

UNIVERSIDADE FEDERAL DE JUIZ DE FORA
INSTITUTO DE CIÊNCIAS EXATAS
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

**Um processo automatizado para análise
estatística apoiado por inteligência artificial
generativa**

Júlia Condé Araújo

JUIZ DE FORA
SETEMBRO, 2024

Um processo automatizado para análise estatística apoiado por inteligência artificial generativa

JÚLIA CONDÉ ARAÚJO

Universidade Federal de Juiz de Fora

Instituto de Ciências Exatas

Departamento de Ciência da Computação

Bacharelado em Ciência da Computação

Orientador: Marco Antônio Pereira Araújo

JUIZ DE FORA

SETEMBRO, 2024

UM PROCESSO AUTOMATIZADO PARA ANÁLISE ESTATÍSTICA
APOIADO POR INTELIGÊNCIA ARTIFICIAL GENERATIVA

Júlia Condé Araújo

MONOGRAFIA SUBMETIDA AO CORPO DOCENTE DO INSTITUTO DE CIÊNCIAS
EXATAS DA UNIVERSIDADE FEDERAL DE JUIZ DE FORA, COMO PARTE INTE-
GRANTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE
BACHAREL EM CIÊNCIA DA COMPUTAÇÃO.

Aprovada por:

Marco Antônio Pereira Araújo
Doutor em Engenharia de Sistemas e Computação

Luiz Maurílio da Silva Maciel
Doutor em Engenharia de Sistemas e Computação

Tássio Ferenzini Martins Sirqueira
Doutor em Informática

JUIZ DE FORA
02 DE SETEMBRO, 2024

*Aos meus pais, Marco Antônio Pereira Araújo
e Verônica Cristina Condé Araújo, base de
toda minha educação e a quem devo tudo que
sou.*

Resumo

A necessidade de realizar análises estatísticas de forma eficiente tem se tornado um grande desafio no contexto de pesquisas acadêmicas. A Inteligência Artificial (IA) surge como uma alternativa promissora para enfrentar essa complexidade, tornando a análise estatística e a análise de dados mais dinâmicas. Tais análises acadêmicas demandam uma abordagem meticulosa para determinar o teste estatístico mais apropriado a ser aplicado em cada situação específica. A IA, com sua habilidade de processar dados, identificar padrões e se adaptar às preferências dos usuários, simplifica o processo de análise estatística e análise de dados. Este trabalho concentra-se na criação de um processo automatizado, baseado em IA generativa e integração de ferramentas, para facilitar o processo de análise estatística. Essa abordagem busca simplificar a ordem e escolha de quais testes estatísticos seriam os mais apropriados em cada contexto. Desta forma, apresenta uma hipótese de que um processo automatizado que combine IA e conhecimento humano pode aumentar a eficiência em análises de dados. Tal abordagem pode, de forma significativa, contribuir para uma maior agilidade no processo de análise estatística, apoiando o pesquisador nessa importante, embora árdua, tarefa.

Palavras-chave: Estatística, Inteligência Artificial, Inteligência Artificial Generativa, Análise de Dados.

Abstract

The need to conduct statistical analyses efficiently has become a major challenge in the context of academic research. Artificial Intelligence (AI) emerges as a promising alternative to tackle this complexity, making statistical research and data analysis more dynamic. Such academic analyses require a meticulous approach to determine the most appropriate statistical test to be applied in each specific situation. AI, with its ability to process data, identify patterns, and adapt to user preferences, simplifies the process of statistical analysis and data analysis. This work focuses on creating an automated process, based on generative AI and tool integration, to facilitate the statistical analysis process. This approach seeks to simplify the order and selection of which statistical tests would be most appropriate in each context. Thus, it presents a hypothesis that an automated process combining AI and human knowledge can increase efficiency in academic analyses. Such an approach can significantly contribute to greater agility in the statistical survey process, supporting the researcher in this important, albeit challenging, task.

Keywords: Statistics, Artificial Intelligence, Generative Artificial Intelligence, Data Analysis.

Agradecimentos

Gostaria de expressar minha sincera gratidão à minha família, que sempre me apoiou e acreditou em mim ao longo de toda esta jornada.

Agradeço especialmente ao meu pai, Marco Antônio Pereira Araújo, que não apenas atuou como orientador deste trabalho, mas também ofereceu um apoio fundamental. Sua orientação e compreensão foram essenciais para a realização deste trabalho.

À minha mãe, Verônica Cristina Condé Araújo, que sempre esteve ao meu lado com um amor imenso e apoio incondicional. Seu carinho constante e encorajamento foram uma fonte essencial de força e inspiração durante toda a minha jornada.

À minha irmã gêmea, Marina Condé Araújo, que percorreu toda a jornada acadêmica ao meu lado. Seu apoio e companheirismo foram fundamentais para meu crescimento acadêmico e pessoal.

Aos professores Luiz Maurílio da Silva Maciel e Tássio Ferenzini Martins Siqueira, membros da banca, pelos seus ensinamentos e aos funcionários da universidade que, durante esses anos, contribuíram de algum modo para o meu enriquecimento pessoal e profissional.

“A persistência é o caminho do êxito”.

(Charles Chaplin)

Conteúdo

Lista de Figuras	8
Lista de Tabelas	10
Lista de Abreviações	11
1 Introdução	12
1.1 Contextualização	12
1.2 Justificativa	13
1.3 Descrição do problema	13
1.4 Questão de pesquisa	14
1.5 Hipótese	14
1.6 Objetivos	14
1.7 Publicações Obtidas	15
1.8 Organização do Trabalho	16
2 Fundamentação Teórica	18
2.1 Estatística	18
2.2 Delineamento Experimental	21
2.3 Inteligência Artificial	22
2.3.1 <i>Machine Learning</i>	23
2.3.2 <i>Deep Learning</i>	24
2.3.3 Inteligência Artificial Generativa	25
2.3.4 <i>Large Language Model</i>	25
2.4 Considerações do Capítulo	26
3 Referencial Teórico	27
3.1 Mapeamento sistemático da literatura	27
3.1.1 PICOC	27
3.1.2 <i>String</i> de Busca	28
3.1.3 Critérios de inclusão e exclusão	28
3.1.4 Critérios de Qualidade	29
3.1.5 Resultado e discussão	29
3.2 Trabalhos Relacionados	30
3.2.1 Discussão dos trabalhos relacionados	30
3.3 Considerações do Capítulo	35
4 Materiais e métodos	36
4.1 Testes Estatísticos	36
4.2 Engenharia de Prompt	38
4.3 Ferramentas utilizadas	41
4.3.1 Ferramenta de LLM	41
4.3.2 Análise de dados	45
4.3.3 Ferramenta de automação	49
4.4 Considerações do capítulo	53

5	Estudo de caso	54
5.1	Automação	54
5.2	Engenharia de <i>prompt</i>	62
5.3	Execução do processo com um fator e dois tratamentos	66
5.3.1	Método paramétrico na análise de variáveis em estudos de um fator e dois tratamentos	68
5.3.2	Método não paramétrico na análise de variáveis em estudos de um fator e dois tratamentos	72
5.4	Execução com um fator e mais de dois tratamentos	75
5.4.1	Método paramétrico na análise de variáveis em estudos de um fator e mais de dois tratamentos	76
5.4.2	Método não paramétrico na análise de variáveis em estudos de um fator e mais de dois tratamentos	81
5.5	Considerações do capítulo	84
6	Conclusões	86
6.1	Limitações e desafios encontrados	86
6.2	Trabalhos Futuros	87
6.3	Aspectos éticos	87
6.4	Considerações Finais	88
	Bibliografia	89

Lista de Figuras

2.1	Representação da área de IA (PEREIRA; MOURA, 2023)	23
3.1	Artigos publicados por ano	31
3.2	Nuvem de palavras por autor	31
4.1	Fluxograma com as etapas da metodologia da pesquisa	36
4.2	Árvore decisão testes estatísticos. Adaptado de Sirqueira et al. (2020)	39
4.3	Iniciar conversa no ChatGPT usando um <i>plugin</i>	46
4.4	Perguntas a serem respondidas usando um <i>plugin</i>	47
4.5	Iniciar conversa no ChatGPT usando <i>plugin</i>	49
5.1	Fluxograma com as etapas do estudo de caso	54
5.2	Configuração do gatilho no Zapier: escolha do aplicativo e evento.	55
5.3	Configuração do gatilho no Zapier: conexão com a conta Google.	55
5.4	Configuração do gatilho no Zapier: configuração da planilha.	56
5.5	Configuração do gatilho no Zapier: teste do gatilho.	56
5.6	Configuração dos parâmetros do Google Sheets: adicionar ação.	57
5.7	Configuração dos parâmetros do Google Sheets: escolha do evento.	57
5.8	Configuração dos parâmetros do Google Sheets: configuração dos parâmetros.	57
5.9	Configuração dos parâmetros do Google Sheets: testar ação.	58
5.10	Configuração dos parâmetros da OpenAI: adicionar ação.	58
5.11	Configuração dos parâmetros da OpenAI: escolha do evento.	59
5.12	Configuração dos parâmetros da OpenAI: configuração dos parâmetros.	59
5.13	Configuração dos parâmetros da OpenAI: testar Ação.	60
5.14	Configuração dos parâmetros da OpenAI: adicionar ação.	60
5.15	Configuração dos parâmetros da OpenAI: escolha do evento.	61
5.16	Configuração dos parâmetros da OpenAI: configuração do documento.	61
5.17	Configuração dos parâmetros da OpenAI: testar ação.	61
5.18	Fluxo da automação completa no Zapier.	62
5.19	Base de dados usada no estudo de caso	67
5.20	Variáveis presentes na base de dados	68
5.21	Agrupamento das variáveis	68
5.22	Estatística descritiva da variável Defeitos (<i>Defects</i>) por grupo	68
5.23	<i>Boxplot</i> da variável Defeitos (<i>Defects</i>) por grupo	69
5.24	Análise de normalidade da variável defeitos	70
5.25	<i>p-value</i> da variável <i>Defects</i>	70
5.26	Teste de Levene	71
5.27	Gráficos para verificar se a variável Defeitos (<i>Defects</i>) apresenta distribuição normal	71
5.28	Teste t de <i>Student</i>	72
5.29	Fragmento do documento gerado para a variável <i>Defects</i>	72
5.30	Estatística descritiva da variável <i>Time</i> por grupo	73
5.31	<i>Boxplot</i> da variável <i>Time</i> por grupo	73
5.32	Análise de normalidade da variável <i>Time</i>	74
5.33	<i>p-value</i> da variável <i>Time</i>	74

5.34	Teste de Mann-Whitney	74
5.35	Fragmento do documento gerado para a variável <i>Time</i>	75
5.36	Base de dados usada no estudo de caso	76
5.37	Variáveis presente na base de dados.	77
5.38	Agrupamento das variáveis.	77
5.39	Estatística descritiva da variável <i>Client</i> por tipo.	78
5.40	<i>Boxplot</i> da variável <i>Client</i> por tipo.	78
5.41	Análise de normalidade da variável <i>Client</i>	79
5.42	<i>p-value</i> da variável <i>Client</i>	79
5.43	Teste de Levene.	79
5.44	Teste ANOVA	80
5.45	Teste de Tukey.	80
5.46	Fragmento do documento gerado para a variável <i>Client</i>	81
5.47	Estatística descritiva da variável <i>Control</i> por tipo	81
5.48	<i>Boxplot</i> da variável <i>Control</i> por tipo.	82
5.49	Teste de Kolmogorov-Smirnov da variável <i>Control</i>	83
5.50	<i>p-value</i> da variável <i>Control</i>	83
5.51	Teste de Kruskal-Wallis	83
5.52	Teste de Mann-Whitney para a variável <i>Control</i>	84
5.53	Fragmento do documento gerado para a variável <i>Control</i>	84

Lista de Tabelas

3.1	PICOC.	28
3.2	Critérios de Inclusão e Exclusão.	29
3.3	Análise de qualidade dos artigos selecionados.	30
3.4	Resultado mapeamento da literatura após critérios de inclusão e exclusão.	30
4.1	Comparação entre ChatGPT, Gemini e LLaMA	43
4.2	Comparação entre as ferramentas de análise estatística.	48
4.3	Comparação entre Zapier, Microsoft Power Automate e UiPath.	51

Lista de Abreviações

ANOVA	Análise de Variância
DCC	Departamento de Ciência da Computação
DL	<i>Deep Learning</i>
IA	Inteligência Artificial
IA GEN	Inteligência Artificial Generativa
ML	<i>Machine Learning</i>
LLMs	Grandes Modelos de Linguagem
UFJF	Universidade Federal de Juiz de Fora

1 Introdução

A importância de realizar análises estatísticas de forma eficiente é um desafio crescente no contexto de pesquisas acadêmicas, segundo Ignácio (2010). Pesquisadores frequentemente se confrontam com a complexidade inerente à condução de análises de dados verdadeiramente relevantes para seus estudos. Nesse cenário, a Inteligência Artificial (IA) emerge como uma abordagem promissora, capaz de enfrentar essa complexidade intrínseca. A IA oferece oportunidades para agilizar e aprimorar o processo de análise de dados, contribuindo para avanços significativos na pesquisa. Sua capacidade de lidar com grandes volumes de dados e identificar padrões é particularmente valiosa. Além disso, a IA pode automatizar tarefas repetitivas e demoradas, liberando tempo para análises mais profundas e interpretações significativas. O desenvolvimento de habilidades em IA e análise estatística é essencial para os pesquisadores enfrentarem os desafios complexos da pesquisa contemporânea. A integração da IA na análise estatística promete revolucionar a forma como é conduzida a pesquisa científica e abre novas oportunidades para avanços significativos em diversas áreas do conhecimento. O foco dessa pesquisa foi na análise estatística para estudos experimentais com um fator e dois ou mais tratamentos.

1.1 Contextualização

Ao longo dos anos, a análise estatística evoluiu para uma disciplina multifacetada, abrangendo uma variedade de métodos e técnicas. Desde a análise descritiva básica até modelos estatísticos avançados, a análise estatística oferece aos pesquisadores um conjunto diversificado de ferramentas para explorar, entender e explicar fenômenos observados em suas áreas de estudo, como evidencia Araújo et al. (2006).

A inteligência artificial generativa (IA Gen), segundo McCarthy (1956), acrescenta uma nova dimensão a esse processo, possibilitando a automação de tarefas analíticas complexas e a geração de *insights* a partir de conjuntos de dados extensos e heterogêneos. Por meio de algoritmos sofisticados, a inteligência artificial generativa é capaz de iden-

tificar padrões, tendências e relações nos dados, permitindo aos pesquisadores formular hipóteses e fazer inferências com maior rapidez e precisão. Além disso, a inteligência artificial generativa desempenha um papel fundamental na validação e verificação dos resultados de pesquisa. A IA generativa ajuda a garantir a robustez e confiabilidade das conclusões tiradas a partir dos dados coletados. Isso é particularmente relevante em um ambiente acadêmico onde a precisão e a integridade dos resultados são imperativos. Portanto, ao integrar a inteligência artificial generativa ao processo de análise estatística, não apenas busca-se otimizar a eficiência e a precisão dessa análise, mas também promover uma maior confiança e credibilidade na pesquisa científica como um todo. É por meio dessa combinação de tecnologia e rigor metodológico que busca-se alcançar novos patamares na investigação científica.

1.2 Justificativa

A pesquisa científica está cada vez mais dependente de análises estatísticas robustas e eficientes, como evidencia Ignácio (2010). Com a crescente complexidade dos dados disponíveis, pesquisadores enfrentam desafios significativos na interpretação e no uso dessas informações para tirar conclusões confiáveis. Além disso, o tempo e os recursos dedicados à análise estatística podem representar uma parte substancial do processo de pesquisa. Nesse contexto, a integração da inteligência artificial generativa como uma ferramenta para apoiar o processo de análise estatística torna-se cada vez mais relevante. Os avanços recentes nessa área oferecem novas oportunidades para automatizar tarefas analíticas complexas e extrair *insights* valiosos dos dados disponíveis. Essas tecnologias têm o potencial não apenas de melhorar a eficiência e precisão das análises, mas também de revelar padrões e relações anteriormente não identificados.

1.3 Descrição do problema

A análise estatística acadêmica desempenha um papel significativo na pesquisa científica, onde a interpretação precisa de dados é fundamental para validar hipóteses e extrair conclusões significativas. No entanto, o crescente volume de dados disponíveis apresenta

desafios substanciais para os pesquisadores, exigindo métodos eficientes e confiáveis para sua análise.

Diante disso, torna-se interessante desenvolver estudos que visem identificar e avaliar abordagens que automatizem e aprimorem o processo de análise estatística. Essa automatização não apenas agiliza o trabalho dos pesquisadores, mas também pode melhorar a consistência e a precisão dos resultados, mitigando potenciais vieses e erros humanos.

Ao investigar e avaliar métodos de automação para análise estatística, os pesquisadores podem otimizar suas investigações, permitindo uma análise mais profunda e detalhada dos conjuntos de dados. Esse avanço contribui não apenas para a descoberta de padrões e correlações relevantes, mas também para o progresso científico em diversas áreas do conhecimento, promovendo uma base sólida para futuras pesquisas e descobertas.

1.4 Questão de pesquisa

Como a IA Generativa pode contribuir para aprimorar as pesquisas, especialmente na análise de dados e de que maneira essas inovações podem impactar a qualidade e eficiência dos resultados científicos e apoiar os pesquisadores na avaliação de seus resultados?

1.5 Hipótese

A automação com apoio de inteligência artificial generativa pode auxiliar na análise estatística de dados.

1.6 Objetivos

O principal objetivo dessa pesquisa é automatizar, com apoio de inteligência artificial generativa, o processo de análise estatística em estudos acadêmicos. Além disso, devem ser analisadas e compreendidas todas as etapas de um processo de exploração de dados e apoiar a automatização desse processo. Para realizar esse estudo, foi feita uma análise de dados, manipulando uma base de dados real. Dessa forma, para automatizar esse processo, foi utilizada uma IA generativa capaz de interagir com um conjunto de dados

apoiando a análise estatística.

Com essa abordagem, é possível importar conjuntos de dados e conduzir análises estatísticas detalhadas para apoiar uma pesquisa. Isso simplifica e amplia a exploração dos dados, incluindo a identificação das variáveis na base, suas categorias correspondentes, a descrição da distribuição dessas variáveis, a construção de gráficos, a realização de testes de normalidade e a determinação do método estatístico aplicado. Além disso, permite a execução de testes de igualdade de variâncias e a seleção justificada de métodos de comparação de médias, com a exportação dos resultados, juntamente com gráficos, em um formato adequado para trabalhos científicos. Dessa maneira, o processo de análise estatística se torna mais ágil e eficiente, gerando um volume significativo de informações de pesquisa de forma automática, o que auxilia o trabalho do pesquisador.

A hipótese subjacente a essa abordagem é que a combinação da IA com o conhecimento humano pode aumentar a eficiência das pesquisas acadêmicas. Ao automatizar e integrar tarefas repetitivas e demoradas, os pesquisadores podem direcionar seu tempo e energia para análises aprofundadas e contribuições originais para o campo científico, gerando novos dados de pesquisa. Portanto, a IA não apenas simplifica e amplia essas tarefas, mas também abre portas para novas descobertas e avanços no conhecimento científico.

1.7 Publicações Obtidas

Durante o 10^o Fórum de Gestão de Dados de Investigação, realizado em novembro de 2023 no Instituto Politécnico de Setúbal, Portugal, foram publicados os artigos reflexivos intitulados "Explorando a Automatização na Análise Estatística com Suporte da Inteligência Artificial Generativa na Gestão de Dados de Investigação" (ARAÚJO et al., 2023b), e "Explorando Automatização na Revisão de Literatura com Suporte de Inteligência Artificial Generativa na Gestão de Dados de Investigação" (ARAÚJO et al., 2023c). Este trabalho destaca a importância e relevância da abordagem proposta, oferecendo uma visão crítica sobre as possíveis implicações da integração da inteligência artificial na análise e gestão de dados de pesquisa. Além disso, no mesmo período, foi publicado o artigo "Em busca de um processo automatizado para levantamento de literatura apoiado por uma inteligência

artificial generativa” (ARAÚJO et al., 2023a), no CIACA, Conferência Ibero-Americana Computação Aplicada em Portugal. Esse estudo evidencia a utilização da IA Generativa em outras etapas do trabalho acadêmico.

1.8 Organização do Trabalho

Este trabalho está estruturado em seis capítulos, cada um focado em um aspecto da pesquisa.

No primeiro capítulo é apresentada a introdução do trabalho, onde foi discutida uma visão geral do tema, destacando sua importância e os objetivos principais. Este capítulo também aborda a motivação por trás da pesquisa e define as questões centrais que foram respondidas ao longo do estudo.

No segundo capítulo é apresentada a fundamentação teórica, onde serão discutidas as bases teóricas que sustentam a pesquisa. Aqui, foram detalhados conceitos fundamentais, teorias relevantes e metodologias que foram utilizadas nesse estudo. Este capítulo estabelece o alicerce teórico sobre o qual o trabalho foi desenvolvido.

No terceiro capítulo é apresentado o referencial teórico da pesquisa, onde podem ser observados estudos anteriores que fornecem o contexto necessário para entender o tema abordado. Este capítulo inclui uma explicação detalhada do mapeamento sistemático da literatura, abordando cada etapa do processo, desde a definição do PICOC, até a construção das *strings* de busca, avaliação da qualidade, seleção e análise dos estudos.

No capítulo quatro é discutido sobre os materiais e métodos escolhidos para realizar esse estudo. Aqui, foram detalhados os testes estatísticos e a escolha das ferramentas utilizadas neste trabalho.

No capítulo cinco é apresentado o estudo de caso do trabalho, detalhando todas as etapas do processo de automatização dos processos estatísticos utilizando IA generativa. Inicialmente, aborda-se a coleta de dados, seguida pela aplicação dos testes estatísticos necessários. O capítulo discute a integração dessas etapas, demonstrando como a IA generativa foi empregada para otimizar e automatizar todo o processo.

Finalmente, o sexto capítulo, a conclusão, sintetiza os principais achados da pesquisa e discute suas contribuições para o campo de estudo. Além disso, são apontadas as

limitações do estudo e sugestões para pesquisas futuras. Este capítulo fecha o trabalho, refletindo sobre os objetivos alcançados e o impacto da pesquisa realizada.

2 Fundamentação Teórica

Neste capítulo, são apresentados os conceitos e tecnologias fundamentais que servem como base para essa pesquisa. Evidenciando áreas como Inteligência Artificial, incluindo *Machine Learning* (Aprendizado de Máquina), *Deep Learning* (Aprendizado Profundo), Inteligência Artificial generativa e Grandes Modelos de Linguagem (LLMs). Para complementar o entendimento, também são abordados termos e teorias de Estatística, buscando uma compreensão mais aprofundada do tema e sua relevância para a proposta da pesquisa.

2.1 Estatística

A Estatística é uma ciência que se dedica ao desenvolvimento e ao uso de métodos para a coleta, resumo, organização, apresentação e análise de dados, segundo Costa (2002).

Os métodos estatísticos fornecem ferramentas poderosas para extrair informações significativas de conjuntos de dados complexos e muitas vezes vastos. Ao utilizar técnicas estatísticas apropriadas, os pesquisadores podem identificar padrões, tendências e relações ocultas nos dados, fornecendo *insights* valiosos que podem informar decisões e políticas em diversas áreas. Além disso, a Estatística desempenha um papel fundamental na formulação de hipóteses, na realização de experimentos e na validação de resultados, garantindo assim a robustez e a confiabilidade das conclusões alcançadas.

O processo de análise estatística é um componente fundamental da pesquisa científica, subdividindo-se em etapas críticas de definição, planejamento, execução e análise, conforme descreve no trabalho de Araújo et al. (2006). Na fase de planejamento, é essencial a formulação de hipóteses e a identificação das variáveis dependentes e independentes, seguida pela seleção de participantes e métodos de análise. Esse estágio envolve ainda o *design* do estudo e a definição de instrumentos, culminando na avaliação de potenciais ameaças à validade do estudo. A análise subsequente foca na interpretação de gráficos e estatísticas descritivas, o tratamento de *outliers*, quando pertinente, e a aplicação de análises estatísticas robustas.

Após a coleta de dados, a estatística descritiva entra em cena para resumir características fundamentais do conjunto de dados, incluindo a utilização de medidas de tendência central como a média, mediana e moda. A média aritmética é obtida pela divisão do somatório dos valores pelo número de observações, enquanto a mediana é o valor que ocupa a posição central em um conjunto de dados ordenado. A moda, por outro lado, é o valor mais frequente dentro do conjunto de dados. Outras métricas, como o valor mínimo e máximo, os percentis e os quartis, fornecem informações adicionais sobre a distribuição dos dados. As medidas de dispersão, incluindo faixa, variância e desvio padrão, são empregadas para avaliar a variação dos dados em torno do valor central, fornecendo *insights* sobre a concentração ou dispersão dos dados, evidenciado na pesquisa de Sirqueira et al. (2020).

Em termos de inferência estatística, uma hipótese estatística é uma suposição sobre um parâmetro desconhecido que pode ser testada para validação ou refutação. A execução de um teste de hipóteses requer a definição de uma hipótese nula (H_0) e pelo menos uma hipótese alternativa (H_1), juntamente com a especificação de um nível de erro estatístico aceitável. Os testes de hipótese são categorizados em paramétricos e não paramétricos. Os primeiros assumem o conhecimento da forma da distribuição dos dados e requerem a verificação de condições como a normalidade e homocedasticidade dos dados. Já os testes não paramétricos não fazem suposições sobre a forma da distribuição, oferecendo uma alternativa quando os dados não satisfazem as condições dos testes paramétricos, evidenciado no trabalho de COOPER e SCHINDLER (2016).

A escolha do teste estatístico adequado depende das características da amostra e dos objetivos do estudo. Diversos testes estatísticos são detalhados na literatura, evidenciado por BATTISTI e BATTISTI (2008), sendo:

- O Teste de Kolmogorov-Smirnov (K-S) é amplamente reconhecido por sua capacidade de avaliar a semelhança entre as distribuições de duas amostras distintas ou para comparar a distribuição de uma amostra com uma distribuição clássica conhecida, como a normal;
- O Teste de Shapiro-Wilk destina-se a verificar se uma amostra segue uma distribuição normal, calculando o valor W , particularmente útil para conjuntos de dados

de menor tamanho;

- O Teste de Levene é empregado para testar a homogeneidade das variâncias entre grupos, assumindo que as variâncias são iguais se comparado ao nível de significância estabelecido, para distribuições normais;
- O Teste T ou Student-T é comumente utilizado para comparar as médias de duas amostras independentes, onde a escolha do teste específico depende das variâncias observadas entre as amostras, sendo um teste paramétrico;
- O método ANOVA ou Análise de Variância, é outra técnica paramétrica, projetada para testar a hipótese de igualdade entre as médias de mais de dois grupos;
- O Teste de Tukey, frequentemente aplicado em conjunto com ANOVA para auxiliar na identificação de grupos cujas médias são significativamente diferentes;
- O Teste de Mann-Whitney apresenta-se como uma opção não paramétrica ao Teste T, exigindo que as amostras sejam independentes e que os dados sejam contínuos ou ordenados em uma escala ordinal, de intervalo ou de razão;
- O Teste Kruskal-Wallis é uma alternativa não paramétrica ao ANOVA, oferecendo um método para realizar análise de variância em dados que não cumprem com os pressupostos dos testes paramétricos, substituindo os valores dos dados por seus *rankings*. Este teste estende a capacidade do ANOVA a conjuntos de dados que não se enquadram em distribuições de probabilidade normais, permitindo uma análise inclusiva de dados não paramétricos.

Em conclusão, a estatística desempenha um papel central na pesquisa científica e no desenvolvimento de conhecimento em diversas áreas. Suas técnicas e métodos fornecem as ferramentas necessárias para a análise rigorosa e objetiva de dados, contribuindo para o avanço do conhecimento humano e para a tomada de decisões informadas em diferentes contextos.

2.2 Delineamento Experimental

O delineamento experimental é um componente essencial para a realização de análises estatísticas em experimentos científicos, garantindo a confiabilidade e a validade dos resultados. Conforme destacado por Montgomery (2017), o delineamento determina a estrutura do experimento e orienta a escolha dos métodos estatísticos a serem utilizados na análise dos dados. Em análise estatística, o fator é uma variável categórica manipulada ou observada para avaliar seu impacto na variável de resposta, com diferentes níveis correspondendo às suas categorias ou valores. O tratamento refere-se a uma combinação específica de níveis dos fatores aplicados às unidades experimentais, permitindo a análise dos efeitos individuais e combinados desses fatores sobre a variável de resposta. Quando o experimento envolve apenas um fator com dois tratamentos, o objetivo principal é comparar as médias dos dois grupos para verificar se há uma diferença significativa entre os grupos. A análise estatística típica utilizada nesse cenário é o teste t de *Student* para amostras independentes ou pareadas, dependendo da natureza dos dados. O teste t de *Student* para amostras independentes é utilizado quando as observações nos dois grupos são independentes entre si, enquanto o teste para amostras pareadas é aplicado quando as observações são dependentes, como no caso de medidas antes e depois de um tratamento no mesmo sujeito. A simplicidade do teste t o torna adequado para a comparação direta de duas médias, sendo amplamente empregado em experimentos científicos.

Quando o experimento envolve um fator com mais de dois tratamentos, a análise estatística torna-se mais complexa. Montgomery (2017) explica que, nesse caso, a comparação direta de médias por meio de múltiplos testes t não é apropriada, pois isso aumentaria o risco de erro, que é o erro de rejeitar uma hipótese nula verdadeira. Para resolver essa questão, utiliza-se a Análise de Variância (ANOVA). A ANOVA de um fator é utilizada para comparar as médias de três ou mais grupos (tratamentos) com base em um único fator. Este teste verifica se pelo menos uma das médias é significativamente diferente das outras. Se a ANOVA indicar uma diferença significativa, testes como o teste de Tukey, são realizados para identificar quais grupos específicos diferem entre si.

Assim, enquanto a análise com dois tratamentos utiliza o teste t para a comparação direta de médias, a presença de mais de dois tratamentos exige o uso da ANOVA

para evitar o aumento de erro e garantir uma análise estatística robusta e adequada ao delineamento experimental.

2.3 Inteligência Artificial

A Inteligência Artificial (IA) é um campo multidisciplinar da ciência da computação voltado para o desenvolvimento de sistemas, tanto físicos quanto digitais, que imitam a capacidade humana de pensar, tomar decisões e solucionar problemas. O termo “Inteligência Artificial” foi introduzido em 1956 pelo pesquisador e matemático McCarthy (1956), que a definiu como “a ciência e engenharia de produzir sistemas inteligentes”.

De acordo com McCarthy, inteligência artificial refere-se à capacidade computacional de alcançar objetivos no mundo, presente em seres humanos, animais e algumas máquinas. Assim, a inteligência não está restrita ao contexto humano nem ao biológico. Além disso, McCarthy ressalta que a IA não se limita apenas à simulação do pensamento e inteligência humanos, mas sim à resolução inteligente de problemas apresentados pelo mundo. Portanto, os pesquisadores têm a liberdade de utilizar diversos métodos para alcançar inteligência ou comportamento considerado inteligente ao resolver problemas do mundo real, mesmo que envolvam habilidades computacionais que os humanos não possuem.

Os sistemas de IA funcionam principalmente através do processamento de dados, aprendendo com as informações que recebem e se adaptando constantemente com a entrada de novos dados. Essa capacidade de aprendizado e adaptação possibilita que os computadores sejam treinados para executar tarefas específicas que anteriormente eram realizadas por humanos ou até mesmo identificar e aplicar novas soluções com base na análise de padrões nos dados.

Além disso, existem conceitos adicionais que são importantes para uma melhor compreensão do funcionamento da inteligência artificial. Estas tecnologias são fundamentais para capacitar um sistema a simular o raciocínio lógico humano.

Na Figura 2.1, pode-se observar a representação da área de IA e as outras tecnologias que serão apresentadas ao longo desta seção.

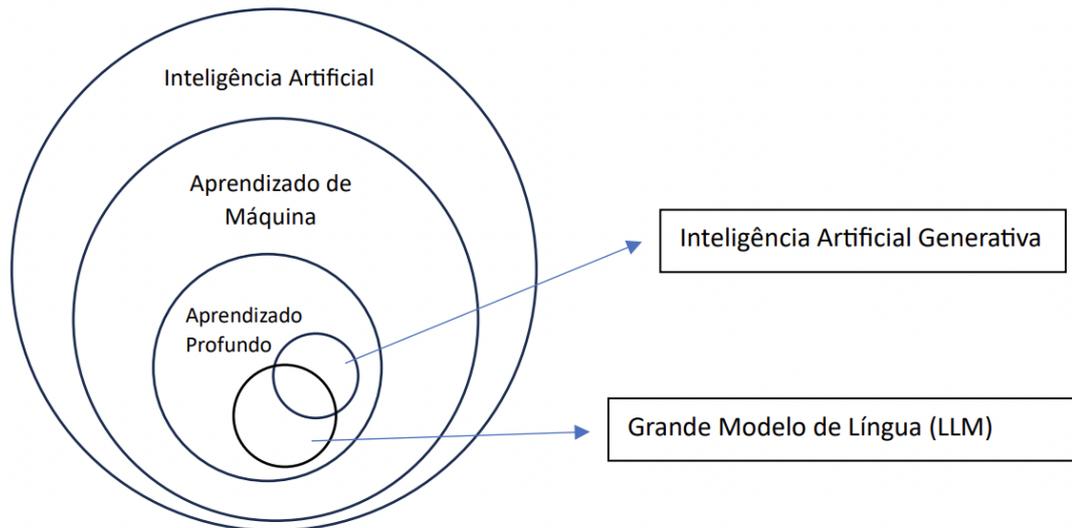


Figura 2.1: Representação da área de IA (PEREIRA; MOURA, 2023)

2.3.1 *Machine Learning*

Machine Learning (ML), ou Aprendizado de Máquina, é uma subárea da Inteligência Artificial focada no desenvolvimento de algoritmos e modelos estatísticos que capacitam os computadores a aprender e aprimorar sua *performance* através da análise de dados, como descrito no trabalho de Wiederhold e McCarthy (1992). Diferentemente da programação tradicional, onde as máquinas são instruídas explicitamente para executar tarefas específicas, os sistemas de ML aprendem com os dados, permitindo-lhes fazer previsões ou tomar decisões de forma autônoma. Essa tecnologia confere aos sistemas a habilidade de aprender e tomar decisões independentes, com base no processamento de dados e na identificação de padrões.

O ciclo de aprendizado de máquina inicia-se com a coleta e preparação de conjuntos de dados pertinentes, seguido pela escolha e treinamento de um modelo apropriado. Durante o treinamento, o modelo é exposto a um conjunto de dados de entrada, ajustando seus parâmetros internos para reconhecer e aprender padrões e relações nos dados. Este processo é iterativo, e o modelo é constantemente aprimorado com a introdução de novos dados, visando aperfeiçoar sua eficácia.

Quanto aos tipos de algoritmos de ML, SALIAN (2018) os categoriza em quatro grupos principais:

- **Aprendizado Supervisionado:** esse método utiliza conjuntos de dados rotulados para treinar o algoritmo, permitindo que aprenda a associar entradas a saídas corretas;
- **Aprendizado Não-Supervisionado:** nesse caso, os dados não são rotulados, e o algoritmo busca identificar padrões ou estruturas intrínsecas nos dados, agrupando-os ou categorizando-os de forma natural;
- **Aprendizado Semi-Supervisionado:** combina elementos dos métodos supervisionado e não-supervisionado, utilizando conjuntos de dados que contêm tanto exemplos rotulados quanto não-rotulados, permitindo ao algoritmo realizar classificações parciais;
- **Aprendizado por Reforço:** esse método utiliza um sistema de recompensa e punição, onde o algoritmo é treinado através de tentativa e erro, recebendo *feedbacks* positivos ou negativos com base em suas ações.

Cada categoria permite a adaptação e implementação adequadas conforme as necessidades e objetivos de cada projeto ou aplicação.

2.3.2 *Deep Learning*

Deep Learning (DL), ou Aprendizado Profundo, termo criado por LeCun Y Bengio (2015), representa uma subárea avançada do campo do *Machine Learning*. Esta abordagem se destaca pela sua capacidade de aprender a partir de grandes volumes de dados, utilizando redes neurais artificiais profundas e complexas. Inspiradas na estrutura e funcionamento dos neurônios humanos, estas redes neurais são capazes de capturar e representar características intrincadas e não-lineares dos dados, permitindo uma modelagem mais precisa e eficaz de problemas complexos.

A arquitetura das redes neurais utilizadas no aprendizado profundo é composta por múltiplas camadas de unidades computacionais, conhecidas como neurônios artificiais. Cada camada processa e transforma os dados de entrada de maneira hierárquica, permitindo que o sistema aprenda representações cada vez mais abstratas e sofisticadas dos dados. Essa capacidade de construir representações hierárquicas e complexas torna o

aprendizado profundo especialmente adequado para tarefas que envolvem reconhecimento de padrões em imagens, áudio, texto e outros tipos de dados.

Além disso, o DL apresenta capacidade de modelar e interpretar dados complexos de maneira eficaz está impulsionando inovações e avanços significativos em diversas áreas, prometendo transformar nossa forma de interagir com a tecnologia e resolver problemas complexos da sociedade moderna.

2.3.3 Inteligência Artificial Generativa

A Inteligência Artificial Generativa (IA Gen) constitui uma subárea da Inteligência Artificial dedicada ao desenvolvimento de sistemas capazes de gerar conteúdo novo e original, como discutido no trabalho de Quintero (2023). Esses sistemas são alimentados com grandes volumes de dados para treinamento, permitindo que aprendam padrões intrínsecos aos dados fornecidos. Através desse processo de treinamento, a IA Generativa torna-se apta a realizar uma ampla gama de tarefas, desde as mais simples até as mais complexas. Uma vez treinada, a IA é capaz de gerar respostas que se alinham aos padrões aprendidos, produzindo conteúdo original de acordo com esses padrões.

2.3.4 *Large Language Model*

Os *Large Language Models* (LLMs), ou Grandes Modelos de Linguagem, representam um avanço significativo no campo da Inteligência Artificial, particularmente na interseção da linguística computacional e do aprendizado de máquina, como evidenciado no trabalho de Minaee et al. (2024). Esses modelos empregam técnicas de aprendizado profundo, uma subdivisão do aprendizado de máquina que utiliza redes neurais com várias camadas para analisar e interpretar grandes conjuntos de dados. Através deste método, os LLMs são capazes de processar e produzir linguagem natural de forma que reflete a complexidade da comunicação humana.

O treinamento desses modelos é realizado com enormes quantidades de texto, abrangendo uma diversidade de fontes e estilos, o que permite aos modelos aprenderem uma ampla gama de padrões linguísticos, vocabulário e gramática. Ao assimilar essas informações, os LLMs desenvolvem uma capacidade notável de compreender o contexto

e gerar respostas que não apenas são gramaticalmente corretas, mas também contextualmente apropriadas.

Esta capacidade de entender e replicar nuances da linguagem humana habilita os LLMs a executar uma variedade impressionante de tarefas. Essa habilidade de simular o discurso humano tem implicações vastas, promovendo uma interação mais natural e eficaz entre humanos e máquinas. Permite que tecnologias baseadas em LLM sejam integradas de maneira mais fluída em diversas aplicações de análise de dados. Ao compreender e gerar texto de maneira semelhante ao humano, esses modelos podem facilitar a comunicação de *insights* complexos, interpretar dados não estruturados e até mesmo gerar relatórios automaticamente. Isso torna a análise de dados mais acessível e eficiente, ajudando as organizações a extrair valor de seus dados de maneira mais rápida e precisa.

2.4 Considerações do Capítulo

Nesse capítulo foram discutidos alguns parâmetros para realizar a análise estatística de dados em estudos experimentais. Além disso, essa pesquisa evidencia quais os tipos de análises estatísticas devem ser realizados em cada etapa da pesquisa. Esse estudo busca demonstrar que diferentes grupos de dados se relacionam melhor com alguns testes estatísticos. Assim, dependendo do grupo de dados que for apresentado a essa discussão sobre análise de dados será possível identificar e compreender qual o teste mais apropriado em cada situação. Dessa forma, nesse capítulo também são apresentadas metodologias que vão ser utilizadas para realizar a análise e comparação dos resultados obtidos nesse estudo. Os conceitos apresentados serão utilizados como fonte de dados e comparações nessa pesquisa, além dos métodos que vamos utilizar para tratar, analisar e comparar os dados, com o objetivo de identificar quais seriam os testes mais adequados para cada situação e realizar sua análise.

3 Referencial Teórico

A pesquisa que foi realizada tem como objetivo automatizar o processo de análise estatística usando uma IA generativa para apoiar o trabalho de pesquisadores.

3.1 Mapeamento sistemático da literatura

Kitchenham, Dyba e Jorgensen (2004) discutem a possibilidade de adoção de pesquisas baseadas em evidências em engenharia de software, citando o mapeamento sistemática da literatura, um tipo de estudo experimental secundário que visa coletar evidências de estudos presentes na literatura para responder perguntas de pesquisas. Logo após essa discussão, Kitchenham e Charters publicaram um relatório técnico com instruções sobre como realizar um mapeamento sistemático da literatura para engenharia de software. Assim, o mapeamento permanece dividido em três etapas principais: (i) Planejamento do mapeamento, (ii) Realização do mapeamento e (iii) Resultados do mapeamento. O planejamento de literatura desejada foi dividido em quatro partes, sendo o PICOC, *string* de busca, critérios de inclusão e exclusão, resultado da pesquisa e critérios de qualidade.

3.1.1 PICOC

A busca de estudos durante a revisão sistemática da literatura é regida por uma *String* de busca booleana, derivada de um conjunto de palavras-chave denominado PICOC, que visa delimitar o escopo da busca. Cada letra da sigla PICOC tem um significado na especialização da busca, o “P” (*population*) refere-se ao grupo específico de pessoas ou entidades que estão sendo estudadas, o “I” (*intervention*) indica as intervenções, tratamentos ou exposições que estão sendo aplicados à população, o “C” (*comparison*), quando deseja comparar métodos ou técnicas eles devem ser listados nessa categoria. O “O” (*outcome*), contém o que deseja analisar no final da pesquisa (resultados) e, por fim, o “C” (*context*), responsável por limitar o contexto ao qual a pesquisa será realizada. A Tabela 3.1, apresenta os elementos do PICOC para esta pesquisa.

Tabela 3.1: PICOC.

PICOC	Palavras- Chave	<i>Key- Words</i>
<i>Population</i>	Estatística	<i>Statistics</i>
<i>Intervention</i>	Análise estatística análise de dados	<i>Statistical analysis, data analysis</i>
<i>Comparison</i>	Não definido	<i>Not defined</i>
<i>Outcome</i>	Ferramenta, aplicação, software	<i>Tool, application, software</i>
<i>Context</i>	Processo automatizado, automação, inteligência artificial	<i>Automated process, automation, artificial intelligence</i>

3.1.2 *String* de Busca

Juntando todos os termos do PICOC é possível criar uma *string* de busca booleana, usada como base para derivar *strings* específicas de mapeamento: (“estatística”) *and* (“análise estatística” *or* “análise de dados”) *and* (“ferramenta” *or* “aplicação” *or* “software”) *and* (“processo automatizado” *or* “automação” *or* “inteligência artificial”).

A *string* de busca também foi gerada na versão em inglês: “*statistics*” *and* (“*statistical analysis*” *or* “*data analysis*”) *and* (“*tool*” *or* “*application*” *or* “*software*”) *and* (“*automated process*” *or* “*automation*” *or* “*artificial intelligence*”).

3.1.3 Critérios de inclusão e exclusão

Neste trabalho foram utilizadas duas máquinas de busca, sendo elas o Acadêmico (2004) e o IEEEExplore (2008), uma vez que constituem duas plataformas de pesquisa essenciais e altamente recomendadas dentro da comunidade acadêmica, particularmente nos domínios da tecnologia, engenharia e ciências. Após executar as *strings* de pesquisa, em português e inglês, nas bases digitais de artigos, são apresentados todos os artigos encontrados pela máquina de busca da base digital utilizada, respeitando as regras definidas na *string* de busca fornecida, a busca foi realizada no mês de maio de 2024. Porém, o mecanismo de busca nem sempre retorna apenas artigos que são do interesse do pesquisador, ou seja, mesmo com a *string* de busca ajustada de acordo com as necessidades, foram necessários critérios para definir quais estudos foram ou não aceitos para participar do mapeamento. Assim, esses artigos foram avaliados por meio dos critérios de inclusão e exclusão, representados na Tabela 3.2.

Tabela 3.2: Critérios de Inclusão e Exclusão.

Inclusão	Artigos contendo estudos experimentais, estudo de caso ou análises sobre automação no contexto de análises estatísticas
	Estudos contendo análise, discussão ou implementação de automação na análise de dados
Exclusão	Capítulos de livro, chamadas a congressos e materiais de ensino básico
	Estudos que não podem ser acessados por inteiro
	Estudos que não abordem o tema na área da Ciência da Computação

3.1.4 Critérios de Qualidade

Em um processo de revisão da literatura, a análise da qualidade dos artigos selecionados é de extrema importância. Normalmente é feito através da construção de um *check-list* com questões a serem respondidas pelos pesquisadores, após a leitura de cada artigo.

Ao analisar as obras apresentadas, é possível concluir que o campo de estudo desta pesquisa já foi abordado por diferentes autores. Porém, é interessante que mesmo abordando temas parecidos cada um apresenta uma metodologia distinta, agregando muito para a pesquisa em questão. A seguir é apresentada uma análise comparativa, conforme mostrado na Tabela 3.3, com um conjunto de questões a serem respondidas como: (1) se apresenta processo automatizado; (2) se apresenta avaliação dos resultados; (3) qual o método de análise dos dados; (4) se apresenta análise estatística completa dos resultados. Observa-se que os trabalhos apresentam contribuições significativas para o assunto deste trabalho, porém nenhuma delas apresentou por completo um modelo automatizado para apoio a análise estatística de dados apoiado por IA generativa. Os artigos descritos na tabela serão apresentados nessa mesma ordem na seção de trabalhos relacionados.

3.1.5 Resultado e discussão

Após definir todos os detalhes do protocolo, é possível realizar a busca na literatura. No contexto desta pesquisa em particular, foram realizadas as seguintes atividades: (i) importação de estudos encontrados em bases de dados digitais, (ii) eliminação de artigos duplicados e (iii) aplicação dos critérios de inclusão e exclusão. A amostra, contendo os artigos encontrados tanto na *string* em português quanto na em inglês, foi composta por 55 artigos, sendo 7 duplicados, 43 removidos pelos critérios de inclusão e exclusão e 20 artigos foram utilizados na pesquisa, representados na Tabela 3.4.

Tabela 3.3: Análise de qualidade dos artigos selecionados.

Questões	Processo automatizado	Avaliação dos resultados	Método de análise de dados	Análise estatística
(MOURA, 2004)	Sim	Qualitativa	Mineração de dados	Sim
(PRADO, 2023)	Sim	Qualitativa	Técnicas de aprendizado de máquina	Não
(NAVARRO, 2021)	Não	Qualitativa	Modelos de IA	Não
(BUENO, 2010)	Sim	Quantitativa	Mineração de dados	Não
(PINHO et al., 2022)	Não	Quantitativa e qualitativa	Estatística e Mineração de dados	Sim

Tabela 3.4: Resultado mapeamento da literatura após critérios de inclusão e exclusão.

Importação de estudos encontrados em bases de dados digitais	55 artigos
Artigos duplicados	7 artigos
Critérios de Inclusão e Exclusão	43 artigos
Artigos utilizados na pesquisa	5 artigos
Critérios de qualidade	5 artigos

3.2 Trabalhos Relacionados

Nesta seção serão apresentadas obras com recorte temático, hipóteses, objetivos e ferramentas metodológicas que apresentam um diálogo e contribuição crítica com o estudo discutido nesta pesquisa. Esses estudos encontrados foram oriundos do mapeamento sistemático.

3.2.1 Discussão dos trabalhos relacionados

Um mapeamento sistemático da literatura utiliza representações visuais cruciais, como gráficos de evolução temporal das publicações, apresentado na Figura 3.1 e nuvens de palavras com os sobrenomes dos autores, Figura 3.2. O gráfico mostra a taxa de crescimento anual de artigos na área, enquanto a nuvem de palavras destaca os pesquisadores proeminentes. Essas representações facilitam a compreensão das tendências de pesquisa e das redes acadêmicas, orientando futuras investigações e identificando lacunas de conhecimento.

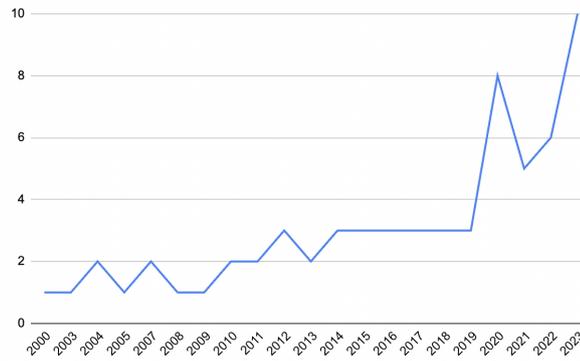


Figura 3.1: Artigos publicados por ano



Figura 3.2: Nuvem de palavras por autor

A pesquisa conduzida por Moura (2004) explora a importância e a aplicabilidade da mineração de textos na seleção, classificação e qualificação de documentos em organizações que possuem grandes acervos de dados não estruturados, como a Embrapa. A autora destaca os desafios enfrentados na gestão de um vasto repositório documental que inclui desde registros de pesquisa até informações administrativas, muitas vezes armazenados sem uma estrutura definida, o que complica a recuperação e o uso eficiente dessas informações. Moura propõe uma metodologia que integra técnicas de mineração de textos com uma forte ênfase na análise de dados para automatizar e otimizar o processo de organização documental. Ela sugere a implementação de ferramentas de software que facilitam a categorização automática dos documentos, promovendo melhorias significativas na produtividade e na confiabilidade dos processos de catalogação existentes. A metodologia proposta inclui a criação de uma base controlada de documentos, a utilização de ferramentas de clusterização de domínio público e a colaboração com especialistas do domínio para a validação dos resultados. Um aspecto central do trabalho de Moura é a aplicação de técnicas estatísticas avançadas na mineração de textos. A autora destaca o uso da

análise estatística para descobrir padrões úteis em grandes coleções de textos. Isso inclui a utilização de métricas como ganho de informação, informação mútua e estimativas de qui-quadrado para a categorização e análise de dados textuais. Essas análises estatísticas são fundamentais para resolver problemas de desambiguação no uso de termos da língua, contribuindo significativamente para a eficácia das ferramentas de mineração de textos. Ao aplicar essas técnicas de análise de dados, espera-se que a Embrapa possa gerenciar seu acervo de forma mais eficaz, melhorando o acesso e a disseminação do conhecimento interno. Isso, por sua vez, deve apoiar decisões estratégicas mais informadas e promover um compartilhamento de conhecimento mais eficiente entre pesquisadores e outras partes interessadas. A pesquisa de Moura é um exemplo significativo de como as técnicas modernas de processamento e análise de dados podem ser aplicadas para resolver problemas tradicionais de gestão de informação em grandes organizações.

O estudo de Prado (2023) explora a integração de sistemas de IA nos processos médicos, demonstrando como a automação suportada por inteligência artificial pode apoiar a análise estatística na medicina. O foco principal do estudo é na melhoria da precisão diagnóstica e na personalização dos tratamentos através da análise de grandes conjuntos de dados, o que é crucial para o desenvolvimento de processos automatizados completos em ambientes médicos. O estudo utiliza técnicas avançadas de IA para automatizar a análise de imagens médicas, onde algoritmos são treinados para identificar padrões complexos inacessíveis ao olho humano. Este processo automatizado facilita diagnósticos mais rápidos e precisos, contribuindo significativamente para melhores prognósticos dos pacientes. A IA também é aplicada na personalização do tratamento, utilizando análises preditivas para adaptar as terapias às necessidades individuais dos pacientes, baseadas em um vasto *array* de dados clínicos e genéticos. Na gestão de recursos de saúde, a IA é usada para otimizar a alocação e utilização de recursos médicos. Sistemas de IA analisam dados operacionais e clínicos em tempo real para melhorar a eficiência dos serviços de saúde, reduzindo custos e melhorando a qualidade do atendimento. Esta aplicação demonstra um exemplo prático de como a análise estatística, apoiada por IA generativa, pode ser integral na reestruturação dos processos médicos. Em conclusão, o artigo ilustra como a inteligência artificial generativa está apoiando a análise estatística na medicina, tornando

os processos mais eficientes e automatizados. Este avanço promete não apenas melhorar os resultados de saúde para os pacientes, mas também otimizar a gestão de recursos em sistemas de saúde cada vez mais complexos.

A tese de Navarro (2021) explora o emprego de técnicas avançadas de inteligência artificial, especificamente processamento de linguagem natural e modelos de aprendizado de máquina, para aprimorar a recuperação de informações em ambientes digitais. Este trabalho se concentra em desenvolver um sistema que, através da utilização de algoritmos como o *Latent Dirichlet Allocation* (LDA) e *Word2Vec*, pode analisar e interpretar grandes volumes de textos digitais, proporcionando uma recuperação de informações mais eficaz e contextualizada. O contexto desta pesquisa se alinha com a busca por processos automatizados para análise estatística, especialmente aqueles apoiados por inteligência artificial generativa. O modelo proposto por Navarro sugere que o uso de inteligência artificial não apenas simplifica a recuperação e análise de dados em grandes repositórios digitais, mas também enriquece a qualidade das informações recuperadas, atribuindo-lhes relevância e contexto semântico. A aplicação de técnicas de inteligência artificial generativa, como as empregadas no estudo, oferece uma base robusta para a automação de análises estatísticas, onde a capacidade de aprender com dados e gerar novas instâncias ou previsões se torna possível através de modelos treinados. Assim, o trabalho de Navarro não apenas avança na aplicação prática de modelos de inteligência artificial em repositórios digitais, mas também enfatiza a análise estatística ao integrar técnicas como LDA e Word2Vec. Isso demonstra como a inteligência artificial pode transformar a análise estatística, movendo-a de métodos tradicionais para processos automatizados e inteligentes que melhoram significativamente a eficiência e a precisão das análises. A utilização desses algoritmos permite um entendimento mais profundo e contextual dos dados, resultando em informações mais relevantes e bem interpretadas para os usuários finais.

O trabalho de Bueno (2010) na sua tese de doutorado propõe um sistema automatizado para a classificação de abelhas baseado em reconhecimento de padrões, que é uma contribuição significativa para os esforços de conservação de biodiversidade e também para o campo da inteligência artificial (IA) aplicada. Esta proposta é especialmente relevante no contexto da busca por um processo automatizado para análise estatística apoiada por

IA generativa, uma vez que aborda a utilização de tecnologias avançadas e metodologias de IA para superar desafios práticos em um campo específico de estudo. Bueno detalha a criação de um modelo de sistema chamado ABeeS (*Automated Bee Identification System*), que incorpora conhecimento especializado para o reconhecimento automatizado de abelhas usando imagens das asas. Este sistema permite a identificação rápida e precisa de espécies de abelhas, o que é crucial para estudos ambientais e conservacionistas, dada a importância das abelhas como polinizadores no ecossistema. Além disso, Bueno enfatiza a necessidade de superar o "impedimento taxonômico" – a falta de especialistas e recursos para classificar todas as espécies conhecidas e desconhecidas. O uso de um sistema automatizado como o ABeeS pode mitigar esse problema ao reduzir a carga de trabalho manual e aumentar a precisão e a velocidade da identificação de espécies. Um aspecto fundamental do trabalho de Bueno é a aplicação de análise estatística para validar a eficiência do sistema ABeeS. A tese detalha o uso de técnicas de morfometria geométrica combinadas com análise estatística discriminante para diferenciar espécies de abelhas. Esse processo é realizado utilizando o software SPSS, alcançando uma taxa de acerto significativa. A análise estatística desempenha um papel crucial ao fornecer uma base quantitativa para avaliar a precisão e a confiabilidade do sistema de identificação automatizada. O uso de IA e reconhecimento de padrões, como aplicado no ABeeS, são exemplos de como a tecnologia pode ser usada para desenvolver métodos inovadores de análise e gestão de grandes volumes de dados biológicos. Essa abordagem não apenas apoia a pesquisa científica, mas também oferece novas ferramentas para a gestão da biodiversidade, permitindo respostas mais rápidas e informadas a questões ambientais críticas.

O artigo de Pinho et al. (2022) investiga o uso de técnicas de Inteligência Artificial, especificamente Processamento de Linguagem Natural (PLN) e Aprendizado de Máquina (ML), para automatizar a análise de textos educacionais, com ênfase na análise estatística dos dados obtidos. O estudo foca na correção de redações do ENEM, um processo que tradicionalmente demanda muito tempo e esforço dos educadores. Com a digitalização do ENEM anunciada pelo MEC em 2019, surgiram novas oportunidades para implementar técnicas avançadas de IA na avaliação de redações. Neste contexto, o artigo detalha um experimento onde 695 redações foram analisadas usando PLN e AM para

identificar desvios dos temas propostos e outros problemas comuns, como falta de coesão textual ou redações muito curtas. A análise estatística desempenhou um papel crucial na interpretação dos dados, utilizando histogramas para entender a distribuição das notas, a quantidade de palavras e sentenças por redação, e a relação entre a quantidade de sentenças e as notas aplicadas. Os resultados do estudo mostraram que as técnicas de IA aplicadas foram eficazes em identificar redações que desviaram do tema e outros erros, facilitando significativamente o trabalho de correção. Além disso, o uso dessas tecnologias, apoiado por uma análise estatística robusta, poderia reduzir o tempo e os custos associados à avaliação das redações, oferecendo *feedback* mais rápido e consistente aos estudantes. Este avanço representa um passo significativo em direção à automatização da análise de textos educacionais, prometendo eficiência e precisão melhoradas no processo de avaliação, o que é especialmente valioso em contextos de grandes exames como o ENEM.

3.3 Considerações do Capítulo

Neste capítulo, foi discutido o referencial teórico utilizado para a realização desta pesquisa. Inicialmente, abordou-se o processo de mapeamento sistemático da literatura, detalhando cada etapa, incluindo o PICOC, a construção da *string* de busca, os critérios de inclusão e exclusão, e os resultados obtidos através das máquinas de busca. Além disso, foram apresentados trabalhos relacionados ao tema do estudo, que contribuíram significativamente para esta pesquisa, porém não apresentam um processo automatizado completo para análise estatística com apoio de IA generativa. Também foram incluídas representações visuais, como gráfico e nuvem de palavras, utilizadas para ilustrar a análise dos artigos revisados. Adicionalmente, foi apresentada uma tabela com a análise de qualidade dos artigos retornados, fornecendo uma visão abrangente da relevância e rigor metodológico dos estudos incluídos. Esses elementos proporcionaram uma base sólida para a compreensão do estado e permitiram uma análise crítica e detalhada das contribuições científicas relevantes para este estudo.

4 Materiais e métodos

Este capítulo é dedicado à explicação detalhada dos materiais e métodos empregados no desenvolvimento do processo automatizado para análise estatística apoiado por IA Generativa. Assim, são apresentadas todas as etapas desse processo, iniciando-se com a definição dos testes estatísticos usados durante o estudo. Em seguida, é explicado como integrar esse passo com a IA Generativa por meio de uma engenharia de *prompt* e, por fim, como automatizar esse processo. Além disso, serão descritas todas as ferramentas utilizadas, detalhando como utilizá-las, como pode-se observar na sequência apresentada na Figura 4.1.

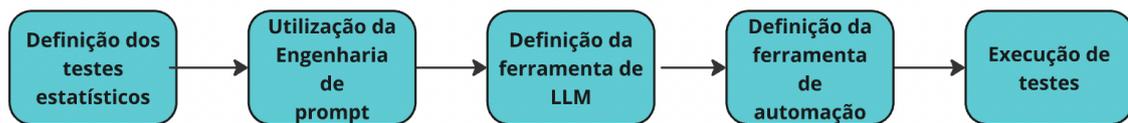


Figura 4.1: Fluxograma com as etapas da metodologia da pesquisa

4.1 Testes Estatísticos

Os testes estatísticos são ferramentas na análise de dados, permitindo a validação de hipóteses e a obtenção de conclusões robustas e fundamentadas. São usadas para verificar a significância das diferenças observadas nos dados, identificar relações entre variáveis e assegurar a confiabilidade dos resultados. Dessa forma, baseado no trabalho de Sirqueira et al. (2020), um conjunto de passos foi estruturado para a condução de uma análise estatística apropriada.

Primeiramente, é importante identificar quais são as variáveis presentes na base de dados utilizada nesse estudo e a categoria de cada uma. Esta etapa inicial estabelece a base para as análises subsequentes, garantindo que todas as variáveis relevantes sejam consideradas e corretamente classificadas. Em seguida, é necessário agrupar essas variáveis e informar o quantitativo de observações em cada grupo. Essa etapa ajuda a entender a

distribuição dos dados e a prepará-los para análises mais detalhadas.

O próximo passo envolve a apresentação das estatísticas descritivas para uma variável de interesse. As estatísticas descritivas, que incluem medidas como média, mediana, desvio padrão e quartis, fornecem uma visão geral dos dados e ajudam a identificar padrões ou *outliers*.

Após as estatísticas descritivas, é útil visualizar a distribuição dos dados através de um *boxplot*. O *boxplot* é uma representação gráfica que mostra a distribuição dos dados com base em seus quartis, destacando possíveis *outliers* e a dispersão dos dados. Essa visualização é crucial para entender melhor a distribuição e a variabilidade dos dados.

Em seguida, deve-se realizar o teste de normalidade para a variável, que verifica se os dados seguem uma distribuição normal. A normalidade refere-se a uma condição em que os dados são distribuídos de maneira simétrica em torno da média, conhecida como distribuição normal ou gaussiana. Esse teste é crucial porque muitos testes estatísticos paramétricos assumem que os dados são normalmente distribuídos. Conforme discutido no trabalho de Sirqueira et al. (2020), se a amostra possui mais de 30 valores, deve-se utilizar o Teste de Kolmogorov-Smirnov (K-S). Para amostras com 30 ou menos valores, o Teste de Shapiro-Wilk é recomendado. Se o teste de normalidade indicar que os dados seguem uma distribuição normal, é possível prosseguir com a verificação para testes paramétricos; caso contrário, devem-se considerar testes não paramétricos.

Se os dados atenderem aos pressupostos de normalidade, pode ser necessário realizar um teste de igualdade de variâncias, conhecido como homocedasticidade, utilizando o Teste de Levene. A homocedasticidade refere-se à condição em que os diferentes grupos de dados possuem variâncias aproximadamente iguais. Este é um pressuposto crucial para muitos testes estatísticos paramétricos.

Na utilização de um teste paramétrico, no caso de dois grupos, utiliza-se o Teste *t*. Caso contrário, ao utilizar um teste não paramétrico, igualmente para dois grupos, pode-se utilizar o teste de Mann-Whitney.

No caso de comparação de médias entre mais de dois grupos, o método paramétrico ANOVA pode ser indicado. Caso contrário, devem ser considerados testes não paramétricos, como o teste de Kruskal-Wallis.

Com base nas análises anteriores, o próximo passo é identificar o teste apropriado para a comparação das médias de uma variável. A escolha do teste (paramétrico ou não paramétrico) depende dos resultados dos testes de normalidade e homocedasticidade, bem como da quantidade de grupos a serem comparados. Testes paramétricos são utilizados quando os dados atendem aos pressupostos de normalidade e homocedasticidade. No caso de comparação entre dois grupos, pode ser utilizado o *t-test* para amostras independentes. Para a comparação de mais de dois grupos, utiliza-se a Análise de Variância (*ANOVA*). Por outro lado, testes não paramétricos, como o teste de Mann-Whitney (para dois grupos) ou o teste de Kruskal-Wallis (para mais de dois grupos), são utilizados quando os pressupostos de normalidade ou homocedasticidade não são atendidos. A escolha do método apropriado garante que a comparação das médias seja realizada de maneira robusta, levando em consideração as características específicas dos dados.

Depois de identificar o teste apropriado, realiza-se o teste de comparação das médias, apresentando o teste de hipóteses e os resultados obtidos, incluindo o valor-p (*p-value*) e o nível de significância. O valor-p indica a probabilidade de obter os resultados observados, assumindo que a hipótese nula é verdadeira. O nível de significância, geralmente fixado em 5%, é o critério usado para decidir se será rejeitada a hipótese nula. Se o valor-p for maior que o nível de significância, aceita-se a hipótese nula, indicando que há uma diferença significativa entre as médias dos grupos comparados. Finalmente, deve-se indicar o resultado final, esclarecendo se existe uma diferença significativa de médias para a variável analisada.

Em resumo, a execução sistemática e estruturada dos testes estatísticos, começando pela identificação e agrupamento das variáveis, passando pelos testes de normalidade e homocedasticidade, e culminando na comparação de médias, é fundamental para garantir a validade e a robustez das conclusões obtidas em um estudo, o fluxograma dos testes está representado na Figura 4.2.

4.2 Engenharia de Prompt

A Engenharia de *Prompt* é uma técnica importante no trabalho com modelos de linguagem avançados e Inteligência Artificial (IA). Este processo envolve a criação e o refinamento

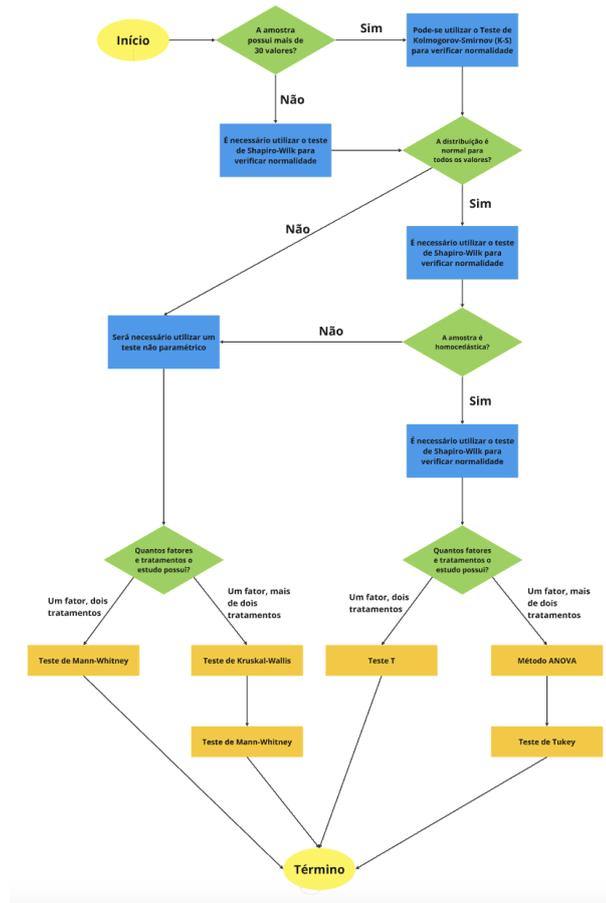


Figura 4.2: Árvore decisão testes estatísticos. Adaptado de Sirqueira et al. (2020)

de entradas de texto (*prompts*) para orientar o comportamento do modelo de IA e obter respostas desejadas de maneira eficiente e precisa. Larguesa (2024) explora em seu livro, de forma detalhada, como essa prática pode ser aplicada, desde a definição básica até suas aplicações práticas em diversos cenários.

O conceito de *prompt* envolve o desenvolvimento de entradas específicas para um modelo de linguagem, que influenciam diretamente a saída gerada pelo sistema. Estes *prompts* não são apenas comandos simples, são estrategicamente formulados para obter informações precisas ou desencadear ações específicas do modelo de IA. A engenharia de *prompt* é essencial, pois permite um controle mais fino sobre a *performance* dos modelos de linguagem, garantindo que a comunicação entre humanos e máquinas seja mais eficiente e relevante.

Em aplicações práticas, a engenharia de *prompt* ajuda a otimizar tarefas como análise de texto, geração automática de respostas e até mesmo programação, onde a

precisão dos *prompts* determina a qualidade dos resultados. Além disso, a criação eficaz de *prompts* pode facilitar a customização de respostas de IA, adequando-as às necessidades específicas do usuário ou contexto, como destacado nos exemplos fornecidos no livro.

A engenharia de *prompt* é, portanto, uma habilidade valiosa para qualquer profissional que deseja aprimorar sua capacidade de interagir e utilizar plenamente as tecnologias de IA. Tornando-se uma ponte entre a linguagem humana e o processamento de linguagem natural, possibilitando que os avanços em IA sejam acessíveis e aplicáveis em variados domínios profissionais e cotidianos.

Em resumo, a engenharia de *prompt* representa uma técnica poderosa para maximizar a eficácia de modelos de IA em tarefas de processamento de linguagem natural. Ao formular entradas de texto claras, específicas e contextualizadas, e iterar com base no *feedback*, é possível orientar o modelo a gerar respostas mais precisas e úteis. Essa prática não apenas melhora a interação com a IA, mas também desbloqueia novas possibilidades de aplicação em diversas áreas, desde a educação até a análise de dados e além.

Neste estudo, foi explorado a aplicação da engenharia de *prompt* para direcionar respostas de uma Inteligência Artificial (IA) generativa em um contexto acadêmico específico. A técnica foi empregada com o intuito de simular a abordagem de um professor universitário e pesquisador, como analista de dados, cientista de dados ou estatístico, visando elevar o nível de formalidade e precisão das respostas fornecidas pela IA. Para isso, o sistema foi configurado para responder a um conjunto pré-determinado de questões focadas em estatística.

A escolha de modelar a IA para agir como um acadêmico foi motivada pela necessidade de garantir que as respostas refletissem um rigor analítico e uma profundidade de conhecimento compatíveis com o esperado em ambientes educacionais e de pesquisa. Isso foi conseguido através de uma cuidadosa elaboração de *prompts* que não apenas guiavam a IA sobre o conteúdo, mas também sobre o estilo de resposta desejado.

Este método de configuração do *prompt* mostrou-se crucial para influenciar efetivamente a performance da IA, destacando como ajustes na formulação de *prompts* podem resultar em melhorias significativas na utilidade das respostas da IA em aplicações específicas. A análise dos resultados sugere que a engenharia de *prompt* representa uma

ferramenta valiosa para aprimorar a interação entre humanos e sistemas inteligentes, especialmente em contextos que exigem um alto grau de especialização e formalidade nas respostas.

4.3 Ferramentas utilizadas

A seleção das ferramentas adequadas é uma etapa fundamental para garantir a eficácia e eficiência em projetos de pesquisa e automação de processos. Nesta seção, foi apresentada a escolha das principais ferramentas para a realização de análises estatísticas e automação de fluxos de trabalho. Assim como, as razões que fundamentaram essas escolhas e como cada uma contribui para a otimização dos processos envolvidos.

4.3.1 Ferramenta de LLM

A escolha da ferramenta de LLM adequada pode impactar significativamente a eficiência e produtividade das operações. Ferramentas como Chatgpt (2022), Gemini (2023), e LLaMA (2023) se destacam, cada uma oferecendo funcionalidades específicas que atendem a diferentes necessidades e níveis de complexidade. Esta seção apresenta uma análise comparativa dessas plataformas, destacando suas principais características e vantagens.

Desenvolvido pela OpenAI, o ChatGPT é um modelo de linguagem baseado na arquitetura GPT (*Generative Pre-trained Transformer*). Esse modelo é conhecido por sua capacidade de gerar texto de forma coerente e relevante, sendo amplamente utilizado em aplicações que envolvem interação conversacional e geração de conteúdo. Uma das vantagens do ChatGPT é sua excelente integração com ferramentas estatísticas, o que facilita a realização de análises complexas diretamente na interface de conversação. O ChatGPT também se destaca pela precisão na interpretação de *prompts*, oferecendo respostas detalhadas e focadas, essenciais para a qualidade das análises estatísticas. A OpenAI oferece suporte contínuo e documentação extensa, além de uma comunidade ativa e recursos educacionais abundantes, tornando-o uma escolha preferencial para muitos projetos de NLP (*Natural Language Processing*).

O Gemini, desenvolvido pelo Google, é um modelo de *dual encoder* que utiliza

duas redes neurais distintas para processar diferentes tipos de dados de entrada de forma separada. Essa abordagem permite que o modelo capture e entenda características específicas de cada tipo de dado (como texto e imagens) antes de combiná-los para formar uma representação mais completa e precisa. A separação e a integração das representações ajudam a melhorar a compreensão e a resposta do modelo, especialmente em contextos que envolvem múltiplas trocas de informações. Além disso, como modelo *multimodal*, o Gemini é projetado para lidar com conversas que se desenrolam ao longo de várias interações. Em vez de tratar cada entrada como um evento isolado, o modelo é capaz de manter e integrar o contexto das trocas anteriores, o que é crucial para gerar respostas coerentes e relevantes em conversas prolongadas. Embora o Gemini apresente algumas limitações em termos de integração contínua e suporte comparado ao ChatGPT, ele demonstra eficácia em tarefas que combinam diversos tipos de dados e tem mostrado bom desempenho em *benchmarks* de raciocínio e compreensão *multimodal*. A combinação dessas características permite que o Gemini mantenha a coerência e a relevância ao longo de conversas mais complexas e dinâmicas.

O LLaMa (*Learning Language Model for Assistance*), desenvolvido pela Meta AI, é um modelo orientado para assistência em tarefas específicas, como recuperação de informações e sistemas de recomendação. Sendo capaz de gerenciar diálogos *multi-turn* e integrar fontes de conhecimento externas para fornecer respostas informativas. A capacidade do LLaMA de lidar com tarefas orientadas a objetivos específicos o torna adequado para aplicações em suporte ao cliente e sistemas de informação. No entanto, o LLaMA apresenta desafios em termos de precisão na interpretação de *prompts* e tem uma comunidade de suporte e recursos educacionais menos ativa comparada ao ChatGPT.

Na escolha de uma ferramenta de LLM (*Large Language Model*) para um processo automatizado de análise estatística, é crucial considerar diversos critérios que garantam a eficácia e a precisão do processo. A facilidade de uso é um fator fundamental, pois a ferramenta deve ser intuitiva e acessível para usuários que podem não possuir experiência técnica avançada, como cita Brown et al. (2020). Isso assegura que a curva de aprendizado não seja um obstáculo significativo, permitindo que os pesquisadores comecem a utilizar a ferramenta de maneira eficiente e rápida. Além disso, a precisão do modelo é vital,

pois determina a qualidade das respostas fornecidas, assegurando que os resultados sejam válidos e úteis para análises estatísticas detalhadas, detalhado no estudo de Radford et al. (2019). A escalabilidade é outro aspecto crítico, visto que a ferramenta deve ser capaz de lidar com volumes crescentes de dados à medida que o projeto avança, adaptando-se a diferentes tamanhos e complexidades de tarefas, como exemplifica Vaswani et al. (2017). A flexibilidade da ferramenta também é essencial, permitindo ajustes e personalizações para diferentes áreas de aplicação e necessidades específicas dos usuários, o que é vital para atender às diversas demandas de diferentes setores, segundo Devlin et al. (2018). Além disso, a capacidade de integração da LLM com outras ferramentas de automação, como plataformas de gerenciamento de dados e sistemas de análise, é fundamental para criar um fluxo de trabalho coeso e eficiente, facilitando a transferência de dados e a coordenação entre diferentes sistemas, conforme Albalak et al. (2024). A segurança é um critério imprescindível, garantindo que os dados sensíveis e as informações coletadas sejam protegidos contra acessos não autorizados, evitando violações de privacidade e mantendo a integridade dos dados. Por fim, o suporte técnico é um fator importante, pois um bom suporte pode resolver problemas rapidamente e minimizar interrupções no processo, garantindo a continuidade e a eficiência do trabalho, conforme Wolf et al. (2020). Esses critérios foram levados em consideração para assegurar que a ferramenta de LLM escolhida não apenas atenda às necessidades técnicas do projeto, mas também proporcione uma experiência de uso eficiente e segura, maximizando a qualidade e a relevância dos resultados obtidos. A comparação entre os modelos de LLMs estão apresentadas na Tabela 4.1.

Critério	ChatGPT	Gemini	LLaMA
Facilidade de Uso	Alta	Alta	Média
Precisão do Modelo	Alta	Alta	Média
Escalabilidade	Alta	Alta	Média
Flexibilidade	Alta	Média	Alta
Capacidade de Integração	Alta	Alta	Média
Segurança	Alta	Média	Média
Suporte Técnico	Alta	Média	Baixa

Tabela 4.1: Comparação entre ChatGPT, Gemini e LLaMA

A escolha do ChatGPT para este trabalho, em comparação com modelos como (Gemini, 2023) e (LLaMA, 2023) LLaMA, baseia-se em diversas considerações específicas

detalhadas na literatura. Primeiramente, enquanto Gemini e LLaMA são modelos promissores desenvolvidos por outras instituições, o ChatGPT destaca-se pela sua integração com ferramentas estatísticas, permitindo a execução fluida e eficiente de análises complexas diretamente na interface de conversação.

Além disso, o ChatGPT demonstra uma capacidade superior de interpretar *prompts* detalhados com alta precisão e coerência, proporcionando respostas mais relevantes e focadas. Esta característica é fundamental para a qualidade das análises estatísticas. A OpenAI oferece suporte contínuo e documentação extensa, garantindo que os usuários tenham acesso às últimas atualizações e inovações, além de uma comunidade ativa e recursos educacionais abundantes. Embora LLaMA e Gemini possuam características avançadas, a robustez, a facilidade de integração e o suporte contínuo do ChatGPT tornam-no a escolha preferencial para este tipo de análise. A decisão foi fundamentada na superioridade do ChatGPT em fornecer uma plataforma completa e confiável para a realização de análises complexas.

O ChatGPT representa um avanço significativo no campo da inteligência artificial e processamento de linguagem natural. Sua relevância para a análise estatística é substancial, oferecendo uma plataforma intuitiva e robusta para processar grandes volumes de dados, interpretar resultados e gerar *insights* valiosos. Capaz de executar tarefas complexas, o ChatGPT facilita a descrição de variáveis, a execução de testes de hipóteses e a visualização de dados, permitindo análises detalhadas e precisas de conjuntos de dados extensos.

Uma das características mais destacadas do ChatGPT é sua capacidade de integração com diversos *plugins* específicos para análise de dados. Esses *plugins* ampliam as funcionalidades da ferramenta, possibilitando a execução de operações estatísticas avançadas diretamente dentro da interface de conversação.

A engenharia de *prompt* é outro aspecto crucial que contribui para a eficácia do ChatGPT em análises estatísticas. Ao elaborar *prompts* detalhados e direcionados, os usuários podem orientar a inteligência artificial para focar em aspectos específicos da análise, garantindo resultados mais relevantes e personalizados. A capacidade do ChatGPT de interpretar e responder a esses *prompts* de maneira precisa é fundamental

para a realização de análises complexas e para a obtenção de *insights* que podem orientar a tomada de decisões baseadas em dados.

4.3.2 Análise de dados

A escolha do método estatístico adequado é fundamental para garantir a validade e a precisão das análises de dados em projetos de pesquisa. No contexto atual, diversas ferramentas e *plugins* facilitam essa escolha, oferecendo funcionalidades específicas para diferentes tipos de análise. Esta seção discute a utilização dos *plugins* Diagrams (2023) e Report (2023), bem como as ferramentas (AKKIO, 2024) e R Core Team (2023), na seleção e aplicação de métodos estatísticos.

O ChatGPT dispõe de uma variedade de *plugins* específicos, concebidos para diversas tarefas, que aumentam significativamente sua funcionalidade. Entre esses *plugins*, destacam-se o *plugin* Diagrams e o *Data analysis*, ambos voltados para a realização de análises estatísticas avançadas. A inclusão desses recursos não apenas expande as capacidades analíticas da plataforma, mas também a transforma em uma ferramenta essencial para estatísticos e pesquisadores, que buscam precisão e eficiência em seus estudos e trabalhos acadêmicos. Para usar um *plugin* basta ir na opção "Explorar GPTs", na barra lateral da ferramenta e escolher qual deseja utilizar, logo após basta iniciar conversa utilizando o *plugin* selecionado, como na Figura 4.3

O *plugin* Diagrams do ChatGPT é uma ferramenta que facilita a criação de visualizações gráficas complexas e dinâmicas. Permite a geração automática de diagramas, gráficos e outras representações visuais de dados, o que torna a interpretação e a comunicação dos resultados mais acessíveis e compreensíveis. A visualização de dados é fundamental na análise estatística, pois ajuda a identificar padrões, tendências e *outliers* que podem não ser evidentes na observação de tabelas de dados brutos. Com o *plugin* Diagrams, os usuários podem rapidamente transformar seus resultados estatísticos em gráficos como histogramas, *boxplots*, *scatter plots* e gráficos de linhas. Essa capacidade de visualizar dados de maneira clara e intuitiva não apenas melhora a compreensão dos resultados, mas também aprimora a capacidade de tomada de decisão baseada em dados.

O *plugin* Data Analysis do ChatGPT é igualmente essencial, amplificando as ca-

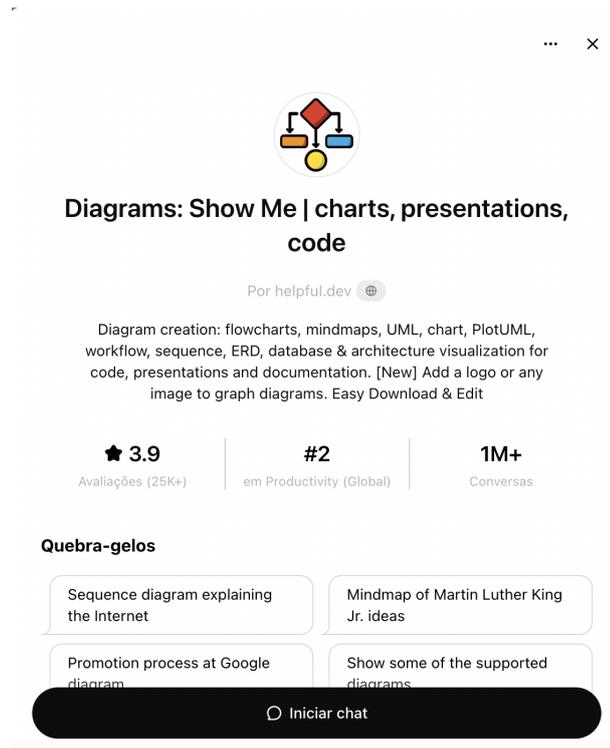


Figura 4.3: Iniciar conversa no ChatGPT usando um *plugin*

pacidades de análise estatística da plataforma. Permitindo a execução de uma vasta gama de análises estatísticas, desde cálculos descritivos básicos, como média, mediana e desvio padrão, até testes de hipóteses avançados, análise de variância, regressões lineares e não lineares, entre outros. A importância do *plugin Data Analysis* reside na sua capacidade de automatizar e simplificar processos que, tradicionalmente, demandariam tempo e esforço significativo. Com ele, os usuários podem rapidamente importar *datasets*, aplicar métodos estatísticos adequados e interpretar os resultados com facilidade. Além disso, o *plugin* oferece funcionalidades de visualização de dados, permitindo que os resultados sejam apresentados de maneira clara e intuitiva.

A integração dos *plugins Diagrams* e *Data Analysis* dentro do ChatGPT otimiza o fluxo de trabalho dos pesquisadores, garantindo maior precisão e eficiência na análise de dados e tomada de decisões informadas. A combinação de análises estatísticas avançadas com visualizações gráficas intuitivas torna o ChatGPT uma solução eficaz na área de análise de dados, oferecendo aos usuários uma ferramenta completa para explorar, analisar e comunicar seus achados de forma eficaz. Para realizar a análise com os dois *plugins* foram fornecidas as mesmas questões e base de dados utilizada na engenharia de

prompt, como pode ser observado na Figura 4.4.

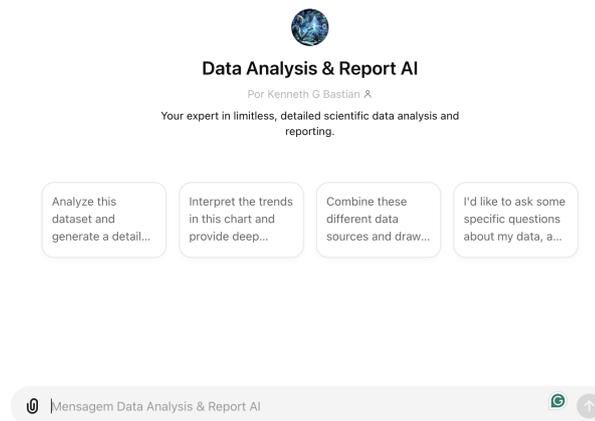


Figura 4.4: Perguntas a serem respondidas usando um *plugin*

O Akkio é uma ferramenta de análise de dados conhecida por sua facilidade de uso e integração com inteligência artificial. Voltada para a análise preditiva, Akkio permite a aplicação de métodos estatísticos sofisticados sem a necessidade de conhecimentos profundos em estatística. A plataforma sugere automaticamente modelos estatísticos com base nos dados fornecidos e oferece funcionalidades de visualização e interpretação de resultados, ajudando os usuários a entenderem as implicações práticas de suas análises.

O pacote estatístico R é uma das ferramentas mais robustas e amplamente utilizadas para análises estatísticas. Com uma vasta gama de pacotes e bibliotecas, R permite a aplicação de praticamente qualquer método estatístico existente. A escolha do método estatístico em R é altamente flexível, pois a ferramenta oferece funções para testes paramétricos e não paramétricos, modelos de regressão, análise de variância e muito mais. Além disso, a comunidade ativa de R fornece suporte contínuo e atualizações regulares, garantindo que os usuários tenham acesso às últimas inovações e melhores práticas em análise de dados.

Na escolha de um método de análise estatística para um processo automatizado de análise, é crucial considerar diversos critérios que garantam a eficácia e a precisão do processo. Primeiramente, a facilidade de uso é fundamental, pois ferramentas com interfaces intuitivas permitem uma rápida adaptação e utilização eficiente, especialmente para usuários com pouca experiência técnica, como evidencia Lewis (2014). Em segundo lugar, a flexibilidade da ferramenta é essencial, medindo sua capacidade de se adaptar a

diferentes tipos de análises e de se integrar com outros sistemas e ferramentas, garantindo que ela possa ser utilizada em diversas situações e para múltiplos propósitos, conforme Wickham e Grolemund (2016). Além disso, as funcionalidades avançadas são críticas, incluindo recursos como modelos de regressão, análise de variância e testes de normalidade, que permitem uma análise detalhada e precisa dos dados, segundo Boulesteix e Strimmer (2007). Por fim, o suporte e a documentação oferecidos pela ferramenta são vitais, pois um bom suporte técnico e uma documentação abrangente ajudam a resolver problemas rapidamente e a maximizar a eficiência e a precisão das análises, de acordo com Wilson et al. (2017). Considerando esses critérios, é possível selecionar a ferramenta mais adequada para atender às necessidades específicas do projeto e garantir a qualidade das análises estatísticas realizadas. A comparação entre as ferramentas seguindo os critérios discutidos pode ser observado na Tabela 4.2.

Critério	<i>Plugins</i>	Akkio	R
Facilidade de uso	Alta	Alta	Média
Flexibilidade	Alta	Média	Alta
Análise estatística	Alta	Média	Alta
Suporte e documentação	Média	Alta	Alta

Tabela 4.2: Comparação entre as ferramentas de análise estatística.

A escolha de usar *plugins* na engenharia de *prompt* do ChatGPT, em vez de ferramentas como Akkio (2024) ou R Core Team (2023), é baseada em várias vantagens significativas, conforme discutido no estudo de Wolf et al. (2020). O ChatGPT permite a formulação de consultas em linguagem natural, tornando a interação mais simples e acessível a usuários com diferentes níveis de experiência técnica. Oferece ainda uma integração robusta com outras ferramentas, melhorando a eficiência e a produtividade ao simplificar fluxos de trabalho complexos. Além disso, a capacidade do ChatGPT de interpretar e responder a *prompts* detalhados com alta precisão garante resultados precisos e personalizados.

A engenharia de *prompt* do ChatGPT é um componente fundamental para maximizar a eficácia da ferramenta em análises estatísticas. Esse processo envolve a criação de instruções claras e específicas que orientam a inteligência artificial a realizar tarefas precisas. Para conduzir uma análise estatística usando o ChatGPT, o primeiro passo é

identificar as variáveis relevantes. Em seguida, elabora-se um *prompt* inicial que descreve o contexto da análise e especifica as ações desejadas, como "Sou um professor universitário, pesquisador e estatístico. A partir da planilha informada, responda aos seguintes tópicos, considerando uma linguagem formal e um nível de significância de 5% em toda a análise". Dessa forma, para usar a engenharia de *prompt* no ChatGPT deve criar um novo GPT e ir em configurações onde será possível ajustar o *prompt* para orientar a IA, como pode ser observado na Figura 4.5

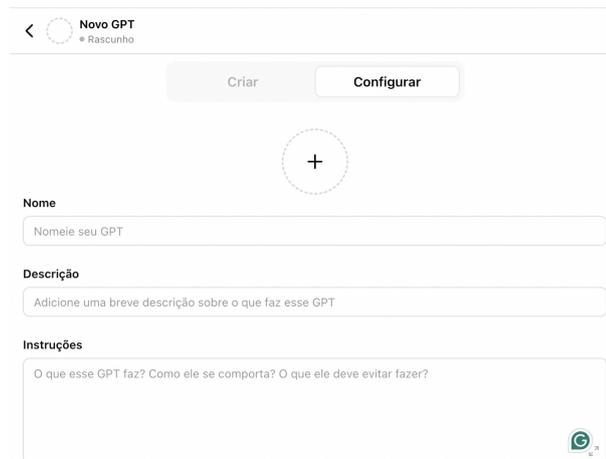


Figura 4.5: Iniciar conversa no ChatGPT usando *plugin*

Além disso, a engenharia de *prompt* permite a integração de *plugins* e ferramentas estatísticas externas, ampliando as capacidades analíticas do ChatGPT. Em resumo, a engenharia de *prompt* é fundamental para orientar o ChatGPT na execução de análises estatísticas complexas, garantindo resultados precisos e relevantes que podem apoiar decisões baseadas em dados.

4.3.3 Ferramenta de automação

No contexto da automação de fluxos de trabalho, a escolha da ferramenta adequada pode impactar significativamente a eficiência e produtividade das operações. Ferramentas como Zapier (2012), Automate (2024) e UiPath (2024) se destacam, cada uma oferecendo funcionalidades específicas que atendem a diferentes necessidades e níveis de complexidade. Ao automatizar e integrar tarefas repetitivas e demoradas, pesquisadores podem redirecionar seu tempo e esforço para análises mais aprofundadas, fazendo contribuições originais ao

campo científico e resultando na geração de novos dados de pesquisa. Esta seção apresenta uma análise comparativa dessas plataformas, ressaltando suas principais características e vantagens.

O Zapier é uma plataforma de automação baseada na nuvem que facilita a integração entre várias aplicações e serviços sem a necessidade de codificação. Com uma interface intuitiva, Zapier permite que os usuários criem automações chamadas "Zaps" para realizar tarefas repetitivas, como transferências de dados e envio de notificações. Suportando mais de 3.000 aplicativos, é ideal para pequenas e médias empresas que buscam simplificar processos rotineiros de maneira rápida e eficiente, mesmo para usuários com pouca experiência técnica.

O Microsoft Power Automate, parte do ecossistema Microsoft Power Platform, permite a criação de fluxos de trabalho automatizados que conectam serviços como Office 365, Dynamics 365 e Azure, além de aplicativos de terceiros. Oferecendo recursos avançados de automação de processos robóticos (RPA - *Robotic Process Automation*), Power Automate é adequado para automatizar tarefas complexas e sistemas legados. Embora tenha uma curva de aprendizado mais acentuada, sua profunda integração com produtos Microsoft torna-o ideal para empresas que já utilizam essas ferramentas e buscam soluções de automação escaláveis e integradas.

O UiPath é uma plataforma líder em automação de processos robóticos (RPA), projetada para automatizar processos empresariais complexos. A plataforma oferece ferramentas abrangentes para desenvolver, gerenciar e executar robôs de software que imitam ações humanas em sistemas digitais. UiPath é altamente escalável e pode ser usado em diversas áreas, como finanças, recursos humanos e atendimento ao cliente. Com capacidades avançadas de inteligência artificial e *machine learning*, é ideal para grandes empresas que necessitam de uma solução robusta e escalável para automação de processos críticos e de alto volume.

Na escolha de uma ferramenta de automação para um processo automatizado, é essencial considerar vários critérios para garantir a eficácia e a eficiência do processo. Primeiramente, a facilidade de uso é fundamental, pois a ferramenta deve ser intuitiva e acessível, permitindo que os usuários a operem sem necessidade de treinamento extensivo,

como descreve Johnson (2022). A capacidade de integração é igualmente importante, pois a ferramenta deve se comunicar de forma eficaz com outras plataformas e sistemas utilizados no processo, como bases de dados acadêmicas, ferramentas de processamento de linguagem natural e softwares de gerenciamento de referências, conforme Smith (2023).

A curva de aprendizado associada à ferramenta é outro critério crucial; uma curva de aprendizado íngreme pode atrasar a implementação do processo automatizado e reduzir a produtividade inicial, segundo Lee (2022). Além disso, a robustez das funcionalidades oferecidas pela ferramenta é vital para assegurar que possa lidar com as várias etapas do processo, desde a geração de palavras-chave até a consolidação e análise dos dados coletados, como Wilson (2023) evidencia em sua pesquisa. A escalabilidade da ferramenta também deve ser considerada, garantindo que ela possa crescer e se adaptar às necessidades futuras do projeto, acompanhando o aumento do volume de dados e a complexidade das tarefas, conforme Taylor (2022).

A segurança e a conformidade com as regulamentações aplicáveis são aspectos imprescindíveis, pois é necessário garantir a proteção dos dados sensíveis e a integridade do processo automatizado, como Anderson (2023). Por fim, o suporte técnico disponível para a ferramenta deve ser avaliado, já que um bom suporte pode resolver problemas rapidamente e minimizar o tempo de inatividade, garantindo a continuidade e a eficiência do processo, segundo Martinez (2022). Esses critérios foram considerados para assegurar que a ferramenta escolhida não só atenda às necessidades técnicas do projeto, mas também proporcione uma experiência de uso eficiente e segura, maximizando a qualidade e a precisão dos resultados obtidos. A comparação entre as ferramentas de automação pode ser observada na Tabela 4.3.

Critério	Zapier	Microsoft Power Automate	UiPath
Facilidade de Uso	Alta	Média	Média
Capacidade de Integração	Alta	Média	Alta
Curva de Aprendizado	Média	Média	Alta
Robustez das Funcionalidades	Alta	Alta	Alta
Escalabilidade	Alta	Alta	Média
Segurança e Conformidade	Alta	Alta	Alta
Suporte Técnico	Média	Alta	Média

Tabela 4.3: Comparação entre Zapier, Microsoft Power Automate e UiPath.

A automação de processos representa um desenvolvimento fundamental na eficiência

operacional em vários domínios, permitindo que tarefas repetitivas sejam executadas com maior precisão e em menor tempo, liberando recursos humanos para atividades que requerem julgamento crítico e pensamento criativo. Neste contexto, a integração de tecnologias de automação em ambientes corporativos e acadêmicos tem se intensificado, facilitando uma transformação digital abrangente que otimiza desde simples tarefas administrativas até complexas operações de dados.

O Zapier foi escolhido em vez de ferramentas como Microsoft Power Automate e UiPath devido à sua habilidade de automatizar fluxos de trabalho, integrando diversas aplicações e ferramentas de forma eficiente. A plataforma permite a criação de "Zaps", automações configuradas para realizar tarefas repetitivas e transferir dados entre diferentes sistemas sem a necessidade de intervenção manual. Essa capacidade é particularmente útil para otimizar processos de coleta, organização e análise de dados, reduzindo significativamente o tempo e o esforço necessários para essas atividades. Além disso, a facilidade de uso do Zapier, aliada à sua ampla compatibilidade com uma vasta gama de aplicações, o torna uma escolha ideal para automatizar fluxos de trabalho em contextos de pesquisa. Em comparação, Microsoft Power Automate e UiPath, embora robustos, frequentemente apresentam uma curva de aprendizado mais acentuada e podem exigir mais configuração para alcançar a mesma fluidez e eficiência proporcionadas pelo Zapier.

Entre as ferramentas disponíveis para automação de processos, o Zapier destaca-se como uma plataforma versátil que interconecta diferentes ferramentas, permitindo a criação de fluxos de trabalho automatizados sem a necessidade de intervenção programática detalhada. A escolha do Zapier é justificada pela sua ampla compatibilidade com uma vasta gama de aplicações e serviços *online*, desde soluções de armazenamento como Google Drive até avançados sistemas de inteligência artificial oferecidos pela OpenAI. Esta compatibilidade, aliada à facilidade de uso através de uma interface gráfica intuitiva, torna o Zapier uma escolha excelente para organizações que buscam implementar soluções de automação robustas sem investir em desenvolvimento de software customizado.

4.4 Considerações do capítulo

Neste capítulo, foram abordadas os métodos e ferramentas utilizados nessa pesquisa, destacando a aplicação de testes estatísticos, a engenharia de *prompt* para interação com IA, a seleção das ferramentas ChatGPT e Zapier e dos *plugins Diagrams* e *Data Analysis*. Os testes estatísticos são fundamentais para validar as hipóteses e analisar a eficácia das intervenções. A engenharia de *prompt* otimizou as respostas da IA, alinhando-as com os objetivos da pesquisa. O uso do ChatGPT justificado por sua capacidade de processamento de linguagem natural, essencial para a análise de dados. A automação tem como objetivo ajudar no fluxo de trabalho, melhorando a eficiência da coleta e análise de dados. A combinação dessas estratégias mostrou-se importante para atingir os objetivos do estudo e sublinhou a importância de abordagens interdisciplinares na pesquisa acadêmica moderna.

5 Estudo de caso

Este capítulo apresenta um estudo de caso cujo principal objetivo é detalhar a automação da análise estatística com o suporte de Inteligência Artificial generativa. O estudo de caso visa demonstrar, passo a passo, desde a aplicação dos métodos estatísticos até a automação por completo, conforme Figura 5.1. Para a demonstração, serão utilizados dados extraídos de uma base real, proporcionando uma visão prática da implementação dos métodos discutidos.

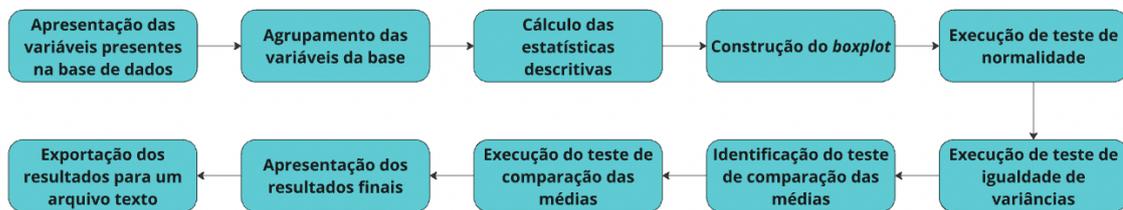


Figura 5.1: Fluxograma com as etapas do estudo de caso

5.1 Automação

O processo de automação utilizando o Zapier envolve quatro etapas principais: detecção de novas entradas em uma planilha do Google Sheets, recuperação de múltiplas linhas dessa planilha, envio de um *prompt* para a API da OpenAI e criação de um documento no Google Docs. Cada uma dessas etapas desempenha um papel importante no processo automatizado, garantindo que os dados sejam processados e documentados de maneira eficiente e precisa.

A primeira etapa do processo é a detecção de novas ou atualizadas entradas em uma planilha do Google Sheets. Esse passo é configurado como o gatilho do fluxo de trabalho no Zapier. Um gatilho, em termos de automação, é um evento que inicia uma sequência de ações pré-determinadas. Nesse caso, o gatilho é ativado sempre que uma nova linha é adicionada ou uma linha existente é atualizada na planilha. Esse gatilho foi escolhido por sua capacidade de monitorar mudanças em tempo real, garantindo que

todas as novas informações sejam imediatamente processadas. Para configurar o gatilho no Zapier foram seguidos os seguintes passos:

- Escolha do Aplicativo e Evento: na interface do Zapier, selecionar Google Sheets como aplicativo e "New or Updated Spreadsheet Row" como o evento do gatilho, como apresentado na Figura 5.2. Assim, toda vez que uma nova linha for adicionada ou uma existente for atualizada na planilha, o fluxo de trabalho será acionado;

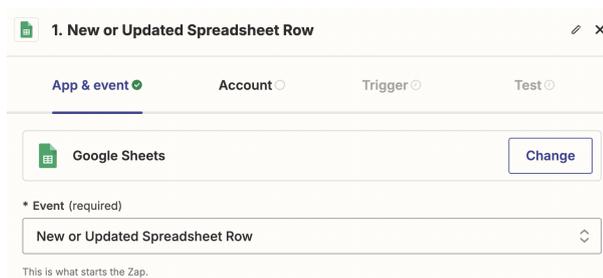


Figura 5.2: Configuração do gatilho no Zapier: escolha do aplicativo e evento.

- Conexão com a Conta Google: a conta Google foi conectada ao Zapier, garantindo que o Zapier tenha permissão para acessar as planilhas necessárias, conforme Figura 5.3;

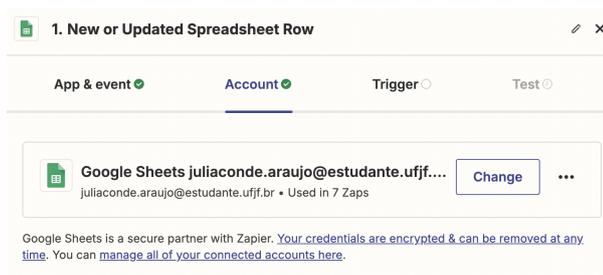


Figura 5.3: Configuração do gatilho no Zapier: conexão com a conta Google.

- Configuração da Planilha: selecionar a planilha específica e a aba (*sheet*) onde as entradas serão monitoradas, como apresentado na Figura 5.4. Esse passo é crucial para direcionar corretamente o gatilho;

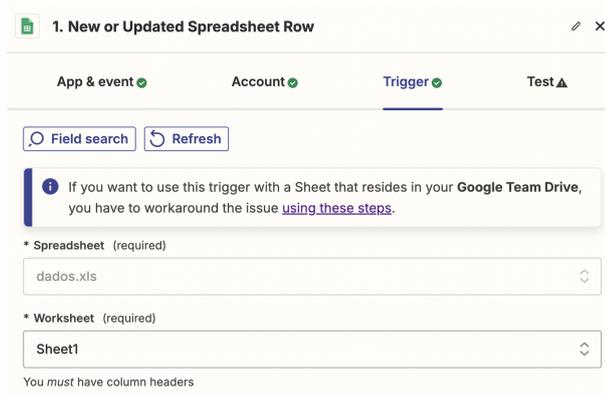


Figura 5.4: Configuração do gatilho no Zapier: configuração da planilha.

- Testar o Gatilho: executar um teste para assegurar que o Zapier consegue detectar novas ou atualizadas entradas na planilha. O teste deve confirmar a capacidade de monitorar mudanças em tempo real, garantindo que todas as novas informações sejam imediatamente processadas, conforme apresentado na Figura 5.5.

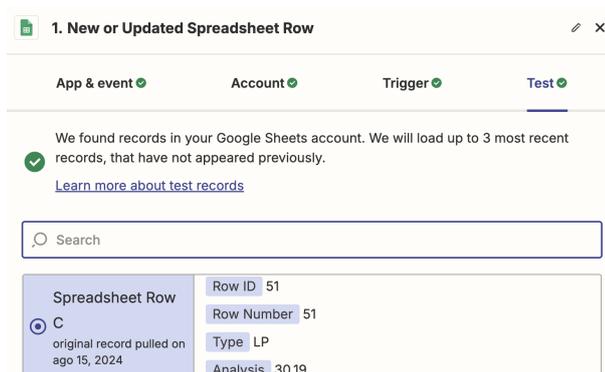


Figura 5.5: Configuração do gatilho no Zapier: teste do gatilho.

Após a ativação do gatilho, a próxima etapa é a recuperação de múltiplas linhas da planilha do Google Sheets. Essa ação utiliza a funcionalidade avançada de obtenção de dados para extrair todas as linhas relevantes, que serão posteriormente analisadas e processadas. A seleção das linhas a serem recuperadas pode ser ajustada conforme a necessidade, permitindo uma flexibilidade na quantidade e no tipo de dados extraídos. Assim, para configurar essa etapa no Zapier, foram seguidos os seguintes passos:

- Adicionar Ação: adicionar uma nova ação ao Zap, selecionando novamente o Google Sheets como aplicativo, conforme Figura 5.6;

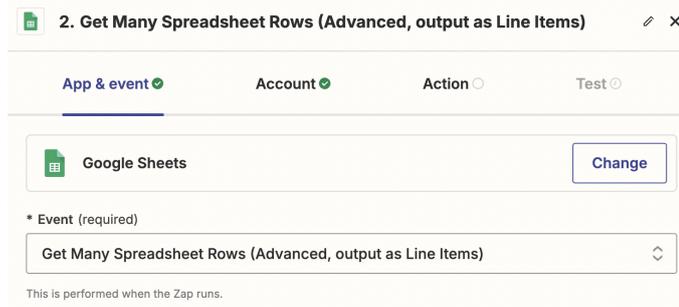


Figura 5.6: Configuração dos parâmetros do Google Sheets: adicionar ação.

- Escolha do Evento: selecionar "Get Many Spreadsheet Rows (Advanced, output as Line Items)" como o evento, Figura 5.7;

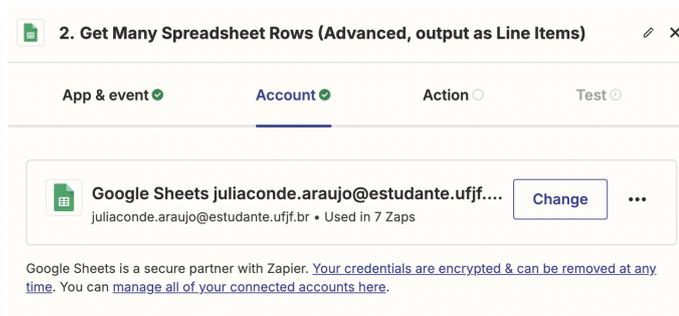


Figura 5.7: Configuração dos parâmetros do Google Sheets: escolha do evento.

- Configuração dos Parâmetros: configurar os parâmetros para definir quais linhas devem ser recuperadas, como apresentado na Figura 5.8;

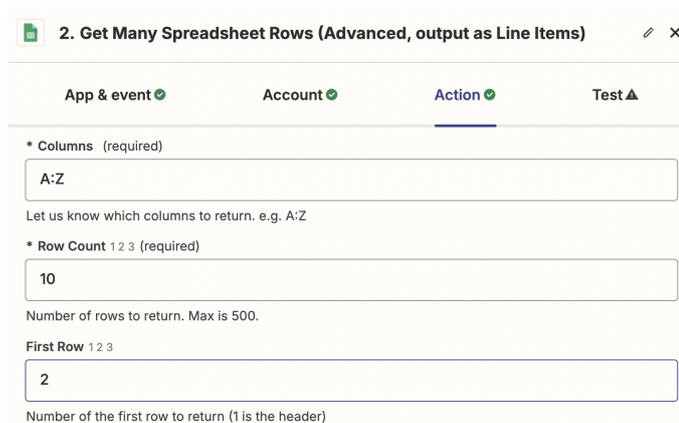


Figura 5.8: Configuração dos parâmetros do Google Sheets: configuração dos parâmetros.

- Testar a Ação: executar um teste para assegurar que as linhas corretas estão sendo recuperadas conforme os critérios estabelecidos, conforme Figura 5.9.

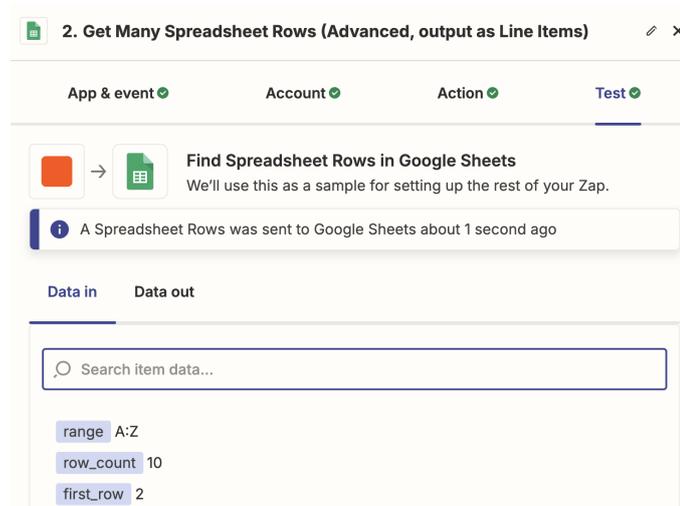


Figura 5.9: Configuração dos parâmetros do Google Sheets: testar ação.

A terceira etapa envolve o envio de um *prompt* para a API da OpenAI. O *prompt* que será formulado posteriormente nessa pesquisa solicita à API que gere uma análise estatística dos dados fornecidos. A API da OpenAI, utilizando modelos avançados de linguagem como o GPT-4o, processa o *prompt* e retorna um texto gerado que sintetiza e interpreta os dados fornecidos. Esse passo é fundamental para transformar dados brutos em informações significativas e bem estruturadas. Nessa etapa os passos seguidos para configurar a ação no Zapier foram:

- Adicionar Ação: adicionar uma nova ação, desta vez selecionando o OpenAI como aplicativo, conforme Figura 5.10;

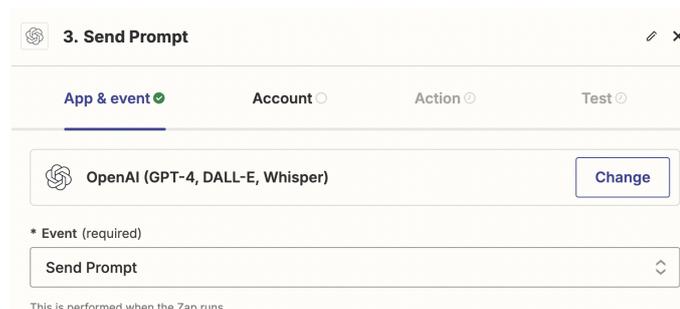


Figura 5.10: Configuração dos parâmetros da OpenAI: adicionar ação.

- Escolha do Evento: selecionar "Send Prompt in OpenAI (GPT-4, DALL-E, Whisper)" como o evento, como apresentado na Figura 5.11;

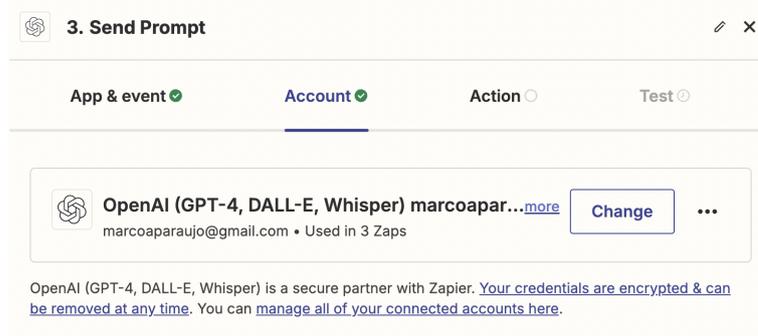


Figura 5.11: Configuração dos parâmetros da OpenAI: escolha do evento.

- Configuração dos Parâmetros: definir o modelo do ChatGPT que será usado e informar o *prompt* e os dados a serem analisados (Figura 5.12). Deve verificar a adequação do modelo e a correta formatação do *prompt*, ajustando os dados de entrada conforme necessário para obter resultados precisos;

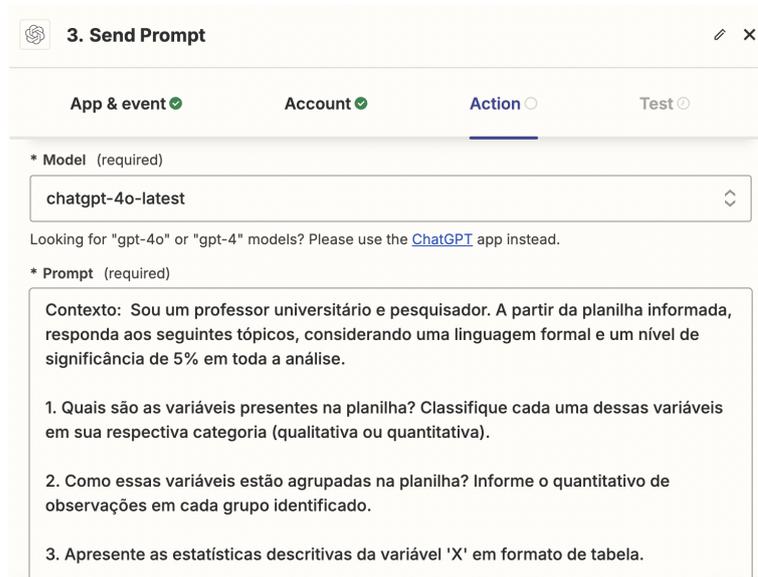


Figura 5.12: Configuração dos parâmetros da OpenAI: configuração dos parâmetros.

- Testar a Ação: executar um teste para garantir que a API da OpenAI receba o *prompt* corretamente e retorne o texto gerado conforme esperado, se certificando de que o teste foi concluído com sucesso antes de prosseguir, conforme apresentado na Figura 5.13.

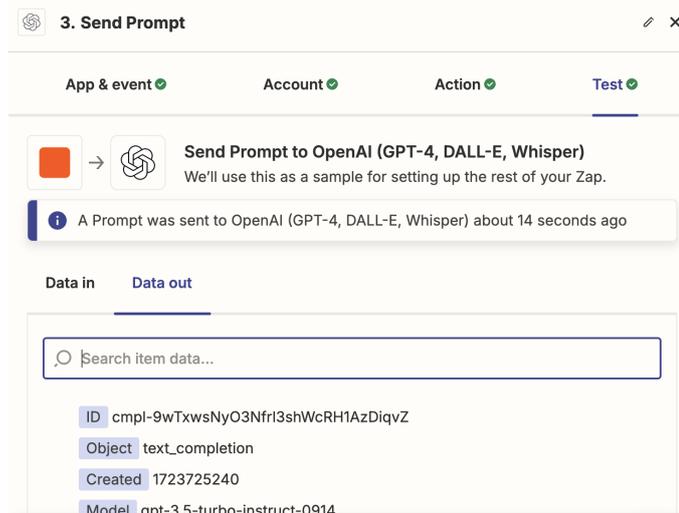


Figura 5.13: Configuração dos parâmetros da OpenAI: testar Ação.

Por fim, os resultados gerados pela IA são automaticamente inseridos em um documento do Google Docs. Essa etapa finaliza o fluxo de automação, consolidando os resultados em um formato organizado e acessível para análise ou compartilhamento posterior. Essa capacidade de documentação automática é especialmente útil para empresas e pesquisadores que necessitam manter registros detalhados de suas operações e resultados de análise. A configuração dessa etapa no Zapier deve-se seguir os passos:

- Adicionar Ação: adicionar uma nova ação, selecionando o Google Docs como aplicativo, como pode ser observado na Figura 5.14;

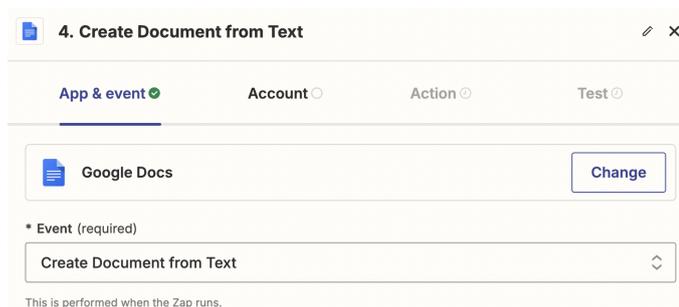


Figura 5.14: Configuração dos parâmetros da OpenAI: adicionar ação.

- Escolha do Evento: selecionar *"Create Document from Text"* como o evento, conforme Figura 5.15;

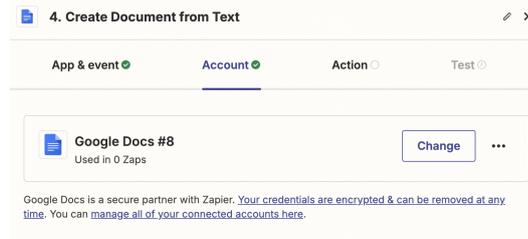


Figura 5.15: Configuração dos parâmetros da OpenAI: escolha do evento.

- Configuração do Documento: informar o título do documento e o texto gerado pela API da OpenAI como o conteúdo do documento, Figura 5.16;

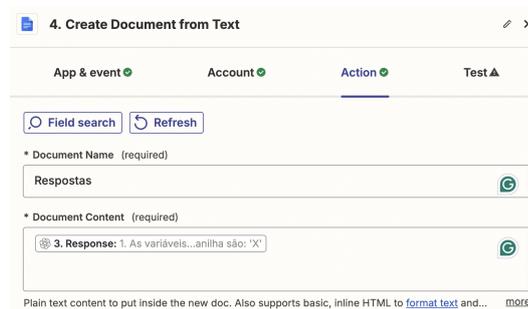


Figura 5.16: Configuração dos parâmetros da OpenAI: configuração do documento.

- Testar a Ação: executar um teste para assegurar que o documento é criado corretamente no Google Docs com o conteúdo gerado pela API, como pode ser observado na Figura 5.17.

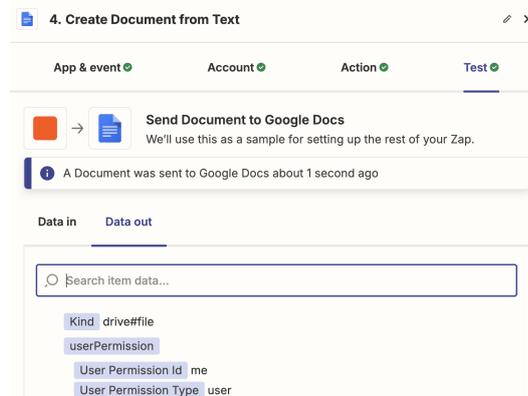


Figura 5.17: Configuração dos parâmetros da OpenAI: testar ação.

Através deste estudo de caso sobre o uso do Zapier para automação de processos interligados com IA Generativa, fica evidente que a automação não apenas simpli-

fica a gestão de tarefas rotineiras, mas também potencializa a utilização de tecnologias avançadas, contribuindo significativamente para a produtividade e inovação dentro das organizações. Assim, o Zapier se apresenta não apenas como uma ferramenta de automação, mas como um facilitador de transformação digital abrangente. Na Figura 5.18 é possível observar o fluxo completo da automação realizada neste estudo.

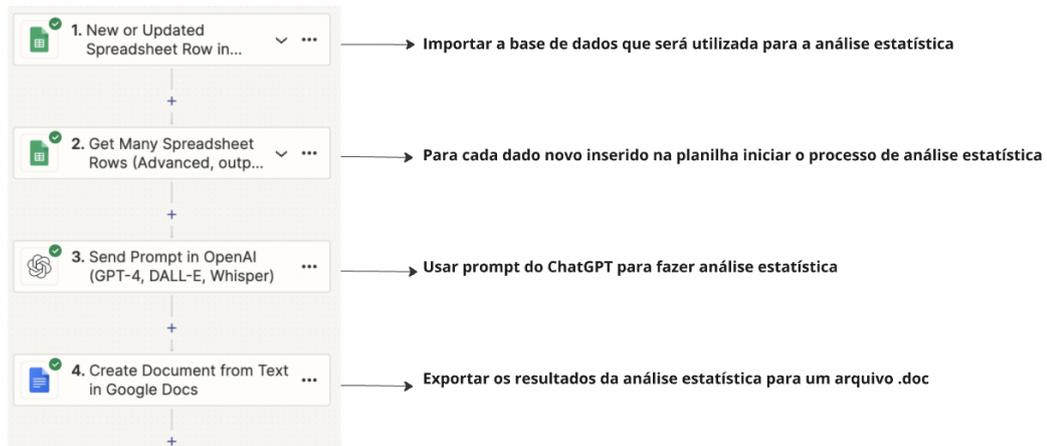


Figura 5.18: Fluxo da automação completa no Zapier.

5.2 Engenharia de *prompt*

A interação com modelos de linguagem avançados, como o GPT-4, tornou-se cada vez mais relevante na condução de análises acadêmicas e científicas. A eficácia dessas interações está intrinsecamente ligada à construção de *prompts* adequados, que servem como o canal de comunicação entre o usuário e o modelo de IA. Essa seção explora a construção de um *prompt* para uma análise estatística específica, justificando cada elemento com base nos princípios de engenharia de *prompt* descritos por Larguesa (2024).

Larguesa (2024) enfatiza que a engenharia de *prompt* é fundamental para maximizar a eficácia das interações entre humanos e sistemas de IA. Um *prompt* bem estruturado deve seguir princípios de clareza, especificidade e relevância contextual. A ausência desses elementos pode resultar em respostas imprecisas ou irrelevantes, o que é particularmente crítico em contextos acadêmicos, onde a precisão é essencial.

O *prompt* construído nesse trabalho, apresenta uma sequência de instruções e foi cuidadosamente construído com base nas diretrizes estabelecidas por Larguesa (2024). A

variável "X" deve ser modificada pela variável que deseja fazer a análise estatística. O *prompt* foi construído da seguinte forma:

Contexto:Sou um professor universitário e pesquisador. A partir da planilha informada, responda aos seguintes tópicos, considerando uma linguagem formal e um nível de significância de 5% em toda a análise.

1. Caracterize as variáveis presentes na planilha? Classifique cada uma dessas variáveis em sua respectiva categoria (qualitativa ou quantitativa).
2. Como essas variáveis estão agrupadas na planilha? Informe o quantitativo de observações em cada grupo identificado.
3. Apresente as estatísticas descritivas da variável 'X' em formato de tabela.
4. Construa um boxplot para a variável 'X', considerando os grupos previamente identificados.
5. Faça o teste de normalidade da variável 'X'. Apresente o teste de hipóteses utilizando o método estatístico adequado à quantidade de dados: Caso a amostra possua mais de 30 valores, utilize o Teste de Kolmogorov-Smirnov; caso contrário, utilize o Teste de Shapiro-Wilk.
6. Caso a análise anterior indique a necessidade, realize o teste de igualdade de variâncias para a variável 'X' utilizando o Teste de Levene. Apresente o teste de hipóteses e o resultado deste teste.
7. Baseando-se nas análises anteriores (normalidade e homocedasticidade), identifique e justifique o teste adequado para a comparação das médias da variável 'X': Caso a distribuição dos dados seja normal e as variâncias entre os grupos sejam homogêneas, utilize o Teste T para comparar dois grupos. ou ANOVA para comparar mais de dois grupos em conjunto com o método de Tukey para identificar as amostras diferentes. Se a distribuição dos dados não for normal ou as variâncias entre os grupos não forem homogêneas, utilize o Teste de Mann-Whitney para a comparação de dois grupos ou o Teste de Kruskal-Wallis para a comparação de mais de dois grupos em conjunto com o método de Mann-Whitney para identificar as amostras diferentes.

8. Realize o teste de comparação de médias utilizando o método adequado. Apresente o teste de hipóteses e os resultados obtidos.
9. Apresente o resultado final, especificando se há uma diferença significativa entre as médias dessa variável.
10. Exporte todos os resultados obtidos, incluindo os gráficos, para um documento em formato .doc, utilizando uma linguagem formal apropriada para sua inclusão em um artigo científico.

A estruturação do *prompt* para a análise estatística seguiu uma lógica rigorosa e meticulosamente planejada. Em primeiro lugar, foi essencial iniciar com uma instrução clara que destacasse o contexto da pesquisa, explicitando a identidade do solicitante como “um professor universitário e pesquisador” (contexto). Essa introdução é fundamental para situar o modelo de IA dentro de um ambiente acadêmico, onde a precisão e a formalidade são primordiais. A instrução, “A partir da planilha informada, responda aos seguintes tópicos, considerando uma linguagem formal e um nível de significância de 5% em toda a análise”, foi projetada para garantir que o modelo compreenda as expectativas e os padrões exigidos em uma análise científica rigorosa. Essa abordagem inicial estabelece o tom da interação e assegura que o modelo aplique uma metodologia adequada, seguindo as normas acadêmicas e respondendo de forma precisa às questões propostas.

Logo depois, foi solicitado que a IA generativa classificasse as variáveis presentes na planilha (instrução 1). Segundo o autor, ao iniciar a interação com um pedido de classificação, estabelece-se um fundamento sólido para todas as análises subsequentes. A solicitação para que o modelo classifique as variáveis como qualitativas ou quantitativas garante que o modelo compreenda a natureza das variáveis desde o início, eliminando ambiguidades e facilitando a análise subsequente.

Em seguida, o *prompt* orienta o modelo a identificar padrões de agrupamento e quantificar as observações dentro de cada grupo (instrução 2). A clareza nesta etapa, conforme defendido pelo pesquisador em seu livro, assegura que o modelo execute a tarefa de maneira ordenada e lógica, minimizando o risco de respostas inadequadas ou incompletas. Essa orientação específica é importante para garantir que o modelo não apenas identifique

os grupos, mas também os quantifique corretamente.

A apresentação de estatísticas descritivas da variável de interesse em formato de tabela foi outro passo essencial na construção do *prompt* (instrução 3). Na construção do *prompt* a variável a ser analisada deve ser evidenciada, deve-se substituir o "X" no *prompt* pela variável desejada. De acordo com o autor, um *prompt* eficaz deve não apenas direcionar o conteúdo da resposta, mas também especificar o formato em que essa resposta deve ser entregue. Ao solicitar explicitamente que as estatísticas descritivas sejam apresentadas em uma tabela, garante-se que o modelo produza resultados que sejam imediatamente úteis para o contexto acadêmico.

Além disso, a construção de um *boxplot* para a variável identificada, considerando os grupos previamente agrupados, segue o princípio da clareza e contextualização defendido pelo autor (instrução 4). Essa parte do *prompt* não apenas instrui o modelo a criar a visualização, mas também a fazê-lo com base no contexto já estabelecido, assegurando que o gráfico reflita corretamente os dados analisados.

Quando se trata de realizar testes de normalidade, o *prompt* foi cuidadosamente estruturado para incorporar a especificidade e a adaptabilidade (instrução 5). A escolha entre o Teste de Kolmogorov-Smirnov ou o Teste de Shapiro-Wilk, dependendo do tamanho da amostra, demonstra um entendimento refinado das práticas estatísticas, garantindo que o modelo utilize o método mais apropriado, o que é fundamental para a validade dos resultados, como salientado por Larguesa (2024).

O teste de igualdade de variâncias, solicitado condicionalmente com base nos resultados anteriores, exemplifica a aplicação da lógica condicional na construção do *prompt* (instrução 6). O autor sugere que antecipar possíveis caminhos analíticos torna os *prompts* mais eficazes. Ao guiar o modelo para executar o teste de Levene apenas se necessário, o *prompt* reflete uma compreensão profunda do fluxo de análise estatística.

Ao solicitar a comparação das médias (instrução 7), o *prompt* foi desenhado para garantir que o modelo escolha o método estatístico adequado, baseado em resultados de normalidade e homocedasticidade anteriores. Caso a distribuição dos dados seja normal e as variâncias entre os grupos sejam homogêneas, utiliza-se o Teste T para comparar dois grupos. Caso a distribuição dos dados seja normal e as variâncias entre os grupos sejam

homogêneas, utiliza-se ANOVA juntamente com o método de Tukey para identificar as amostras diferentes para comparar mais de dois grupos. Se a distribuição dos dados não for normal ou as variâncias entre os grupos não forem homogêneas, utiliza-se o Teste de Mann-Whitney para a comparação de dois grupos. Se a distribuição dos dados não for normal ou as variâncias entre os grupos não forem homogêneas no caso de comparar mais de dois grupos utiliza-se o Teste de Kruskal-Wallis juntamente com o método de Mann-Whitney para identificar as amostras diferentes. O autor argumenta que pedir ao modelo para justificar a escolha do método promove um maior engajamento analítico, resultando em respostas que são tanto metodologicamente sólidas quanto cientificamente justificadas.

Finalmente, o *prompt* conclui orientando o modelo a realizar o teste de comparação de médias apropriado (instrução 8) e a apresentar um resumo final (instrução 9), especificando se há ou não diferença significativa entre as médias. Essa conclusão clara, como preconizado pelo autor, ajuda a evitar respostas ambíguas, garantindo que o modelo produza um resumo final útil e alinhado às expectativas acadêmicas. Por fim, o resultado obtido deve ser exportado para um arquivo .doc (instrução 10).

Em resumo, a construção de um *prompt* eficaz exige uma combinação de clareza, especificidade e contexto, como detalhado por Larguesa (2024). Aplicando esses princípios na estruturação de *prompts* para análise estatística, é possível guiar modelos de linguagem avançados a gerar respostas que não apenas atendam às expectativas acadêmicas, mas que também promovam uma interação produtiva e metodologicamente sólida.

5.3 Execução do processo com um fator e dois tratamentos

A utilização da IA Generativa possibilita a automação da análise de dados, por meio de ferramentas como Chatgpt (2022). Com essa abordagem, é possível manipular grandes conjuntos de dados e conduzir análises estatísticas detalhadas para apoiar uma pesquisa, simplificando e apoiando a exploração de dados. Assim, nessa seção foi realizada uma análise estatística, manipulando uma base de dados real. O principal objetivo dos dados

dessa base, é verificar se existem diferenças significativas na utilização de uma técnica identificação de defeitos de software nas versões inglês e português, em relação ao tempo, número de discrepâncias (falsos-positivos) e defeitos encontrados. Para automatizar esse processo foi utilizado o ChatGPT, capaz de interagir com um conjunto de dados apoiando a análise estatística. Primeiramente, foi importado um conjunto de dados no Google Sheets, seguindo o primeiro passo do processo de automação no Zapier, como descrito na seção anterior, que foi utilizado como base dessa pesquisa, como pode ser observado na Figura 5.19. Conforme o segundo passo da automação no Zapier o processo irá se repetir caso haja alguma informação nova na base de dados fornecida.

Group	Participant	Time	Discrepan	Defects
EP	EP01	540	13	4
EP	EP02	900	17	5
EP	EP03	160	25	5
EP	EP04	630	25	5
EP	EP05	106	19	7
EP	EP06	73	15	4
EP	EP07	200	8	1
EP	EP08	607	3	2
EP	EP09	200	24	4
EP	EP10	240	9	3
EP	EP11	240	28	7
PP	PP02	540	20	2
PP	PP03	167	37	4
PP	PP04	450	13	5
PP	PP05	80	7	2
PP	PP07	667	27	9
PP	PP08	450	9	2
PP	PP10	180	14	6
PP	PP11	193	8	3

Figura 5.19: Base de dados usada no estudo de caso

Seguindo com o processo a terceira etapa foi conectar a base de dados com o *prompt* do ChatGPT, fornecendo os tópicos a serem respondidos. Assim, o primeiro passo desse estudo foi definir quais seriam as variáveis presentes na base utilizada e a categoria de cada uma delas, como evidenciado no *prompt* descrito na seção anterior, essa etapa é de extrema importância em uma análise estatística para entender os dados e definir como serão tratados (instrução 1). O resultado dessa busca pode ser observado na Figura 5.20, que apresenta as variáveis identificadas sendo *Group*, *Participant*, *Time*, *Discrepancies* e *Defects*. A variável *Group* apresenta dois grupos sendo participantes de

língua inglesa (EP) e participantes de língua portuguesa (PP).

- **Group:** Categoria, representando o grupo ao qual o participante pertence.
- **Participant:** Categoria, identificando individualmente cada participante.
- **Time:** Quantitativa, representando o tempo gasto (presumivelmente em segundos).
- **Discrepancies:** Quantitativa, representando o número de discrepâncias.
- **Defects:** Quantitativa, representando o número de defeitos.

Figura 5.20: Variáveis presentes na base de dados

Além disso, as variáveis foram automaticamente categorizadas, como demonstrado na Figura 5.21 (instrução 2).

- **Group:** Divide os dados em grupos, cada um representando uma técnica de leitura diferente.
- **Participant:** Serve como identificador único para os indivíduos dentro de cada grupo.
- **Time, Discrepancies, Defects:** Estas variáveis são medidas para cada participante e estão agrupadas sob os identificadores de "Group" e "Participant".

Figura 5.21: Agrupamento das variáveis

5.3.1 Método paramétrico na análise de variáveis em estudos de um fator e dois tratamentos

Seguindo com a abordagem proposta, uma das variáveis presente na base foi escolhida para continuar a análise, a variável Defeitos (*Defects*), que representa o número de defeitos na realização de uma tarefa. Assim, para começar essa análise, foi solicitado no *prompt* que a IA Generativa apresentasse as estatísticas descritivas para a variável em questão por grupo (instrução 3), usada para descrever, organizar e resumir conjunto de dados de uma variável, como demonstrado na Figura 5.22. Essa visualização permite que pesquisadores realizem uma análise mais detalhada dos dados.

Group	count	mean	std	min	25%	50%	75%	max
EP	11.0	4.2727272727272725	1.8488325554743508	1.0	3.5	4.0	5.0	7.0
PP	8.0	4.125	2.4748737341529163	2.0	2.0	3.5	5.25	9.0

Figura 5.22: Estatística descritiva da variável Defeitos (*Defects*) por grupo

O próximo passo desse estudo foi construir um *boxplot* da variável Defeitos (*Defects*) por grupo (instrução 4), fornecendo uma representação gráfica dos dados, sendo: mínimo, primeiro quartil, mediana, terceiro quartil e máximo, sendo de extrema importância para pesquisadores seguirem com as análises estatísticas, como pode ser observado na Figura 5.23.

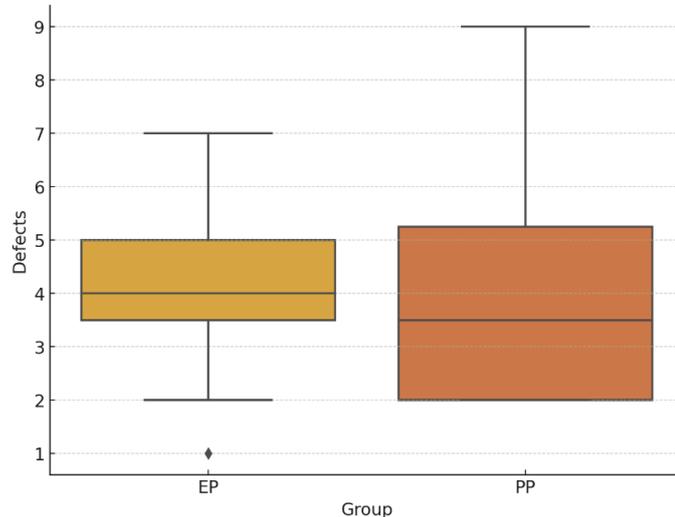


Figura 5.23: *Boxplot* da variável Defeitos (*Defects*) por grupo

Seguindo na análise foi feito o teste de normalidade da variável em questão (instrução 5), utilizado para verificar se a distribuição de probabilidade associada a um conjunto de dados pode ser aproximada a uma distribuição normal. Assim será utilizado o teste de Shapiro-Wilk, uma vez que a tabela possui menos que 30 amostras. Deve-se considerar as seguintes hipóteses com nível de significância de 5%:

- H0 (hipótese nula): as amostras apresentam distribuição normal;
- H1 (hipótese alternativa): as amostras não apresentam distribuição normal.

Assim, seguindo tópicos a serem respondidos, no *prompt* o próximo passo é apresentar o teste de normalidade, figura 5.24. Como pode ser observado na figura 5.25, o *p-value* é 0.3876 e é superior ao nível de significância estabelecido de 0,05 então a hipótese nula é aceita e os dados seguem uma distribuição normal.

Logo depois, foi realizado o teste de igualdade de variâncias, para a variável *Defects* (instrução 6), considerando seus grupos. Esse teste só é realizado caso seja necessário

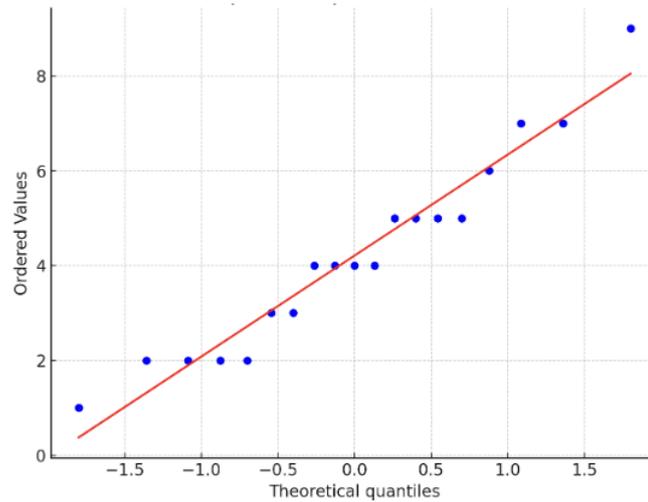


Figura 5.24: Análise de normalidade da variável defeitos

O teste de normalidade de Shapiro-Wilk foi realizado para a variável "Defects". O valor do teste de Shapiro-Wilk é 0.9495 e o valor-p é 0.3876.

Figura 5.25: *p-value* da variável *Defects*

avaliar a igualdade de variâncias de uma variável calculada para dois ou mais grupos, em uma distribuição normal. Para isso, o Teste de Levene é utilizado e considera as seguintes hipóteses com nível de significância de 5%:

- H0 (hipótese nula): as amostras são homocedásticas;
- H1 (hipótese alternativa): as amostras não são homocedásticas.

Se o *p-value* resultante do teste for maior que o nível de significância estabelecido (foi considerado 0,05), aceita-se a hipótese nula de que existe igualdade de variâncias, devendo-se utilizar um método paramétrico para a comparação dos dados. Em caso contrário, um método não paramétrico deve ser utilizado. Conforme pode ser observado na Figura 5.26, o *p-value* (valor-p) é superior ao nível de significância estabelecido, aceitando-se a hipótese nula de que os grupos possuem igualdade de variâncias e, com isso, um teste paramétrico deve ser utilizado na comparação dos grupos. O próximo passo foi realizar o teste apropriado para a comparação das médias da variável Defeitos (*Defects*), evidenciado no *prompt* utilizado neste trabalho (instrução 7). Esse teste foi realizado considerando as análises anteriores, uma vez que, com base nos resultados já obtidos, observa-se que

a variável estudada segue uma distribuição normal e as variâncias são iguais entre os grupos. O resultado dessa busca pode ser observado na Figura 5.27, pelo histograma e gráfico associado, concluindo que a variável segue distribuição normal.

Resultado do Teste de Levene:

- **Estatística do Teste:** 0.6351
- **Valor-p:** 0.4365

Figura 5.26: Teste de Levene

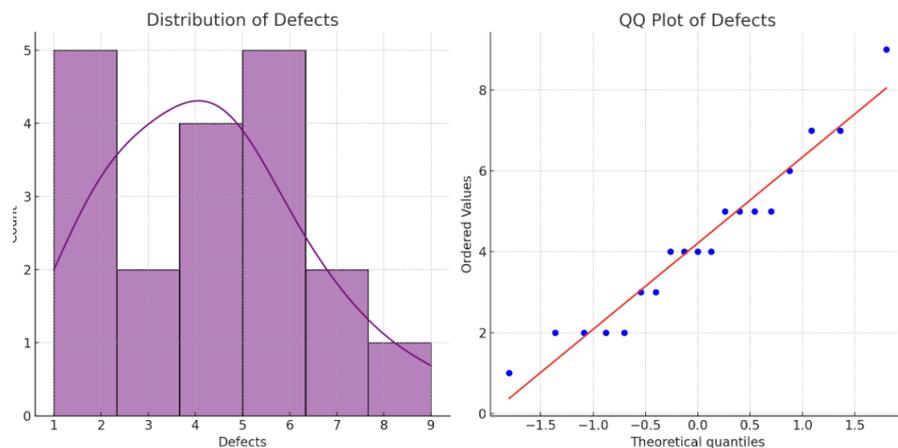


Figura 5.27: Gráficos para verificar se a variável Defeitos (*Defects*) apresenta distribuição normal

Com base nessas condições, a IA Generativa identificou corretamente qual seria o teste mais apropriado para a continuidade da análise, conforme detalhado nas instruções do *prompt* construído na seção anterior (instrução 8). Dessa forma, o teste t de *Student* deve ser aplicado para amostras independentes e foi utilizado para comparar as médias entre os dois grupos, considerando as seguintes hipóteses com nível de significância de 5%:

- H0 (hipótese nula): não há diferença entre as médias;
- H1 (hipótese alternativa): há diferença entre as médias.

Assim, conforme pode ser observado na Figura 5.28 (instrução 9), não existe uma diferença significativa nas médias da variável entre seus grupos, um vez que o *p-value* (valor-p) é superior ao nível de significância estabelecido e a hipótese nula foi aceita.

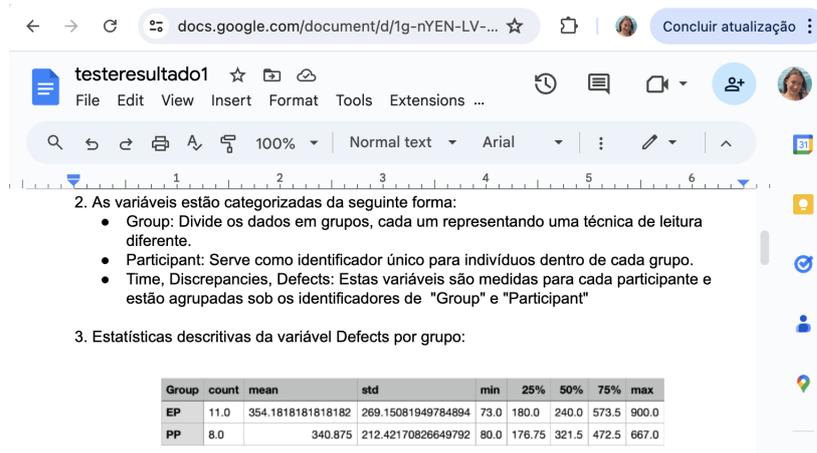
Resultado do Teste t de Student:

- **Estatística do Teste (t):** 0.1493
- **Valor-p:** 0.8831

Figura 5.28: Teste t de *Student*

Assim, pode-se concluir, que a variável *Defects*, estatisticamente, não apresenta diferença significativa das amostras da variável entre os grupos, considerando um nível de significância de 5%.

Por fim, no *prompt*, foi solicitado que exportasse todos os resultados das análises anteriores para um arquivo .doc (instrução 10), conforme figura 5.29, incluindo todos os gráficos realizados durante a análise.

Figura 5.29: Fragmento do documento gerado para a variável *Defects*

5.3.2 Método não paramétrico na análise de variáveis em estudos de um fator e dois tratamentos

Seguindo com a abordagem proposta, outra variáveis presente na base foi escolhida para continuar a análise, a variável Tempo (*Time*), que representa o tempo gasto na realização de uma tarefa. Assim, para começar essa análise, foi solicitado no *prompt* que a IA Generativa apresentasse as estatísticas descritivas para a variável em questão por grupo (instrução 3) usada para descrever, organizar e resumir conjunto de dados de uma variável, como demonstrado na Figura 5.30. Essa visualização permite que pesquisadores realizem

uma análise mais detalhada dos dados.

Group	count	mean	std	min	25%	50%	75%	max
EP	11.0	354.1818181818182	269.15081949784894	73.0	180.0	240.0	573.5	900.0
PP	8.0	340.875	212.42170826649792	80.0	176.75	321.5	472.5	667.0

Figura 5.30: Estatística descritiva da variável *Time* por grupo

O próximo passo desse estudo foi construir um *boxplot* da variável Tempo (*Time*) por grupo (instrução 4) fornecendo uma representação gráfica dos dados, sendo: mínimo, primeiro quartil, mediana, terceiro quartil e máximo, sendo de extrema importância para pesquisadores seguirem com as análises estatísticas, como pode ser observado na Figura 5.31.

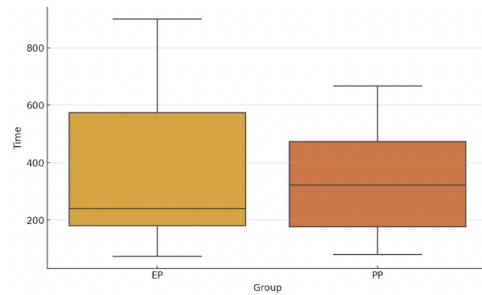
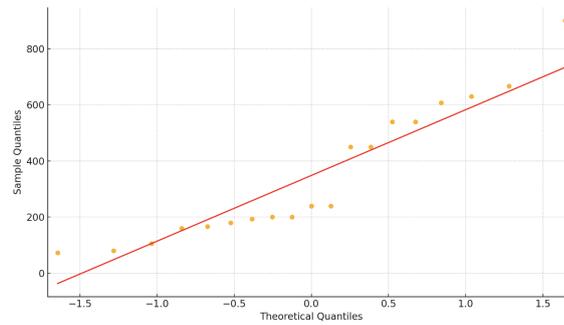


Figura 5.31: Boxplot da variável *Time* por grupo

Seguindo na análise foi feito o teste de normalidade da variável em questão (instrução 5) utilizado para verificar se a distribuição de probabilidade associada a um conjunto de dados pode ser aproximada a uma distribuição normal. Assim será utilizado o teste de Shapiro-Wilk, uma vez que a tabela possui menos que 30 amostras. Deve-se considerar as seguintes hipóteses com nível de significância de 5%:

- H0 (hipótese nula): as amostras apresentam distribuição normal;
- H1 (hipótese alternativa): as amostras não apresentam distribuição normal.

Assim, seguindo tópicos a serem respondidos, no *prompt* o próximo passo é apresentar o teste de normalidade, figura 5.32. Como pode ser observado na figura 5.33, o *p-value* é inferior ao nível de significância estabelecido de 0,05 então a hipótese nula é rejeitada e os dados não seguem uma distribuição normal.

Figura 5.32: Análise de normalidade da variável *Time*

O teste de Shapiro-Wilk foi realizado para a variável "Time". O valor do teste de Shapiro-Wilk é 0,887 e o valor do p-value é 0,028.

Figura 5.33: *p-value* da variável *Time*

Seguindo com a análise, como os dados não apresentam distribuição normal não é necessário verificar sua homodasticidade. Nesse caso deve ser aplicado um teste não paramétrico para amostras independentes (instrução 7), assim foi utilizado o teste de Mann-Whitney para comparar as médias entre os dois grupos (instrução 8), considerando as seguintes hipóteses com nível de significância de 5%:

- H0 (hipótese nula): não há diferença entre as médias;
- H1 (hipótese alternativa): há diferença entre as médias.

Assim, conforme pode ser observado na Figura 5.34 (instrução 9), não existe uma diferença significativa nas médias da variável entre seus grupos, um vez que o *p-value* (valor-p) é superior ao nível de significância estabelecido e a hipótese nula foi aceita.

- **Estadística do teste (U):** 46.5
- **Valor-p:** 0.869

Figura 5.34: Teste de Mann-Whitney

Assim, pode-se concluir, que a variável *Time*, estatisticamente, não apresenta diferença significativa das amostras da variável entre os grupos, considerando um nível de significância de 5%.

Por fim, no *prompt*, foi solicitado que exportasse todos os resultados das análises anteriores para um arquivo .doc (instrução 10), conforme figura 5.35, incluindo todos os gráficos realizados durante a análise.

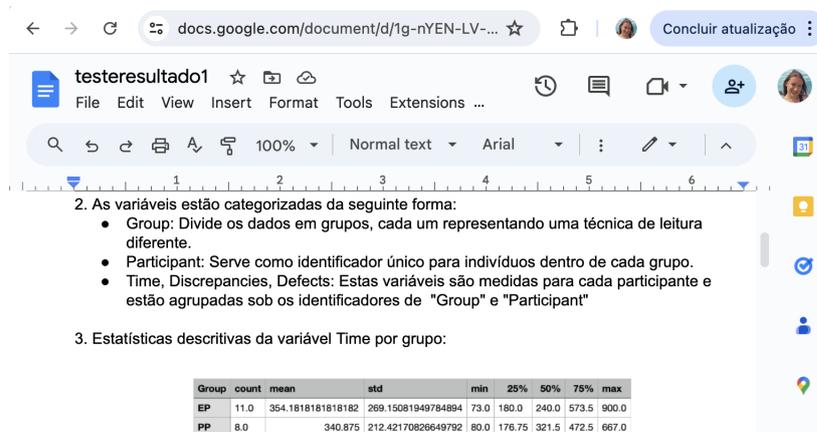


Figura 5.35: Fragmento do documento gerado para a variável *Time*

5.4 Execução com um fator e mais de dois tratamentos

Nesta seção, é apresentado um outro estudo de caso envolvendo um único fator e mais de dois tratamentos, abordando tanto métodos paramétricos quanto não paramétricos. O objetivo deste estudo é demonstrar, de forma detalhada, o procedimento para a aplicação dos métodos indicados na árvore de decisão. Para essa demonstração, utilizam-se dados reais extraídos de uma base originada de um questionário de levantamento de riscos, elaborado como uma técnica para a quantificação de riscos. A base de dados usada nesta seção pode ser observada na Figura 5.36.

O primeiro passo deste estudo foi a definição das variáveis presentes na base de dados utilizada e a categorização de cada uma delas (instrução 1). Essa etapa é crucial em uma análise estatística, pois permite compreender a estrutura dos dados e definir as abordagens analíticas adequadas. O resultado dessa etapa pode ser visualizado na Figura 5.37, onde estão listadas as variáveis identificadas: *Type*, *Analysis*, *Design*, *Imp*, *Test*, *Plan*, *Control*, *Equi*, *Contract*, *PolEst* e *Client*. A variável *Type* está dividida nos seguintes grupos: projetos pequenos (SP), projetos médios (MP) e projetos grandes (LP).

Type	Analysis	Design	Imp	Test	Plan	Control	Equi	Contract	Pol_Est	Client
SP	0,00	16,65	3,17	11,11	22,20	7,94	9,52	1,59	28,55	23,80
SP	20,61	20,61	16,74	12,88	23,18	10,30	0,00	1,29	2,58	6,44
SP	11,50	6,91	3,00	9,19	15,95	11,50	11,50	11,50	3,00	15,95
SP	17,08	13,12	4,59	1,32	22,35	22,35	9,19	10,50	30,25	30,25
SP	11,12	5,25	1,40	10,10	18,18	11,12	17,10	8,07	6,54	11,12
SP	12,49	8,98	7,41	12,49	18,48	10,48	29,98	0,50	35,94	24,25
SP	11,25	1,38	1,10	4,69	15,96	6,10	11,25	12,19	22,05	14,03
SP	18,97	6,33	0,85	11,24	33,54	31,62	15,46	5,07	12,64	25,29
SP	13,94	7,84	7,84	6,31	0,00	0,00	7,84	6,31	6,31	20,91
SP	13,72	13,72	9,81	5,88	11,76	5,88	21,58	1,97	3,92	11,76
SP	0,00	22,51	14,17	20,67	0,00	5,01	0,00	1,00	8,45	36,68
SP	37,00	22,30	11,15	22,30	37,00	37,00	18,54	3,68	37,00	37,00
SP	0,00	7,62	8,31	9,33	15,23	5,62	21,44	0,69	0,69	0,00
SP	26,37	18,14	8,10	3,11	26,37	16,49	32,97	8,10	6,50	14,84
SP	20,75	15,28	3,27	16,37	23,99	21,85	8,16	1,09	25,12	25,12
SP	10,84	8,13	9,50	6,78	15,82	12,25	5,92	0,94	15,82	14,00
SP	10,50	4,38	1,75	7,01	12,32	11,38	10,50	4,38	19,33	18,45
MP	42,47	31,03	13,07	24,51	29,40	26,14	32,66	1,63	27,77	34,30
MP	0,00	16,21	2,16	12,96	31,33	19,45	20,54	1,08	10,80	15,13
MP	15,15	8,46	7,35	3,79	14,70	14,48	13,14	0,22	10,91	11,80
MP	11,06	7,54	4,02	0,50	13,07	12,56	11,56	12,06	14,06	13,57
MP	7,14	3,38	1,87	0,38	13,14	12,14	12,76	0,00	8,26	21,02
MP	11,25	2,49	6,92	7,79	17,31	13,85	4,15	2,50	20,76	12,98
MP	20,30	13,52	9,66	17,40	21,75	22,25	17,89	0,55	14,51	23,17
MP	16,71	0,00	0,81	7,95	32,62	9,53	32,62	3,20	5,96	28,53
MP	17,07	11,11	1,22	14,62	17,07	34,08	6,15	4,88	35,31	19,50
MP	19,09	5,73	2,79	5,73	12,76	11,79	14,03	0,71	10,09	17,28
MP	14,17	11,93	13,41	12,67	9,69	20,87	13,41	0,74	29,08	35,03
MP	13,29	7,81	0,00	10,94	13,29	8,59	12,50	1,55	16,41	0,00
MP	20,72	11,81	7,39	1,50	32,51	29,53	23,64	47,29	45,76	42,84
MP	47,71	21,15	1,79	10,57	37,11	31,80	17,67	24,75	24,75	45,71
MP	21,82	20,13	16,73	11,72	26,85	10,06	26,85	1,67	3,35	21,82
MP	23,94	0,00	12,78	3,69	15,65	14,73	23,02	0,92	20,25	22,07
MP	16,26	7,31	0,90	11,38	11,38	18,69	26,00	16,26	28,43	24,38
MP	14,94	7,92	6,50	14,94	16,23	16,23	14,94	22,25	23,51	23,51
LP	12,15	5,93	2,54	8,48	14,41	13,84	12,99	0,28	14,12	15,26
LP	24,31	12,49	3,94	0,66	19,72	18,40	13,80	27,60	17,08	22,99
LP	20,13	4,46	11,17	11,17	20,13	15,67	15,67	11,17	31,31	20,13
LP	18,92	5,06	1,90	5,06	30,32	26,52	18,92	11,37	12,64	30,32
LP	13,95	6,98	1,16	9,30	11,64	10,47	9,30	9,30	13,95	13,95
LP	20,59	3,37	5,71	13,75	45,74	6,86	38,92	3,39	62,91	61,75
LP	26,90	23,43	19,96	14,75	49,52	49,50	22,57	0,87	21,70	33,80
LP	24,87	9,45	2,62	5,36	24,89	24,89	24,89	1,87	17,28	24,87
LP	11,06	9,40	8,29	20,48	12,74	16,07	19,27	0,56	32,12	31,01
LP	17,68	17,68	9,37	10,40	14,39	23,39	19,75	0,52	20,27	27,55
LP	10,87	8,89	0,98	1,98	0,00	27,66	12,85	22,72	19,75	17,77
LP	29,01	16,10	16,10	22,54	22,54	29,01	41,90	5,26	48,31	32,24
LP	25,54	17,04	10,97	1,21	38,95	37,71	23,14	24,33	34,22	49,89
LP	25,93	22,67	3,76	25,93	38,87	22,67	48,60	9,70	32,43	32,43
LP	30,19	17,25	17,25	17,25	30,19	30,19	43,08	4,31	30,19	43,08

Figura 5.36: Base de dados usada no estudo de caso

Logo depois, as variáveis foram categorizadas, conforme Figura 5.38 (instrução 2).

5.4.1 Método paramétrico na análise de variáveis em estudos de um fator e mais de dois tratamentos

Seguindo com a abordagem proposta, uma das variáveis presente na base foi escolhida para continuar a análise, a variável *Client* (instrução 3). Assim, para começar essa análise, foi solicitado no *prompt* que a IA Generativa apresentasse as estatísticas descritivas para

- **Analysis:** Quantitativa, representando problemas relacionados com o levantamento dos requisitos, sua estabilidade, nível de dificuldade de implementação, validação e complexidade do sistema.
- **Design:** Quantitativa, representando problemas relacionados à correta concepção da arquitetura, interfaces, algoritmos e mecanismos que facilitem a implementação do sistema.
- **Imp:** Quantitativa, representando problemas relacionados à complexidade de implementação dos algoritmos, inadequação de linguagem ou hardware e reutilização de código.
- **Test:** Quantitativa, representando problemas relacionados ao planejamento e execução, condições de realização, tipos e abrangência dos testes do sistema.
- **Plan:** Quantitativa, representando problemas relacionados à experiência dos gerentes, capacidade de elaboração de planejamento e estimativas do projeto, bem como aspectos de definição, utilização e adequação do processo de desenvolvimento de sistemas.
- **Control:** Quantitativa, representando problemas relacionados à condução do projeto, aprovação de artefatos, resolução de conflitos e apoio à equipe de desenvolvimento, bem como as atividades de acompanhamento e replanejamento ao longo do projeto e a avaliação do processo de desenvolvimento.
- **Equi:** Quantitativa, representando problemas relacionados à capacidade, estabilidade, treinamento, maturidade e forma de trabalhar da equipe, bem como o ambiente de desenvolvimento, e o grau com que a equipe segue os planejamentos e os processos.
- **Pol_Est:** Quantitativa, representando problemas relacionados à Política e à Estrutura Organizacional, apoio da Alta Gerência ao projeto, metas e conflitos de interesses.
- **Client:** Quantitativa, representando problemas relacionados ao envolvimento do cliente no projeto, número de usuários e nível de mudanças que serão provocadas pelo sistema.

Figura 5.37: Variáveis presente na base de dados.

- **Análise e Planejamento:** Inclui as variáveis **Analysis** e **Plan**, com um total de 74 observações.
- **Design e Implementação:** Abrange as variáveis **Design** e **Imp**, com um total de 74 observações.
- **Teste e Controle:** Inclui as variáveis **Test** e **Control**, com um total de 74 observações.
- **Equipe e Estrutura Organizacional:** Cobre as variáveis **Equi** e **Pol_Est**, com um total de 74 observações.
- **Clientes:** Contém a variável **Client**, com um total de 74 observações.

Figura 5.38: Agrupamento das variáveis.

a variável em questão por grupo, usada para descrever, organizar e resumir conjunto de dados de uma variável, como demonstrado na Figura 5.39. Essa visualização permite que pesquisadores realizem uma análise mais detalhada dos dados.

O próximo passo desse estudo foi construir um *boxplot* da variável (*Client*) por grupo (instrução 4), fornecendo uma representação gráfica dos dados, sendo: mínimo, primeiro quartil, mediana, terceiro quartil e máximo, sendo de extrema importância para pesquisadores seguirem com as análises estatísticas, como pode ser observado na Figura 5.40.

Seguindo na análise foi feito o teste de normalidade da variável em questão (instrução 5) utilizado para verificar se a distribuição de probabilidade associada a um con-

Type	count	mean	std	min	25%	50%	75%	max
LP	15.0	30.469120000000007	13.09678318226055	13.95	21.5579	30.3163	33.1117	61.7524
MP	18.0	22.923838888888888	11.301464784975694	0.0	15.670500000000002	21.944300000000005	27.49075	45.709399999999995
SP	17.0	19.405376470588234	10.004478448145699	0.0	14.0	18.45	25.116	37.0041

Figura 5.39: Estatística descritiva da variável *Client* por tipo.

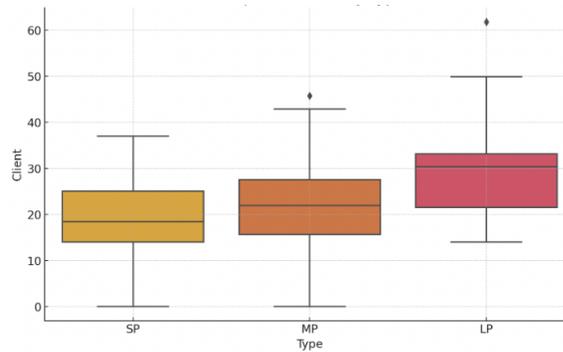


Figura 5.40: *Boxplot* da variável *Client* por tipo.

junto de dados pode ser aproximada a uma distribuição normal. Assim será utilizado o teste de Kolmogorov-Smirnov, uma vez que a tabela possui mais que 30 amostras. Deve-se considerar as seguintes hipóteses com nível de significância de 5%:

- H0 (hipótese nula): as amostras apresentam distribuição normal;
- H1 (hipótese alternativa): as amostras não apresentam distribuição normal.

Assim, seguindo tópicos a serem respondidos, no *prompt* o próximo passo é apresentar o teste de normalidade, Figura 5.41. Como pode ser observado na Figura 5.42, o *p-value* é superior ao nível de significância estabelecido de 0,05 então a hipótese nula é aceita e os dados seguem uma distribuição normal.

Logo depois, foi realizado o teste de igualdade de variâncias, para a variável *Client* (instrução 6) considerando seus grupos. Esse teste só é realizado caso seja necessário avaliar a igualdade de variâncias de uma variável calculada para dois ou mais grupos, em uma distribuição normal. Para isso, o Teste de Levene é utilizado e considera as seguintes hipóteses com nível de significância de 5%:

- H0 (hipótese nula): grupos possuem igualdade de variâncias;
- H1 (hipótese alternativa): grupos não possuem igualdade de variâncias.

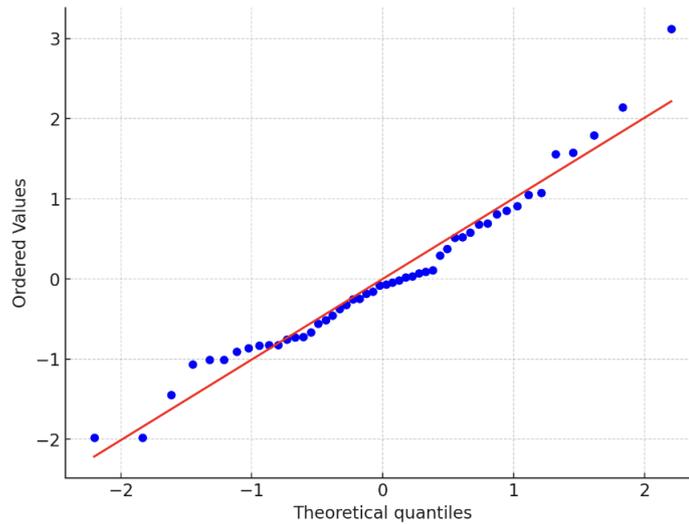


Figura 5.41: Análise de normalidade da variável *Client*.

O teste de Kolmogorov-Smirnov (KS) foi realizado para a variável "Client". O valor do teste de Kolmogorov-Smirnov 0.1172 e o p-value é 0.4636.

Figura 5.42: *p-value* da variável *Client*

Se o *p-value* resultante do teste for maior que o nível de significância estabelecido (foi considerado 0,05), aceita-se a hipótese nula de que existe igualdade de variâncias, devendo-se utilizar um método paramétrico para a comparação dos dados. Em caso contrário, um método não paramétrico deve ser utilizado. Conforme pode ser observado na Figura 5.43, o *p-value* (valor-p) é superior ao nível de significância estabelecido, aceitando-se a hipótese nula de que os grupos possuem igualdade de variâncias e, com isso, um teste paramétrico deve ser utilizado na comparação dos grupos. O próximo passo foi realizar o teste apropriado para a comparação das médias da variável (*Client*), evidenciado no *prompt* utilizado neste trabalho, instrução 7. Esse teste foi realizado considerando as análises anteriores, uma vez que, com base nos resultados já obtidos, observa-se que a variável estudada segue uma distribuição normal e as variâncias são iguais entre os grupos. Assim, conclui-se que a variável apresenta uma distribuição normal.

- **Estatística de Levene:** 0,1884
- **Valor p:** 0,8289

Figura 5.43: Teste de Levene.

Com base nessas condições, a IA Generativa identificou corretamente qual seria o teste mais apropriado para a continuidade da análise (instrução 7). Dessa forma, o ANOVA deve ser aplicado para realizar a análise de um fator com mais de 2 tratamentos (instrução 8), considerando as seguintes hipóteses com nível de significância de 5%:

- H0 (hipótese nula): grupos possuem médias iguais;
- H1 (hipótese alternativa): Ao menos um grupo possui média diferente.

Assim, conforme pode ser observado na Figura 5.44 (instrução 9), existe ao menos um grupo que possui média diferente, um vez que o *p-value* (valor-p) é inferior ao nível de significância estabelecido e a hipótese alternativa foi aceita.

- **Estatística F: 3,8346**
- **Valor p: 0,0287**

Figura 5.44: Teste ANOVA

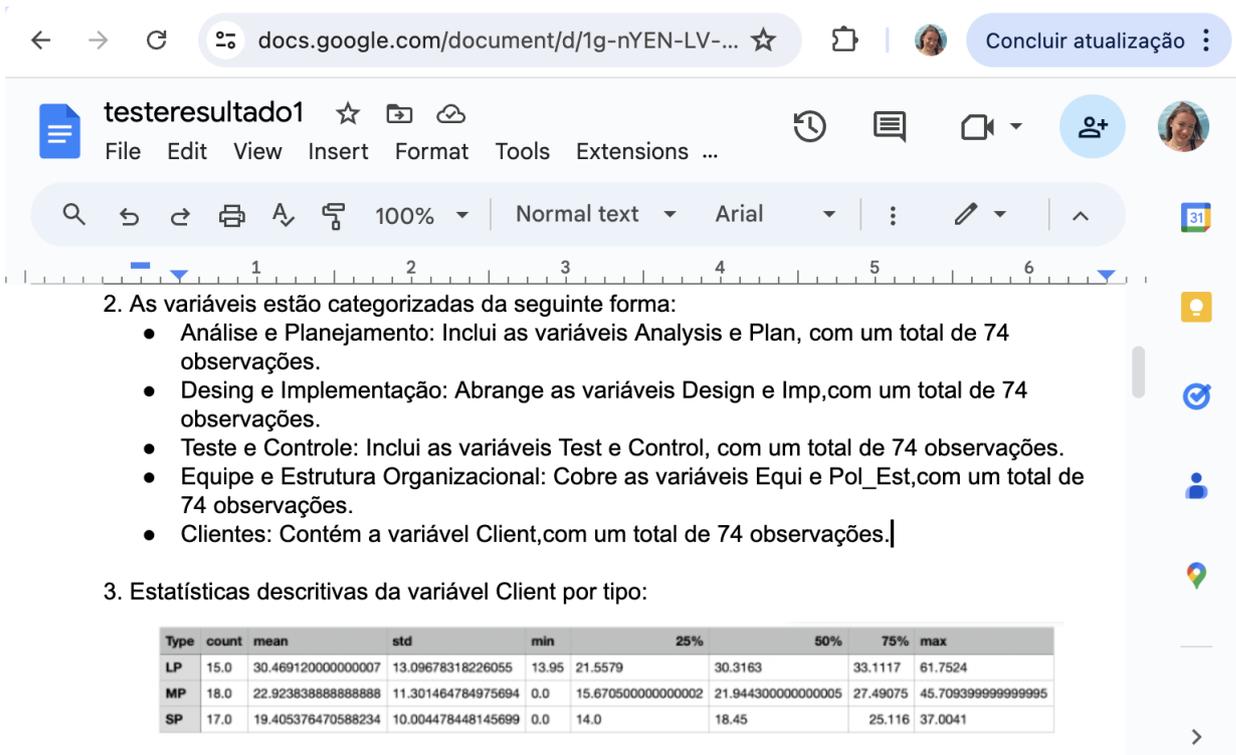
Dessa forma, foi necessário realizar o teste de Tukey, conforme Figura 5.45.

group1	group2	meandiff	p-adj	lower	upper	reject
LP	MP	-7.5453	0.1548	-17.2426	2.152	False
LP	SP	-11.0637	24	-20.8898	-1.2377	True
MP	SP	-3.5185	0.6382	-12.8994	5.8625	False

Figura 5.45: Teste de Tukey.

Assim, pelo teste de Tukey, pode-se observar que não existe diferença significativa entre os projetos médios e grandes e nem entre os projetos pequenos e médios. Mas há diferença significativa entre projetos pequenos e grandes (coluna *reject* igual a *True*).

Por fim, no *prompt*, foi solicitado que exportasse todos os resultados das análises anteriores para um arquivo .doc (instrução 10), conforme Figura 5.46, incluindo todos os gráficos realizados durante a análise.



2. As variáveis estão categorizadas da seguinte forma:

- Análise e Planejamento: Inclui as variáveis Analysis e Plan, com um total de 74 observações.
- Desing e Implementação: Abrange as variáveis Design e Imp, com um total de 74 observações.
- Teste e Controle: Inclui as variáveis Test e Control, com um total de 74 observações.
- Equipe e Estrutura Organizacional: Cobre as variáveis Equi e Pol_Est, com um total de 74 observações.
- Clientes: Contém a variável Client, com um total de 74 observações.]

3. Estatísticas descritivas da variável Client por tipo:

Type	count	mean	std	min	25%	50%	75%	max
LP	15.0	30.469120000000007	13.09678318226055	13.95	21.5579	30.3163	33.1117	61.7524
MP	18.0	22.923838888888888	11.301464784975694	0.0	15.670500000000002	21.944300000000005	27.49075	45.709399999999995
SP	17.0	19.405376470588234	10.004478448145699	0.0	14.0	18.45	25.116	37.0041

Figura 5.46: Fragmento do documento gerado para a variável *Client*

5.4.2 Método não paramétrico na análise de variáveis em estudos de um fator e mais de dois tratamentos

Seguindo com a abordagem proposta, outra variável presente na base foi escolhida para continuar a análise, a variável *Control*. Assim, para começar essa análise, foi solicitado no *prompt* que a IA Generativa apresentasse as estatísticas descritivas para a variável em questão por grupo (instrução 3), usada para descrever, organizar e resumir conjunto de dados de uma variável, como demonstrado na Figura 5.47. Essa visualização permite que pesquisadores realizem uma análise mais detalhada dos dados.

Type	count	mean	std	min	25%	50%	75%	max
LP	15.0	23.52352	10.90374033323297	6.864299999999999	15.866550000000002	23.3933	28.33435	49.4966
MP	18.0	18.154344444444444	7.844112758002746	8.59	12.245000000000001	15.480150000000002	21.904050000000005	34.083700000000001
SP	17.0	13.346170588235296	9.786946469262306	0.0	6.1	11.12	16.4864	37.0041

Figura 5.47: Estatística descritiva da variável *Control* por tipo

O próximo passo desse estudo foi construir um *boxplot* da variável *Control* por tipo (instrução 4), fornecendo uma representação gráfica dos dados, sendo: mínimo, primeiro quartil, mediana, terceiro quartil e máximo, sendo de extrema importância para

pesquisadores seguirem com as análises estatísticas, como pode ser observado na Figura 5.48.

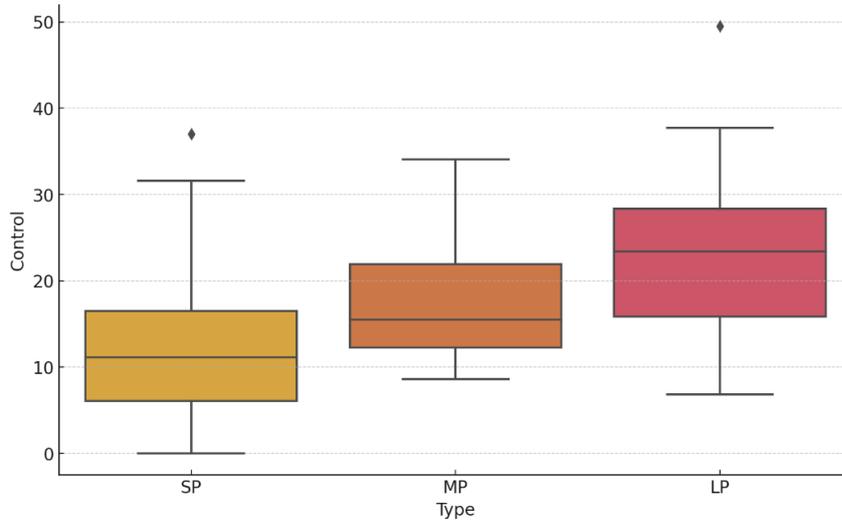


Figura 5.48: *Boxplot* da variável *Control* por tipo.

Seguindo na análise, foi feito o teste de normalidade da variável em questão (instrução 5), utilizado para verificar se a distribuição de probabilidade associada a um conjunto de dados pode ser aproximada a uma distribuição normal. Assim foi utilizado o teste de Kolmogorov-Smirnov, uma vez que a tabela possui mais que 30 amostras. Deve-se considerar as seguintes hipóteses com nível de significância de 5%:

- H0 (hipótese nula): dados apresentam distribuição normal;
- H1 (hipótese alternativa): dados não apresentam distribuição normal.

Assim, seguindo tópicos a serem respondidos, no *prompt* o próximo passo é apresentar o teste de normalidade, Figura 5.49. Como pode ser observado na Figura 5.50, apesar de ser superior ao nível de significância estabelecido de 0,05, nessa pesquisa foi considerado que a hipótese nula é rejeitada e os dados não seguem uma distribuição normal, para executar todos os cenários de teste o *prompt* foi alterado para caso o *p-value* for muito próximo ao nível de significância a hipótese nula deve ser rejeitada.

Seguindo com a análise, como os dados não apresentam uma distribuição normal não é necessário verificar sua homodasticidade. Nesse caso deve ser aplicado um teste não paramétrico (instrução 7), como a análise em questão é de um fator e mais de dois

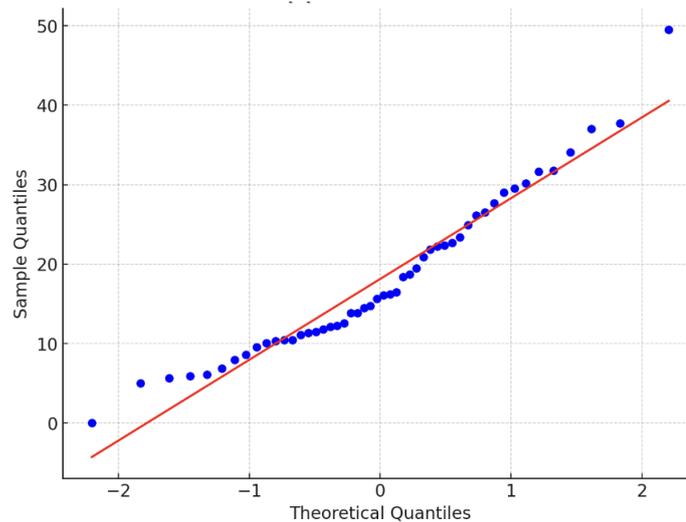


Figura 5.49: Teste de Kolmogorov-Smirnov da variável *Control*.

O teste de Kolmogorov-Smirnov (KS) foi realizado para a variável "Control". O valor do teste de Kolmogorov-Smirnov foi 0,124 e o p-value é 0,052.

Figura 5.50: p-value da variável *Control*.

tratamentos, o teste de Kruskal-Wallis é o mais apropriado (instrução 8), considerando as seguintes hipóteses com nível de significância de 5%:

- H0 (hipótese nula): as médias dos grupos são iguais;
- H1 (hipótese alternativa): existe ao menos um grupo com média diferente.

Assim, conforme pode ser observado na Figura 5.51 (instrução 9), existe ao menos um grupo possui média diferente, um vez que o *p-value* (valor-p) é inferior ao nível de significância estabelecido e a hipótese alternativa foi aceita.

- Estatística do teste: 9.1645
- Valor-p: 0.0102

Figura 5.51: Teste de Kruskal-Wallis

Dessa forma, foi necessário realizar o teste de Mann-Whitney, conforme figura 5.52.

Assim, pelo teste de Mann-Whitney, pode-se observar que existe diferença significativa entre os projetos médios e pequenos e nem entre os projetos pequenos e médios.

Comparison	Statistic	P-Value	Significant (p < 0.05)
SP vs MP	91.0	0.0424	Yes
SP vs LP	55.0	0.0065	Yes
MP vs LP	94.0	0.1431	No

Figura 5.52: Teste de Mann-Whitney para a variável *Control*

Mas há não há diferença significativa entre projetos médios e grandes (coluna *Significant*).

Por fim, no *prompt*, foi solicitado que exportasse todos os resultados das análises anteriores para um arquivo .doc (instrução 10), conforme figura 5.53, incluindo todos os gráficos realizados durante a análise.

The screenshot shows a Google Docs interface with the following content:

2. As variáveis estão categorizadas da seguinte forma:

- Análise e Planejamento: Inclui as variáveis Analysis e Plan, com um total de 74 observações.
- Desing e Implementação: Abrange as variáveis Design e Imp, com um total de 74 observações.
- Teste e Controle: Inclui as variáveis Test e Control, com um total de 74 observações.
- Equipe e Estrutura Organizacional: Cobre as variáveis Equi e Pol_Est, com um total de 74 observações.
- Clientes: Contém a variável Client, com um total de 74 observações.

3. Estatísticas descritivas da variável Control por tipo:

Type	count	mean	std	min	25%	50%	75%	max
LP	15.0	23.52352	10.90374033323297	6.8642999999999999	15.8665500000000002	23.3933	28.33435	49.4966
MP	18.0	18.154344444444444	7.844112758002746	8.59	12.245000000000001	15.480150000000002	21.904050000000005	34.08370000000001
SP	17.0	13.346170588235296	9.786946469262306	0.0	6.1	11.12	16.4864	37.0041

Figura 5.53: Fragmento do documento gerado para a variável *Control*

5.5 Considerações do capítulo

Neste capítulo foi apresentado um estudo de caso que segue detalhadamente os passos delineados no capítulo anterior. O estudo enfatiza a aplicação de testes estatísticos, a utilização de engenharia de *prompt* para otimizar a interação com o ChatGPT, e um passo a passo minucioso da implementação da automação com o Zapier. Ao integrar essas ferramentas e técnicas, o estudo ilustra de forma prática como é possível aprimorar

a análise de dados, tornando o processo mais eficiente e preciso, ao mesmo tempo em que se automatizam tarefas complexas e repetitivas. Por fim, foi solicitado que a IA generativa exportasse todos os resultados obtidos para um documento em formato *doc*, em uma linguagem formal para que possa ser utilizado em um trabalho acadêmico.

6 Conclusões

Com este trabalho será possível apoiar a eficiência das pesquisas acadêmicas, combinando IA com o conhecimento humano. Ao automatizar e integrar tarefas repetitivas e demoradas, pesquisadores podem redirecionar seu tempo e esforço para análises estatísticas mais aprofundadas e para fazer contribuições originais ao campo científico, resultando na geração de novos dados de pesquisa. Assim, a IA não apenas simplifica e otimiza essas tarefas, mas também abre caminho para novas descobertas e avanços no conhecimento científico.

6.1 Limitações e desafios encontrados

A integração das diversas ferramentas utilizadas neste projeto, incluindo o Zapier, a API da OpenAI, e outras plataformas auxiliares, apresentou uma série de desafios que evidenciam a complexidade desse processo. Um dos principais desafios foi a integração dos diferentes sistemas, cada um com suas próprias características, limitações e requisitos de configuração. A necessidade de garantir que todas as ferramentas se comunicassem de maneira eficaz e fluida demandou muitos ajustes e testes para evitar inconsistências e falhas na automação.

Embora o processo automatizado desenvolvido tenha demonstrado eficiência, ele apresenta limitações significativas devido à dependência de ferramentas pagas, como o ChatGPT e o Zapier. Essas ferramentas, em suas versões gratuitas, possuem restrições que limitam a funcionalidade completa do sistema, exigindo a versão paga para acessar recursos avançados. Além disso, a confiabilidade do sistema está atrelada à continuidade desses serviços terceirizados, o que pode representar um risco. Por fim, a necessidade de subscrição pode tornar os custos proibitivos em contextos de grande volume de dados ou uso simultâneo, limitando a escalabilidade e a acessibilidade da solução.

6.2 Trabalhos Futuros

Diante dos resultados obtidos com a automação dos processos utilizando o Zapier em conjunto com a API da OpenAI, é fundamental considerar aprimoramentos contínuos para maximizar a eficácia dessa abordagem. Assim como, a realização de revisões periódicas das configurações do Zapier, bem como dos parâmetros aplicados na API da OpenAI, para assegurar que o sistema continue a atender às necessidades específicas do projeto. Além disso, a integração com outras ferramentas e plataformas poderá expandir significativamente as capacidades de automação, permitindo análises mais robustas e um controle mais detalhado sobre os dados processados. O desenvolvimento contínuo das habilidades das equipes envolvidas na operação e manutenção desses fluxos automatizados é igualmente essencial, garantindo que os profissionais estejam preparados para implementar melhorias e adaptarem-se às novas funcionalidades tecnológicas.

Além disso, destaca-se a questão da escalabilidade da solução proposta. À medida que o volume de dados e a complexidade dos processos aumentam, será imperativo avaliar a capacidade do Zapier e da API da OpenAI de processarem grandes quantidades de dados de forma eficiente, sem comprometer a qualidade dos resultados. A capacidade de adaptação às mudanças tecnológicas contínuas e às novas exigências do mercado também representa um desafio significativo, exigindo uma abordagem flexível e proativa na gestão dessas tecnologias.

Outra oportunidade de trabalho futuro é ampliar a proposta para abranger soluções envolvendo técnicas de análise qualitativa e mista.

6.3 Aspectos éticos

O uso de inteligência artificial na análise estatística, levanta questões éticas que não podem ser ignoradas. A privacidade e a proteção dos dados utilizados são prioridades, devendo-se assegurar que todas as práticas estejam em conformidade com as legislações vigentes, como a Lei Geral de Proteção de Dados (LGPD, 2018). A transparência nos processos e a explicabilidade dos resultados gerados pela IA são igualmente cruciais para garantir a confiança e a replicabilidade das análises. Além disso, a mitigação de vieses e a

integridade na divulgação dos resultados são elementos fundamentais para a manutenção da credibilidade científica.

Neste estudo, o ChatGPT foi utilizado para o refinamento da escrita.

6.4 Considerações Finais

O estudo apresentado demonstra que a automação de processos para análise estatística, quando integrada a ferramentas avançadas como o Zapier e a API da OpenAI, não apenas simplifica a gestão de tarefas rotineiras, mas também amplia as possibilidades de inovação nas organizações. O uso combinado dessas tecnologias se revela como um potente catalisador para a produtividade e a transformação digital. Para assegurar que os benefícios dessa automação sejam plenamente realizados, é crucial manter um compromisso constante com o aprimoramento das soluções implementadas e com a adaptação às evoluções tecnológicas emergentes.

Bibliografia

ACADÊMICO, G. 2004. Disponível em: <https://scholar.google.com.br/?hl=pt>.

AKKIO, I. *Akkio: Easy AI for Everyone*. 2024. Accessed: 2024-07-04. Disponível em: <https://www.akkio.com/>.

ALBALAK, A.; ELAZAR, Y.; XIE, S. M.; LONGPRE, S.; LAMBERT, N.; WANG, X.; MUENNIGHOFF, N.; HOU, B.; PAN, L.; JEONG, H. et al. A survey on data selection for language models. *arXiv preprint arXiv:2402.16827*, 2024.

ANDERSON, J. Data security in automated workflows. *Journal of Cybersecurity and Privacy*, v. 9, n. 3, p. 203–217, 2023.

ARAÚJO, J. C.; ARAÚJO, M. C.; ; ARAÚJO, M. A. P. Em busca de um processo automatizado para levantamento de literatura apoiado por uma inteligência artificial generativa. 2023.

ARAÚJO, J. C.; ARAÚJO, M. C.; ; ARAÚJO, M. A. P. Explorando a automatização na análise estatística com suporte da inteligência artificial generativa na gestão de dados de investigação. 2023.

ARAÚJO, J. C.; ARAÚJO, M. C.; ; ARAÚJO, M. A. P. Explorando automatização na revisão de literatura com suporte de inteligência artificial generativa na gestão de dados de investigação. 2023.

ARAÚJO, M.; BARROS, M.; TRAVASSOS, G.; MURTA, L. Métodos estatísticos aplicados em engenharia de software experimental. *XXI SBBD-XX SBES*, 2006.

AUTOMATE, M. P. *Microsoft Power Automate: Automate Your Way*. 2024. Disponível em: <https://flow.microsoft.com/>.

BATTISTI, I. D. E.; BATTISTI, G. *Métodos estatísticos*. 2008.

BOULESTEIX, A.-L.; STRIMMER, K. Partial least squares: a versatile tool for the analysis of high-dimensional genomic data. *Briefings in Bioinformatics*, Oxford University Press, v. 8, n. 1, p. 32–44, 2007.

BROWN, T.; MANN, B.; RYDER, N.; SUBBIAH, M.; KAPLAN, J. D.; DHARIWAL, P.; NEELAKANTAN, A.; SHYAM, P.; SASTRY, G.; ASKELL, A. et al. Language models are few-shot learners. *Advances in neural information processing systems*, v. 33, p. 1877–1901, 2020.

BUENO, J. F. *Sistema automatizado de classificação de abelhas baseado em reconhecimento de padrões*. 2010. Disponível em: <https://doi.org/10.11606/T.3.2010.tde-10012011-114124>.

CHATGPT. 2022. Disponível em: <https://chat.openai.com/>.

COOPER, D. R.; SCHINDLER, P. S. Métodos de pesquisa em administração. *12^a Edição*. McGraw Hill Brasil., 2016.

COSTA, P. L. d. O. C. N. *Estatística*. [S.l.]: Editora Blucher, 2002.

DEVLIN, J.; CHANG, M.-W.; LEE, K.; TOUTANOVA, K. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.

DIAGRAMS. 2023. Disponível em: <https://chatgpt.com/g/g-5QhhsfDj-diagrams-show-me-ideas-presentations-code>.

Gemini. *Gemini: Um Modelo de Linguagem de Grande Escala*. 2023. Accessed: 2024-07-04.

IEEEXPLORE. 2008. Disponível em: <https://ieeexplore.ieee.org/Xplore/home.jsp>.

IGNÁCIO, S. A. Importância da estatística para o processo de conhecimento e tomada de decisão. *Revista Paranaense de Desenvolvimento-RPD*, n. 118, p. 175–192, 2010.

JOHNSON, E. Ease of use in workflow automation tools. *Journal of Automation Studies*, v. 15, n. 3, p. 45–67, 2022.

KITCHENHAM, B.; DYBA, T.; JORGENSEN, M. Evidence-based software engineering, in ‘proceedings of icse. In: IEEE. [S.l.], 2004. p. 273–281.

LARGUESA, R. *Engenharia de Prompt para Devs: Um guia para aprender a usar a IA antes que a IA aprenda a usar você*. Casa do Código, 2024. ISBN 9788555193712. Disponível em: <https://books.google.com.br/books?id=Qj79EAAAQBAJ>.

LECUN Y BENGIO, G. H. Y. *Deep Learnig*. 2015.

LEE, D. Learning curves in the adoption of automation tools. *Automation Research and Practice*, v. 18, n. 4, p. 89–101, 2022.

LEWIS, J. R. Usability: Lessons learned ... and yet to be learned. *International Journal of Human-Computer Interaction*, Taylor Francis, v. 30, n. 9, p. 663–684, 2014.

LGPD. 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm.

LLaMA. *LLaMA: Large Language Model Architecture*. 2023. Accessed: 2024-07-04.

MARTINEZ, C. Technical support for automation tools. *Journal of Technical Support and Maintenance*, v. 11, n. 4, p. 150–162, 2022.

MCCARTHY, J. *What is artificial intelligence?* 1956.

MINAEE, S.; MIKOLOV, T.; NIKZAD, N.; CHENAGHLU, M.; SOCHER, R.; AMATRIAIN, X.; GAO, J. *Large Language Models: A Survey*. 2024.

MONTGOMERY, D. C. *Design and analysis of experiments*. [S.l.]: John wiley & sons, 2017.

MOURA, M. F. *Proposta de Utilização de Mineração de Textos para Seleção, Classificação e Qualificação de Documentos*. 2004. Disponível em: <chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.infoteca.cnptia.embrapa.br/bitstream/doc/2327/1/doc47.pdf>.

- NAVARRO, F. P. Uso da inteligência artificial para recuperação da informação com abordagem semântica: modelo de aplicação para documentos textuais em ambientes digitais. Universidade Estadual Paulista (Unesp), 2021.
- PEREIRA, I. da S. D.; MOURA, S. A. de. Compreendendo a inteligência artificial generativa na perspectiva da língua. *SciELO Preprints*, 2023.
- PINHO, C. M. d. A.; MOURA, A. F. d.; GASPAR, M. A.; NAPOLITANO, D. M. R. Identificação de deficiências em textos educacionais com a aplicação de processamento de linguagem natural e aprendizado de máquina. *ETD Educação Temática Digital*, UNICAMP, v. 24, n. 2, p. 350–372, 2022.
- PRADO, G. R. B. e. *Estudo sobre aplicação de técnicas de inteligência artificial na medicina*. 2023. Disponível em: <http://repositorio.unitau.br/jspui/handle/20.500.11874/6956>.
- QUINTERO, J. S. V. The importance of artificial intelligence (ai) and its limitations. *SCT Proceedings in Interdisciplinary Insights and Innovations*, v. 1, p. 70–70, 2023.
- R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2023. Disponível em: <https://www.R-project.org/>.
- RADFORD, A.; WU, J.; CHILD, R.; LUAN, D.; AMODEI, D.; SUTSKEVER, I. et al. Language models are unsupervised multitask learners. *OpenAI blog*, v. 1, n. 8, p. 9, 2019.
- REPORT, D. A. . 2023. Disponível em: <https://chatgpt.com/g/g-2OebMtWeG-data-analysis-report-ai>.
- SALIAN, I. *SuperVize Me: What's the Difference Between Supervised, Unsupervised, Semi-Supervised and Reinforcement Learning?* 2018. Disponível em: <https://blogs.nvidia.com/blog/supervised-unsupervised-learning/>.
- SIRQUEIRA, T. F. M.; MIGUEL, M. A.; DALPRA, H. L. de O.; ARAÚJO, M. A. P. Aplicação de métodos estatísticos em engenharia de software: Teoria e prática. 2020.
- SMITH, R. Integration capabilities of modern automation platforms. *International Journal of Workflow Automation*, v. 12, n. 1, p. 22–34, 2023.
- TAYLOR, M. Scalability in workflow automation tools. *Journal of Scalable Computing*, v. 14, n. 5, p. 77–90, 2022.
- UIPATH. *UiPath: Automation for All*. 2024. Disponível em: <https://www.uipath.com/>.
- VASWANI, A.; SHAZEER, N.; PARMAR, N.; USZKOREIT, J.; JONES, L.; GOMEZ, A. N.; KAISER, L.; POLOSUKHIN, I. Attention is all you need. *Advances in neural information processing systems*, v. 30, 2017.
- WICKHAM, H.; GROLEMUND, G. *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. [S.l.]: O'Reilly Media, 2016.
- WIEDERHOLD, G.; MCCARTHY, J. Arthur samuel: Pioneer in machine learning. *IBM Journal of Research and Development*, v. 36, p. 329 – 331, 06 1992.
- WILSON, G.; BRYAN, J.; CRANSTON, K.; KITZES, J.; NEDERBRAGT, L.; TEAL, T. K. Good enough practices in scientific computing. *PLOS Computational Biology*, Public Library of Science, v. 13, n. 6, p. e1005510, 2017.

WILSON, M. Functionality and robustness of automation software. *Software Engineering Journal*, v. 20, n. 2, p. 112–126, 2023.

WOLF, T.; DEBUT, L.; SANH, V.; CHAUMOND, J.; DELANGUE, C.; MOI, A.; CISTAC, P.; RAULT, T.; LOUF, R.; FUNTOWICZ, M. et al. Transformers: State-of-the-art natural language processing. In: *Proceedings of the 2020 conference on empirical methods in natural language processing: system demonstrations*. [S.l.: s.n.], 2020. p. 38–45.

ZAPIER. 2012. Disponível em: <https://zapier.com/app/home>.