?                               |   |   |
| -  | Considerando sua capacidade de leitura                 |   |   |
| 24 | em Língua Estrangeira, em que situação                 | 3 | Leio apenas em Frânces                            |
|    | você melhor se enquadra?                               |   |   |
|    | Considerando sua capacidade de leitura                 |   |   |
| 24 | em Língua Estrangeira, em que situação                 | 4 | Leio apenas em Inglês                             |
|    | você melhor se enquadra?                               |   |   |
|    |  |   |   |

|    | Considerando sua capacidade de leitura     |   |   |
|----|--|---|---|
| 24 | em Língua Estrangeira, em que situação     | 5 | Leio apenas uma outra Língua Estrangeira            |
|    | você melhor se enquadra?                   |   |   |
|    | Considerando sua capacidade de leitura     |   |   |
| 24 | em Língua Estrangeira, em que situação     | 6 | Leio duas ou mais línguas estrangeiras              |
|    | você melhor se enquadra?                   |   |   |
| 25 | Você já iniciou algum curso superior?      | 1 | Sim, mas não concluí                                |
| 25 | Você já iniciou algum curso superior?      | 2 | Sim, estou cursando                                 |
| 25 | Você já iniciou algum curso superior?      | 3 | Sim, mas já concluí                                 |
| 25 | Você já iniciou algum curso superior?      | 4 | Não   |
|    | Se já iniciou um curso universitário, qual |   |   |
| 26 | sua principal razão para participar deste  | 1 | Por não ter gostado                                 |
|    | Processo Seletivo?                         |   |   |
|    | Se já iniciou um curso universitário, qual |   |   |
| 26 | sua principal razão para participar deste  | 2 | Por estar mais decidido e preparado                 |
|    | Processo Seletivo?                         |   |   |
|    | Se já iniciou um curso universitário, qual |   |   |
| 26 | sua principal razão para participar deste  | 3 | Por ter maior disponibilidade para fazer este curso |
|    | Processo Seletivo?                         |   |   |
|    | Se já iniciou um curso universitário, qual |   |   |
| 26 | sua principal razão para participar deste  | 4 | Porque este curso complementa os estudos já feitos  |
|    | Processo Seletivo?                         |   |   |
|    | Se já iniciou um curso universitário, qual |   |   |
| 26 | sua principal razão para participar deste  | 5 | Por desejar outra formação                          |
|    | Processo Seletivo?                         |   |   |

|    | Se já iniciou um curso universitário, qual |   |   |
|----|--|---|---|
| 26 | sua principal razão para participar deste  | 6 | Não me enquadro nesta situação  |
|    | Processo Seletivo?                         |   |   |
| 27 | Quanto à sua escolha por este curso,       | 1 | Absolutamente decidido  |
| 21 | você se considera:                         |   | Absolutamente decidido  |
| 27 | Quanto à sua escolha por este curso,       | 2 | Muito decidido  |
|    | você se considera:                         | 2 | Walto decidido  |
| 27 | Quanto à sua escolha por este curso,       | 3 | Decidido  |
| 21 | você se considera:                         | 3 | Decidido  |
| 27 | Quanto à sua escolha por este curso,       | 4 | Indecise (entre a encão que foz e uma eutra)                            |
| 21 | você se considera:                         | 4 | Indeciso (entre a opção que fez e uma outra)                            |
| 27 | Quanto à sua escolha por este curso,       | 5 | Muito indeciso (entre a opção que fez e                                 |
| 21 | você se considera:                         | 3 | várias outras)  |
|    | No caso de ter respondido indeciso         |   | Tor divides quento às sues babilidades nere s                           |
| 28 | ou muito indeciso, o que contribuiu        | 1 | Ter dúvidas quanto às suas habilidades para a                           |
|    | para sua indecisão?                        |   | opção feita   |
|    | No caso de ter respondido indeciso         |   | Ter dúvidas quanto ao seu real interesse pela                           |
| 28 | ou muito indeciso, o que contribuiu        | 2 | área escolhida  |
|    | para sua indecisão?                        |   | area esconiida  |
|    | No caso de ter respondido indeciso         |   | Tou décidos accorto às mossibilidades de incousão                       |
| 28 | ou muito indeciso, o que contribuiu        | 3 | Ter dúvidas quanto às possibilidades de inserção no mercado de trabalho |
|    | para sua indecisão?                        |   | no mercado de trabamo   |
|    | No caso de ter respondido indeciso         |   |   |
| 28 | ou muito indeciso, o que contribuiu        | 4 | Falta de maior conhecimento sobre o curso                               |
|    | para sua indecisão?                        |   |   |
|    |  |   |   |

|    | No caso de ter respondido indeciso      |   |   |
|----|---|---|---|
| 28 | ou muito indeciso, o que contribuiu     | 5 | Falta de maior conhecimento sobre a profissão escolhida |
|    | para sua indecisão?                     |   |   |
|    | No caso de ter respondido indeciso      |   |   |
| 28 | ou muito indeciso, o que contribuiu     | 6 | Não me enquadro nesta situação                          |
|    | para sua indecisão?                     |   |   |
| 29 | Que recursos você utilizou para         | 1 | Conversas com familiares                                |
| 29 | fazer sua escolha?                      | 1 | Conversas com familiares                                |
| 29 | Que recursos você utilizou para         | 2 | Converses com professores                               |
| 29 | fazer sua escolha?                      | 2 | Conversas com professores                               |
| 29 | Que recursos você utilizou para         | 3 | Conversas com profissionais que atuam na área           |
| 29 | fazer sua escolha?                      | 3 |   |
| 29 | Que recursos você utilizou para         | 4 | Consulta a material informativo sobre cursos e          |
| 29 | fazer sua escolha?                      | 4 | profissões (guias de profissões, revistas, livros, etc) |
| 29 | Que recursos você utilizou para         | 5 | Orientação profissional                                 |
| 29 | fazer sua escolha?                      | 3 | Orientação profissional                                 |
| 29 | Que recursos você utilizou para         | 6 | Nonhum om osnosial                                      |
| 29 | fazer sua escolha?                      | Ü | Nenhum em especial                                      |
| 30 | Qual o motivo que o levou a escolher o  | 1 | Margada da trabalha a magibilidadas salarisis           |
| 30 | curso para o qual está se candidatando? | 1 | Mercado de trabalho e possibilidades salariais          |
| 30 | Qual o motivo que o levou a escolher o  | 2 | Descibilidado do contribuir para a coniedado            |
| 30 | curso para o qual está se candidatando? | 2 | Possibilidade de contribuir para a sociedade            |
| 20 | Qual o motivo que o levou a escolher o  | 3 | Descibilidado do cursor algo do que gosto               |
| 30 | curso para o qual está se candidatando? | 3 | Possibilidade de cursar algo de que gosta               |
|    |   |   |   |

| 30 | Qual o motivo que o levou a escolher o  | 4 | Por ter habilidades relacionadas ao curso |
|----|---|---|---|
|    | curso para o qual está se candidatando? | 4 |   |
| 30 | Qual o motivo que o levou a escolher o  | 5 | Gosto pelas matérias do curso             |
| 30 | curso para o qual está se candidatando? | J | Gosto pelas materias do cuiso             |
| 30 | Qual o motivo que o levou a escolher o  | 6 | Baixa concorrência pelas vagas            |
| 30 | curso para o qual está se candidatando? | U | Baixa concorrencia pelas vagas            |
| 30 | Qual o motivo que o levou a escolher o  | 7 | Permite conciliar aula e trabalho         |
| 30 | curso para o qual está se candidatando? | 1 | remite concinar auta e trabamo            |
| 30 | Qual o motivo que o levou a escolher o  | 8 | Outro motivo                              |
| 30 | curso para o qual está se candidatando? | 0 | Outro motivo                              |
| 31 | Quem ou o que mais influenciou          | 1 | A família                                 |
| 31 | na escolha do curso?                    | 1 |   |
| 31 | Quem ou o que mais influenciou          | 2 | Colegas e amigos                          |
| 31 | na escolha do curso?                    | ۷ |   |
| 31 | Quem ou o que mais influenciou          | 3 | Professor ou escola                       |
| 31 | na escolha do curso?                    | 3 |   |
| 31 | Quem ou o que mais influenciou          | 4 | Teste vocacional                          |
| 31 | na escolha do curso?                    | 4 |   |
| 31 | Quem ou o que mais influenciou          | 5 | Improved a talevisão                      |
| 31 | na escolha do curso?                    | 5 | Imprensa e televisão                      |
| 31 | Quem ou o que mais influenciou          | 6 | Duoficcionais de éves                     |
| 21 | na escolha do curso?                    | U | Profissionais da área                     |
| 21 | Quem ou o que mais influenciou          | 7 | Trabalba atual                            |
| 31 | na escolha do curso?                    | 7 | Trabalho atual                            |
|    |   |   |   |

| 31 | Quem ou o que mais influenciou                        | 8 | Outros  |
|----|---|---|---|
|    | na escolha do curso?                                  | O | Outros  |
| 32 | O que você espera, EM PRIMEIRO                        | 1 | Aquisição de cultura geral ampla                        |
| 32 | LUGAR, de um curso universitário?                     | 1 | Aquisição de cultura gerai ampia                        |
| 32 | O que você espera, EM PRIMEIRO                        | 2 | Farmana and salarah and a salarah and a salarah a       |
| 32 | LUGAR, de um curso universitário?                     | 2 | Formação profissional, voltada para o trabalho          |
| 32 | O que você espera, EM PRIMEIRO                        |   | Farmana takina waltada asaa a maanisa                   |
| 32 | LUGAR, de um curso universitário?                     | 3 | Formação teórica, voltada para a pesquisa               |
| 20 | O que você espera, EM PRIMEIRO                        | 4 | Formação acadêmica para melhorar a atividade            |
| 32 | LUGAR, de um curso universitário?                     | 4 | prática que já estou desempenhando                      |
| 20 | O que você espera, EM PRIMEIRO                        | 5 | Aquisição de conhecimentos que me permitam              |
| 32 | LUGAR, de um curso universitário?                     |   | compreender melhor o mundo em que vivemos               |
| 32 | O que você espera, EM PRIMEIRO                        | 6 | Aquisição de conhecimentos que permitam                 |
| 32 | LUGAR, de um curso universitário?                     |   | melhorar meu nível de instrução                         |
| 32 | O que você espera, EM PRIMEIRO                        | - | Aquisição de conhecimentos que permitam                 |
| 32 | LUGAR, de um curso universitário?                     | 7 | melhorar meu nível social/financeiro e de minha família |
| 32 | O que você espera, EM PRIMEIRO                        |   | Diploma de nível superior                               |
| 32 | LUGAR, de um curso universitário?                     | 8 |   |
| 33 | Dentre os fatores que terá que enfrentar/solucionar   | 1 | Habilidadaa aaraa (Caaraa iiridaa mala ayyaa            |
| 33 | durante o curso, qual você considera mais importante? | 1 | Habilidades específicas exigidas pelo curso             |
| 33 | Dentre os fatores que terá que enfrentar/solucionar   | 2 | Relacionamentos   |
| 33 | durante o curso, qual você considera mais importante? | ۷ | пенасионалнениох  |
| 22 | Dentre os fatores que terá que enfrentar/solucionar   |   | Downietânsie a bébitas de estuda                        |
| 33 | durante o curso, qual você considera mais importante? | 3 | Persistência e hábitos de estudo                        |
|    |   |   |   |

| 33 | Dentre os fatores que terá que enfrentar/solucionar   | 4  | Informações da profissão                      |
|----|---|----|---|
| 33 | durante o curso, qual você considera mais importante? | 4  |   |
| 33 | Dentre os fatores que terá que enfrentar/solucionar   | 5  | Conciliar estudos com vida pessoal e familiar |
| 33 | durante o curso, qual você considera mais importante? | J  | Concinal estudos com vida pessoal e familial  |
| 33 | Dentre os fatores que terá que enfrentar/solucionar   | 6  | Não se enquadra nas alternativas anteriores   |
| 33 | durante o curso, qual você considera mais importante? | U  | ivao se enquadra has alternativas anteriores  |
| 34 | Como você se informou sobre o Processo Seletivo?      | 1  | TV  |
| 34 | Como você se informou sobre o Processo Seletivo?      | 2  | Rádio   |
| 34 | Como você se informou sobre o Processo Seletivo?      | 3  | Jornal  |
| 34 | Como você se informou sobre o Processo Seletivo?      | 4  | Internet                                      |
| 34 | Como você se informou sobre o Processo Seletivo?      | 5  | Folder  |
| 34 | Como você se informou sobre o Processo Seletivo?      | 6  | Cartaz em ônibus                              |
| 34 | Como você se informou sobre o Processo Seletivo?      | 7  | Outro cartaz                                  |
| 34 | Como você se informou sobre o Processo Seletivo?      | 8  | No colégio/cursinho                           |
| 34 | Como você se informou sobre o Processo Seletivo?      | 9  | Feira de Profissões UFPR                      |
| 34 | Como você se informou sobre o Processo Seletivo?      | 10 | Outros  |
| 35 | A sua cor ou raça é:                                  | 1  | Branca  |
| 35 | A sua cor ou raça é:                                  | 2  | Preta   |
| 35 | A sua cor ou raça é:                                  | 3  | Amarela                                       |
| 35 | A sua cor ou raça é:                                  | 4  | Parda   |
| 35 | A sua cor ou raça é:                                  | 5  | Indígena                                      |
| 36 | Se você é indígena, qual é a sua etnia?               | 1  | Candidato não se autoidentifica como indígena |
| 36 | Se você é indígena, qual é a sua etnia?               | 2  | Ava Guarani                                   |
| 36 | Se você é indígena, qual é a sua etnia?               | 3  | Bakairi                                       |
| 36 | Se você é indígena, qual é a sua etnia?               | 4  | Charua  |
|    |   |    |   |

| 36 | Se você é indígena, qual é a sua etnia? | 5  | Chipirá      |
|----|---|----|--------------|
| 36 | Se você é indígena, qual é a sua etnia? | 6  | Kadiwéu      |
| 36 | Se você é indígena, qual é a sua etnia? | 7  | Kaiapó       |
| 36 | Se você é indígena, qual é a sua etnia? | 8  | Kaigang      |
| 36 | Se você é indígena, qual é a sua etnia? | 9  | Kainxana     |
| 36 | Se você é indígena, qual é a sua etnia? | 10 | Kaiowá       |
| 36 | Se você é indígena, qual é a sua etnia? | 11 | Kambela      |
| 36 | Se você é indígena, qual é a sua etnia? | 12 | Kanamari     |
| 36 | Se você é indígena, qual é a sua etnia? | 13 | Kariri       |
| 36 | Se você é indígena, qual é a sua etnia? | 14 | Krenak       |
| 36 | Se você é indígena, qual é a sua etnia? | 15 | M´byá        |
| 36 | Se você é indígena, qual é a sua etnia? | 16 | Nãndeva      |
| 36 | Se você é indígena, qual é a sua etnia? | 17 | Potiguara    |
| 36 | Se você é indígena, qual é a sua etnia? | 18 | Suruí        |
| 36 | Se você é indígena, qual é a sua etnia? | 19 | Tariano      |
| 36 | Se você é indígena, qual é a sua etnia? | 20 | Terena       |
| 36 | Se você é indígena, qual é a sua etnia? | 21 | Tikuna       |
| 36 | Se você é indígena, qual é a sua etnia? | 22 | Tucano       |
| 36 | Se você é indígena, qual é a sua etnia? | 23 | Tupi Guarani |
| 36 | Se você é indígena, qual é a sua etnia? | 24 | Tupinikim    |
| 36 | Se você é indígena, qual é a sua etnia? | 25 | Tuxá         |
| 36 | Se você é indígena, qual é a sua etnia? | 26 | Umutina      |
| 36 | Se você é indígena, qual é a sua etnia? | 27 | Xakriaba     |
| 36 | Se você é indígena, qual é a sua etnia? | 28 | Xavante      |
| 36 | Se você é indígena, qual é a sua etnia? | 29 | Xerente      |
|    |   |    |              |

| 36 | Se você é indígena, qual é a sua etnia? | 30 | Xetá   |
|----|---|----|--|
| 36 | Se você é indígena, qual é a sua etnia? | 31 | Xokleng/laklano  |
| 36 | Se você é indígena, qual é a sua etnia? | 32 | Outra Etnia  |
| 37 | Pertence a uma comunidade quilombola?   | 1  | Sim  |
| 37 | Pertence a uma comunidade quilombola?   | 2  | Não  |
| 38 | Dentre os problemas da sociedade atual, | 1  | HIV e doenças sexualmente transmissíveis                 |
| 30 | qual você considera mais importante?    |    |  |
| 38 | Dentre os problemas da sociedade atual, | 2  | Gravidez não planejada                                   |
| 30 | qual você considera mais importante?    |    |  |
| 38 | Dentre os problemas da sociedade atual, | 3  | Desestruturação familiar e perda de valores morais       |
| 30 | qual você considera mais importante?    |    |  |
| 38 | Dentre os problemas da sociedade atual, | 1  | Violência urbana   |
| 36 | qual você considera mais importante?    | 4  |  |
| 38 | Dentre os problemas da sociedade atual, | 5  | Desigualdade social e desemprego                         |
| 30 | qual você considera mais importante?    |    |  |
| 38 | Dentre os problemas da sociedade atual, | 6  | Uso de drogas lícitas (álcool e tabaco)                  |
| 30 | qual você considera mais importante?    |    |  |
| 38 | Dentre os problemas da sociedade atual, | 7  | Uso de drogas ilícitas (maconha, cocaína, solvente etc.) |
| 30 | qual você considera mais importante?    |    |  |
| 20 | Dentre os problemas da sociedade atual, | 8  | Deterioração do meio ambiente                            |
| 38 | qual você considera mais importante?    |    |  |
| 38 | Dentre os problemas da sociedade atual, | 0  | Outros   |
| 30 | qual você considera mais importante?    | 9  |  |
|    |   | ı  | C:   |
| 39 | Você apresenta alguma necessidade       | 1  | Sim  |

| 39 | Você apresenta alguma necessidade educacional especial? | 2  | Não   |
|----|---|----|---|
| 39 | Você apresenta alguma necessidade                       | 3  | Não dispõe da informação                    |
| 03 | educacional especial?                                   | ŭ  |   |
| 40 | Tipo de necessidade educacional especial:               | 1  | Cegueira                                    |
| 40 | Tipo de necessidade educacional especial:               | 2  | Visão subnormal ou baixa visão              |
| 40 | Tipo de necessidade educacional especial:               | 3  | Surdez                                      |
| 40 | Tipo de necessidade educacional especial:               | 4  | Deficiência Auditiva                        |
| 40 | Tipo de necessidade educacional especial:               | 5  | Deficiência Neuromotora                     |
| 40 | Tipo de necessidade educacional especial:               | 6  | Surdocegueira                               |
| 40 | Tipo de necessidade educacional especial:               | 7  | Deficiência Múltipla                        |
| 40 | Tipo de necessidade educacional especial:               | 8  | Deficiência Intelectual                     |
| 40 | Tipo de necessidade educacional especial:               | 9  | Altas Habilidades/Superdotação              |
| 40 | Tipo de necessidade educacional especial:               | 10 | Transtorno do Espectro do Autismo           |
| 40 | Tipo de necessidade educacional especial:               | 11 | Dislexia, dislalia, disgrafia, discalculia  |
| 40 | Tipo de necessidade educacional especial:               | 12 | TDHA  |
| 40 | Tipo de necessidade educacional especial:               | 13 | Não se enquadra nas alternativas anteriores |
|    |   |    |   |