

State of Data 2022

Análise de Descobertas relacionadas com a
Satisfação dos Profissionais da Área de Dados



Introdução

Este trabalho contempla o desafio chamado State of Data 2022, criado dentro da plataforma Kaggle por [Data Hackers + Bain], cuja finalidade é explorar o dataset mais completo do mapeamento do mercado brasileiro de dados.

Objetivo

- Fazer descobertas sobre os motivos que tornam o profissional da área de dados satisfeito.
- Explorar formas de plotar gráficos utilizando a biblioteca matplotlib da linguagem python.

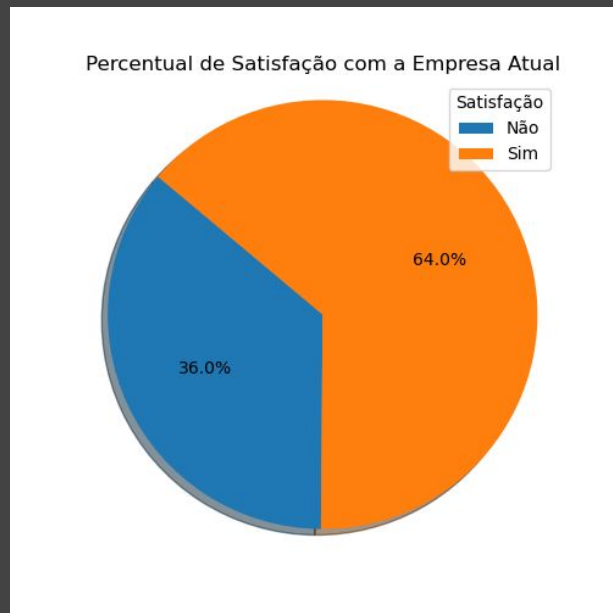


Metodologia

- Para este projeto, foi utilizado o dataset 'State_of_data_2022.csv' que cobre o mapeamento do mercado brasileiro de dados entre 10/10/22 e 11/28/22.
- Os gráficos foram plotados utilizando a linguagem de programação Python e as bibliotecas pandas, numpy e matplotlib.pyplot. A plataforma ChatGPT teve papel fundamental para fins de aprendizado e plotagem de gráficos mais complexos.
- O dataset foi dividido em 8 partes, mantendo a coluna de identificação para futura consulta em SQL, sendo as partes 5, 6, 7 e 8 irrelevantes para este trabalho. Vários gráficos foram plotados com a finalidade de fazer exploração individual dos temas contidos nos dados.
- Temas utilizados neste trabalho: Dados Demográficos, Carreira, Desafios e Conhecimento.
- Nenhum dado foi filtrado da base original, pois o nível de satisfação, ainda que focado nos profissionais que já trabalham, não pode ser fator determinante para excluir aqueles que ainda não conseguiram se posicionar na mesma e se sentem insatisfeitos.

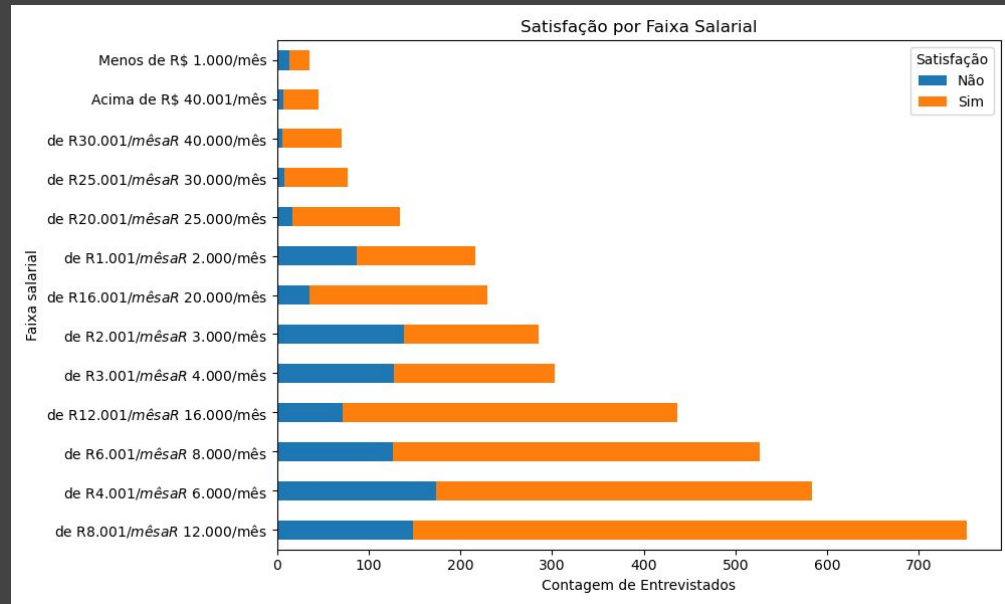
Base das Descobertas

Todo minha análise e descobertas para trazer insights teve como base a satisfação do profissional com a empresa atual, o que significa que, seja da área de dados ou não, sua satisfação com as questões abordadas a seguir dão suporte e/ou refletem no resultado final. Isso para trazer uma perspectiva maior para além dos profissionais já atuantes. As plotagens seguintes são baseadas na satisfação e nas contagens dos resultados dos tópicos respondidos neste mapeamento dos profissionais da área de dados.



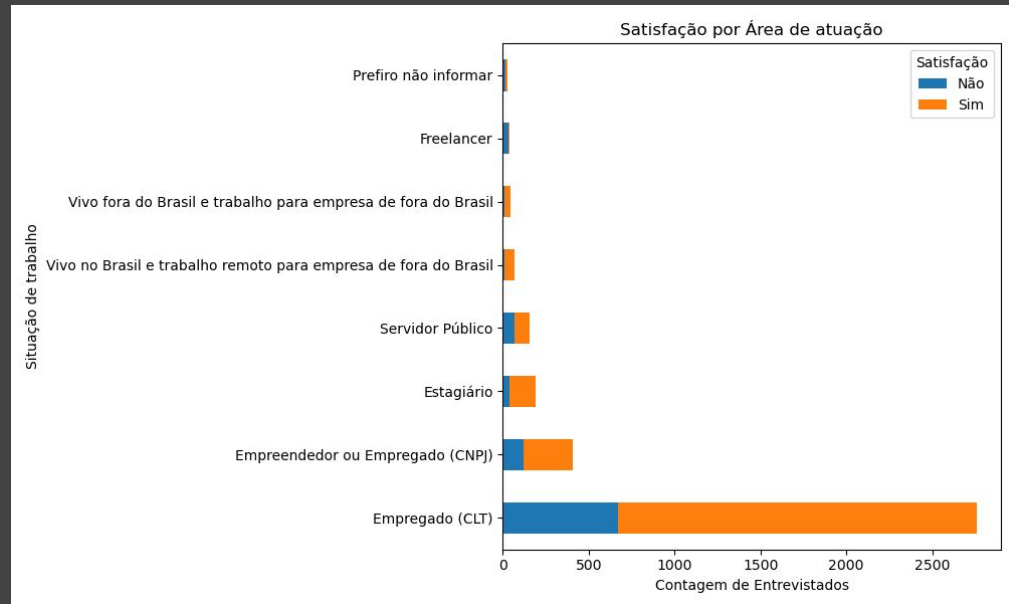
Faixa Salarial

Pessoas com salários na faixa de 1000 a 4000 reais por mês, tendem a apresentar um índice de satisfação menor do que pessoas com salários mais elevados.



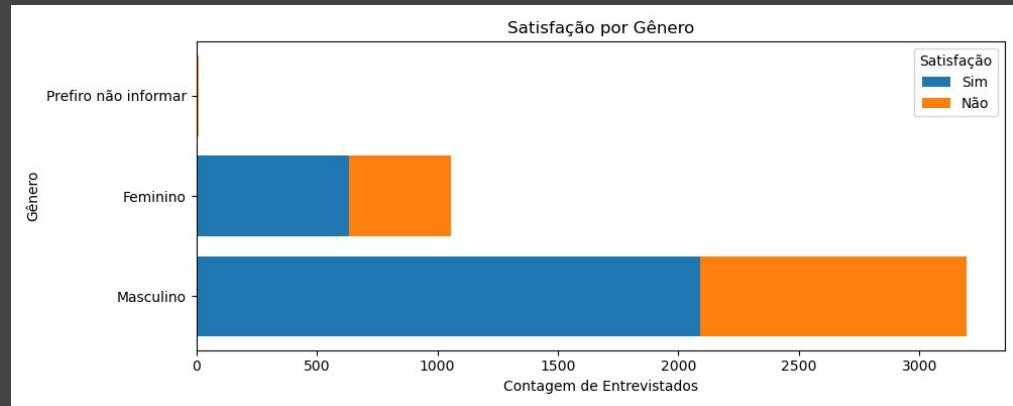
Área de Atuação

Os profissionais de dados mais satisfeitos são, em proporção, os que trabalham para/ou fora do país. Isso pode ser sustentado pelos altos salários. Em contrapartida, aqueles que trabalham como freelancer ou são servidores públicos estão em maioria insatisfeitos com a situação de trabalho atual. Estagiários, pelo tempo na área, não representam uma variável confiável o bastante para definir nível de satisfação com a área de dados.



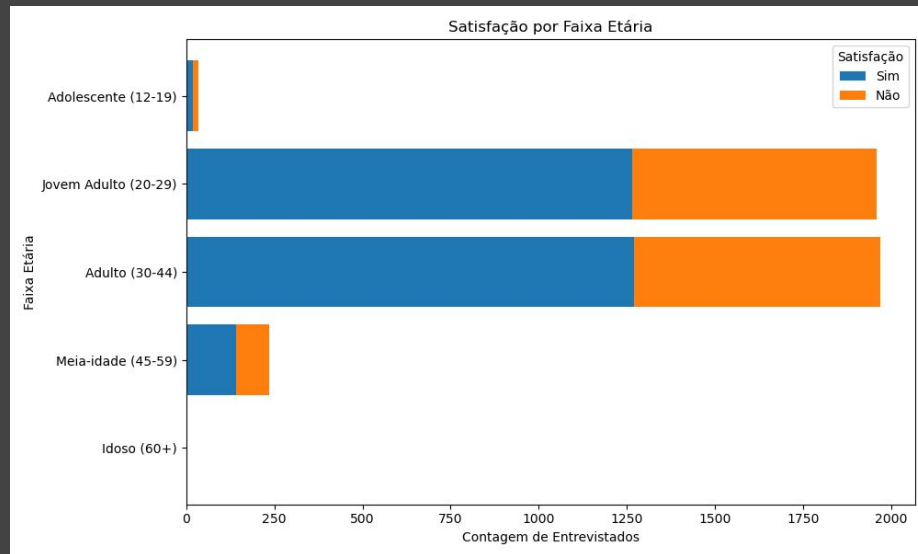
Gênero

O gênero dos entrevistados não representa alteração do índice de satisfação com a área de dados. Mostrando que, se existir, há pouca variação no grau de satisfação com base no gênero. Apontando que a área de dados é receptiva para qualquer gênero.



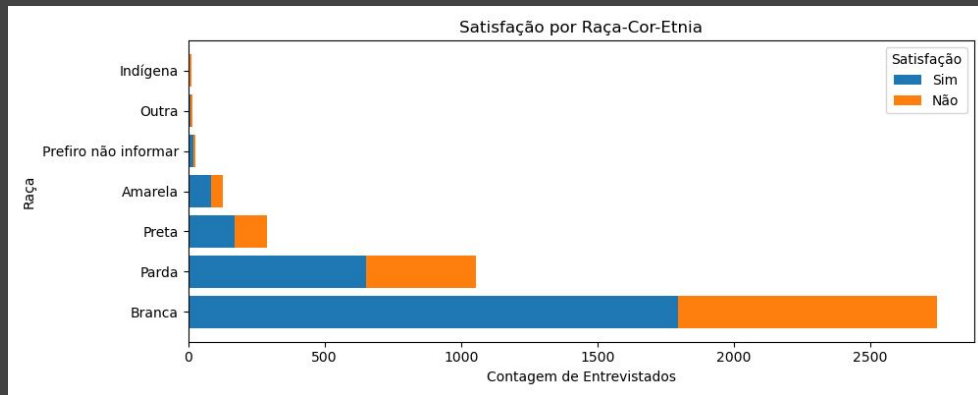
Faixa Etária

Entrevistados entre 20 a 59 anos estão satisfeitos acima da média em suas empresas atuais. Porém, em parâmetros totais, 1/3 dessa categoria permanece insatisfeito.



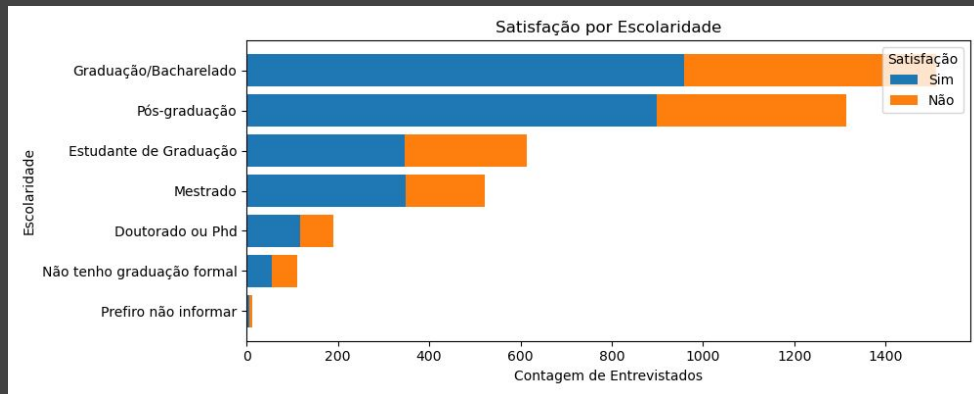
Raça-Cor-Etnia

Apesar de todas as raças e etnias se manterem acima da média, entrevistados brancos, além de representar a grande maioria dos entrevistados, tendem a estar ligeiramente mais satisfeitos em suas empresas.



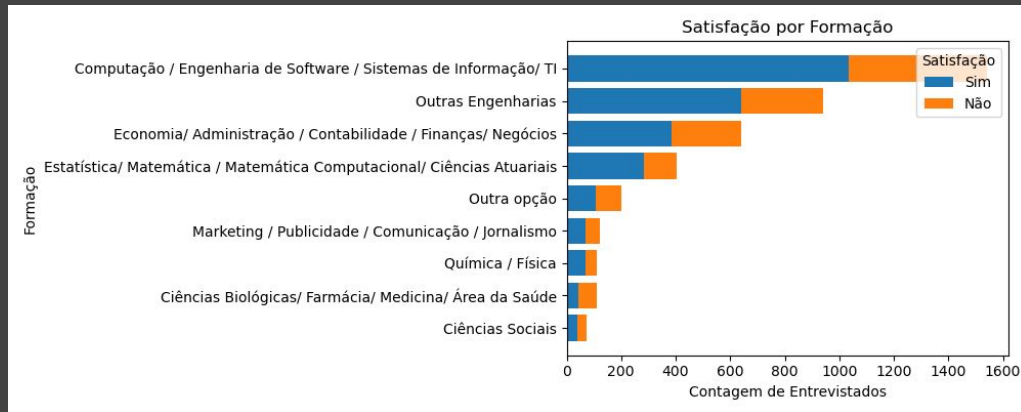
Escolaridade

Uma parcela muito pequena da amostra não possui nenhum tipo de graduação formal.



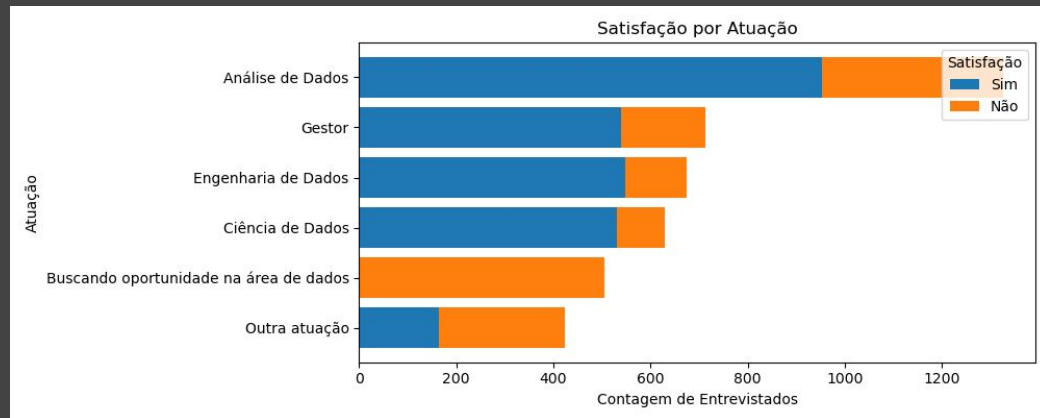
Formação

Pessoas que antes pertenciam a áreas que não estejam relacionadas com exatas, tendem a ser menos satisfeitas na área de dados.



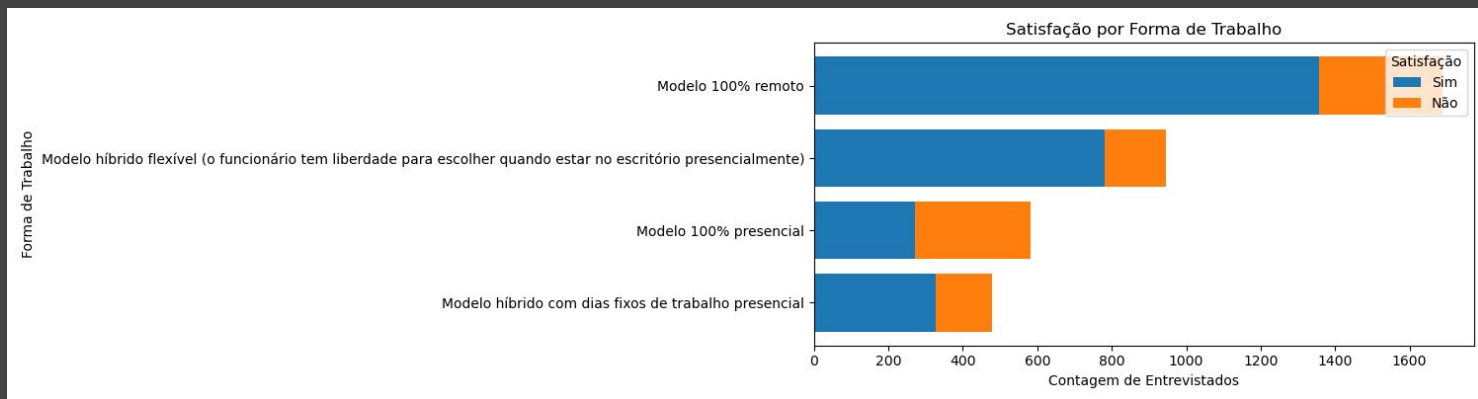
Atuação

Um ponto que destoa é que cerca de 500 entrevistados da amostra total não estão satisfeitos e isto está diretamente ligado ao fato de ainda estarem procurando oportunidades na área.



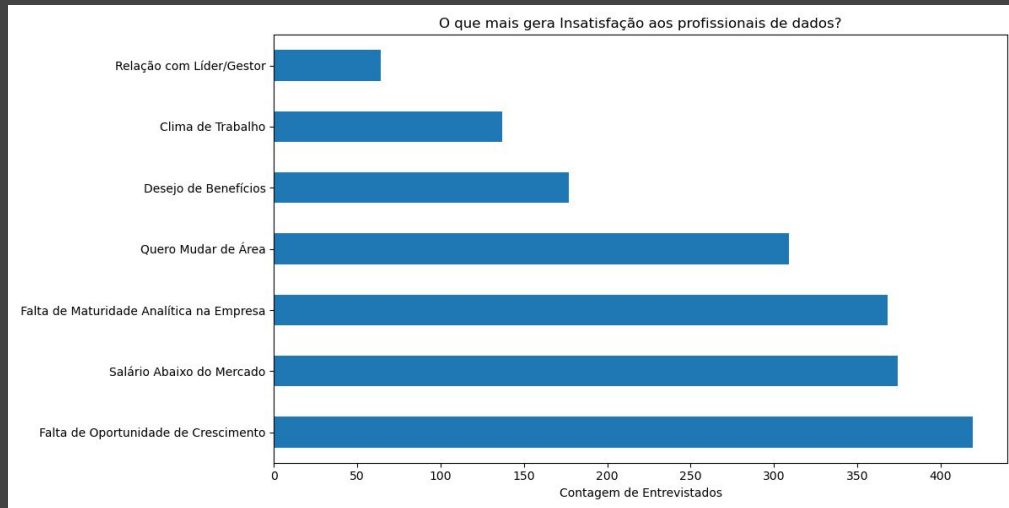
Forma de Trabalho

Profissionais de trabalho que estão no regime 100% presencial são amplamente mais insatisfeitos que qualquer outro profissional que trabalhe em regime diferente.



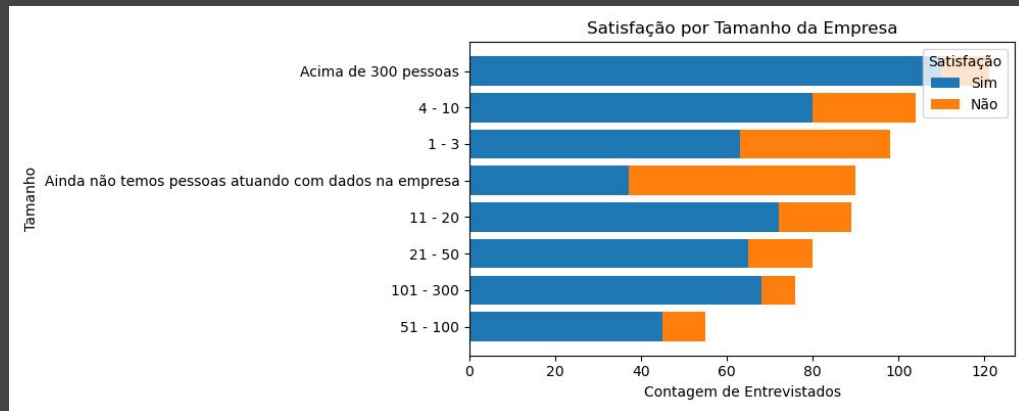
Fatores de Insatisfação

Dentre toda a amostra, pessoas que atuam fora da área de dados e pessoas que buscam oportunidades para ingressar na área de dados são as mais insatisfeitas, o que descaracteriza a área de atuação como fator determinante para a satisfação ou não do profissional, mas sim não estar inserido nela.



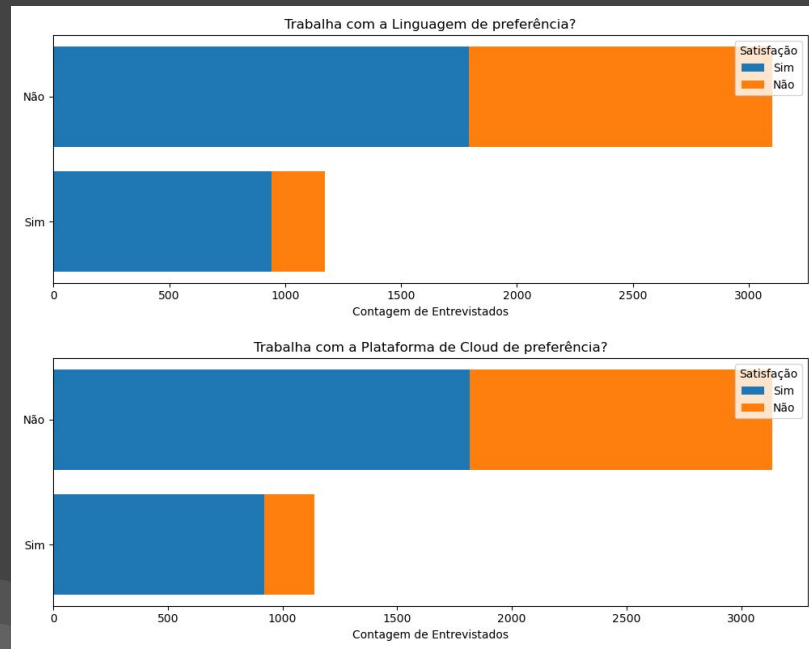
Tamanho da Empresa

Quanto maior a empresa, menor o nível de insatisfação do profissional de dados



Ferramentas

Profissionais que não trabalham com suas ferramentas de preferência tendem a ser mais insatisfeitos que aqueles que unem o útil ao agradável em termos de trabalho com dados



Considerações Finais

- Os profissionais mais satisfeitos trabalhando com dados tendem a ter uma ótima remuneração e trabalhar no regime 100% remoto.
- Trabalhar em empresas com grande número de funcionários, que apresentem oportunidade de crescimento e também possuam maturidade analítica tornam o trabalhador de dados mais satisfeito.
- O profissional que fez a transição de carreira é menos satisfeito que aqueles que vieram de uma formação nativa em exatas. Em complemento, o nível de escolaridade não tem influência direta na satisfação daqueles que atuam com dados, mantendo os índices dentro da média em todas as categorias, de estudante a Phd.
- Dentro da sua área de atuação, os mais insatisfeitos são aqueles que trabalham em outras funções ou ainda não trabalham na área. Porém, cabe um insight sobre os analistas de dados serem mais insatisfeitos que os cientistas de dados ou engenheiros de dados, isso pode ser reflexo de questões já associadas aos índices de satisfação, como remuneração e oportunidade de crescimento.

Agradecimento

“O Sucesso é a soma de pequenos esforços repetidos dia após dia” - Robert Collier

Muito obrigado pelo seu tempo e paciência!

