

NOVA

IMS

Information
Management
School

MDDM

Mestrado em
Data-Driven Marketing

Desempenho de IA Generativa no Copywriting: Um Estudo Comparativo

Raphael Diversi de Matos

Dissertação

apresentada como requisito parcial para obtenção do grau de Mestre em *Data-Driven Marketing*

NOVA Information Management School

Instituto Superior de Estatística e Gestão de Informação

Universidade Nova de Lisboa

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

Desempenho de IA Generativa no Copywriting: Um Estudo Comparativo

por

Raphael Diversi de Matos

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre em
Data-Driven Marketing, com especialização em *Digital Marketing and Analytics*

Orientada por: Flávio L. Pinheiro, PhD em Física, Universidade Nova de Lisboa.

Co-Orientada por: Diana Orghian, PhD em Psicologia, Universidade Nova de Lisboa.

Julho 2024

DECLARAÇÃO DE INTEGRIDADE

Declaro ter realizado o presente trabalho académico com integridade. Confirmando que não recorri à prática de plágio ou de qualquer outra forma de utilização indevida de informação ou de falsificação de resultados durante o processo de elaboração deste trabalho. Declaro ainda que tenho conhecimento das Regras de Conduta e do Código de Honra da NOVA Information Management School.

Raphael Diversi de Matos

Lisboa, 02/07/2024

DEDICATÓRIA

Dedico esta tese à minha querida mãe, que sempre me incentivou a estudar e enfatizou a importância da busca incessante pelo conhecimento e a minha avó. Elas me ensinaram que o saber é uma força transformadora e libertadora na vida das pessoas.

A minha amada esposa, Alicia, minha eterna fonte de apoio e coragem. Desde a graduação em Economia até este momento no mestrado, caminhamos lado a lado e enfrentamos juntos os desafios de um país estrangeiro, seguimos na reconstrução de nossas vidas como imigrantes, a trabalhar arduamente, a estudar intensamente, a desenvolver inúmeros projetos e, finalmente, a escrever esta tese. Quem nunca deixou que eu desistisse dos meus sonhos e esteve sempre ao meu lado, a dar-me apoio em todos os momentos difíceis e a celebrar cada pequena vitória.

À minha irmã e ao meu cunhado, por sempre acreditar em nós e expressar com orgulho o quanto nos admira.

Agradeço aos meus sogros, cujo incentivo constante nos deu força para seguir em frente, mesmo diante das incertezas do futuro.

Aos nossos amigos, que acreditaram em nossa capacidade de alcançar nossos objetivos e nos ofereceram palavras de encorajamento quando mais precisávamos. Nos momentos em que o medo ameaçava nos paralisar, suas palavras de apoio foram fundamentais para continuarmos firmes.

Agradeço ao meu orientador e minha co-orientadora por seus ensinamentos, todo apoio e paciência.

Por fim, mas não menos importante, aos amigos que a NOVA IMS nos trouxe, os quais tornaram esta jornada muito mais agradável e enriquecedora. Também agradeço a Portugal e aos amigos portugueses que tão calorosamente nos acolheram e que, ao longo do tempo, se tornaram parte de nossa família em terras lusitanas.

Com imensa gratidão e amor,

Raphael Diversi de Matos

RESUMO

Nos últimos anos, o marketing digital tem experimentado uma evolução significativa, impulsionada por avanços tecnológicos e um crescente investimento empresarial. Este estudo visa comparar o desempenho de quatro diferentes Modelos de Linguagem de Grande Escala (LLMs) na criação de copywriting para campanhas digitais em diversos setores. A investigação procura entender quais modelos geram os textos mais eficazes, medindo a qualidade e a eficiência do conteúdo produzido.

A investigação centra-se na aplicação da IA generativa no copywriting, um elemento crucial para o sucesso das campanhas de marketing. Ao examinar o output destes LLMs na criação de textos publicitários para produtos e serviços nos setores financeiro, educacional, tecnológico e de saúde, esta investigação fornece perspetivas sobre o seu potencial da IA para revolucionar as práticas de marketing digital.

A metodologia envolve uma análise comparativa do copywriting gerado por cada LLM, avaliando fatores como persuasão, relevância e seu potencial de envolvimento. A investigação também considera a adaptabilidade destes modelos a diferentes requisitos específicos da indústria e públicos-alvo.

Esta investigação contribui para a compreensão do impacto do conteúdo gerado por IA no envolvimento do consumidor e fornece uma base para futuras estratégias na publicidade digital. Ao avaliar os pontos fortes e as limitações de cada LLM na produção de textos de marketing, o estudo oferece perspetivas para profissionais de marketing, publicitários e empresas que procuram aproveitar as tecnologias de IA nos seus esforços de marketing digital.

As conclusões desta investigação têm implicações para o futuro do copywriting e da criação de conteúdo no marketing digital. Elas lançam luz sobre o seu potencial da IA para melhorar a criatividade, aumentar a eficiência e entregar mensagens de marketing mais direcionadas e eficazes.

PALAVRAS-CHAVE

Marketing Digital; IA Generativa; Copywriting; Modelos de Linguagem de Grande Escala; Eficácia Publicitária.

ABSTRACT

In recent years, digital marketing has experienced significant evolution, driven by technological advances and increasing business investment. This study aims to compare the performance of four different Large Language Models (LLMs) in creating copywriting for digital campaigns across various sectors. The investigation seeks to understand which models generate the most effective texts, measuring the quality and efficiency of the produced content.

The research focuses on the application of generative AI in copywriting, a crucial element for the success of marketing campaigns. By examining the output of these LLMs in creating advertising texts for products and services in the financial, educational, technological, and health sectors, this investigation provides insights into AI's potential to revolutionize digital marketing practices.

The methodology involves a comparative analysis of the copywriting generated by each LLM, evaluating factors such as persuasion, relevance, and potential for engagement. The research also considers the adaptability of these models to different industry-specific requirements and target audiences.

This investigation contributes to the understanding of the impact of AI-generated content on consumer engagement and provides a foundation for future strategies in digital advertising. By assessing the strengths and limitations of each LLM in producing marketing texts, the study offers insights for marketing professionals, advertisers, and businesses looking to leverage AI technologies in their digital marketing efforts.

The conclusions of this research have implications for the future of copywriting and content creation in digital marketing. They shed light on AI's potential to enhance creativity, increase efficiency, and deliver more targeted and effective marketing messages.

KEYWORDS

Digital Marketing, Generative AI, Copywriting, Large Language Models, Advertising Effectiveness

Objetivos do Desenvolvimento Sustentável (ODS):



ÍNDICE

Declaração de Integridade	i
Dedicatória	ii
Resumo	iii
Abstract	iv
Índice de Figuras	vii
Lista de Siglas e Abreviaturas	viii
1. Introdução	1
2. Revisão da Literatura	2
2.1. Modelos Generativos com foco na sua aplicação na criação de texto	2
2.2. IA generativa no Marketing	4
2.3. Modelos de Linguagem para Copywriting.	6
2.4. Copywriting Persuasivo	8
2.5. A importância dos títulos na captura da atenção do leitor	10
2.6. Engenharia de prompt	12
2.7. Falta de investigação sobre a percepção do consumidor em relação ao conteúdo gerado por IA, especialmente no contexto de anúncios	14
3. Metodologia	16
3.1. Desenho da investigação	17
3.2. Seleção de IAs	17
3.3. Desenvolvimento das campanhas	18
3.4. Estrutura do inquérito	24
3.5. Parâmetros de Avaliação	26
3.6. Relação com a eficácia do copywriting	26
3.7. Influência da emoção	27
3.8. Recolha de dados	27

4. Resultados e Discussão	28
5. Conclusões e Investigação Futura	41
Referências Bibliográficas	42
Apêndice A	48
Apêndice B	49
Apêndice C	52
Apêndice D	59
Apêndice E	61

ÍNDICE DE FIGURAS

Figura 4.1 - Comparação IAGs no setor Financeiro	28
Figura 4.2 - Comparação IAGs no setor Educação	28
Figura 4.3 - Comparação IAGs no setor Tecnologia	29
Figura 4.4 - Comparação IAGs no setor Saúde	29
Figura 4.5 - Comparação Profissionais e Não Profissionais do marketing no setor Financeiro	32
Figura 4.6 - Comparação Profissionais e Não Profissionais do marketing no setor Educação	32
Figura 4.7 - Comparação Profissionais e Não Profissionais do marketing no setor Tecnologia	33
Figura 4.8 - Comparação Profissionais e Não Profissionais do marketing no setor Saúde	33
Figura 4.9 - Correlação entre Gemini 1.5 Pro e a percepção de que as criou no setor Financeiro	35
Figura 4.10 - Correlação entre ChatGPT e a percepção de que as criou no setor Financeiro	35
Figura 4.11 - Correlação entre Llama 3 e a percepção de que as criou no setor Educação	36
Figura 4.12 - Correlação entre ChatGPT e a percepção de que as criou no setor Tecnologia	36
Figura 4.13 - Correlação entre Claude 3 Opus e a percepção de que as criou no setor Financeiro	37
Figura 4.14 - Correlação entre Claude 3 Opus e a percepção de que as criou no setor Tecnologia	37
Figura 4.15 - Correlação entre Gemini 1.5 Pro e a percepção de que as criou no setor Saúde	38
Figura 4.16 - Correlação entre Gemini 1.5 Pro e a percepção de que as criou no setor Saúde	38

LISTA DE SIGLAS E ABREVIATURAS

- Fin-ChatGPT: Média dos parâmetros do setor Financeiro com ChatGPT (QID153)
- Fin-Claude: Média dos parâmetros do setor Financeiro com Claude 3 Opus (QID161)
- Fin-Gemini: Média dos parâmetros do setor Financeiro com Gemini (QID157)
- Fin-Llama: Média dos parâmetros do setor Financeiro com Llama 3 (QID165)
- Edu-ChatGPT: Média dos parâmetros do setor Educação com ChatGPT (QID234)
- Edu-Claude: Média dos parâmetros do setor Educação com Claude 3 Opus (QID222)
- Edu-Gemini: Média dos parâmetros do setor Educação com Gemini (QID226)
- Edu-Llama: Média dos parâmetros do setor Educação com Llama 3 (QID230)
- Tec-ChatGPT: Média dos parâmetros do setor Tecnologia com ChatGPT (QID252)
- Tec-Claude: Média dos parâmetros do setor Tecnologia com Claude 3 Opus (QID248)
- Tec-Gemini: Média dos parâmetros do setor Tecnologia com Gemini (QID240)
- Tec-Llama: Média dos parâmetros do setor Tecnologia com Llama 3 (QID244)
- Sau-ChatGPT: Média dos parâmetros do setor Saúde com ChatGPT (QID270)
- Sau-Claude: Média dos parâmetros do setor Saúde com Claude 3 Opus (QID266)
- Sau-Gemini: Média dos parâmetros do setor Saúde com Gemini (QID262)
- Sau-Llama: Média dos parâmetros do setor Saúde com Llama 3 (QID258)
- P-Fin-ChatGPT: Percepção do setor Financeiro com ChatGPT (QID156)
- P-Fin-Claude: Percepção do setor Financeiro com Claude 3 Opus (QID164)
- P-Fin-Gemini: Percepção do setor Financeiro com Gemini (QID160)
- P-Fin-Llama: Percepção do setor Financeiro com Llama 3 (QID168)
- P-Edu-ChatGPT: Percepção do setor Educação com ChatGPT (QID237)

P-Edu-Claude: Percepção do setor Educação com Claude 3 Opus (QID225)

P-Edu-Gemini: Percepção do setor Educação com Gemini (QID229)

P-Edu-Llama: Percepção do setor Educação com Llama 3 (QID233)

P-Tec-ChatGPT: Percepção do setor Tecnologia com ChatGPT (QID255)

P-Tec-Claude: Percepção do setor Tecnologia com Claude 3 Opus (QID251)

P-Tec-Gemini: Percepção do setor Tecnologia com Gemini (QID243)

P-Tec-Llama: Percepção do setor Tecnologia com Llama 3 (QID247)

P-Sau-ChatGPT: Percepção do setor Saúde com ChatGPT (QID273)

P-Sau-Claude: Percepção do setor Saúde com Claude 3 Opus (QID269)

P-Sau-Gemini: Percepção do setor Saúde com Gemini (QID265)

P-Sau-Llama: Percepção do setor Saúde com Llama 3 (QID261)

LLMs - Modelos de Linguagem de Grande Escala

IA - Inteligência artificial

IAG - Inteligência artificial generativa

1. INTRODUÇÃO

Nos últimos anos, o marketing digital tem sofrido uma evolução significativa, impulsionada por avanços tecnológicos e um investimento crescente por parte das empresas. Esse aumento de investimento é evidenciado pelos dados da IAB Brasil & Kantar IBOPE Media (2023), que revelam que R\$ 35 bilhões foram investidos em publicidade digital entre janeiro e dezembro de 2023, em comparação com R\$ 32,4 bilhões em 2022, R\$ 30,2 bilhões em 2021 e R\$ 23,7 bilhões em 2020. Esse crescimento contínuo reflete a importância cada vez maior do marketing digital nas estratégias de negócios modernos.

Um dos fatores que têm impulsionado essa transformação é a evolução das tecnologias de inteligência artificial (IA), especialmente a IA generativa. Essas tecnologias têm revolucionado a forma como as campanhas de marketing são criadas e executadas, permitindo um nível de personalização e eficácia sem precedentes. Ferramentas baseadas em IA generativa, como os Modelos de Linguagem de Grande Escala (LLMs), são capazes de produzir conteúdo altamente persuasivo e adaptado a diferentes segmentos de público, otimizando o alcance e o impacto das campanhas publicitárias.

Dentro desse contexto tecnológico, o copywriting emerge como um elemento crucial para o sucesso das campanhas de marketing. A qualidade do texto publicitário é determinante para captar a atenção do público, transmitir mensagens de forma eficaz e, conseqüentemente, gerar conversões. O desenvolvimento de textos envolventes e relevantes exige uma combinação de criatividade, conhecimento do público-alvo e habilidades de escrita, e é aqui que a IA generativa pode desempenhar um papel transformador. Modelos de IA avançados podem auxiliar na criação de conteúdo que ressoe com os consumidores, mantendo a consistência da marca e economizando tempo e recursos.

Este estudo tem como objetivo comparar a performance de quatro diferentes Modelos de Linguagem de Grande Escala no âmbito da criação de copywriting para campanhas digitais. A investigação procura entender quais modelos geram os textos mais eficazes, medindo a qualidade e a eficiência do conteúdo produzido. A análise comparativa entre estes modelos oferece informações relevantes sobre as capacidades específicas de cada um e o seu potencial para revolucionar o campo do marketing digital.

2. REVISÃO DA LITERATURA

2.1. Modelos Generativos com foco na sua aplicação na criação de texto.

A Inteligência Artificial Generativa (IAG) marca um avanço notável no domínio da aprendizagem de máquina, abrangendo diversos algoritmos capazes de produzir conteúdo original. Estes sistemas, nutridos por extensos bancos de dados, demonstram a habilidade de criar textos, imagens, áudio e até códigos de programação que se aproximam notavelmente das produções humanas (Zhang et al., 2023). Esta capacidade de gerar variado conteúdo e de qualidade está a remodelar vários setores, da produção de conteúdo ao desenvolvimento de software.

Na esfera da geração textual, os Modelos de Linguagem de Grande Escala (LLMs) têm se destacado por sua eficácia surpreendente. Sistemas como GPT, BERT e T5, fundamentados em arquiteturas de transformers, são alimentados por vastos corpora textuais, permitindo-lhes aprender sutilezas linguísticas complexas e produzir texto coeso e contextualmente adequado (Brown et al., 2020). Estes modelos exibem uma versatilidade notável, sendo capazes de executar uma ampla gama de tarefas de processamento de linguagem natural sem necessidade de treino específico para cada tarefa, um fenómeno denominado aprendizagem em contexto ou aprendizagem com poucos exemplos.

O rápido desenvolvimento e aperfeiçoamento contínuo dos LLMs resultaram em modelos cada vez mais refinados, como o ChatGPT-4, Gemini, Claude-3 e Copilot, que demonstram capacidades impressionantes em várias tarefas, incluindo aquelas que envolvem raciocínio espacial (Hochmair et al., 2024). Estes progressos ampliaram significativamente o espectro de aplicações práticas da IAG, particularmente na criação de conteúdo textual. Rani et al. (2023) conduziram uma análise comparativa de modelos de IAG, ressaltando suas potencialidades e limitações em diversos contextos de aplicação.

Contudo, a adoção generalizada destes modelos não está livre de desafios. Um dos problemas mais significativos é o fenómeno das "alucinações", onde os modelos geram informações factualmente incorretas, mas apresentadas de maneira convincente. Este fenómeno suscita questões importantes sobre a confiabilidade e aplicabilidade dos LLMs em áreas que demandam alta precisão (Nah et al., 2023). Além disso, emergem preocupações éticas

relacionadas à autoria, direitos autorais e o potencial de disseminação de desinformação em larga escala. Wahid et al. (2023) abordam estas questões no âmbito do marketing de conteúdo, enfatizando a necessidade de diretrizes éticas para o uso de IAG na criação de conteúdo.

Apesar destes desafios, o potencial transformador da IAG é incontestável. Investigadores e profissionais de diversas áreas têm explorado suas aplicações em campos como educação personalizada, assistência à comunicação para pessoas com deficiências e aumento da produtividade em indústrias criativas (Rudolph et al., 2023). À medida que a tecnologia progride, torna-se crucial desenvolver estruturas éticas e legais sólidas para regular o desenvolvimento e a aplicação responsável destes poderosos modelos generativos.

Uma das áreas onde a IAG tem demonstrado um impacto particularmente significativo é o marketing. A capacidade destes modelos de gerar texto persuasivo, personalizado e contextualmente relevante oferece oportunidades sem precedentes para revolucionar as estratégias de marketing digital. Kshetri et al. (2024) destacam as diversas aplicações da IAG no marketing, que vão desde a criação de conteúdo personalizado até a otimização de campanhas publicitárias. Huang e Rust (2021) propõem uma estrutura estratégica para a implementação de IA no marketing, enfatizando a importância de alinhar as capacidades da IA com os objetivos de marketing. A IAG reduz drasticamente o tempo necessário para produzir conteúdo de marketing (Kshetri et al., 2023).

Ma e Sun (2020) investigam como a aprendizagem de máquina e a IA estão conectando o poder computacional às percepções humanas no marketing, criando novas oportunidades para personalização e engajamento do cliente. Yau et al. (2021) discutem especificamente como o marketing de IA pode ser utilizado para aprimorar as relações com os clientes, oferecendo experiências mais personalizadas e relevantes. A personalização de serviços e conteúdos em websites e aplicações, o que, segundo Haleem et al. (2022), constitui o passo inicial para impulsionar campanhas de marketing personalizadas e criar um envolvimento significativo com o consumidor.

A transição da compreensão geral da IAG para suas aplicações específicas no marketing representa um passo crucial na exploração do potencial desta tecnologia. Shaik (2023) examina o impacto da IA no marketing, destacando como ela está transformando práticas tradicionais e criando novas oportunidades para as empresas se conectarem com os seus

públicos-alvo. Tauheed et al. (2024) exploram o papel da IA em estratégias de marketing digital, enfatizando como pode ser utilizada para melhorar a eficácia das campanhas e a experiência do cliente.

Yudiyanto e Purnomo (2023) oferecem observações perspicazes sobre como estratégias de marketing digital potencializadas por IA podem aumentar as taxas de conversão em plataformas de e-commerce, demonstrando o valor prático dessas tecnologias no cenário competitivo atual.

À medida que avançamos para uma era onde a IAG se torna cada vez mais integrada às estratégias de marketing, é fundamental que profissionais e investigadores continuem a explorar seu potencial, ao mesmo tempo em que abordam de forma proativa os desafios éticos e práticos associados a seu uso.

2.2. IA generativa no Marketing

No âmbito do marketing digital, os LLMs estão provocando uma revolução no campo do copywriting. Esses modelos exibem uma notável habilidade de produzir textos persuasivos, títulos cativantes e conteúdo personalizado em larga escala (Gao et al., 2023), adaptando-se a diversos tons de voz e estilos de marca. Bubeck et al. (2023) ressaltam que esses sistemas de IA, como o GPT-4, têm a capacidade de emular a escrita humana de forma convincente, possibilitando uma comunicação mais eficaz e alinhada com a identidade da marca.

A IA generativa tem-se revelado particularmente útil no copywriting, oferecendo uma vantagem considerável sobre as abordagens convencionais. Kshetri et al. (2024) salientam que essas ferramentas podem gerar rapidamente variações de textos publicitários, permitindo testes A/B mais eficientes e uma otimização constante das mensagens de marketing. Huang e Rust (2021) apresentam uma estrutura estratégica para a implementação de IA no marketing, enfatizando a capacidade da IA de criar experiências altamente individualizadas. No contexto do copywriting, isto traduz-se em conteúdo que pode ser automaticamente ajustado para diferentes segmentos de público, elevando a relevância e o impacto das mensagens de marketing.

Além da redação publicitária, a IA generativa está a ser empregue na criação de títulos e anúncios digitais de maneira eficiente. Gołąb-Andrzejak (2023) analisou o impacto do ChatGPT na elaboração de campanhas publicitárias digitais, destacando como estas ferramentas podem produzir anúncios personalizados em grande escala. Nah et al. (2023) sublinham a importância da supervisão humana para assegurar precisão, relevância da marca e alinhamento estratégico. Isso indica que, apesar dos avanços na IA generativa, o papel do redator publicitário humano continua a ser fundamental, especialmente em termos de revisão, edição e garantia de qualidade. Estes modelos não só produzem textos eficazes em termos de engajamento, mas também podem ser utilizados para aprimorar a estratégia de conteúdo ao longo do tempo com base em métricas de desempenho (Chen et al., 2023).

A aplicação da IA generativa na criação de publicidade digital em redes sociais representa um avanço significativo na personalização e otimização de campanhas. Wahid et al. (2023) enfatizaram o uso de IA generativa para o marketing de conteúdo, onde modelos como o ChatGPT podem redigir e personalizar posts para plataformas de redes sociais, ajustando o tom e o estilo de acordo com o público-alvo. Esta capacidade de adaptação e personalização permite que as empresas atinjam o seu público de maneira mais eficaz, aumentando o engajamento e as taxas de conversão, enquanto mantêm a consistência da marca em todas as suas comunicações digitais. No campo da otimização para SEO, Yudiyanto e Purnomo (2023) discutem como as estratégias de marketing digital potencializadas por IA podem aumentar as taxas de conversão em plataformas de e-commerce, sugerindo que o conteúdo gerado por IA, quando otimizado para SEO, pode contribuir significativamente para o desempenho do marketing digital.

A IA generativa está a transformar o marketing digital ao possibilitar a criação rápida e eficiente de conteúdo altamente personalizado e persuasivo. Desde a criação de títulos atraentes até a elaboração de campanhas publicitárias completas, esses modelos de IA não aumentam apenas a produtividade, mas também melhoram a eficácia das estratégias de marketing. Contudo, a supervisão humana continua a ser essencial para garantir a qualidade e a relevância das mensagens, demonstrando que a colaboração entre IA e profissionais de marketing é fundamental para o sucesso contínuo nesta área.

2.3. Modelos de Linguagem para Copywriting.

O ChatGPT-4, desenvolvido pela OpenAI, representa um marco significativo na evolução dos modelos de linguagem de grande escala. Fundamentado na arquitetura transformer, com aprimoramentos substanciais em relação às versões precedentes, as suas capacidades abrangem a compreensão e geração de texto em múltiplos idiomas, raciocínio complexo e execução de tarefas multifacetadas. No contexto do marketing digital, o ChatGPT-4 pode ser utilizado na criação de textos persuasivos e personalizados, títulos atraentes, e na criação de conteúdo de alta qualidade em larga escala. A sua aptidão para compreensão contextual viabiliza a elaboração de mensagens de marketing alinhadas com o tom e o estilo da marca, possibilitando uma comunicação mais eficaz e personalizada com diferentes segmentos de público (Brown et al., 2020).

Na área de copywriting, o ChatGPT-4 destaca-se pela produção de textos nuançados e personalizados, possibilitando a criação rápida de múltiplas versões de conteúdo. Além disso, a IA pode automatizar tarefas repetitivas, permitindo que os redatores humanos se concentrem em aspectos mais estratégicos e criativos do marketing. Estudos de Nah et al. (2023) evidenciam que a revisão e supervisão humanas são cruciais para assegurar a precisão e a relevância do conteúdo, mitigando a tendência ocasional de gerar informações imprecisas. A capacidade do ChatGPT-4 de se adaptar a diferentes tons e estilos de marca pode aumentar consideravelmente a eficiência e a eficácia das campanhas de marketing, conforme discutido por Huang e Rust (2021).

O Claude 3 Opus, criado pela Anthropic, é um modelo de linguagem avançado reconhecido pela sua precisão e capacidade de processamento de contexto extenso. Concebido para lidar com tarefas complexas e manter a coerência em conversas longas, este modelo é particularmente útil para análise e síntese de informações detalhadas. No marketing digital, o Claude 3 Opus pode ser utilizado para gerar campanhas publicitárias que consideram implicações morais e éticas de longo prazo, tornando-o valioso em contextos onde a tomada de decisões sensíveis é crucial. Investigações conduzidas por Chen et al. (2023) sugerem que este modelo pode gerar textos publicitários que não são apenas persuasivos, mas também alinhados com valores éticos, aumentando a confiança e a lealdade do consumidor.

No copywriting, a habilidade do Claude 3 Opus de produzir respostas que navegam por dilemas éticos e consideram valores humanos adiciona uma dimensão extra à criação de

conteúdo. Contudo, críticos como Patel e Nguyen (2024) ressaltam que a definição de valores éticos em sistemas de IA suscita questões sobre viés cultural, indicando que a colaboração contínua entre desenvolvedores,eticistas e comunidades diversas é fundamental para a implementação responsável dessas tecnologias.

O Gemini 1.5 Pro, desenvolvido pela Google, é um modelo multimodal capaz de processar e gerar conteúdo em diversos formatos, como texto, imagem e áudio. Um modelo multimodal integra diferentes tipos de dados numa única estrutura, permitindo a compreensão e criação de conteúdo que combina texto com imagens, áudio e outros formatos. Esta capacidade torna o Gemini 1.5 Pro particularmente valioso para criar campanhas de marketing integradas que aproveitam múltiplos canais de mídia. Estudos de Wu et al. (2023) demonstram que este modelo pode interpretar gráficos complexos e extrair percepções significativas de dados visuais, como a análise de gráficos de desempenho de campanhas e a correlação com dados textuais de feedback do cliente. Isso é especialmente útil para criar campanhas de marketing baseadas em dados, onde a combinação de diferentes tipos de informações podem levar a decisões mais informadas e estratégias mais eficazes.

Para copywriting, o Gemini 1.5 Pro oferece a vantagem de gerar títulos e conteúdo otimizados para múltiplos formatos de mídia, permitindo uma abordagem integrada às campanhas de marketing digital. Investiga de Kim e Rodrigues (2024) enfatizam a necessidade de métodos robustos para explicar e auditar as decisões tomadas por sistemas multimodais complexos como o Gemini 1.5 Pro, garantindo que as campanhas sejam eficazes, transparentes e confiáveis.

O Llama 3, desenvolvido pela Meta, é um modelo de código aberto projetado para democratizar o acesso a tecnologias avançadas de linguagem. A sua arquitetura baseada em transformers é otimizada para eficiência computacional, permitindo um desempenho competitivo mesmo com um tamanho de modelo relativamente menor. No marketing digital, o Llama 3 pode ser utilizado para criar campanhas personalizadas e acessíveis, facilitando a inovação e a investigação colaborativa. Estudos conduzidos por Gupta et al. (2024) mostram que a comunidade de desenvolvedores tem utilizado o Llama 3 como base para criar modelos especializados em domínios específicos, desde a análise de sentimentos em redes sociais até à assistência em programação.

Para copywriting, o Llama 3 é particularmente eficaz devido à sua abordagem de "modelo base + ajuste fino", permitindo a personalização do conteúdo para diferentes domínios específicos. Esta abordagem é eficiente em cenários onde os recursos computacionais ou dados de treino são limitados. Além disso, o Llama 3 tem sido instrumental no avanço da investigação em interpretabilidade e robustez de modelos de linguagem, permitindo a criação de campanhas que não são apenas eficazes, mas também transparentes e confiáveis. Isso é corroborado por Yudiyanto e Purnomo (2023), que destacam a importância de modelos interpretáveis na otimização de estratégias de marketing digital, melhorando a confiança e a aceitação dos consumidores.

2.4. Copywriting Persuasivo

Os fundamentos do copywriting persuasivo são essenciais para uma comunicação de marketing eficaz, especialmente no ambiente digital. Esses princípios baseiam-se num entendimento profundo da psicologia do consumidor e em técnicas de persuasão desenvolvidas ao longo de décadas de investigações em publicidade e marketing. Kumar (2015) na sua investigação sobre neuromarketing revelou que os anúncios de texto demonstraram ser mais eficazes do que os anúncios gráficos.

Cialdini (2007) propõe que a persuasão eficaz se alicerça em seis princípios fundamentais: reciprocidade, escassez, autoridade, consistência, afinidade e consenso social. Quando aplicados ao copywriting, esses princípios podem aumentar significativamente a eficácia da mensagem. Por exemplo, o princípio da escassez pode ser utilizado para criar um senso de urgência numa oferta, enquanto a autoridade pode ser utilizada para estabelecer credibilidade. Entre os elementos cruciais do copywriting persuasivo, destacam-se a proposta de valor única (UVP), a segmentação precisa do público-alvo e a implementação estratégica de chamadas para ação (CTAs). Osterwalder et al. (2014) enfatizam que uma UVP eficaz deve comunicar claramente como o produto resolve um problema específico do cliente ou melhorar a sua situação, entregando benefícios concretos. Isso é particularmente importante num mercado saturado, onde a diferenciação é crucial.

A segmentação do público-alvo permite a personalização da mensagem, aumentando a sua relevância e, conseqüentemente, a sua eficácia. Kotler e Keller (2012) ressaltam que a

compreensão profunda das necessidades, desejos e comportamentos do público-alvo é crucial para o desenvolvimento de estratégias de marketing eficazes. Esse entendimento permite que os copywriters adaptem sua linguagem e argumentos para ressoar de uma forma mais intensa com seu público-alvo.

As chamadas para ação (CTAs) são elementos críticos que guiam o consumidor para o próximo passo desejado. Estudos em neuromarketing, como os conduzidos por Renvoisé e Morin (2007), sugerem que CTAs claras e diretas, que apelam para o "cérebro primitivo" do consumidor, são mais propensas a gerar respostas imediatas e desejadas. Isso destaca a importância de usar linguagem clara e direta ao solicitar uma ação específica do consumidor.

A narrativa e o apelo emocional também desempenham um papel crucial no copywriting persuasivo. Zak (2015) demonstra que histórias bem construídas podem aumentar a produção de oxitocina no cérebro, facilitando a conexão emocional e a empatia. Essa conexão emocional é fundamental para superar a resistência natural dos consumidores à publicidade e criar engajamento duradouro com a marca.

Kahneman (2011) sugere que as decisões de compra são frequentemente impulsionadas por respostas emocionais rápidas, seguidas de justificativas racionais. Neste contexto, copywriters eficazes utilizam técnicas de storytelling para criar narrativas que ressoam com as experiências e aspirações do público-alvo, enquanto comunicam simultaneamente os benefícios do produto ou serviço.

Berger (2013), nas suas investigações sobre conteúdo viral, indica que mensagens que evocam emoções de alta excitação, como admiração ou ansiedade, são mais propensas a serem compartilhadas, aumentando o alcance orgânico da comunicação de marketing. Portanto, a integração habilidosa de elementos narrativos e emocionais no copywriting não aumenta apenas a persuasão imediata, mas também potencializa a disseminação e o impacto de longo prazo da mensagem de marketing.

Todos esses elementos - os princípios de persuasão de Cialdini (2007), a proposta de valor única, a segmentação precisa, as chamadas para ação eficazes, e o uso de narrativa e apelo emocional - convergem na criação de copy persuasivo. No entanto, o primeiro ponto de contato entre o consumidor e a mensagem de marketing é frequentemente o título. Um título

persuasivo deve incorporar muitos desses princípios numa frase concisa e impactante, capturando imediatamente a atenção do público-alvo e incentivando-o a continuar a ler ou a clicar no anúncio. Compreender o que faz um título ser persuasivo é, portanto, um aspecto crítico do copywriting eficaz.

2.5. A importância dos títulos na captura da atenção do leitor.

Na investigação científica sobre a eficácia das manchetes em atrair a atenção dos leitores, um elemento fundamental do copywriting persuasivo é examinado, especialmente no contexto digital repleto de informações. A base psicológica de manchetes efetivas está enraizada em princípios de cognição e comportamento do consumidor. Kahneman (2011) argumenta que o "Sistema 1" do cérebro, responsável pelo processamento rápido e intuitivo, é ativado primeiramente ao encontrar novas informações, tornando as primeiras impressões cruciais para o engajamento subsequente.

Os gatilhos mentais são cruciais na elaboração de manchetes cativantes. Cialdini (2007) identificou seis princípios de influência aplicáveis a manchetes: escassez, autoridade, consistência, reciprocidade, prova social e afinidade. Por exemplo, manchetes que destacam a escassez ou a autoridade podem aumentar consideravelmente as taxas de clique.

A concisão e o impacto da linguagem são elementos essenciais para manchetes eficazes. Ogilvy (1983) apresentou que manchetes com menos de 10 palavras tendem a ter melhor desempenho, alinhando-se com os princípios contemporâneos de SEO e otimização para dispositivos móveis. O uso de palavras emocionalmente carregadas, como demonstrado por estudos de neuromarketing de Renvoisé e Morin (2007), pode ativar respostas emocionais rápidas, aumentando o engajamento. Corroborando com essa ideia, Berger et al. (2023) demonstram que a linguagem emocional tem um forte impacto na atenção do público, tornando o conteúdo mais envolvente e memorável.

A personalização das manchetes com base em dados demográficos e comportamentais do público-alvo, como proposto por estudos de marketing, pode aumentar consideravelmente a sua relevância e eficácia, resultando em maiores taxas de conversão. Neste contexto, é importante considerar as descobertas de investigações recentes sobre como diferentes emoções afetam a atenção. Emoções como surpresa, alegria, curiosidade e empatia tendem a aumentar a atenção, enquanto estados emocionais intensos negativos podem diminuí-la.

No seu livro, Maslen (2015) ressalta que a emoção é um elemento fundamental na arte da persuasão em copywriting. Ele propõe que as decisões de compra são, na sua essência, emocionais, com a razão servindo apenas para justificar escolhas já feitas pelo coração. Maslen (2015) enfatiza a importância de recorrer a sentimentos básicos como o medo, a ambição, a culpa e o orgulho, argumentando que estes exercem uma influência poderosa no comportamento do consumidor. O autor sugere que redatores eficazes devem identificar o sentimento predominante do público-alvo em relação ao produto ou serviço e, em seguida, elaborar um conteúdo que dialogue com essa emoção, intensificando-a ou oferecendo uma solução para aliviar sensações negativas.

Maslen (2015) também destaca a relevância da clareza e simplicidade na linguagem. Ele defende o uso de palavras do dia a dia, frases curtas e parágrafos concisos para maximizar o entendimento e o impacto. O autor enfatiza que o objetivo do copywriting não é impressionar o leitor com um vocabulário rebuscado, mas comunicar a mensagem de forma clara e convincente. Maslen (2015) também ressalta a eficácia de técnicas como o uso de verbos ativos, a eliminação de calão desnecessário e a estruturação do texto de maneira que guie naturalmente o leitor para a ação desejada.

Por fim, Maslen (2015) sublinha a importância de conhecer profundamente o público-alvo e adaptar o tom e o estilo da escrita de acordo. Ele propõe que os copywriters devem realizar pesquisas detalhadas sobre os seus potenciais clientes, incluindo as suas necessidades, desejos, medos e aspirações. Com base nesse conhecimento, o texto pode ser personalizado para ressoar com as vivências e valores específicos do público-alvo. Além disso, Maslen (2015) ressalta a importância de estabelecer credibilidade através do uso estratégico de dados, depoimentos e histórias de caso, bem como a necessidade de criar um sentido de urgência para motivar a ação imediata do leitor. Ele argumenta que a combinação desses elementos - emoção, clareza, conhecimento do público e urgência - forma a base de um copywriting verdadeiramente persuasivo e eficaz na venda.

O modelo AIDA (Atenção, Interesse, Desejo, Ação) permanece como uma estrutura essencial no copywriting persuasivo, conforme abordado por Maslen (2015). Este framework, inicialmente proposto por E. St. Elmo Lewis no final do século XIX, mantém a sua relevância na era digital devido à sua eficácia em organizar mensagens persuasivas. Maslen (2015) ressalta que o AIDA reflete o processo psicológico experienciado pelos consumidores antes de tomarem uma decisão de compra. Inquéritos recentes em neurociência do consumidor

corroboram a validade deste modelo, evidenciando que o cérebro processa informações de marketing em etapas semelhantes às propostas pelo AIDA.

Na fase de Atenção, Maslen (2015) argumenta que técnicas como manchetes impactantes e uso estratégico de elementos visuais são cruciais. Isso está em consonância com estudos que enfatizam a importância da relevância e clareza na criação de textos persuasivos.

Para despertar Interesse, recomenda-se o uso de narrativas cativantes e fatos surpreendentes, o que se relaciona com a eficácia da curiosidade e surpresa em capturar a atenção, como apontado por investigações recentes.

O estágio de Desejo envolve a intensificação das emoções positivas e a mitigação de possíveis objeções. Isso alinha-se com a ideia de combinar o apelo emocional e a lógica para uma persuasão mais eficaz, como destacado em estudos contemporâneos.

Por fim, na etapa de Ação, Maslen (2015) recomenda o uso de chamadas para ação (CTAs) claras e urgentes, o que se relaciona diretamente com a importância de criar um sentido de urgência e fornecer instruções precisas, elementos-chave de textos persuasivos identificados na literatura mais recente.

2.6. Engenharia de prompt

A engenharia de prompts é uma área fundamental no desenvolvimento e aplicação de LLMs, definida por Liu et al. (2023a) como "a prática de projetar e otimizar instruções ou consultas para LLMs com o objetivo de melhorar o seu desempenho em tarefas específicas". Esta prática tem-se mostrado crucial para aprimorar a eficácia dos LLMs em diversas aplicações, incluindo o copywriting e o marketing digital.

A criação de prompts eficazes envolve vários aspectos importantes. A especificidade e clareza são fundamentais, como demonstrado por Wei et al. (2022), que observaram que prompts mais detalhados e explícitos tendem a produzir respostas mais precisas e relevantes. No contexto do copywriting, isso implica incluir detalhes específicos sobre o público-alvo, o tom desejado e os objetivos da mensagem publicitária.

A estruturação adequada do prompt também é crucial. Kojima et al. (2022) introduziram o conceito de "chain-of-thought prompting", que envolve guiar o modelo através de um processo de raciocínio passo a passo para melhorar o desempenho em tarefas complexas. Esta

abordagem pode ser particularmente útil em tarefas de marketing que requerem análise e planejamento estratégico.

White et al. (2023) propõem um catálogo de padrões de prompts, incluindo "Persona Prompting" e "Chain of Thought Prompting", que podem ser particularmente eficazes em várias tarefas. Estes padrões oferecem estruturas predefinidas que podem ser adaptadas para diferentes contextos de marketing e publicidade. Kshetri (2023) observa que com os prompts adequados, esta tecnologia pode criar 10 abordagens publicitárias diferentes para um produto, baseando-se em avaliações de clientes ou outros fatores.

A iteração e refinamento são aspectos importantes do processo de engenharia de prompts. Ekin (2023) enfatiza a importância de uma abordagem de "tentativa e erro", onde os prompts são continuamente ajustados com base nos resultados obtidos. Esta abordagem iterativa permite otimizar os prompts para obter os melhores resultados possíveis em diferentes contextos de marketing.

A consideração do contexto e conhecimento prévio é outro aspecto crucial. Ekin (2023) sugere que a inclusão de informações de fundo relevantes pode melhorar significativamente a qualidade e relevância das respostas geradas. Isto é particularmente importante no marketing, onde o conhecimento do mercado, da marca e do público-alvo é essencial para criar mensagens eficazes.

O estímulo à criatividade e ao pensamento crítico também é um aspecto importante na elaboração de prompts. Giray (2023) destaca a importância de estruturar prompts de forma a estimular a criatividade e o pensamento crítico do LLM. Isto pode levar a ideias mais inovadoras e a soluções criativas para desafios de marketing. Mogaji et al. (2023) salientam a importância de uma abordagem sinérgica entre profissionais e tecnologia. Os investigadores destacam, em particular, a necessidade de os profissionais desenvolverem novas competências, com ênfase especial na engenharia de prompt. Segundo Mogaji et al. (2023), o domínio desta habilidade pode reduzir significativamente a sensação de inadequação e a relutância em adotar ferramentas de IAG. Esta competência permite aos profissionais aprimorar a sua interação com sistemas de IA, potencializando a eficácia e a criatividade na produção de conteúdo publicitário.

Além da estrutura e conteúdo do prompt, podem ser ajustados diversos parâmetros para otimizar as respostas dos LLMs. A temperatura, por exemplo, controla a aleatoriedade das

respostas. Uma temperatura mais baixa torna as respostas mais determinísticas, enquanto uma temperatura mais alta aumenta a criatividade e diversidade. O Top P, ou amostragem de núcleo, é outra técnica que permite um equilíbrio entre diversidade e qualidade nas respostas geradas. Existem outros parâmetros como Presence Penalty, Frequency Penalty, e Max Tokens, que podem ser ajustados para aperfeiçoar ainda mais as respostas dos LLMs.

A relevância de criar prompts eficazes é multifacetada. Brown et al. (2020) demonstraram que prompts bem elaborados podem melhorar significativamente o desempenho de um LLM em tarefas complexas, mesmo com poucos ou nenhum exemplo prévio (few-shot learning). Além disso, prompts eficazes permitem adaptar LLMs a tarefas específicas sem a necessidade de retreinamento extensivo.

A qualidade, relevância e utilidade das respostas geradas pelos modelos de IA são diretamente influenciadas pela formulação do prompt. Liu et al. (2023b) observaram que o uso de prompts que incluem exemplos relevantes (few-shot prompting) pode melhorar a precisão das respostas em tarefas de classificação de texto em até 30% em comparação com prompts simples.

A mitigação de vieses é outro aspecto crucial na engenharia de prompts. Sheng et al. (2020) demonstraram que a formulação cuidadosa de prompts pode ajudar a reduzir vieses de gênero e outros preconceitos nas respostas geradas por modelos de linguagem. Isto é particularmente importante no marketing, onde a inclusividade e a equidade são cada vez mais valorizadas.

Em suma, a engenharia de prompts é uma área crítica no desenvolvimento e aplicação de LLMs para marketing e publicidade. Prompts bem elaborados, combinados com a otimização cuidadosa de parâmetros, podem melhorar significativamente a qualidade, relevância e utilidade das respostas geradas, além de mitigar vieses e permitir um maior controlo sobre o estilo e tom das saídas. À medida que a tecnologia de IA continua a evoluir, a importância da engenharia de prompts provavelmente aumentará, tornando-se uma habilidade essencial para profissionais de marketing que trabalham com IA e processamento de linguagem natural.

2.7. Falta de investigação sobre a percepção do consumidor em relação ao conteúdo gerado por IA, especialmente no contexto de anúncios.

Apesar dos avanços tecnológicos, existe uma lacuna significativa na compreensão de como o conteúdo gerado por IA é percebido e recebido pelos consumidores, bem como as suas implicações para a eficácia publicitária.

Esta lacuna de conhecimento é particularmente evidente em três áreas principais. Primeiramente, há uma escassez de inquéritos empíricos que examinem as diferenças na eficácia e percepção entre conteúdo gerado por modelos diferentes de IA. Em segundo lugar, pouco se sabe sobre as variações na avaliação desse conteúdo entre profissionais de marketing e consumidores comuns. Por fim, a relação entre a qualidade percebida do conteúdo gerado por IA e sua atribuição à autoria humana permanece largamente inexplorada.

Para abordar essas lacunas, são propostas as seguintes hipóteses:

H1: Haverá diferenças significativas na avaliação geral entre as quatro IAs usadas para gerar as campanhas.

Esta hipótese procura investigar se existem diferenças perceptíveis na qualidade e eficácia do conteúdo gerado por diferentes modelos de IA. A sua investigação é crucial para compreender as nuances e capacidades específicas de cada modelo no contexto do copywriting publicitário.

H2: Profissionais de marketing e não-profissionais terão avaliações significativamente diferentes sobre qual IA produziu as melhores campanhas.

Esta hipótese explora a potencial divergência entre as percepções de especialistas e consumidores comuns. A sua investigação pode revelar informações sobre a eficácia prática versus a percepção do público-alvo, informando estratégias mais eficazes de implementação de IA no marketing.

H3: As campanhas criadas pela IA com as maiores pontuações gerais serão mais frequentemente percebidas como criadas por humanos.

Esta hipótese aborda questões fundamentais de autenticidade e confiança no conteúdo gerado por IA. A sua investigação pode lançar luz sobre os fatores que influenciam a percepção de

autenticidade em conteúdo publicitário e as implicações para a eficácia da comunicação de marca.

Estas hipóteses, fundamentadas na literatura existente e nas lacunas identificadas, visam contribuir significativamente para o campo do marketing digital e o uso de IA em publicidade.

3. METODOLOGIA

3.1. Desenho da investigação

A avaliação humana é crucial para determinar a eficácia e qualidade real dos anúncios de texto gerados por IA. Embora as métricas automatizadas forneçam dados úteis, elas não capturam completamente as nuances da linguagem, relevância e atratividade que os humanos podem perceber. Wang et al. (2021) demonstram a importância de utilizar juízes humanos para avaliar anúncios em termos de qualidade e atratividade. O processo de avaliação humana proposto por Wang et al. (2021) é rigoroso e abrangente, focando-se em quatro critérios essenciais: fluência linguística, semelhança com textos escritos por humanos, relevância para a página de destino do produto e precisão das informações. Esta abordagem multifacetada garante que os anúncios não sejam apenas tecnicamente corretos, mas também eficazes do ponto de vista do marketing e da comunicação. A metodologia de avaliação humana descrita no estudo também incorpora medidas para reduzir vieses, como a randomização da ordem dos anúncios e o posicionamento aleatório nas comparações lado a lado. Isso aumenta a confiabilidade e a objetividade das avaliações. Portanto, enquanto as IAs generativas podem criar conteúdo em larga escala, a avaliação humana permanece indispensável para garantir que esse conteúdo atenda aos padrões de qualidade, relevância e eficácia necessários no campo competitivo da publicidade digital. Esta abordagem híbrida, combinando a eficiência da IA com o discernimento humano, representa o estado da arte na avaliação de conteúdo publicitário criado por IA.

3.2. Seleção de IAs

A seleção das quatro IAs generativas - Gemini 1.5 Pro, ChatGPT 4, Claude 3 Opus e Llama 3 - para este estudo baseia-se em critérios que refletem o estado atual da tecnologia de IA e sua aplicabilidade no campo do copywriting para marketing digital. Estas IAs representam o que há de mais avançado em processamento de linguagem natural, sendo desenvolvidas por líderes da indústria como Google, OpenAI, Anthropic e Meta, respectivamente. A diversidade das suas origens garante uma variedade de abordagens e práticas, essencial para uma comparação abrangente e significativa. Além disso, estes modelos demonstram capacidades excepcionais em tarefas de criação de texto, adaptabilidade a diferentes estilos de escrita e compreensão de contexto, características cruciais para o copywriting eficaz (Zhang et al., 2023; Rani et al., 2023). Conforme destacado por Aydın e Karaarslan (2023), estas IAs são

consideradas concorrentes diretas do ChatGPT, cada uma oferecendo recursos únicos e potencialmente superiores em determinados aspectos, o que as torna particularmente relevantes para uma análise comparativa no campo do copywriting para marketing digital.

A relevância destas IAs para o copywriting em marketing digital é multifacetada. Primeiramente, elas possuem a capacidade de criar conteúdo persuasivo e adaptável a diferentes públicos-alvo, um requisito fundamental no marketing digital contemporâneo. Sua habilidade em processar e sintetizar grandes volumes de dados permite a criação de copy informado por tendências de mercado e comportamento do consumidor. Além disso, estas IAs têm demonstrado potencial para inovação criativa, podendo criar ideias frescas e abordagens únicas para campanhas publicitárias. Essa combinação de habilidades torna-as ferramentas potencialmente poderosas para profissionais de marketing, capazes de complementar e amplificar a criatividade humana (Gołąb-Andrzejak, 2023; Kshetri et al., 2024).

A inclusão destas IAs específicas no estudo não reflete apenas o estado atual da tecnologia, mas também antecipa futuras tendências no campo do marketing digital. Ao comparar os seus desempenhos em tarefas de copywriting, o estudo fornecerá informações sobre as capacidades e limitações da IA generativa neste domínio. Isso é particularmente relevante num momento em que a indústria de marketing está a adotar cada vez mais soluções baseadas em IA para otimizar e personalizar campanhas publicitárias. Os resultados deste estudo poderão informar decisões estratégicas sobre a integração de IA no processo criativo de marketing, bem como orientar o desenvolvimento futuro destas tecnologias para melhor atender às necessidades específicas do copywriting em marketing digital (Hartmann et al., 2023; Osadchaya et al., 2024).

3.3. Desenvolvimento das campanhas

O desenvolvimento das campanhas para cada setor (Financeiro, Educação, Tecnologia, Saúde) seguiu um processo meticuloso de criação de prompts, visando extrair o melhor desempenho das IAs generativas selecionadas (Gemini 1.5 Pro, ChatGPT 4, Claude 3 Opus e Llama 3). O processo baseou-se num template de prompt padronizado, inspirado nas melhores práticas de engenharia de prompts, como sugerido por White et al. (2023). Este template foi projetado para fornecer instruções claras e contexto suficiente para cada IA, permitindo o desenvolvimento de conteúdo de copywriting relevante e eficaz para cada setor.

O prompt padrão foi estruturado para incluir elementos cruciais como o nome do produto ou serviço, benefícios principais, descrição do público-alvo, requisitos específicos dos textos, e instruções para interação com uma imagem. Esta estrutura garantiu que cada IA recebesse informações consistentes e abrangentes para criar o conteúdo solicitado. Além disso, o prompt incluía diretrizes específicas sobre o comprimento dos textos (125 caracteres para o texto principal e 40 caracteres para o título), alinhando-se às boas práticas do Facebook Ads, onde "O comprimento de texto recomendado para a maioria dos posicionamentos é de 125 caracteres para o campo Texto principal, 40 caracteres para o campo Título e 25 caracteres para o campo Descrição".

Para cada setor, o prompt foi adaptado com informações específicas. Por exemplo, para o setor Financeiro, o foco foi na "Conta Digital Jovem", destacando benefícios como a ausência de taxas de manutenção e a facilidade de uso. No setor de Educação, o prompt centrou-se na "Pós Graduação em Ciência de Dados", enfatizando a capacitação em metodologias avançadas de análise de dados. Para o setor de Tecnologia, o produto em destaque foi o "celular Z50 Pro", com ênfase nas suas características avançadas e tecnologia de ponta. No setor de Saúde, o foco foi na "Mentoria Equilibrium", destacando o acompanhamento nutricional personalizado e a transformação saudável.

A seleção dos produtos e serviços para cada setor foi realizada para abranger uma ampla gama de público-alvo, garantindo uma diversidade significativa nas campanhas de marketing a serem analisadas. Esta abordagem permite uma avaliação mais abrangente da capacidade dos LLMs em adaptar-se a diferentes contextos e audiências.

No setor Financeiro, a escolha da "Conta Digital Jovem" visa o público jovem, tipicamente e Geração Z. Esta seleção reflete a crescente tendência de digitalização bancária e a preferência deste grupo demográfico por soluções financeiras ágeis e sem burocracia. Este público é geralmente mais receptivo a inovações tecnológicas e valoriza a praticidade e a ausência de custos, de acordo com o inquérito realizado pela RADAR Febraban (2023), 71% dos jovens no Brasil, na faixa dos 18 aos 24 anos preferem os bancos digitais.

Para o setor de Educação, a "Pós Graduação em Ciência de Dados" é direcionada a profissionais em busca de especialização ou transição de carreira, geralmente na faixa etária de 25 a 45 anos. A escolha deste curso reflete a crescente demanda por habilidades em análise de dados no mercado de trabalho. O público-alvo aqui é mais maduro, focado em

desenvolvimento profissional e provavelmente já estabelecido nas suas carreiras. O Graduate Management Admission Council (GMAC, 2021) mostrou que 47% de todos os programas de mestrados em análise de dados reportaram aumento nas aplicações.

No setor de Tecnologia, o "Celular Z50 Pro" visa consumidores de produtos premium, com maior poder aquisitivo. A escolha de um smartphone de alta gama permite explorar como os LLMs lidam com a comunicação de características técnicas avançadas e benefícios de produtos de luxo. O relatório da Counterpoint (2023) demonstra a importância crescente dos smartphones premium no mercado global. Em 2022, pela primeira vez, a receita gerada por esses dispositivos ultrapassou 50% do total, evidenciando a preferência do consumidor por aparelhos com mais recursos e tecnologia avançada, mesmo com preços mais altos.

Para o setor de Saúde, a "Mentoria Equilibrium" é voltada para adultos preocupados com saúde e bem-estar, tipicamente na faixa dos 25 aos 45 anos. A escolha reflete a crescente tendência de busca por soluções personalizadas de saúde e bem-estar. Os dados do relatório da Global Wellness Economy Monitor (2023) revelam o crescimento consistente do setor de nutrição, alimentação saudável e perda de peso. Em 2019, esse setor representava 20.9% da economia global de bem-estar (US\$ 911.3 bilhões de um total de US\$ 4.9317 trilhões). Em 2022, esse percentual subiu para 21.4% (US\$ 1,079.3 bilhões de um total de US\$ 5,611.6 trilhões).

Esta seleção diversificada permite avaliar a versatilidade dos LLMs em criar conteúdo para diferentes faixas etárias, variados níveis de poder aquisitivo, distintos níveis de conhecimento técnico e diversas motivações de compra. Ao escolher produtos e serviços tão distintos, o estudo pode avaliar a capacidade dos LLMs de adaptar o tom, o estilo e o conteúdo das mensagens para ressoar com públicos-alvo específicos, testando assim a flexibilidade e eficácia destes modelos em cenários de marketing variados e realistas.

Cada prompt setorial foi cuidadosamente elaborado para comunicar com o público-alvo específico. Por exemplo, o prompt para a Conta Digital Jovem foi direcionado a jovens de até 25 anos, enquanto o da Pós Graduação em Ciência de Dados se focou em profissionais de 25 a 35 anos buscando aperfeiçoamento. O prompt do celular Z50 Pro foi voltado para consumidores de rendimentos elevados que valorizam tecnologia de ponta, e o da "Mentoria Equilibrium" se focou em mulheres adultas de 25 a 45 anos interessadas em saúde e bem-estar.

É importante notar que os prompts foram criados especificamente para campanhas no Facebook Ads, levando em consideração que o conteúdo criado sempre será acompanhado por uma imagem. Isso reflete as melhores práticas de publicidade nesta plataforma, onde a combinação de texto conciso e imagens impactantes é crucial para o engajamento do público. A limitação de caracteres nos prompts não só se alinha às recomendações do Facebook para evitar truncamento em diferentes dispositivos e posicionamentos, mas também desafia as IAs a criar mensagens impactantes e persuasivas dentro de restrições realistas de espaço, simulando as condições reais enfrentadas por copywriters no ambiente de marketing digital.

Os parâmetros como "temperatura" e "top_P" são ajustáveis em modelos de linguagem e influenciam diretamente a natureza dos outputs gerados. A "temperatura" controla a aleatoriedade do texto criado: um valor mais baixo (próximo de 0) torna o output mais determinístico e conservador, limitando as opções a palavras de alta probabilidade, enquanto um valor mais alto (como 1 ou acima) aumenta a criatividade e variedade, permitindo a inclusão de palavras menos prováveis. O parâmetro "top_P," por outro lado, ajusta a probabilidade acumulativa da seleção de palavras. Com $top_P = 0.9$, por exemplo, o modelo considera apenas as palavras cujas probabilidades acumuladas somam 90%, filtrando assim opções menos prováveis e focando nas mais prováveis. Ajustar estes parâmetros permite um controle preciso sobre o equilíbrio entre a criatividade e a exatidão nos outputs gerados, adaptando-os às necessidades específicas de diferentes tarefas e contextos. Ao optar por não utilizar parâmetros específicos como "temperatura" e "top_P", a investigação simula de forma mais autêntica como os profissionais de marketing interagiriam com as IAs generativas em seu cotidiano. Esta decisão metodológica é crucial, pois permite que as IAGs ajustem os seus outputs de forma mais orgânica aos diferentes contextos e públicos-alvo de cada setor, sem as restrições artificiais que poderiam ser impostas por configurações predeterminadas.

Esta liberdade concedida às IAGs é particularmente relevante considerando a diversidade dos setores abordados (Financeiro, Educação, Tecnologia e Saúde) e os seus respectivos públicos-alvo. Por exemplo, permite que a IAG adapte seu tom e estilo de forma mais natural, podendo ser mais ousada e criativa para uma campanha voltada para jovens no setor financeiro, enquanto adota uma abordagem mais conservadora e informativa para uma campanha de pós-graduação no setor de educação. Esta flexibilidade é essencial para avaliar a verdadeira capacidade das IAGs de compreenderem e de se adaptarem a diferentes contextos de marketing.

O processo de formação dos prompts para este estudo foi meticulosamente desenvolvido com o objetivo de ser específico e claro, seguindo as melhores práticas de engenharia de prompts. Esta abordagem é corroborada pelo trabalho de White et al. (2023), que propõe um catálogo de padrões de prompts para melhorar a engenharia de prompts com ChatGPT. A especificidade e clareza nos prompts são cruciais para obter resultados relevantes e de alta qualidade das IAs generativas, especialmente em contextos complexos como o marketing digital. Conforme destacado por Chen et al. (2023) na sua investigação sobre avaliação de qualidade de texto usando grandes modelos de linguagem, prompts bem estruturados e detalhados podem melhorar significativamente o desempenho das IAGs em tarefas específicas.

Nesta investigação, foi utilizado o mesmo prompt para todas as IAGs, assegurando uma base comparativa consistente. Este prompt padrão foi adaptado para cada setor específico, de modo a refletir as particularidades e exigências de cada área de atuação. Dentro de cada setor, foi garantido que o mesmo prompt fosse utilizado para todas as IAGs, permitindo uma avaliação justa e equilibrada do desempenho de cada tecnologia na criação de campanhas publicitárias. Esta abordagem padronizada foi crucial para manter a uniformidade e a comparabilidade dos resultados obtidos ao longo do estudo.

A estrutura dos prompts utilizados neste estudo, que inclui elementos como benefícios do produto, descrição do público-alvo, requisitos específicos dos textos e instruções para interação com imagens, alinha-se com as recomendações de Ding et al. (2023) sobre a avaliação de co-criação de texto humano-IA. Esta abordagem abrangente nos prompts não orienta apenas as IAs a criar conteúdo mais relevante e contextualizado, mas também facilita uma avaliação mais precisa de suas capacidades em tarefas de copywriting para marketing digital. Além disso, a inclusão de instruções específicas sobre o comprimento do texto e a otimização para diferentes plataformas de mídia social reflete as condições reais enfrentadas pelos profissionais de marketing, aproximando o estudo das práticas quotidianas do setor, como sugerido por Osadchaya et al. (2024) na sua análise dos paradoxos da IA generativa na indústria publicitária.

A escolha dos setores Financeiro, Educação, Tecnologia e Saúde para este inquérito comparativo de desempenho de IAs generativas em copywriting é fundamentada em dados significativos sobre investimento em publicidades online e geração de PIB. Estes setores não

representam apenas áreas cruciais da economia moderna, mas também são líderes na adoção de tecnologias digitais e investimento em marketing online.

O setor Financeiro, incluindo fintechs e bancos digitais, têm demonstrado um crescimento expressivo nos últimos anos. Segundo um relatório da eMarketer (2022), o setor financeiro foi responsável por 14,6% do total de gastos com publicidade digital nos Estados Unidos em 2021, ficando atrás apenas do retalho. Este investimento significativo reflete a importância crescente dos serviços financeiros digitais, especialmente para o público jovem, alvo da campanha da "Conta Digital Jovem" no nosso estudo. Além disso, o setor financeiro contribui substancialmente para o PIB global, com o setor de serviços financeiros representando aproximadamente 20-25% do PIB em economias desenvolvidas, de acordo com dados do World Bank (2023).

O setor de Educação, particularmente o ensino superior, a escolha de uma Pós-Graduação em Ciência de Dados como foco reflete a crescente demanda por habilidades em análise de dados, com o Bureau of Labor Statistics (2023) dos EUA projetando um crescimento de 31% nas vagas para cientistas de dados entre 2020 e 2030.

O setor de Tecnologia, representado no nosso estudo pelo celular Z50 Pro, é um dos maiores investidores em publicidade digital. De acordo com um relatório da Statista (2023), as empresas de tecnologia foram responsáveis por aproximadamente 20% dos gastos globais em publicidade digital em 2022. Além disso, o setor de tecnologia é um dos principais motores do crescimento econômico moderno, com empresas de tecnologia representando uma parcela significativa do PIB em muitas economias avançadas. Por exemplo, nos Estados Unidos, o setor de tecnologia contribuiu com aproximadamente 10,5% do PIB em 2022, segundo dados do CompTIA (2021).

O setor de Saúde, incluindo serviços de bem-estar e nutrição como a "Mentoria Equilibrium" em nosso estudo, é um contribuinte substancial para o PIB, representando em média 10% do PIB em países da OCDE (2023), com alguns países a chegar aos 18%.

Estes setores não representam apenas áreas de significativo investimento em marketing digital, mas também são cruciais para a economia global, contribuindo substancialmente para o PIB. Além disso, são setores que estão a passar por rápidas transformações digitais, tornando-os particularmente relevantes para um estudo sobre a eficácia de IAs generativas em copywriting para marketing digital. A diversidade destes setores também permite uma

avaliação mais abrangente das capacidades das IAs em adaptar-se a diferentes contextos e públicos-alvo, fornecendo observações úteis sobre a aplicabilidade dessas tecnologias em várias indústrias.

3.4. Estrutura do inquérito

A estrutura do inquérito foi projetado para recolher dados abrangentes sobre a percepção e eficácia das campanhas criadas por IA, bem como para compreender o perfil dos inquiridos e as suas opiniões sobre o impacto da IA no marketing. O inquérito é organizado em várias seções distintas, começando com perguntas introdutórias que avaliam a familiaridade dos participantes com IA generativa e marketing digital. Essas questões iniciais utilizam uma escala Likert de 5 pontos, variando de "Extremamente familiar" a "Não é familiar de maneira alguma", proporcionando uma base para contextualizar as respostas subsequentes.

A seção principal do inquérito apresenta as campanhas geradas por IA para cada um dos quatro setores: Financeiro, Educação, Tecnologia e Saúde. É importante notar que para cada setor, foram criadas quatro campanhas diferentes, cada uma gerada por uma das quatro IAs selecionadas (Gemini 1.5 Pro, ChatGPT 4, Claude 3 Opus e Llama 3). Isso permite uma comparação direta entre as capacidades de cada IA em diferentes contextos setoriais. Para cada campanha, os participantes avaliam seis parâmetros cruciais: Qualidade, Criatividade, Clareza, Originalidade, Atratividade e Emoção, utilizando uma escala Likert de 7 pontos, que varia de "Muito baixa" a "Muito alta". Esta escala de 7 pontos foi escolhida para oferecer uma granularidade maior nas avaliações, permitindo uma análise mais detalhada das nuances nas percepções dos participantes.

A estrutura do inquérito foi elaborada para fornecer aos participantes um contexto rico e realista antes de apresentar as campanhas geradas pelas IAs. A inclusão de descrições detalhadas dos produtos ou serviços para cada setor antes da apresentação das campanhas é um elemento crucial do design da investigação. Estas introduções fornecem aos participantes as informações necessárias para avaliar as campanhas de forma mais informada e contextualizada, simulando o tipo de briefing que um profissional de marketing receberia antes de avaliar ou criar uma campanha.

Por exemplo, no caso da "Conta Digital Jovem" do setor financeiro, a descrição não apenas define o produto e seu público-alvo, mas também destaca benefícios específicos e sugere uma imagem mental para complementar o texto. Isso permite que os participantes avaliem não apenas a criatividade e a qualidade do texto gerado pela IA, mas também a sua relevância e adequação ao contexto específico do produto e do público-alvo. O mesmo se aplica aos outros setores, onde as descrições fornecem um panorama claro dos benefícios, características e contexto visual das campanhas.

Além das avaliações específicas, o inquérito inclui perguntas sobre o interesse despertado pela campanha, a adequação da linguagem à faixa etária do participante, e a percepção sobre a autoria da campanha (humano vs. IA). Essas questões utilizam escalas Likert de 5 pontos, oferecendo um equilíbrio entre detalhamento e facilidade de resposta. Ao final de cada seção setorial, os participantes são solicitados a escolher a sua campanha preferida entre as quatro apresentadas, fornecendo conclusões importantes sobre as preferências gerais.

A parte final do questionário é dedicada à recolha de informações demográficas e profissionais dos participantes. Esta seção inclui perguntas sobre gênero, idade, nacionalidade, país de residência, escolaridade e rendimentos, adaptadas para contextos brasileiros e portugueses. Além disso, há um conjunto de perguntas específicas para profissionais de marketing, explorando as suas opiniões sobre o impacto da IA no campo. Estas questões abrangem temas como a percepção da qualidade do conteúdo gerado por IA, funções de marketing em risco, ferramentas de IA utilizadas, e a importância percebida da IA para atividades profissionais.

As perguntas relacionadas ao marketing são particularmente importantes para visualizar o perfil do profissional e as suas perspectivas sobre a integração da IA no campo. Elas exploram preocupações sobre o impacto da IA, potenciais benefícios da IA generativa no marketing, e áreas onde a IA pode ser mais útil. Essas questões utilizam uma combinação de escalas Likert e opções de múltipla escolha, permitindo uma análise detalhada das atitudes e expectativas dos profissionais de marketing em relação à IA. Esta estrutura abrangente do inquérito permite uma análise profunda não apenas da eficácia das campanhas geradas por IA, mas também do contexto mais amplo de como a IA está a ser percebida e integrada no campo do marketing digital.

Esta abordagem na estrutura do inquérito - combinando prompts não restritivos para as IAs com contextualizações ricas para os participantes - cria um ambiente de avaliação que se aproxima muito das condições reais de trabalho em marketing digital. Isso aumenta significativamente a validade do estudo, permitindo observações mais relevantes e aplicáveis sobre o desempenho das IAs generativas em cenários de marketing do mundo real.

Além disso, a inclusão de perguntas específicas sobre a percepção dos participantes quanto à autoria das campanhas (humano vs. IA) e as suas opiniões sobre o impacto da IA no campo do marketing adiciona uma camada valiosa de análise. Essas questões não avaliam apenas a qualidade percebida do conteúdo gerado por IA, mas também exploram as atitudes e expectativas dos profissionais de marketing em relação à integração da IA nas suas práticas de trabalho. Isso fornece um contexto mais amplo para interpretar os resultados, permitindo uma compreensão mais abrangente do papel atual e futuro da IA generativa no marketing digital.

3.5. Parâmetros de Avaliação

Os parâmetros escolhidos para avaliar a eficácia do copywriting são Qualidade, Criatividade, Clareza, Originalidade, Atratividade e Emoção. A Qualidade engloba aspectos como fluência linguística, correção gramatical e precisão das informações, essenciais para estabelecer credibilidade. Criatividade e Originalidade são cruciais para destacar o texto em meio à saturação de conteúdo, enquanto a Clareza assegura que a mensagem seja compreendida facilmente. A Atratividade refere-se ao apelo visual e conceitual do texto, e a Emoção é fundamental para criar uma conexão com o público-alvo, influenciando diretamente o engajamento e a ação desejada.

3.6. Relação com a eficácia do copywriting

Estes parâmetros relacionam-se diretamente com a eficácia do copywriting ao abordar diferentes aspectos que influenciam a percepção e resposta do público. A Qualidade e Clareza formam a base para uma comunicação efetiva, garantindo que a mensagem seja transmitida corretamente. Criatividade e Originalidade capturam a atenção e diferenciam o texto da concorrência. A Atratividade aumenta o engajamento inicial, enquanto a Emoção fortalece a conexão com o leitor, aumentando a probabilidade de conversão. Wang et al. (2021) também

se serviu dessa abordagem, destacando a importância da qualidade e atratividade na avaliação de anúncios de texto.

3.7. Influência da emoção

Diversos estudos científicos reforçam a importância da emoção na eficácia do copywriting. Por exemplo, Zulkifly & Firdaus (2014) e Akgun et al. (2017) demonstraram que conteúdos emocionais têm maior impacto na tomada de decisão do consumidor. Chang et al. (2015) e Jang et al. (2013) evidenciaram que a emoção positiva em anúncios aumenta a intenção de compra. Wang et al. (2021) e Ding et al. (2023) exploraram como diferentes tipos de apelos emocionais afetam o engajamento em mídias sociais. Hossain et al. (2021) destacou a importância da emoção na criação de conteúdo viral. Esses estudos reforçam que a incorporação efetiva de elementos emocionais no copywriting pode melhorar significativamente o seu impacto e eficácia.

3.8. Recolha de dados

O inquérito obteve um total de 251 respostas, das quais 82 foram consideradas completas e válidas para análise. O público-alvo consistiu em falantes nativos de língua portuguesa, predominantemente brasileiros e portugueses, todos maiores de 18 anos, que consentiram voluntariamente participar no estudo. A recolha de dados foi realizada através da plataforma Qualtrics. Para garantir uma amostra diversificada, o inquérito foi amplamente divulgado em diversos grupos de redes sociais, partilhado por via de aplicativos de mensagens instantâneas e distribuído por e-mail para a comunidade universitária.

4. RESULTADOS E DISCUSSÃO

H1: Haverá diferenças significativas na avaliação geral entre as quatro IAs usadas para gerar as campanhas.

Para testar esta hipótese H1, foi escolhida uma Análise de Variância (ANOVA) de medidas repetidas, considerando que o estudo envolveu múltiplas comparações entre as IAs em diferentes setores (Financeiro, Educação, Tecnologia e Saúde), com os mesmos participantes avaliando todas as condições. O teste de esfericidade de Mauchly indicou uma violação da suposição de esfericidade ($p < 0.05$), necessitando uma correção nos graus de liberdade. A correção de Greenhouse-Geisser foi aplicada a todos os efeitos para garantir uma abordagem conservadora e consistente na análise. Esta correção ajusta os graus de liberdade para produzir um valor F mais preciso, reduzindo o risco de erro Tipo I.

Os resultados da ANOVA revelaram um efeito principal significativo para o tipo de IA utilizada [$F(2.961, 239.880) = 5.687, p < .001, \eta^2 = 0.015$]. Este resultado suporta diretamente a H1, indicando que existem, de fato, diferenças significativas na avaliação geral entre as quatro IAs. O tamanho do efeito ($\eta^2 = 0.015$), embora pequeno, sugere que estas diferenças são detectáveis e relevantes para a análise. Além disso, a análise revelou uma interação significativa entre o tipo de IA e os setores [$F(7.526, 609.623) = 5.920, p < .001, \eta^2 = 0.038$]. Este resultado acrescenta uma nuance importante à H1, indicando que as diferenças entre as IAs não são uniformes em todos os setores, mas variam dependendo do contexto específico da campanha.

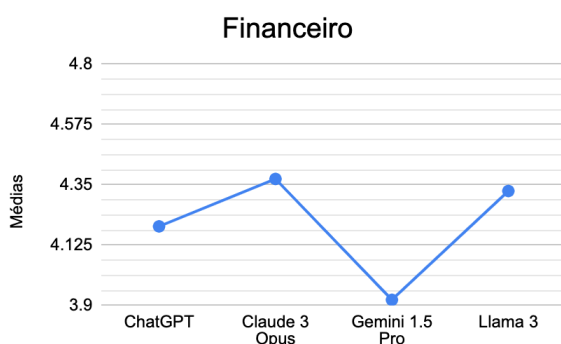


Figura 4.1 - Comparação IAGs no setor Financeiro

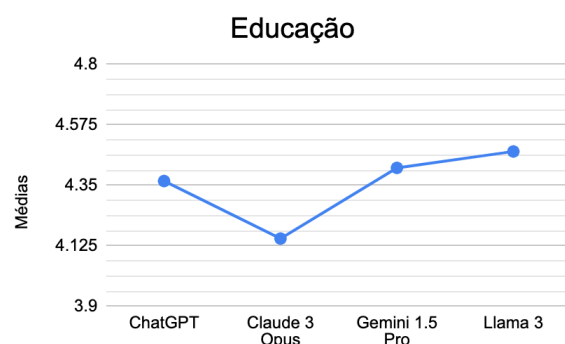


Figura 4.2 - Comparação IAGs no setor Educação

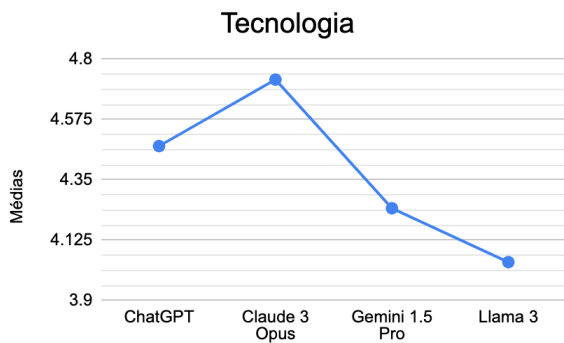


Figura 4.3 - Comparação IAGs no setor Tecnologia

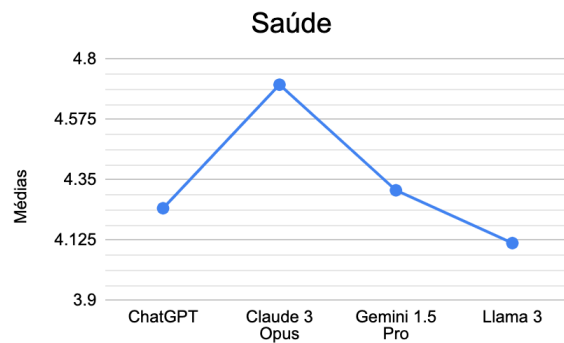


Figura 4.4 - Comparação IAGs no setor Saúde

A análise post-hoc forneceu detalhes adicionais sobre essas diferenças. No setor Financeiro, Claude 3 Opus demonstrou uma tendência de melhor desempenho em comparação com as outras IAs, embora as diferenças não tenham atingido o nível convencional de significância estatística ($p < 0.05$). Embora Claude 3 Opus tenha apresentado no setor Financeiro médias mais altas em comparação com Llama 3 (diferença média = 0.045, $p = 1.000$), ChatGPT (diferença de 0.177, $p = 1.000$) e Gemini 1.5 Pro (diferença de 0.451, $p = 0.087$) as diferenças em relação às outras IAs não foram estatisticamente significativas ao nível convencional de $p < 0.05$. Isso sugere que, embora haja uma tendência de melhor desempenho do Claude 3 Opus neste setor, a variabilidade nas avaliações e o tamanho da amostra não permitem afirmar com certeza estatística que seu desempenho foi superior.

No setor de Tecnologia, o Claude 3 Opus superou o ChatGPT (diferença média = 0.248, $p = 1.000$), mas essa diferença não foi estatisticamente significativa. Entretanto, o Claude 3 Opus superou o Gemini 1.5 Pro (diferença média = 0.480, $p = 0.041$) e o Llama 3 (diferença média = 0.681, $p < 0.001$), mostrando haver significância estatística nas duas últimas comparações.

No setor de Educação, os resultados da análise post-hoc indicaram que não houve diferenças estatisticamente significativas no desempenho das IAs avaliadas. As comparações entre ChatGPT e Claude 3 Opus (diferença média = 0.213, $p = 0.971$), ChatGPT e Gemini 1.5 Pro (diferença média = -0.049, $p = 1.000$), e ChatGPT e Llama 3 (diferença média = -0.110, $p = 1.000$) não revelaram significância estatística. Da mesma forma, as comparações entre Claude 3 Opus e Gemini 1.5 Pro (diferença média = -0.262, $p = 0.853$), Claude 3 Opus e Llama 3 (diferença média = -0.323, $p = 0.548$), e Gemini 1.5 Pro e Llama 3 (diferença média = -0.061, $p = 1.000$) não mostraram diferenças significativas. Esses resultados sugerem que, no

contexto educacional, as IAs analisadas apresentam desempenhos equivalentes, sem evidências de superioridade estatística de uma sobre as outras.

No setor de Saúde, o ChatGPT foi superado pelo Claude 3 Opus (diferença média = -0.461, $p = 0.067$), mas essa diferença não foi estatisticamente significativa. O Claude 3 Opus também superou o Gemini 1.5 Pro (diferença média = 0.394, $p = 0.338$), mas essa diferença não foi estatisticamente significativa. No entanto, o Claude 3 Opus superou o Llama 3 (diferença média = 0.591, $p = 0.001$), mostrando haver significância estatística nesta última comparação.

As variações de desempenho entre setores ressaltam a complexidade das diferenças entre as IAs e a importância do contexto setorial na avaliação de seu desempenho. Em conclusão, a hipótese H1, que postulava diferenças significativas no desempenho das IAs na criação de campanhas, foi parcialmente suportada pelos resultados da ANOVA e da análise post-hoc. Embora a ANOVA tenha revelado um efeito principal significativo e uma interação significativa entre o tipo de IA e os setores, indicando variações no desempenho das IAs, muitas das comparações específicas entre pares de IAs dentro dos setores não apresentaram significância estatística. No setor Financeiro, Claude 3 Opus demonstrou uma tendência de melhor desempenho em comparação com as outras IAs, embora sem significância estatística. Nos setores de Tecnologia e Saúde, Claude 3 Opus superou significativamente algumas IAs, especialmente em comparação com Llama 3. No setor de Educação, não houve diferenças estatisticamente significativas entre as IAs analisadas. Estes resultados têm implicações importantes para a seleção e aplicação de IAs em diferentes contextos setoriais, sugerindo que a escolha da IA mais apropriada pode depender significativamente do setor específico para o qual o conteúdo está sendo gerado.

H2: Profissionais de marketing e não-profissionais terão avaliações significativamente diferentes sobre qual IA produziu as melhores campanhas.

A hipótese sugere que a experiência e o conhecimento específico dos profissionais de marketing poderiam levar a uma percepção distinta das campanhas geradas por IAG em comparação com pessoas sem essa formação especializada.

As variáveis utilizadas neste estudo (por exemplo, Fin-ChatGPT, Tec-Claude, Edu-Llama, Sau-Gemini) representam as médias das avaliações de campanhas específicas para cada combinação de setor (Financeiro, Tecnologia, Educação, Saúde) e IAG (ChatGPT, Claude 3 Opus, Llama 3, Gemini).

Para criar as variáveis, calculou-se a média das avaliações dos parâmetros Qualidade, Criatividade, Clareza, Originalidade, Atratividade e Emoção para cada combinação de setor e IAG. Assim, por exemplo, Fin-ChatGPT representa a média das avaliações da campanha gerada pelo ChatGPT para o setor Financeiro, enquanto Tec-Claude representa a média das avaliações da campanha gerada pelo Claude 3 Opus para o setor de Tecnologia, e assim por diante. Isso fornece uma medida composta da percepção geral de cada campanha. Para testar esta hipótese H2, foi realizado um teste t-student de amostras independentes para cada campanha (representada pelas variáveis criadas), comparando as avaliações dos dois grupos. O teste t-student analisa se há diferenças estatisticamente significativas entre as médias dos dois grupos. É importante notar que, embora o teste de Igualdade de Variâncias (Levene's) para o Sau-Gemini tenha resultado em $p = 0.049$, que é ligeiramente abaixo do limiar convencional de 0.05, optou-se por utilizar o teste t-student para esta variável assim como para todos os outros. Esta decisão foi tomada considerando que o valor está muito próximo de 0.05 e para manter a consistência na análise com as demais variáveis. A maioria dos p-valores são maiores que 0.05, indicando que não há diferenças estatisticamente significativas entre os grupos para a maioria das campanhas. A única exceção é o Tec-Claude com $p = 0.021 < 0.05$. O d de Cohen, que mede o tamanho do efeito, é geralmente pequeno para a maioria das campanhas (< 0.5). Novamente, o Tec-Claude se destaca com um efeito médio ($d = 0.690$). Observando as médias, nota-se que as médias foram geralmente menores para o grupo de profissionais do marketing (grupo 1) do que para não profissionais (grupo 2), com exceção das médias de Edu-Claude e Sau-Llama. Esta tendência sugere que os profissionais de marketing tenderam a avaliar as campanhas de forma mais positiva na maioria dos casos. Isso pode indicar que sua experiência e conhecimento específico os levaram a apreciar mais os aspectos técnicos e criativos das campanhas geradas por IAG, possivelmente reconhecendo nuances e qualidades que podem não ser imediatamente evidentes para não profissionais.

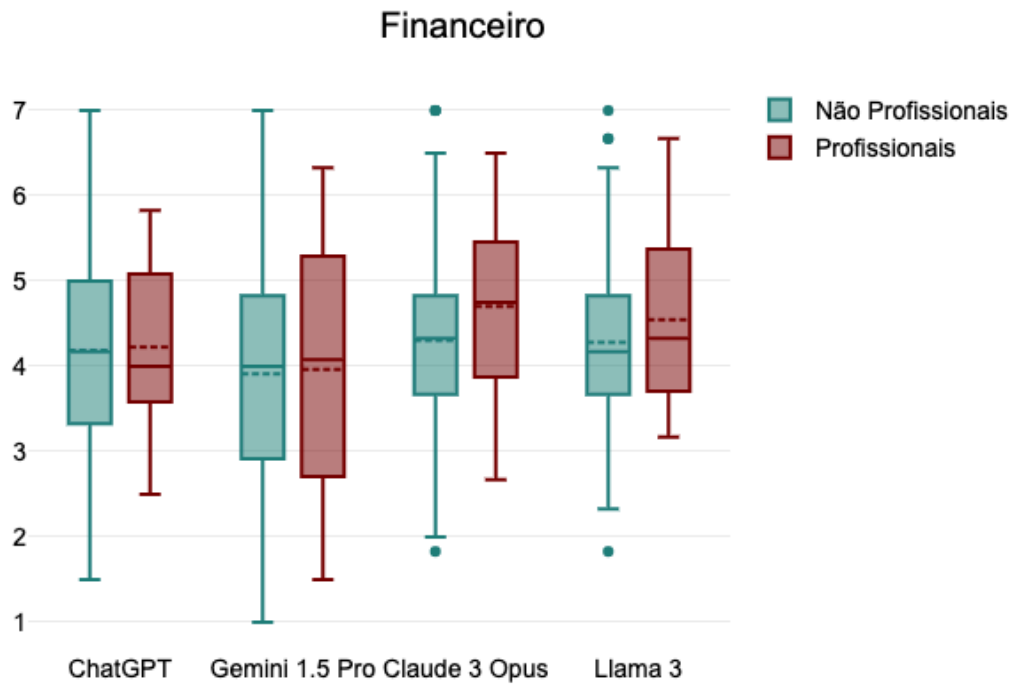


Figura 4.5 - Comparação Profissionais e Não Profissionais do marketing no setor Financeiro

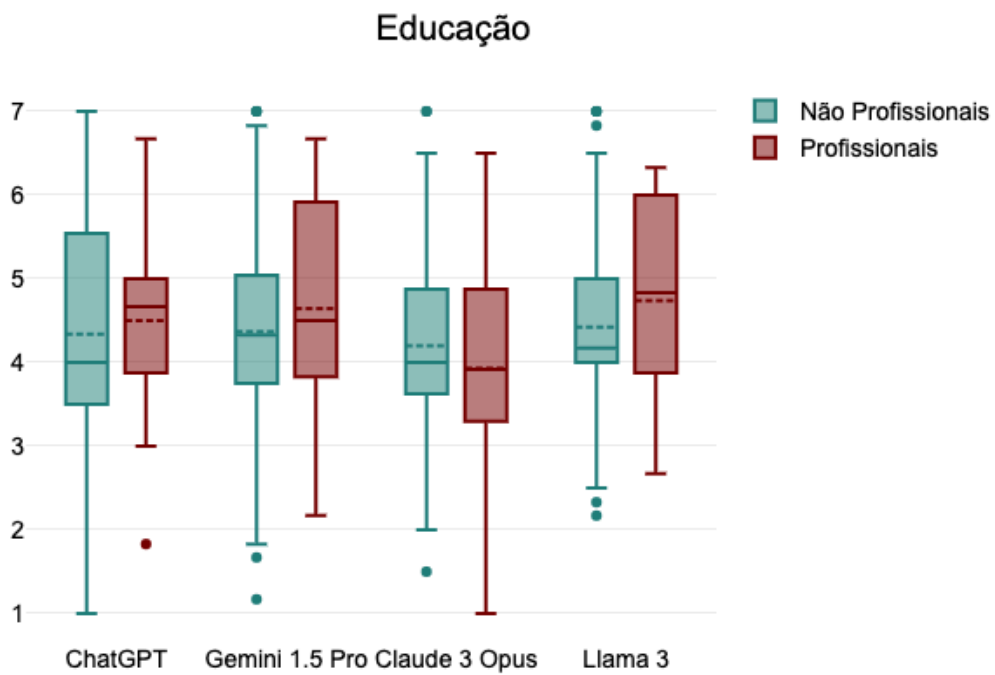


Figura 4.6 - Comparação Profissionais e Não Profissionais do marketing no setor Educação

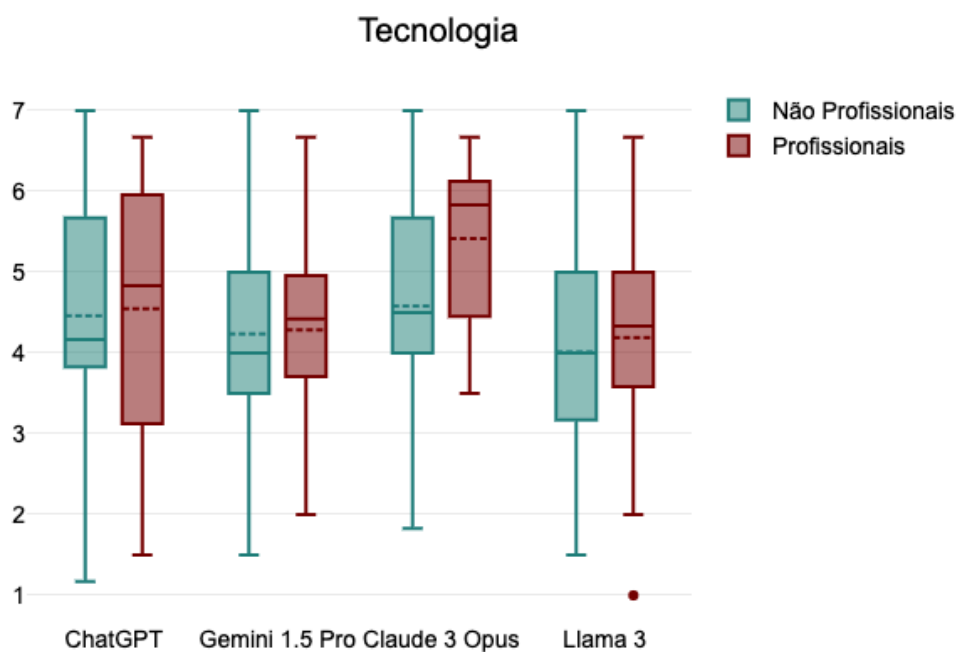


Figura 4.7 - Comparação Profissionais e Não Profissionais do marketing no setor Tecnologia

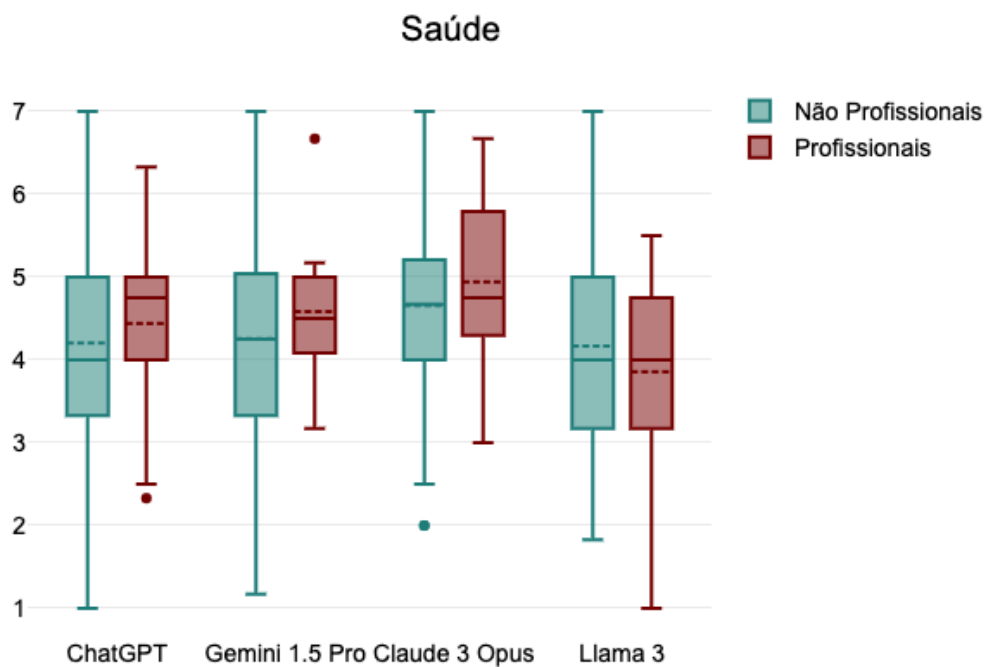


Figura 4.8 - Comparação Profissionais e Não Profissionais do marketing no setor Saúde

Com base nesses resultados, a hipótese H2 não é fortemente suportada pelos dados. Na maioria dos casos, não há diferenças significativas entre as avaliações de profissionais de marketing e não-profissionais, apesar da tendência de avaliações ligeiramente mais altas pelos profissionais de marketing. Sendo a mais notável a campanha de Tec-Claude, onde os profissionais tenderam a avaliar bem mais positivamente que os não-profissionais de marketing. Isso sugere que, em geral, a percepção da qualidade das campanhas geradas por IAG é relativamente consistente entre profissionais de marketing e não-profissionais. As diferenças de formação e experiência profissional não parecem influenciar substancialmente as avaliações na maioria dos casos, com os profissionais tendendo a avaliá-las ligeiramente de forma mais positiva. As pequenas variações observadas nas médias podem refletir nuances na apreciação das campanhas, possivelmente devido à experiência profissional, mas não são suficientes para indicar uma diferença sistemática na percepção entre os dois grupos. A exceção do Tec-Claude merece atenção especial, pois indica que para esta campanha específica, a expertise em marketing pode ter levado a uma avaliação significativamente diferente. Isso pode sugerir que em certos contextos ou para certos tipos de campanhas, a formação profissional pode influenciar a percepção da qualidade do conteúdo gerado por IA. No entanto, é importante notar que este estudo tem limitações, como o tamanho desigual dos grupos (14 profissionais vs. 68 não-profissionais), o que pode afetar a robustez dos resultados.

H3: As campanhas criadas pela IA com as maiores pontuações gerais serão mais frequentemente percebidas como criadas por humanos.

A hipótese H3 propõe que as campanhas criadas por IA com as maiores pontuações gerais serão mais frequentemente percebidas como criadas por humanos. Para analisar esta hipótese, vamos examinar as correlações entre as variáveis já mencionadas (Fin-ChatGPT, Tec-Claude, Edu-Llama, Sau-Gemini) e as percepções de quem as criou (representadas pelos P-Fin-ChatGPT, P-Tec-Claude, P-Edu-Llama, P-Sau-Gemini).

Todas as correlações são negativas, o que significa que quanto maior a pontuação geral da campanha, menor o valor na escala de percepção de criação (lembrando que 1 representa "definitivamente criado por humano" e 5 "definitivamente criado por IA").

A maioria das correlações é estatisticamente significativa ($p < 0.05$), com muitas sendo altamente significativas ($p < 0.001$ ou $p < 0.01$).

Correlações fortes:

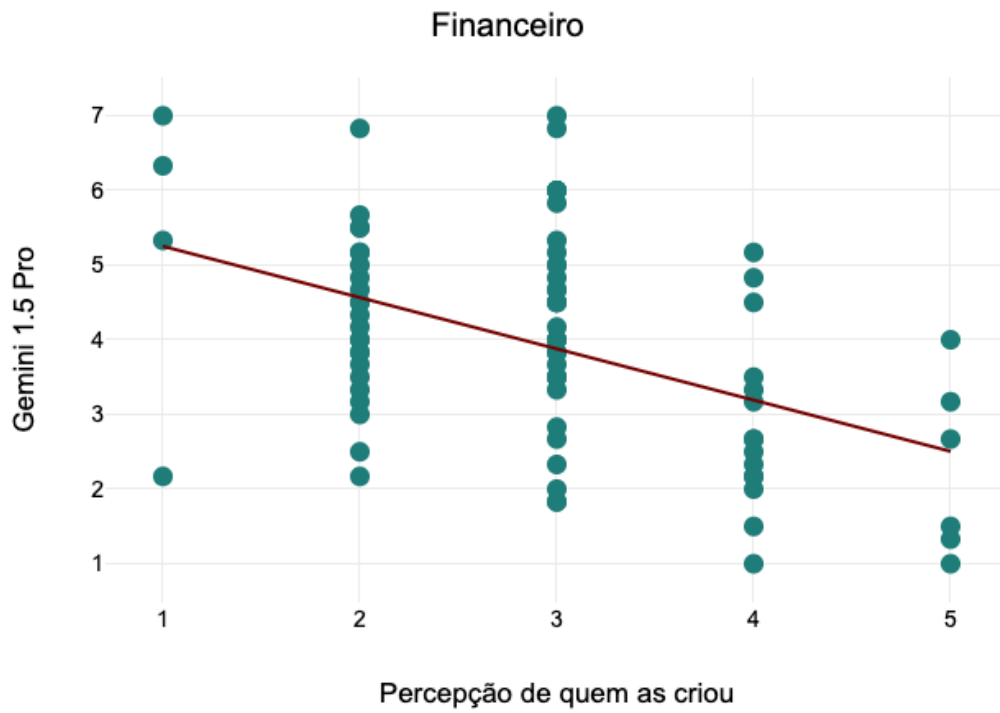


Figura 4.9 - Correlação entre Gemini 1.5 Pro e a percepção de que as criou no setor Financeiro

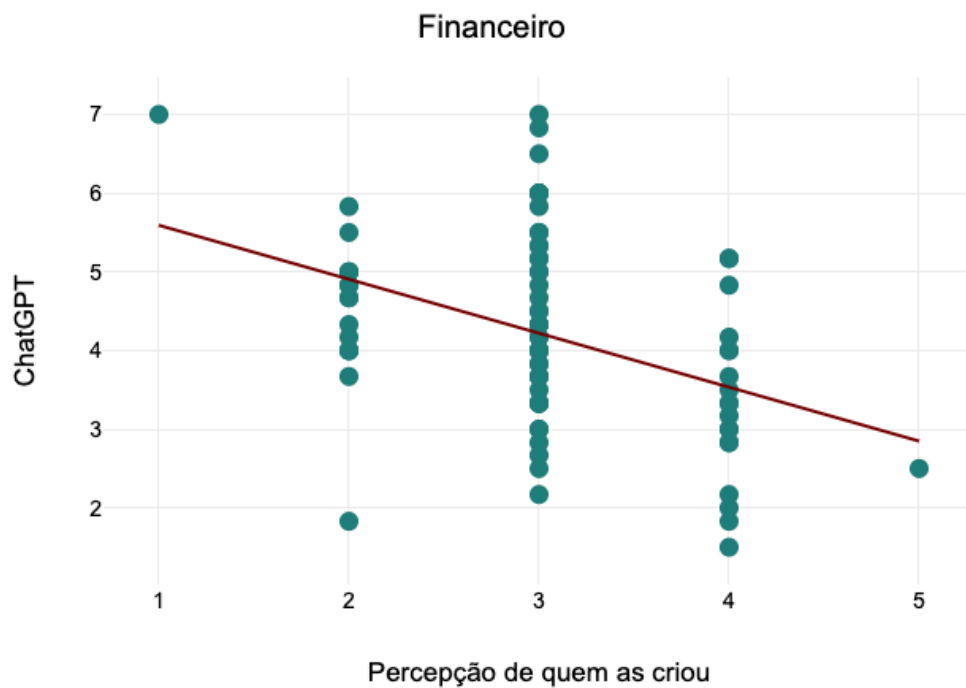


Figura 4.10 - Correlação entre ChatGPT e a percepção de que as criou no setor Financeiro

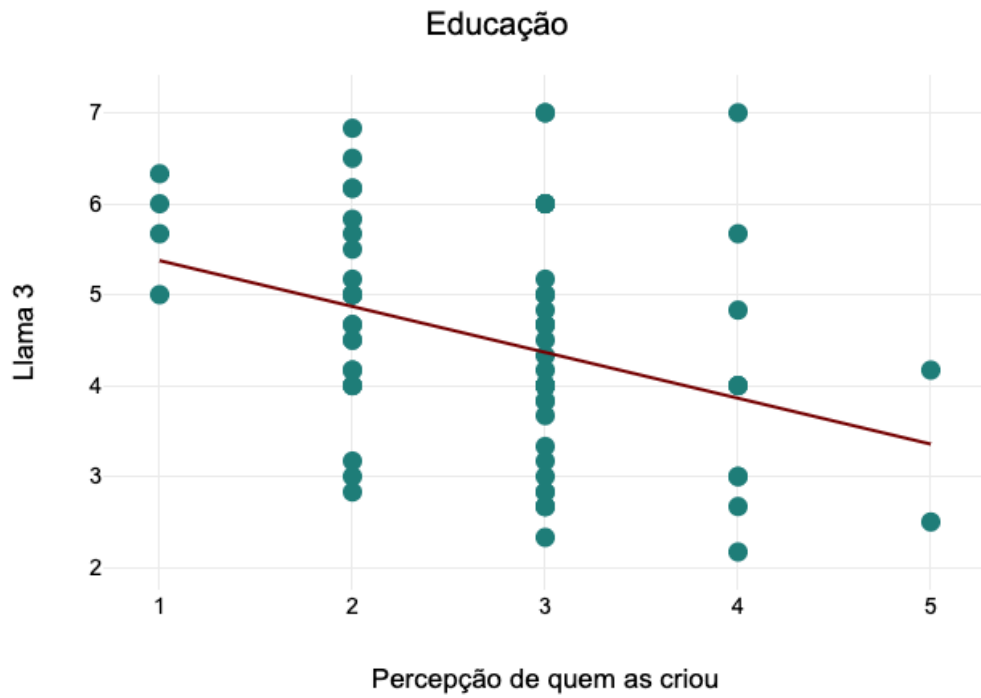


Figura 4.11 - Correlação entre Llama 3 e a percepção de que as criou no setor Educação

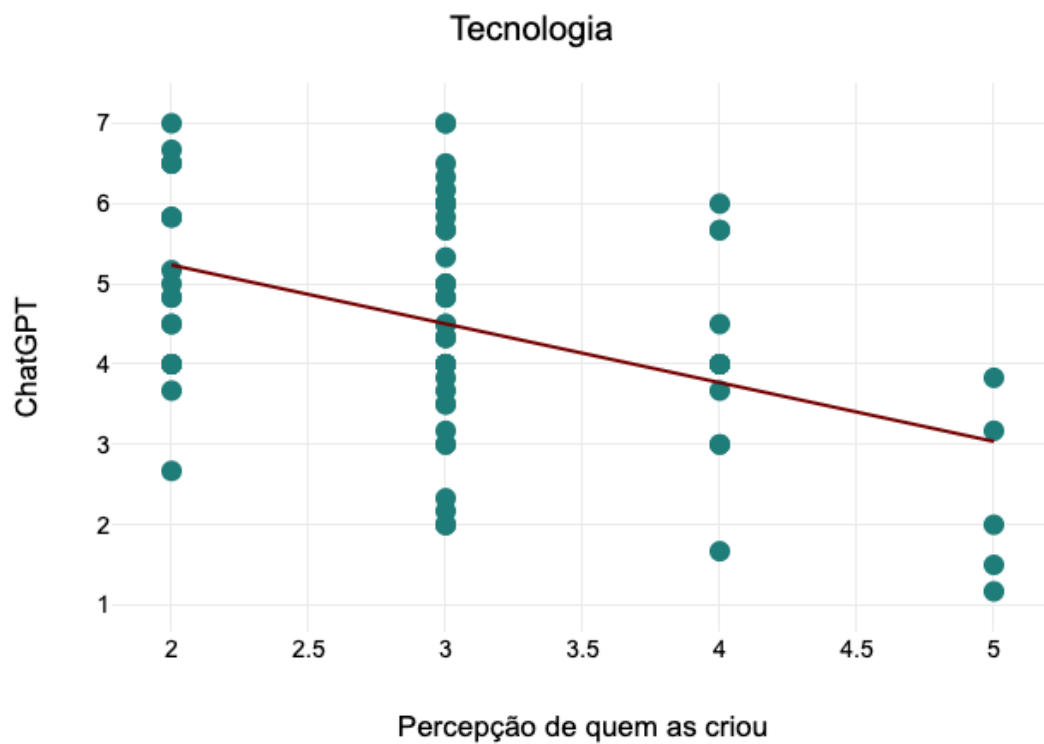


Figura 4.12 - Correlação entre ChatGPT e a percepção de que as criou no setor Tecnologia

Correlações fracas:

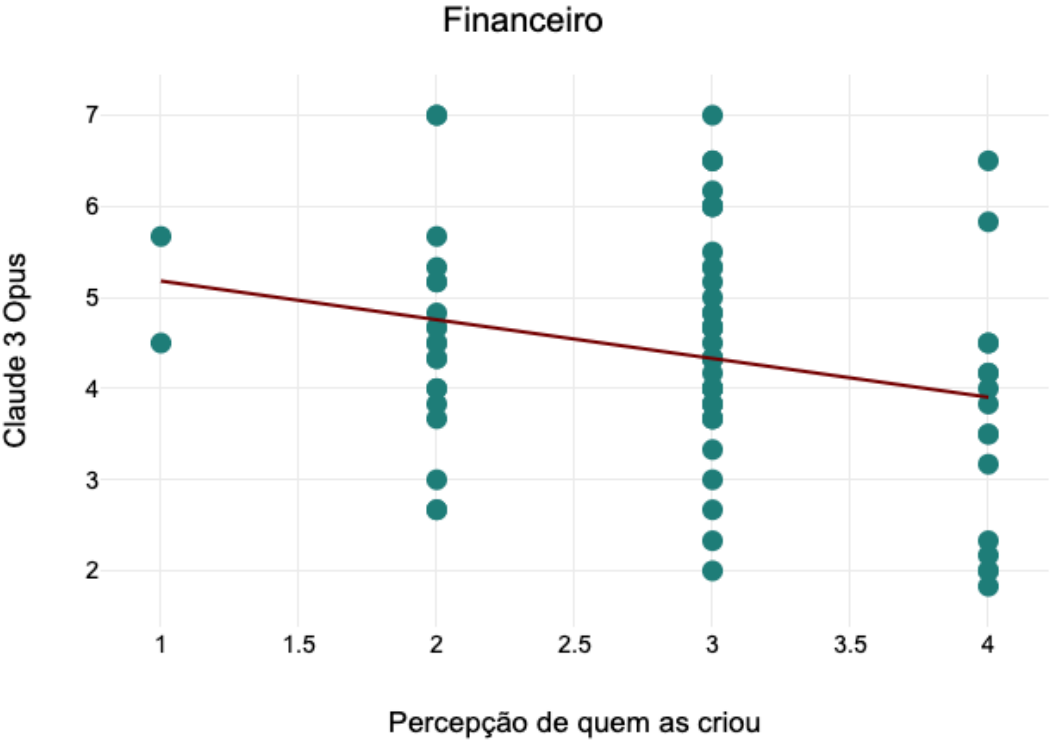


Figura 4.13 - Correlação entre Claude 3 Opus e a percepção de que as criou no setor Financeiro

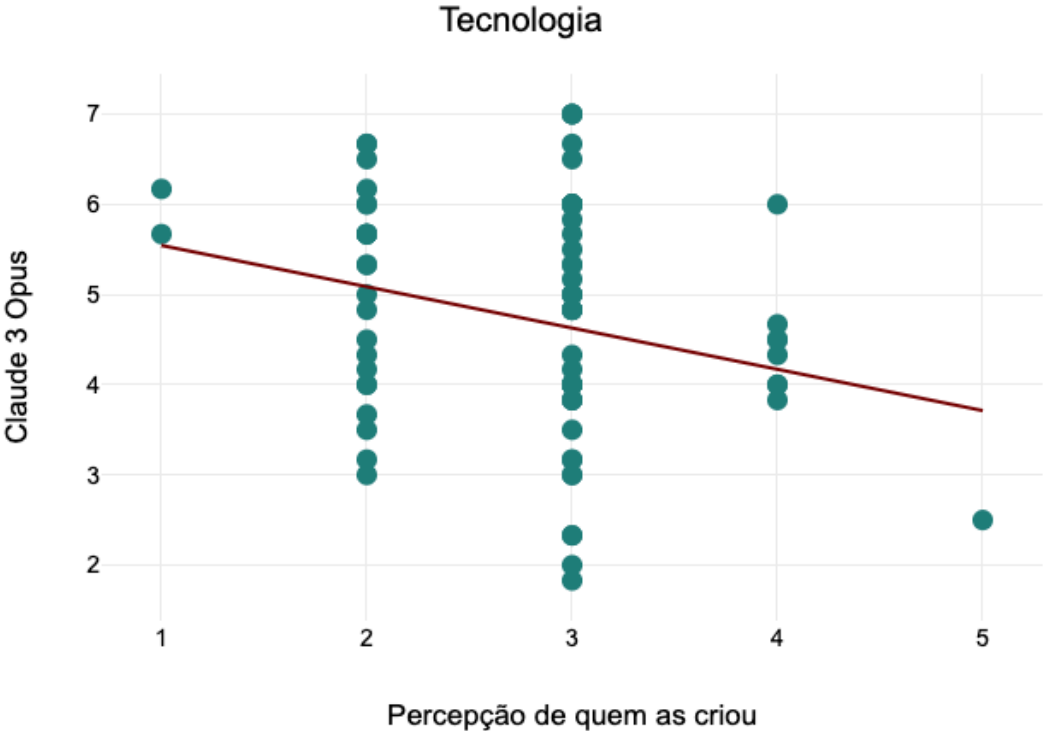


Figura 4.14 - Correlação entre Claude 3 Opus e a percepção de que as criou no setor Tecnologia

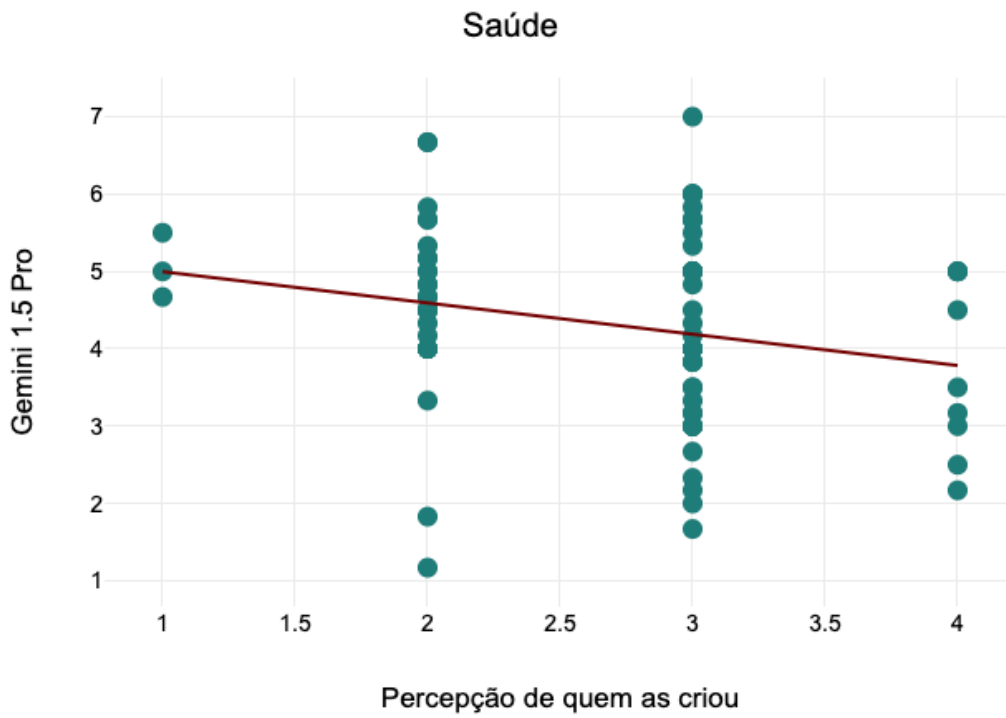


Figura 4.15 - Correlação entre Gemini 1.5 Pro e a percepção de que as criou no setor Saúde

Correlação não significativa:

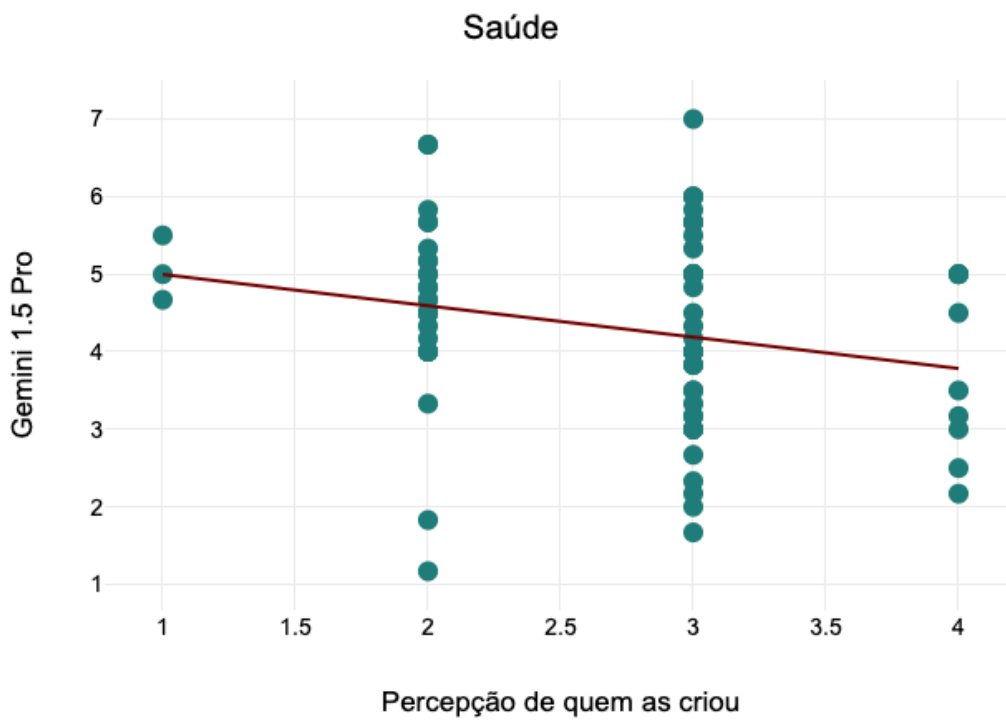


Figura 4.16 - Correlação entre Gemini 1.5 Pro e a percepção de que as criou no setor Saúde

As correlações variam em força, mas a maioria está na faixa de fraca a moderada. As correlações mais fortes incluem: Gemini 1.5 Pro no setor Financeiro ($r = -0.465$, $p < 0.001$), ChatGPT 4 no setor Financeiro ($r = -0.405$, $p < 0.001$), Claude 3 Opus no setor de Educação ($r = -0.401$, $p < 0.001$) e ChatGPT 4 no setor de Tecnologia ($r = -0.415$, $p < 0.001$). As correlações mais fracas foram observadas para Claude 3 Opus no setor Financeiro ($r = -0.256$, $p = 0.020$), Claude 3 Opus no setor de Tecnologia ($r = -0.256$, $p = 0.020$) e Gemini 1.5 Pro no setor de Saúde ($r = -0.238$, $p = 0.031$). No setor de Saúde, a correlação para ChatGPT 4 não foi significativa ($r = -0.186$, $p = 0.094$).

Os testes de normalidade (Shapiro-Wilk) indicam que a maioria dos dados segue uma distribuição normal, o que fortalece a confiabilidade das correlações de Pearson utilizadas.

Com base nesses resultados, podemos concluir que a hipótese H3 é suportada pelos dados. As correlações negativas indicam que, em geral, quanto maior a pontuação geral de uma campanha, mais provável é que ela seja percebida como criada por humanos. No entanto, é importante notar que a força dessa relação varia entre as diferentes IAGs e setores. Algumas campanhas, como as do Gemini 1.5 Pro e ChatGPT 4 no setor financeiro, mostram uma relação mais forte entre a qualidade percebida e a atribuição à criação humana. Por outro lado, algumas campanhas, como as do Claude 3 Opus nos setores financeiro e de tecnologia, mostram uma relação mais fraca. A única exceção notável é a campanha do ChatGPT 4 no setor de saúde, onde não foi encontrada uma correlação significativa. Isso pode sugerir que, neste caso específico, outros fatores além da qualidade geral da campanha influenciaram a percepção de quem a criou. Em resumo, a hipótese H3 é geralmente suportada, indicando que campanhas de maior qualidade tendem a ser percebidas como criadas por humanos. No entanto, a força dessa relação varia, sugerindo que outros fatores, como o setor da campanha ou características específicas das IAs, também podem influenciar essa percepção. Como visto por Ford (2023), as campanhas geradas por IAG foram percebidas relativamente com menos entusiasmo, o que pode ser visto pela hipótese H3 que quanto menor a nota dada para uma campanha mais era percebida como criada pela IAG.

Durante a condução desta investigação, diversos desafios significativos foram enfrentados, o que impactou os resultados e as conclusões finais. Primeiramente, a quantidade limitada de dados recolhidos representou uma limitação crucial. A amostra restrita não permitiu uma análise robusta e abrangente, comprometendo a generalização dos resultados. Além disso, a

participação escassa de profissionais do marketing limitou a diversidade e a profundidade das percepções obtidas, uma vez que esses especialistas possuem informações relevantes sobre a eficácia das campanhas publicitárias.

Outro desafio substancial foi a complexidade e a extensão do inquérito aplicado. O inquérito longo e detalhado tornou-se um obstáculo para os respondentes, resultando em desistências e respostas incompletas, onde somente 32,67% concluíram o inquérito. Um inquérito mais simples e objetivo poderia ter facilitado a participação, garantindo dados mais completos e representativos. Para futuros estudos, seria benéfico focar em apenas um ou dois setores específicos e comparar um número reduzido de IAGs, o que simplificaria o processo de recolha de dados e potencialmente aumentaria a taxa de resposta.

5. CONCLUSÕES E INVESTIGAÇÃO FUTURA

Os resultados sugerem haver diferenças na percepção da qualidade do conteúdo gerado pelas diferentes IAGs, porém não foram significantes estatisticamente, sendo assim suportando parcialmente a primeira hipótese (H1). Essas diferenças variam dependendo do setor analisado, evidenciando a importância do contexto na avaliação da qualidade da IAGs. A segunda hipótese (H2), que postulava diferenças significativas entre as avaliações de profissionais de marketing e não profissionais, não foi fortemente corroborada pelos dados. As percepções entre os grupos foram similares na maioria dos casos, sugerindo que a qualidade do conteúdo gerado por IAGs é percebida de forma relativamente uniforme por diferentes públicos. A terceira hipótese (H3) foi a que encontrou maior respaldo nos dados. Constatou-se uma forte correlação entre a qualidade percebida das campanhas e a atribuição de autoria humana, indicando que as IAs já são capazes de gerar conteúdo persuasivo que se assemelha à escrita humana.

Para investigações futuras, recomenda-se aprimorar a metodologia, buscando uma amostragem de dados mais robusta e representativa, com maior participação de profissionais de marketing. Simplificar o questionário, diminuindo o número de perguntas e tornando-o mais objetivo, pode aumentar a taxa de resposta e a qualidade dos dados.

Além disso, seria interessante realizar testes A/B com campanhas reais, utilizando textos criados por IAGs e por humanos, para comparar diretamente os seus desempenhos em termos de engajamento e conversão. Analisar o impacto de parâmetros específicos de cada IAG, como "temperatura" e "top_P", na percepção do conteúdo gerado também seria um caminho promissor para investigações futuras.

Portanto, a investigação demonstra o potencial das IAGs como ferramentas valiosas para o marketing digital, mas ressalta a importância de uma análise crítica e contextualizada dos seus resultados. A investigação futura, com metodologias aprimoradas e focadas em dados reais, será crucial para aprofundar a compreensão do impacto real da IAG nessa área e para o desenvolvimento de estratégias ainda mais eficazes.

REFERÊNCIAS BIBLIOGRÁFICAS

Akgun, A., Ayar, H., Etliloglu, T., & Keskin, H. (2017). Persuasive messages and emotional responses in social media marketing. *Pressacademia*, 4(3), 202-208.

Aydın, Ö., & Karaarslan, E. (2023). Is ChatGPT leading generative AI? What is beyond expectations?. *Academic Platform Journal of Engineering and Smart Systems*, 11(3), 118-134.

Berger, J. (2013). *Contagious: Why things catch on*. Simon and Schuster.

Berger, J., Moe, W. W., & Schweidel, D. A. (2023). What holds attention? Linguistic drivers of engagement. *Journal of Marketing*, 87(5), 793-809.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. *Advances in neural information processing systems*, 33, 1877-1901.

Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., ... & Zhang, Y. (2023). Sparks of artificial general intelligence: Early experiments with GPT-4. arXiv preprint arXiv:2303.12712.

Bureau of Labor Statistics. (2021). *Occupational Outlook Handbook: Data Scientists*. U.S. Department of Labor. <https://www.bls.gov/ooh/math/data-scientists.htm> (acessado em 15 de julho de 2024).

Chang, Y. T., Yu, H., & Lu, H. P. (2015). Persuasive messages, popularity cohesion, and message diffusion in social media marketing. *Journal of Business Research*, 68(4), 777-782.

Chen, Y., Wang, R., Jiang, H., Shi, S., & Xu, R. (2023). Exploring the use of large language models for reference-free text quality evaluation: An empirical study. arXiv preprint arXiv:2304.00723.

Cialdini, R. B., (2007). *Influence: The psychology of persuasion* (Vol. 55, p. 339). New York: Collins.

CompTIA. (2023). *State of the Tech Workforce*. <https://www.cyberstates.org/> (acessado em 15 de julho de 2024).

Counterpoint. (2023). Premium market captures half of global smartphone revenue in 2022 for the first time. <https://www.counterpointresearch.com/insights/premium-market-captures-half-global-smartphone-revenue-2022-first-time/> (acessado em 15 de julho de 2024).

Ding, Z., Smith-Renner, A., Zhang, W., Tetreault, J. R., & Jaimes, A. (2023). Harnessing the power of LLMs: Evaluating human-AI text co-creation through the lens of news headline generation. In Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (pp. 3321-3339).

Ekin, S. (2023). Prompt engineering for ChatGPT: a quick guide to techniques, tips, and best practices. Authorea Preprints.

eMarketer. (2022). US Digital Ad Spending by Industry 2021. <https://www.emarketer.com/content/us-digital-ad-spending-by-industry-2021> (acessado em 15 de julho de 2024).

Ford, J., Jain, V., Wadhvani, K., & Gupta, D. G. (2023). AI advertising: An overview and guidelines. *Journal of Business Research*, 166.

Gao, B., Wang, Y., Xie, H., & Hu, Y. (2023). Artificial intelligence in advertising: Advancements, challenges, and ethical considerations in targeting, personalization, content creation, and ad optimization. *SAGE Open*, 13(4), 1-15.

Giray, L. (2023). Prompt engineering with ChatGPT: a guide for academic writers. *Annals of biomedical engineering*, 51(12), 2629-2633.

Global Wellness Institute. (2023). *GWI-WE-Monitor-2023*. https://globalwellnessinstitute.org/wp-content/uploads/2023/11/GWI-WE-Monitor-2023_FIN_AL.pdf (acessado em 15 de julho de 2024).

Gołąb-Andrzejak, E. (2023). The impact of generative AI and ChatGPT on creating digital advertising campaigns. *Cybernetics and Systems*.

Graduate Management Admission Council. (2021). Application trends survey. <https://www.gmac.com/market-intelligence-and-research/market-research/application-trends-survey> (acessado em 15 de julho de 2024).

- Haleem, A., Javaid, M., Qadri, M. A., Singh, R. P., & Suman, R. (2022). Artificial intelligence (AI) applications for marketing: A literature-based study. *International Journal of Intelligent Networks*, 3, 119-132.
- Hartmann, J., Exner, Y., & Domdey, S. (2023). The power of generative marketing: Can generative AI reach human-level visual marketing content? Available at SSRN: <https://ssrn.com/abstract=4597899>
- Hochmair, H. H., Juhász, L., & Kemp, T. (2024). Correctness comparison of ChatGPT-4, Gemini, Claude-3, and Copilot for spatial tasks. arXiv preprint arXiv:2401.02404.
- Hossain, A., Karimuzzaman, M., Hossain, M. M., & Rahman, A. (2021). Text mining and sentiment analysis of newspaper headlines. *Information*, 12(10), 423.
- Huang, M. H., & Rust, R. T. (2021). A strategic framework for artificial intelligence in marketing. *Journal of the Academy of Marketing Science*, 49(1), 30-50.
- IAB Brasil & Kantar IBOPE Media. (2023). Digital AdSpend 2023 H1. 2023. <https://iabbrasil.com.br/pesquisa-digital-adspend-2023-h1/> (acessado em 15 de julho de 2024).
- Jang, H. J., Sim, J., Lee, Y., & Kwon, O. (2013). Deep sentiment analysis: Mining the causality between personality-value-attitude for analyzing business ads in social media. *Expert Systems with Applications*, 40(18), 7492-7503.
- Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.
- Kojima, T., Gu, S. S., Reid, M., Matsuo, Y., & Iwasawa, Y. (2022). Large language models are zero-shot reasoners. arXiv preprint arXiv:2205.11916.
- Kotler, P., & Keller, K. L. (2012). *Marketing management (14th ed.)*. Prentice Hall.
- Kshetri, N. (2023). Generative artificial intelligence in marketing. *IT Professional*, 25(5), 71-75.
- Kshetri, N., Dwivedi, Y. K., Davenport, T. H., & Panteli, N. (2024). Generative artificial intelligence in marketing: Applications, opportunities, challenges, and research agenda. *International Journal of Information Management*, 75, 102754.

- Kumar, S. (2015). Neuromarketing: The new science of advertising. *Universal Journal of Management*, 3(12), 524-531.
- Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., & Neubig, G. (2023a). Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. *ACM Computing Surveys*, 55(9), 1-35.
- Liu, X., Zheng, Y., Du, Z., Ding, M., Qian, Y., Yang, Z., & Tang, J. (2023b). GPT understands, too. *AI Open*.
- Maslen, A. (2015). *Persuasive copywriting: Using psychology to engage, influence and sell*. Kogan Page Publishers.
- Mogaji, E., Olaleye, S., & Ukpabi, D. (2023). To ChatGPT, or not to ChatGPT: Navigating the paradoxes of generative AI in the advertising industry. *Business Horizons*.
- Nah, F. F. H., Zheng, R., Cai, J., Siau, K., & Chen, L. (2023). Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration. *Journal of Information Technology Case and Application Research*, 25(3), 277-304.
- OECD. (2023). *Health at a Glance 2023: OECD Indicators*. OECD Publishing. <https://www.oecd.org/health/health-at-a-glance.htm> (acessado em 15 de julho de 2024).
- Ogilvy, D. (1983). *Ogilvy on advertising*. Crown.
- Osadchaya, E., Marder, B., Yule, J. A., Yau, A., Lavertu, L., Stylos, N., ... & Li, J. (2024). To ChatGPT, or not to ChatGPT: Navigating the paradoxes of generative AI in the advertising industry. *Business Horizons*, 67(1), 108045.
- Osterwalder, A., Pigneur, Y., Bernarda, G., & Smith, A. (2014). *Value proposition design: How to create products and services customers want*. John Wiley & Sons.
- Rani, G., Singh, J., & Khanna, A. (2023). Comparative analysis of generative AI models. In *2023 International Conference on Advances in Computation, Communication and Information Technology (ICAICCIT)* (pp. 760-765). IEEE.
- Renois , P., & Morin, C. (2007). *Neuromarketing: Understanding the buy buttons in your customer's brain*. Thomas Nelson.

Rudolph, J., Tan, S., & Tan, S. (2023). War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education. *Journal of Applied Learning and Teaching*, 6(1), 364-389.

Shaik, M. (2023). Impact of artificial intelligence on marketing. *East Asian Journal of Multidisciplinary Research*, 2(3), 993-1004.

Sheng, E., Chang, K. W., Natarajan, P., & Peng, N. (2019). The woman worked as a babysitter: On biases in language generation. arXiv preprint arXiv:1909.01326.

Statista. (2023). Distribution of digital advertising spending worldwide in 2022, by industry. <https://www.statista.com/statistics/276671/distribution-of-digital-ad-spend-worldwide-by-industry/> (acessado em 15 de julho de 2024).

Tauheed, J., Shabbir, A., & Pervez, M. S. (2024). Exploring the role of artificial intelligence in digital marketing strategies. *Journal of Business, Communication & Technology*, 54-65.

Wahid, R., Mero, J., & Ritala, P. (2023). Editorial: Written by ChatGPT, illustrated by Midjourney: generative AI for content marketing. *Asia Pacific Journal of Marketing and Logistics*, 35(8), 1813-1822.

Wang, X., Gu, X., Cao, J., Zhao, Z., Yan, Y., Middha, B., & Xie, X. (2021). Reinforcing pretrained models for generating attractive text advertisements. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining* (pp. 3697-3707).

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q., & Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems*, 35, 24824-24837.

White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., ... & Schmidt, D. C. (2023). A prompt pattern catalog to enhance prompt engineering with ChatGPT. arXiv preprint arXiv:2302.11382.

World Bank. (2023). World Development Indicators: Structure of Output. <https://data.worldbank.org/indicator/NV.SRV.TOTL.ZS> (acessado em 15 de julho de 2024).

Wu, J., Gan, W., Chen, Z., Wan, S., & Yu, P. S. (2023). Multimodal Large Language Models: A Survey. arXiv preprint arXiv:2311.13165.

Yau, A. K. L., Saad, N. M., & Chong, Y. W. (2021). Artificial intelligence marketing (AIM) for enhancing customer relationships. *Applied Sciences*, 11(18), 8592.

Yudiyanto, & Purnomo, J. (2023). Digital marketing strategy to increase sales conversion on e-commerce platforms. *Journal of Contemporary Administration and Management (ADMAN)*, 1(2), 54-62.

Zak, P. J. (2015, January). Why inspiring stories make us react: The neuroscience of narrative. In *Cerebrum: the Dana forum on brain science* (Vol. 2015). Dana Foundation.

Zhang, C., Zhang, C., Zheng, S., Qiao, Y., Li, C., Zhang, M., ... & Hong, C. S. (2023). A complete survey on generative AI (AIGC): Is ChatGPT from GPT-4 to GPT-5 all you need? arXiv preprint arXiv:2303.11717.

Zulkifly, H. Z., & Firdaus, N. (2014). Persuasion and the online consumers: Investigating copywriting strategies in native advertisements. *International Journal of Social Science and Humanity*, 4(6), 430-434.

APÊNDICE A

This is to certify that

Project No.: **DDMKT2024-6-95788**

Project Title: **Performance of Generative AI in Copywriting: A Comparative Study**

Principal Researcher: **Raphael Diversi de Matos**

according to the regulations of the Ethics Committee of NOVA IMS and MagIC Research Center this project was considered to meet the requirements of the NOVA IMS Internal Review Board, being considered **APPROVED** on 6/9/2024.

It is the Principal Researcher's responsibility to ensure that all researchers and stakeholders associated with this project are aware of the conditions of approval and which documents have been approved.

The Principal Researcher is required to notify the Ethics Committee, via amendment or progress report, of

- Any significant change to the project and the reason for that change;
- Any unforeseen events or unexpected developments that merit notification;
- The inability of the Principal Researcher to continue in that role or any other change in research personnel involved in the project.

Lisbon, 6/9/2024

NOVA IMS Ethics Committee

ethicscommittee@novaims.unl.pt

APÊNDICE B

Pergunta	Escala
Qual é o seu nível de familiaridade com Inteligência Artificial (IA) generativa?	Escala Likert de 5 pontos (Extremamente familiar - Não é familiar de maneira alguma)
Qual é o seu nível de familiaridade com marketing digital?	Escala Likert de 5 pontos (Extremamente familiar - Não é familiar de maneira alguma)
Você já utilizou ou trabalhou com ferramentas de IA generativa (como ChatGPT)?	Sim/Não
Como você avaliaria sua satisfação geral com as IA generativas?	Escala Likert de 7 pontos (Bem acima da média - Bem abaixo da média)
Para cada campanha (4 por setor: Financeiro, Educação, Tecnologia e Saúde):	
Qualidade	Escala Likert de 7 pontos (Muito baixa - Muito alta)
Criatividade	Escala Likert de 7 pontos (Muito baixa - Muito alta)
Clareza	Escala Likert de 7 pontos (Muito baixa - Muito alta)
Originalidade	Escala Likert de 7 pontos (Muito baixa - Muito alta)
Atratividade	Escala Likert de 7 pontos (Muito baixa - Muito alta)
Emoção	Escala Likert de 7 pontos (Muito baixa - Muito alta)
O título e o texto desta campanha despertam seu interesse em [ação desejada - abrir conta, matricular-se, comprar, participar]?	Escala Likert de 5 pontos (Definitivamente sim - Definitivamente não)

O título e o texto desta campanha têm uma linguagem adequada para a sua faixa etária?	Escala Likert de 5 pontos (Definitivamente sim - Definitivamente não)
Com base no título e o texto desta campanha, qual é a sua percepção sobre quem as criou?	Escala de 5 pontos (Definitivamente foram criadas por um humano - Definitivamente foram criadas por uma IA)
Para cada setor:	
Qual das campanhas você prefere?	Escolha única entre as 4 opções
Como você se identifica?	Múltipla escolha (Homem, Mulher, Outro / Não binário, Prefiro não dizer)
Qual é a sua idade?	Múltipla escolha (faixas etárias)
Qual é a sua nacionalidade?	Múltipla escolha (Brasileira, Portuguesa, Outra)
Em qual país reside atualmente?	Múltipla escolha (Brasil, Portugal, Outro)
Qual é a sua escolaridade?	Múltipla escolha (níveis de escolaridade - Brasil)
Qual é a sua escolaridade?	Múltipla escolha (níveis de escolaridade - Portugal)
Qual é a sua faixa de renda mensal?	Múltipla escolha (faixas de renda - Brasil)
Qual é o seu rendimento mensal?	Múltipla escolha (faixas de renda - Portugal)
Qual das seguintes opções melhor descreve sua ocupação atual?	Múltipla escolha (opções de ocupação - Brasil)
Qual das seguintes opções melhor descreve sua ocupação atual?	Múltipla escolha (opções de ocupação - Portugal)
É profissional do marketing?	Sim/Não
Em qual dos seguintes setores você atua predominantemente?	Múltipla escolha (setores de marketing)
O conteúdo escrito por IA é melhor do que o conteúdo escrito por humanos.	Escala Likert de 5 pontos (Concordo totalmente - Discordo totalmente)

Quais funções de marketing você acredita estarem mais em risco com o avanço da IA? (Selecione até 3 opções)	Múltipla escolha (funções de marketing)
Quais ferramentas de IA você já utilizou ou está familiarizado? (Selecione todas que se aplicam)	Múltipla escolha (ferramentas de IA)
O quão importante você acredita que a IA é para realizar suas atividades profissionais?	Escala Likert de 5 pontos (Extremamente importante - Sem qualquer importância)
Qual a possibilidade que IA irá substituir o seu trabalho?	Escala Likert de 5 pontos (Totalmente possível - Totalmente impossível)
Quais são as principais preocupações sobre o impacto da IA no Marketing? (Selecione até 3 opções)	Múltipla escolha (preocupações sobre IA)
Quais você acredita serem os principais benefícios do uso de IA generativa no Marketing? (Selecione até 3 opções)	Múltipla escolha (benefícios da IA)
Em quais áreas do Marketing você acredita que a IA generativa possa ser mais útil? (Selecione todas as que se aplicam)	Múltipla escolha (áreas de aplicação da IA)

APÊNDICE C

Prompt padrão:

Aja como um copywriter experiente e crie um texto principal com no máximo 125 caracteres e um título de no máximo 40 caracteres, ambos cativantes, para uma campanha de marketing digital que promova [nome do produto ou serviço].

****Benefícios****: Destaque os principais benefícios [benefícios principais].

****Público-Alvo****: Direcione o texto ao público-alvo específico [descrição do público-alvo].

****Requisitos dos Textos****:

- Devem ser atraentes, emocionais e claros.
- Otimizados para conversões tanto em redes sociais (Facebook, Instagram, Twitter) quanto em plataformas de ads (Google Ads, Meta Ads).
- Inclua um Call to Action. Incentive o público a agir com verbos, como por exemplo, "compre", "clique", "descubra" ou "experimente".
- Utilizar uma linguagem que ressoe com [perfil demográfico ou psicográfico específico].

****Interação com a Imagem****:

- O texto e o título devem complementar uma imagem [descrever a imagem].
- Conectar-se à emoção ou ação mostrada na imagem.

O objetivo da campanha é gerar interesse e levar o público a visitar nossa página de produtos para saber mais.

Prompt Financeiro:

Aja como um copywriter experiente e crie um texto principal com no máximo 125 caracteres e um título de no máximo 40 caracteres, ambos cativantes, para uma campanha de marketing digital que promova Conta Digital Jovem.

****Benefícios****: Destaque os principais benefícios: A conta 100% digital e sem taxa de manutenção da conta

Poupa, faz pagamentos e divide as tuas despesas. Só fazer o download da app e abrir a conta num instante. É fácil e num instante, para começar a usá-la na hora.

Todas as contas, uma só app

Veja todas as tuas contas bancárias no app e tenha uma visão 360° do teu dinheiro

Poupe por Objetivos - Crie os seus objetivos e coloque dinheiro de lado sem esforço

Defina Limites de Despesas - Organiza as suas despesas por categoria, visualiza onde é que gastas mais e define limites para os teus gastos

Todos são bem-vindos. Sem montante mínimo de abertura, oferecemos uma solução acessível a todos.

Oferece um serviço simples! Tudo aquilo de que precisa: pagar e ser pago. Nem mais, nem menos.

Focamos em oferecer aos nossos clientes apenas o que precisam, quando precisam e pagando apenas por isso. Controle o seu dinheiro com toda a conveniência: pode gerir a sua conta online pelo App e pela Área de Cliente desktop ou no seu comércio local.

Oferece o mesmo produto independentemente dos rendimentos e origem do cliente. Trabalhamos com o objetivo de facilitar o seu dia a dia, contribuindo para a inclusão financeira e social.

****Público-Alvo****: Direcione o texto ao público-alvo específico são jovens de até 25 anos.

****Requisitos dos Textos****:

- Devem ser atraentes, emocionais e claros.
- Otimizados para conversões tanto em redes sociais (Facebook, Instagram, Twitter) quanto em plataformas de ads (Google Ads, Meta Ads).
- Inclua um Call To Action. Incentive o público a agir com verbos, como por exemplo, "compre", "clique", "descubra" ou "experimente".
- Utilizar uma linguagem que ressoe com jovens que precisem de uma conta sem custos e com cartão de crédito, que vão começar a gerenciar os seus próprios gastos.

****Interação com a Imagem****:

- O texto e o título devem complementar uma imagem de um jovem fazendo um pagamento sem contato usando seu smartphone em um terminal de pagamento com tecnologia NFC. Ele

está em um ambiente descontraído, possivelmente uma loja ou café, e sua expressão concentrada ressalta a facilidade e conveniência dos pagamentos móveis para a geração mais jovem.

- Conectar-se à emoção ou ação mostrada na imagem.

****Objetivo da campanha****: O objetivo da campanha é gerar interesse e levar o público a visitar nossa página da Conta Digital Jovem para saber mais.

Prompt Educação:

Aja como um copywriter experiente e crie um texto principal com no máximo 125 caracteres e um título de no máximo 40 caracteres, ambos cativantes, para uma campanha de marketing digital que promova a Pós Graduação em Ciência de Dados.

****Benefícios****: Destaque os principais benefícios: A Pós-Graduação em Ciência de Dados é projetada para capacitar profissionais a utilizarem metodologias e ferramentas avançadas para transformar dados brutos em informações estratégicas. Este curso é essencial para aqueles que buscam orientar decisões empresariais com base em análises de dados, seja para entrar em novos mercados, lançar novas linhas de produtos ou serviços, otimizar processos, transformar modelos de negócios ou competir de maneira eficaz em um mercado cada vez mais guiado por decisões fundamentadas em dados.

Este programa também tem como objetivo a qualificação dos profissionais na área da ciência dos dados e na análise avançada, com um foco especial na criação de modelos preditivos e sua aplicação prática em cenários empresariais. A formação oferecida neste curso é abrangente e rigorosa, visando proporcionar aos participantes um sólido domínio teórico e prático dos conhecimentos mais recentes na área, abordando-a de forma interdisciplinar e multidisciplinar, o que é imprescindível para garantir uma inserção eficaz no mercado de trabalho.

Prover competências e habilidades necessárias para processar e analisar grandes volumes de dados (Big Data), extraindo valor e conhecimento para a tomada de decisões.

Desenvolver competências avançadas em Data & Text Mining, Machine Learning, e suas aplicações para solucionar problemas específicos de forma eficiente.

Dotar os estudantes com metodologias, procedimentos e técnicas de investigação que permitam a identificação, formulação e resolução de problemas (e projetos) de maneira crítica, criativa e autônoma.

****Público-Alvo****: Direcione o texto ao público-alvo específico: a todos os profissionais interessados em aprofundar os seus conhecimentos nos domínios da ciência dos dados e analítica avançada, bem como na criação de modelos preditivos e a sua aplicação prática em cenários empresariais. Este curso é ideal para aqueles que desejam inovar e liderar em um ambiente empresarial cada vez mais dependente de dados, qualificando-os para enfrentar os desafios contemporâneos com competência e confiança.

****Requisitos dos Textos****:

- Devem ser atraentes, emocionais e claros.
- Otimizados para conversões tanto em redes sociais (Facebook, Instagram, Twitter) quanto em plataformas de ads (Google Ads, Meta Ads).
- Inclua um Call to Action. Incentive o público a agir com verbos, como por exemplo, "compre", "clique", "descubra" ou "experimente".
- Utilizar uma linguagem que ressoe com pessoas de 25 a 35 anos, que provavelmente já estejam trabalhando e queiram se aperfeiçoar fazendo uma pós graduação.

****Interação com a Imagem****:

- O texto e o título devem complementar uma imagem de uma pessoa de 30 anos subindo as escadas de uma universidade, com um notebook na mão e sorrindo por ter a oportunidade de subir profissionalmente devido aos estudos.
- Conectar-se à emoção ou ação mostrada na imagem.

O objetivo da campanha é gerar interesse e levar o público a visitar nossa página do curso de Pós Graduação para saber mais.

Prompt Tecnologia:

Aja como um copywriter experiente e crie um texto principal com no máximo 125 caracteres e um título de no máximo 40 caracteres, ambos cativantes, para uma campanha de marketing digital que promova o celular Z50 Pro.

****Benefícios****: Destaque os principais benefícios celular lançou o novo celular mais avançado, topo da gama, com melhor estabilização de câmera para vídeos, resolução 4K, fotos com o melhor zoom do mercado, tela AMOLED de 6 polegadas, com resolução 4K e taxa de atualização de 120 Hz. O aparelho conta com o chipset mais rápido do mercado, além de entregar câmera principal de 50 MP. Completam o conjunto, a bateria de 5.000 mAh, com carregamento ultra rápido. Ainda possui IA que vai se tornar fundamental nas tarefas do cotidiano, como a tradução de voz em tempo real, além de funções inteligentes tais como o resumo de reuniões e de notas. A IA também possui identificação e pesquisa de conteúdos, basta apontar a câmera para um objeto que já pode verificar onde comprar na internet. Este celular oferece aos consumidores o que há de mais avançado no mercado de smartphones. Com tecnologia de última geração, os smartphones aliam design sofisticado a recursos de hardware e software.

****Público-Alvo****: Direcione o texto ao público-alvo específico: Usuários que buscam o máximo de desempenho, buscam o melhor custo-benefício a longo prazo, valorizam o investimento em um produto de alta qualidade, consumidores que apreciam exclusividade, criadores de conteúdo, entusiastas de tecnologia.

****Requisitos dos Textos****:

- Devem ser atraentes, emocionais e claros.
- Otimizados para conversões tanto em redes sociais (Facebook, Instagram, Twitter) quanto em plataformas de ads (Google Ads, Meta Ads).
- Inclua um Call to Action. Incentive o público a agir com verbos, como por exemplo, "compre", "clique", "descubra" ou "experimente".
- Utilizar uma linguagem que ressoe com pessoas que possuem alta renda disponível.

****Interação com a Imagem****:

- O texto e o título devem complementar uma imagem de uma pessoa sorridente editando uma foto em seu celular, usando um aplicativo intuitivo com recursos de IA que facilitam o processo, mesmo para iniciantes.
- Conectar-se à emoção ou ação mostrada na imagem.

O objetivo da campanha é gerar interesse e levar o público a visitar nossa página do celular Z50 Pro para saber mais.

Prompt Saúde:

Aja como um copywriter experiente e crie um texto principal com no máximo 125 caracteres e um título de no máximo 40 caracteres, ambos cativantes, para uma campanha de marketing digital que promova Mentoria Equilibrium.

****Benefícios****: Destaque os principais benefícios: A Mentoria Equilibrium oferece um acompanhamento nutricional personalizado com consultas individuais e planos alimentares adaptados às suas necessidades. Você terá acesso a recursos exclusivos como e-books de receitas e encontros regulares, tudo para promover uma transformação saudável e duradoura.

****Público-Alvo****: Direcione o texto ao público-alvo específico: Mentoria Equilibrium é ideal para mulheres que buscam perder peso de forma saudável, sem restrições extremas. É especialmente indicada para quem deseja melhorar a autoestima e manter hábitos alimentares equilibrados.

****Requisitos dos Textos****:

- Devem ser atraentes, emocionais e claros.
- Otimizados para conversões tanto em redes sociais (Facebook, Instagram, Twitter) quanto em plataformas de ads (Google Ads, Meta Ads).
- Inclua um Call to Action. Incentive o público a agir com verbos, como por exemplo, "compre", "clique", "descubra" ou "experimente".
- Utilizar uma linguagem que ressoe com: Mulheres adultas, de 25 a 45 anos. Psicograficamente, são pessoas que prezam pela saúde e bem-estar, buscam equilíbrio entre prazer e nutrição, e têm alto engajamento em cuidados pessoais e estética.

****Interação com a Imagem****:

- O texto e o título devem complementar uma imagem: Manhã radiante, sorriso vibrante. Mulher corre na praia, conquistando o peso ideal e a vida dos sonhos. Sem dietas, apenas amor próprio e equilíbrio. A cada passo, a felicidade floresce.
- Conectar-se à emoção ou ação mostrada na imagem.

O objetivo da campanha é gerar interesse e levar o público a visitar nossa página da Mentoria Equilibrium para saber mais.

APÊNDICE D

Financeiro	IAG
Título: "Liberte-se com a Conta Digital Jovem"	ChatGPT 4
Texto: "Gerencie seu dinheiro com um clique! Abra sua Conta Digital Jovem agora e descubra a liberdade!"	
Título: "Leve sua grana no bolso."	Gemini 1.5 Pro
Texto: "Sem taxas, 100% digital e com tudo na mão pra você controlar seu dinheiro. Abra sua conta agora!"	
Título: "Conta Jovem: Seu Futuro Financeiro Agora"	Claude 3 Opus
Texto: "Abra sua conta digital em instantes, sem taxas! Pague, poupe e divida despesas em um app. Clique e inicie seu futuro!"	
Título: "Ganhe controle sobre o dinheiro!"	Llama 3
Texto: "Conta Digital Jovem: poupe, faça pagamentos e organize despesas com facilidade. Sem custos! Experimente Já"	
Educação	IAG
Título: "Domine a ciência dos dados e conquiste!"	Claude 3 Opus
Texto: "Pós em Ciência de Dados: transforme dados em decisões estratégicas. Suba de nível na carreira! Saiba mais"	
Título: "Domine a Ciência de Dados"	Gemini 1.5 Pro
Texto: "Suba na carreira com a Pós-Graduação em Ciência de Dados. Inove e lidere com o poder da análise! Saiba mais."	
Título: "Aprenda Ciência de Dados"	Llama 3
Texto: "Desbloqueie o poder dos dados. Aprenda a transformar dados em conhecimento com a Pós-Graduação em Ciência de Dados."	
Título: "Escalando para o Sucesso com Dados"	ChatGPT 4
Texto: "Transforme sua carreira: domine Big Data e Machine Learning! Clique e conheça nossa Pós em Ciência de Dados hoje mesmo."	
Tecnologia	IAG
Título: "Evolua sua Criatividade"	Gemini 1.5 Pro

Texto: "A IA do Z50 Pro faz da edição profissional algo fácil e intuitivo. Descubra um novo mundo de possibilidades."	
Título: "Desbloqueie o Futuro com o Z50 Pro"	Llama 3
Texto: "Desbloqueie a tecnologia mais avançada. Com o Z50 Pro, IA para todos. Compre agora!"	
Título: "Edição de Fotos Simplificada com Z50 Pro"	Claude 3 Opus
Texto: "Crie memórias perfeitas com o Z50 Pro. IA intuitiva para edições incríveis. Descubra o poder em suas mãos. Saiba mais!"	
Título: "Revolucione sua Experiência Digital!"	ChatGPT 4
Texto: "Redefina a criatividade. O Z50 Pro traz o futuro até você. Clique e descubra o extraordinário!"	
Saúde	IAG
Título: "Desbloqueie o seu equilíbrio!"	Llama 3
Texto: "Perca peso saudável com a Mentoria Equilibrium! Consultas individuais, planos adaptados e recursos exclusivos. Clique aqui!"	
Título: "Equilibre-se e Floresça!"	Gemini 1.5 Pro
Texto: "Encontre a sua melhor versão! 🍌 Descubra o prazer de uma vida leve e feliz com a Mentoria Equilibrium. Clique e saiba mais!"	
Título: "Abraze sua jornada de transformação!"	Claude 3 Opus
Texto: "Descubra a liberdade de uma vida saudável e feliz com a Mentoria Equilibrium. Seu caminho para o equilíbrio começa aqui!"	
Título: "Transforme-se com Equilibrium!"	ChatGPT 4
Texto: "Alcance seu peso ideal com amor próprio! Experimente a Mentoria Equilibrium e transforme sua vida. Clique aqui!"	

APÊNDICE E

H1 - Repeated Measures ANOVA

Within Subjects Effects

Cases	Sphericity Correction	Sum of Squares	df	Mean Square	F	p	η^2
AI	None	14.507	3.000	4.836	5.687	< .001	0.015
	Greenhouse-Geisser	14.507	2.961	4.899	5.687	< .001	0.015
Residuals	None	206.613	243.000	0.850			
	Greenhouse-Geisser	206.613	239.880	0.861			
Sectors	None	5.813	3.000	1.938	2.287	0.079	0.006
	Greenhouse-Geisser	5.813	2.972	1.956	2.287	0.080	0.006
Residuals	None	205.863	243.000	0.847			
	Greenhouse-Geisser	205.863	240.716	0.855			
AI * Sectors	None	37.706 ^a	9.000 ^a	4.190 ^a	5.920 ^a	< .001 ^a	0.038
	Greenhouse-Geisser	37.706	7.526	5.010	5.920	< .001	0.038
Residuals	None	515.917	729.000	0.708			
	Greenhouse-Geisser	515.917	609.623	0.846			

Note. Type III Sum of Squares

^a Mauchly's test of sphericity indicates that the assumption of sphericity is violated ($p < .05$).

Between Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p
Residuals	1165.511	81	14.389		

Note. Type III Sum of Squares

Descriptives

AI	Sectors	N	Mean	SD	SE	Coefficient of variation
ChatGPT	Financial	82	4.195	1.232	0.136	0.294
	Education	82	4.366	1.378	0.152	0.316
	Technology	82	4.476	1.426	0.157	0.319
	Health	82	4.244	1.368	0.151	0.322
Claude 3 Opus	Financial	82	4.372	1.214	0.134	0.278
	Education	82	4.152	1.268	0.140	0.305
	Technology	82	4.724	1.244	0.137	0.263
	Health	82	4.705	1.150	0.127	0.244
Gemini 1.5 Pro	Financial	82	3.921	1.456	0.161	0.371
	Education	82	4.415	1.306	0.144	0.296
	Technology	82	4.244	1.179	0.130	0.278
	Health	82	4.311	1.214	0.134	0.282
Llama 3	Financial	82	4.327	1.079	0.119	0.249
	Education	82	4.476	1.201	0.133	0.268
	Technology	82	4.043	1.295	0.143	0.320
	Health	82	4.114	1.268	0.140	0.308

Assumption Checks

Test of Sphericity

	Mauchly's W	Approx. X ²	df	p-value	Greenhouse-Geisser ε	Huynh-Feldt ε	Lower Bound ε
AI	0.981	1.507	5	0.912	0.987	1.000	0.333
Sectors	0.986	1.108	5	0.953	0.991	1.000	0.333
AI * Sectors	0.423	66.948	44	0.015	0.836	0.930	0.111

Post Hoc Tests

Post Hoc Comparisons - AI * Sectors

		Mean Difference	SE	t	Ptukey	Pholm
ChatGPT, Financial	Claude 3 Opus, Financial	-0.177	0.135	-1.313	0.995	1.000
	Gemini 1.5 Pro, Financial	0.274	0.135	2.038	0.803	1.000
	Llama 3, Financial	-0.132	0.135	-0.981	1.000	1.000
	ChatGPT, Education	-0.171	0.135	-1.269	0.997	1.000
	Claude 3 Opus, Education	0.043	0.138	0.310	1.000	1.000
	Gemini 1.5 Pro, Education	-0.220	0.138	-1.593	0.970	1.000
	Llama 3, Education	-0.280	0.138	-2.036	0.805	1.000
	ChatGPT, Technology	-0.280	0.135	-2.084	0.775	1.000
	Claude 3 Opus, Technology	-0.528	0.138	-3.836	0.013	0.015
	Gemini 1.5 Pro, Technology	-0.049	0.138	-0.354	1.000	1.000
	Llama 3, Technology	0.152	0.138	1.106	0.999	1.000
	ChatGPT, Health	-0.049	0.135	-0.362	1.000	1.000
	Claude 3 Opus, Health	-0.510	0.138	-3.703	0.020	0.024
	Gemini 1.5 Pro, Health	-0.116	0.138	-0.841	1.000	1.000
Llama 3, Health	0.081	0.138	0.590	1.000	1.000	
Claude 3 Opus, Financial	Gemini 1.5 Pro, Financial	0.451	0.135	3.351	0.065	0.087
	Llama 3, Financial	0.045	0.135	0.332	1.000	1.000
	ChatGPT, Education	0.006	0.138	0.044	1.000	1.000
	Claude 3 Opus, Education	0.220	0.135	1.631	0.962	1.000
	Gemini 1.5 Pro, Education	-0.043	0.138	-0.310	1.000	1.000
	Llama 3, Education	-0.104	0.138	-0.752	1.000	1.000
	ChatGPT, Technology	-0.104	0.138	-0.752	1.000	1.000
	Claude 3 Opus, Technology	-0.352	0.135	-2.613	0.391	0.821
	Gemini 1.5 Pro, Technology	0.128	0.138	0.929	1.000	1.000
	Llama 3, Technology	0.329	0.138	2.390	0.556	1.000
	ChatGPT, Health	0.128	0.138	0.929	1.000	1.000
	Claude 3 Opus, Health	-0.333	0.135	-2.477	0.490	1.000
	Gemini 1.5 Pro, Health	0.061	0.138	0.443	1.000	1.000
	Llama 3, Health	0.258	0.138	1.874	0.888	1.000

Gemini 1.5 Pro, Financial	Llama 3, Financial	-0.407	0.135	-3.019	0.163	0.258	
	ChatGPT, Education	-0.445	0.138	-3.231	0.092	0.130	
	Claude 3 Opus, Education	-0.232	0.138	-1.682	0.951	1.000	
	Gemini 1.5 Pro, Education	-0.494	0.135	-3.670	0.023	0.028	
	Llama 3, Education	-0.555	0.138	-4.028	0.006	0.007	
	ChatGPT, Technology	-0.555	0.138	-4.028	0.006	0.007	
	Claude 3 Opus, Technology	-0.803	0.138	-5.827	< .001	< .001	
	Gemini 1.5 Pro, Technology	-0.323	0.135	-2.401	0.547	1.000	
	Llama 3, Technology	-0.122	0.138	-0.885	1.000	1.000	
	ChatGPT, Health	-0.323	0.138	-2.346	0.589	1.000	
	Claude 3 Opus, Health	-0.785	0.138	-5.695	< .001	< .001	
	Gemini 1.5 Pro, Health	-0.390	0.135	-2.900	0.217	0.367	
	Llama 3, Health	-0.193	0.138	-1.402	0.991	1.000	
	Llama 3, Financial	ChatGPT, Education	-0.039	0.138	-0.280	1.000	1.000
Claude 3 Opus, Education		0.175	0.138	1.269	0.997	1.000	
Gemini 1.5 Pro, Education		-0.087	0.138	-0.634	1.000	1.000	
Llama 3, Education		-0.148	0.135	-1.103	0.999	1.000	
ChatGPT, Technology		-0.148	0.138	-1.077	1.000	1.000	
Claude 3 Opus, Technology		-0.396	0.138	-2.877	0.228	0.389	
Gemini 1.5 Pro, Technology		0.083	0.138	0.605	1.000	1.000	
Llama 3, Technology		0.285	0.135	2.114	0.756	1.000	
ChatGPT, Health		0.083	0.138	0.605	1.000	1.000	
Claude 3 Opus, Health		-0.378	0.138	-2.744	0.304	0.580	
Gemini 1.5 Pro, Health		0.016	0.138	0.118	1.000	1.000	
Llama 3, Health		0.213	0.135	1.586	0.971	1.000	
ChatGPT, Education		Claude 3 Opus, Education	0.213	0.135	1.585	0.971	1.000
		Gemini 1.5 Pro, Education	-0.049	0.135	-0.362	1.000	1.000
	Llama 3, Education	-0.110	0.135	-0.815	1.000	1.000	
	ChatGPT, Technology	-0.110	0.135	-0.816	1.000	1.000	
	Claude 3 Opus, Technology	-0.358	0.138	-2.597	0.402	0.849	
	Gemini 1.5 Pro, Technology	0.122	0.138	0.885	1.000	1.000	
	Llama 3, Technology	0.323	0.138	2.346	0.589	1.000	
	ChatGPT, Health	0.122	0.135	0.906	1.000	1.000	
	Claude 3 Opus, Health	-0.339	0.138	-2.464	0.500	1.000	
	Gemini 1.5 Pro, Health	0.055	0.138	0.398	1.000	1.000	
	Llama 3, Health	0.252	0.138	1.829	0.906	1.000	

Claude 3 Opus, Education	Gemini 1.5 Pro, Education	-0.262	0.135	-1.947	0.853	1.000
	Llama 3, Education	-0.323	0.135	-2.400	0.548	1.000
	ChatGPT, Technology	-0.323	0.138	-2.346	0.589	1.000
	Claude 3 Opus, Technology	-0.571	0.135	-4.244	0.003	0.003
	Gemini 1.5 Pro, Technology	-0.091	0.138	-0.664	1.000	1.000
	Llama 3, Technology	0.110	0.138	0.797	1.000	1.000
	ChatGPT, Health	-0.091	0.138	-0.664	1.000	1.000
	Claude 3 Opus, Health	-0.553	0.135	-4.108	0.004	0.005
	Gemini 1.5 Pro, Health	-0.159	0.138	-1.151	0.999	1.000
	Llama 3, Health	0.039	0.138	0.280	1.000	1.000
Gemini 1.5 Pro, Education	Llama 3, Education	-0.061	0.135	-0.453	1.000	1.000
	ChatGPT, Technology	-0.061	0.138	-0.443	1.000	1.000
	Claude 3 Opus, Technology	-0.309	0.138	-2.242	0.667	1.000
	Gemini 1.5 Pro, Technology	0.171	0.135	1.269	0.997	1.000
	Llama 3, Technology	0.372	0.138	2.700	0.332	0.655
	ChatGPT, Health	0.171	0.138	1.239	0.998	1.000
	Claude 3 Opus, Health	-0.291	0.138	-2.110	0.759	1.000
	Gemini 1.5 Pro, Health	0.104	0.135	0.770	1.000	1.000
Llama 3, Education	Llama 3, Health	0.301	0.138	2.183	0.709	1.000
	ChatGPT, Technology	4.877×10^{-11}	0.138	3.540×10^{-10}	1.000	1.000
	Claude 3 Opus, Technology	-0.248	0.138	-1.800	0.916	1.000
	Gemini 1.5 Pro, Technology	0.232	0.138	1.682	0.951	1.000
	Llama 3, Technology	0.433	0.135	3.217	0.096	0.135
	ChatGPT, Health	0.232	0.138	1.682	0.951	1.000
	Claude 3 Opus, Health	-0.230	0.138	-1.667	0.955	1.000
	Gemini 1.5 Pro, Health	0.165	0.138	1.195	0.998	1.000
ChatGPT, Technology	Llama 3, Health	0.362	0.135	2.688	0.340	0.672
	Claude 3 Opus, Technology	-0.248	0.135	-1.842	0.901	1.000
	Gemini 1.5 Pro, Technology	0.232	0.135	1.721	0.941	1.000
	Llama 3, Technology	0.433	0.135	3.215	0.096	0.135
	ChatGPT, Health	0.232	0.135	1.722	0.941	1.000
	Claude 3 Opus, Health	-0.230	0.138	-1.667	0.955	1.000
	Gemini 1.5 Pro, Health	0.165	0.138	1.195	0.998	1.000
	Llama 3, Health	0.362	0.138	2.626	0.381	0.797
Claude 3 Opus, Technology	Gemini 1.5 Pro, Technology	0.480	0.135	3.562	0.033	0.041
	Llama 3, Technology	0.681	0.135	5.057	< .001	< .001
	ChatGPT, Health	0.480	0.138	3.482	0.043	0.055
	Claude 3 Opus, Health	0.018	0.135	0.136	1.000	1.000
	Gemini 1.5 Pro, Health	0.413	0.138	2.995	0.173	0.275
	Llama 3, Health	0.610	0.138	4.426	0.001	0.001
Gemini 1.5 Pro, Technology	Gemini 1.5 Pro, Health	0.413	0.138	2.995	0.173	0.275
	Llama 3, Health	0.610	0.138	4.426	0.001	0.001
	Llama 3, Technology	0.201	0.135	1.494	0.983	1.000
	ChatGPT, Health	8.537×10^{-11}	0.138	6.196×10^{-10}	1.000	1.000
	Claude 3 Opus, Health	-0.461	0.138	-3.349	0.065	0.087
Llama 3, Technology	Gemini 1.5 Pro, Health	-0.067	0.135	-0.498	1.000	1.000
	Llama 3, Health	0.130	0.138	0.944	1.000	1.000
	ChatGPT, Health	-0.201	0.138	-1.461	0.986	1.000
	Claude 3 Opus, Health	-0.663	0.138	-4.809	< .001	< .001
ChatGPT, Health	Gemini 1.5 Pro, Health	-0.268	0.138	-1.947	0.853	1.000
	Llama 3, Health	-0.071	0.135	-0.529	1.000	1.000
	Claude 3 Opus, Health	-0.461	0.135	-3.427	0.051	0.067
Claude 3 Opus, Health	Gemini 1.5 Pro, Health	-0.067	0.135	-0.498	1.000	1.000
	Llama 3, Health	0.130	0.135	0.966	1.000	1.000
	Gemini 1.5 Pro, Health	0.394	0.135	2.928	0.203	0.338
Gemini 1.5 Pro, Health	Llama 3, Health	0.591	0.135	4.393	0.001	0.001
	Llama 3, Health	0.197	0.135	1.464	0.986	1.000

Note. P-value adjusted for comparing a family of 120

H2 - Independent Samples T-Test

Independent Samples T-Test

	t	df	p	Cohen's d	SE Cohen's d
QID153	0.103	80	0.918	0.030	0.294
QID157	0.122	80	0.903	0.036	0.294
QID161	1.121	80	0.266	0.329	0.300
QID222	-0.723	80	0.472	-0.212	0.296
QID226	0.716	80	0.476	0.210	0.296
QID230	0.897	80	0.372	0.263	0.298
QID234	0.398	80	0.692	0.117	0.294
QID240	0.145	80	0.885	0.043	0.294
QID244	0.467	80	0.642	0.137	0.295
QID248	2.352	80	0.021	0.690	0.321
QID252	0.206	80	0.837	0.061	0.294
QID258	-0.830	80	0.409	-0.244	0.297
QID262	0.921	80	0.360 ^a	0.270	0.298
QID266	0.839	80	0.404	0.246	0.297
QID270	0.588	80	0.558	0.173	0.295

Note. Student's t-test.

^a Levene's test is significant ($p < .05$), suggesting a violation of the equal variance assumption

Assumption Checks

Test of Normality (Shapiro-Wilk)

		W	p
QID153	1	0.960	0.724
	2	0.987	0.682
QID157	1	0.937	0.377
	2	0.983	0.498
QID161	1	0.980	0.973
	2	0.966	0.060
QID222	1	0.988	0.998
	2	0.973	0.139
QID226	1	0.951	0.575
	2	0.982	0.449
QID230	1	0.917	0.197
	2	0.963	0.044
QID234	1	0.979	0.968
	2	0.972	0.133
QID240	1	0.972	0.902
	2	0.979	0.314
QID244	1	0.965	0.809
	2	0.974	0.167
QID248	1	0.890	0.080
	2	0.977	0.243
QID252	1	0.924	0.248
	2	0.967	0.068
QID258	1	0.933	0.338
	2	0.972	0.129
QID262	1	0.929	0.293
	2	0.981	0.394
QID266	1	0.961	0.737
	2	0.974	0.164
QID270	1	0.915	0.188
	2	0.973	0.146

Note. Significant results suggest a deviation from normality.

Test of Equality of Variances (Levene's)

	F	df ₁	df ₂	p
QID153	0.517	1	80	0.474
QID157	0.855	1	80	0.358
QID161	0.012	1	80	0.912
QID222	0.201	1	80	0.655
QID226	0.698	1	80	0.406
QID230	0.399	1	80	0.529
QID234	0.994	1	80	0.322
QID240	0.020	1	80	0.888
QID244	0.227	1	80	0.635
QID248	0.115	1	80	0.735
QID252	1.862	1	80	0.176
QID258	0.333	1	80	0.565
QID262	4.003	1	80	0.049
QID266	0.397	1	80	0.530
QID270	1.160	1	80	0.285

Descriptives

Group Descriptives

	Group	N	Mean	SD	SE	Coefficient of variation
QID153	1	14	4.226	1.022	0.273	0.242
	2	68	4.189	1.278	0.155	0.305
QID157	1	14	3.964	1.573	0.420	0.397
	2	68	3.912	1.443	0.175	0.369
QID161	1	14	4.702	1.090	0.291	0.232
	2	68	4.304	1.234	0.150	0.287
QID222	1	14	3.929	1.418	0.379	0.361
	2	68	4.199	1.242	0.151	0.296
QID226	1	14	4.643	1.416	0.378	0.305
	2	68	4.368	1.288	0.156	0.295
QID230	1	14	4.738	1.269	0.339	0.268
	2	68	4.422	1.189	0.144	0.269
QID234	1	14	4.500	1.223	0.327	0.272
	2	68	4.338	1.414	0.171	0.326
QID240	1	14	4.286	1.168	0.312	0.273
	2	68	4.235	1.189	0.144	0.281
QID244	1	14	4.190	1.470	0.393	0.351
	2	68	4.012	1.266	0.154	0.315
QID248	1	14	5.417	1.111	0.297	0.205
	2	68	4.581	1.229	0.149	0.268
QID252	1	14	4.548	1.699	0.454	0.374
	2	68	4.461	1.377	0.167	0.309
QID258	1	14	3.857	1.182	0.316	0.306
	2	68	4.167	1.287	0.156	0.309
QID262	1	14	4.583	0.857	0.229	0.187
	2	68	4.255	1.273	0.154	0.299
QID266	1	14	4.940	1.012	0.270	0.205
	2	68	4.657	1.177	0.143	0.253
QID270	1	14	4.440	1.079	0.288	0.243
	2	68	4.203	1.424	0.173	0.339

H3 - Correlations

Pearson's Correlations

			Pearson's r	p
QID161	-	QID164	-0.256*	0.020

* p < .05, ** p < .01, *** p < .001

Assumption checks

Shapiro-Wilk Test for Multivariate Normality

Shapiro-Wilk	p
0.992	0.880

Pearson's Correlations

			Pearson's r	p
QID153	-	QID156	-0.405***	< .001

* p < .05, ** p < .01, *** p < .001

Assumption checks

Shapiro-Wilk Test for Multivariate Normality

Shapiro-Wilk	p
0.983	0.348

Pearson's Correlations

			Pearson's r	p
QID157	-	QID160	-0.465***	< .001

* p < .05, ** p < .01, *** p < .001

Assumption checks

Shapiro-Wilk Test for Multivariate Normality

Shapiro-Wilk	p
0.990	0.769

Pearson's Correlations

			Pearson's r	p
QID226	-	QID229	-0.306**	0.005

* p < .05, ** p < .01, *** p < .001

Assumption checks

Shapiro-Wilk Test for Multivariate Normality

Shapiro-Wilk	p
0.981	0.258

Pearson's Correlations

			Pearson's r	p
QID222	-	QID225	-0.401***	< .001

* p < .05, ** p < .01, *** p < .001

Assumption checks

Shapiro-Wilk Test for Multivariate Normality

Shapiro-Wilk	p
0.976	0.123

Pearson's Correlations

			Pearson's r	p
QID165	-	QID168	-0.282*	0.010

* p < .05, ** p < .01, *** p < .001

Assumption checks

Shapiro-Wilk Test for Multivariate Normality

Shapiro-Wilk	p
0.973	0.087

Pearson's Correlations

			Pearson's r	p
QID230	-	QID233	-0.343**	0.002

* p < .05, ** p < .01, *** p < .001

Assumption checks

Shapiro-Wilk Test for Multivariate Normality

Shapiro-Wilk	p
0.972	0.066

Pearson's Correlations

			Pearson's r	p
QID234	-	QID237	-0.331**	0.002

* p < .05, ** p < .01, *** p < .001

Assumption checks

Shapiro-Wilk Test for Multivariate Normality

Shapiro-Wilk	p
0.981	0.248

Pearson's Correlations

			Pearson's r	p
QID240	-	QID243	-0.348**	0.001

* p < .05, ** p < .01, *** p < .001

Assumption checks

Shapiro-Wilk Test for Multivariate Normality

Shapiro-Wilk	p
0.982	0.320

Pearson's Correlations

			Pearson's r	p
QID244	-	QID247	-0.348**	0.001

* p < .05, ** p < .01, *** p < .001

Assumption checks

Shapiro-Wilk Test for Multivariate Normality

Shapiro-Wilk	p
0.991	0.865

Pearson's Correlations

			Pearson's r	p
QID248	-	QID251	-0.256*	0.020

* p < .05, ** p < .01, *** p < .001

Assumption checks

Shapiro-Wilk Test for Multivariate Normality

Shapiro-Wilk	p
0.970	0.053

Pearson's Correlations

			Pearson's r	p
QID252	-	QID255	-0.415***	< .001

* p < .05, ** p < .01, *** p < .001

Assumption checks

Shapiro-Wilk Test for Multivariate Normality

Shapiro-Wilk	p
0.962	0.017

Pearson's Correlations

		Pearson's r	p
QID258	-	QID261	-0.344** 0.002

* p < .05, ** p < .01, *** p < .001

Assumption checks

Shapiro-Wilk Test for Multivariate Normality

Shapiro-Wilk	p
0.984	0.405

Pearson's Correlations

		Pearson's r	p
QID262	-	QID265	-0.238* 0.031

* p < .05, ** p < .01, *** p < .001

Assumption checks

Shapiro-Wilk Test for Multivariate Normality

Shapiro-Wilk	p
0.974	0.090

Pearson's Correlations

		Pearson's r	p
QID266	-	QID269	-0.323** 0.003

* p < .05, ** p < .01, *** p < .001

Assumption checks

Shapiro-Wilk Test for Multivariate Normality

Shapiro-Wilk	p
0.981	0.262

Pearson's Correlations

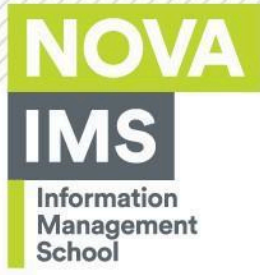
		Pearson's r	p
QID270	-	QID273	-0.186 0.094

* p < .05, ** p < .01, *** p < .001

Assumption checks

Shapiro-Wilk Test for Multivariate Normality

Shapiro-Wilk	p
0.964	0.023



NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação

Universidade Nova de Lisboa