

Curso: Ciência de Dados
Semestre: 2º semestre de 2023
Componente Curricular/Tema: Projeto Aplicado II
Nome Completo dos Aluno: Thainá Vieira dos Santos TIA: 22500081 Vinicius Caumo Segatto TIA: 22506861 Leonardo dos Reis Olher TIA: 22510249 Nicolas Pinotti TIA: 22514112 Vinícius Vieira da Cunha Oliveira TIA: 22505865
Nome dos Professores: Anderson Adaime de Borba

1) Aplicando Conhecimento

Definição da empresa, área de atuação e apresentação dos dados que serão utilizados (imagem ou texto).

Entregar um documento estruturado, conforme a apresentação da videoaula e do texto de apoio, contendo a definição da organização escolhida, sua área de atuação, a apresentação dos dados que serão utilizados (metadados) e o link para o Github do projeto. Não se esqueça de observar os processos avaliativos disponíveis no Moodle para este componente curricular.

Nesta etapa do projeto, você deve entregar um documento conforme os itens:

- Grupo de trabalho.
- Premissas do projeto: definição da organização escolhida, área de atuação e apresentação dos dados que serão utilizados (imagem ou texto).
- Objetivos e metas.
- Cronograma de atividades (Estimativa).

SUMÁRIO

1. APRESENTAÇÃO DO PROJETO.....	
1.1. DEFINIÇÃO DA ORGANIZAÇÃO ESCOLHIDA.....	
1.2. ÁREA DE ATUAÇÃO.....	
1.3. DADOS UTILIZADOS.....	
2. OBJETIVOS E METAS.....	
3. CRONOGRAMA DE ATIVIDADES.....	
4. REFERÊNCIAS BIBLIOGRÁFICAS.....	

1) APRESENTAÇÃO DO PROJETO

1.1. DEFINIÇÃO DA ORGANIZAÇÃO ESCOLHIDA

Nome da Empresa: DataTrend Insights
Missão: Nossa missão é fornecer análises de dados de alta qualidade para empresas que desejam tomar decisões estratégicas com base nas tendências salariais em Ciência de Dados 2023.
Visão: Tornar-se a principal fonte de insights e recomendações de remuneração em Ciência de Dados, ajudando empresas a atrair, reter e motivar os melhores talentos.

DataTrend Insights está enfrentando um problema interno relacionado à retenção de Analistas de Dados Júnior. Nossa empresa está perdendo talentos promissores para concorrentes devido a diferenças salariais e benefícios. Precisamos identificar a faixa salarial competitiva para Analistas de Dados Júnior em 2023 e criar uma estratégia de retenção eficaz.

1.2. AREA DE ATUAÇÃO

A empresa "DataTrend Insights" atua principalmente nos ramos de consultoria em remuneração oferecendo serviços de consultoria para outras empresas, ajudando-as a compreender e ajustar suas políticas de remuneração em relação às tendências salariais em Ciência de Dados, envolvendo análises detalhadas do mercado, identificação de faixas salariais competitivas e recomendações específicas. E no ramo soluções de software, onde desenvolve e vende soluções de software especializadas para a gestão de remuneração em Ciência de Dados. Isso incluiria ferramentas de análise de dados, modelagem de salários e previsão de tendências salariais.

1.3. DADOS UTILIZADOS

A DataTrend Insights baseia suas análises de tendências salariais em Ciência de Dados em 2023 em um conjunto de dados cuidadosamente selecionado e abrangente. Este conjunto de dados serve como a espinha dorsal de nossos esforços analíticos, fornecendo

informações essenciais para compreender a dinâmica salarial na indústria de Ciência de Dados. Abaixo estão os principais componentes desse conjunto de dados:

- Ano de Trabalho: Esta coluna representa o ano específico da coleta de dados salariais.
- Nível de Experiência: Os funcionários são categorizados de acordo com seu nível de experiência, incluindo iniciantes, experientes, de nível médio e sêniores.
- Tipo de Emprego: Cada profissional é rotulado com seu tipo de emprego, que pode ser tempo integral, contratado, freelancer ou meio período.
- Cargo: Registramos os cargos dos funcionários, abrangendo uma variedade de títulos, como "Cientista Aplicado" e "Analista de Qualidade de Dados".
- Salário: Esta coluna contém os valores salariais, expressos em suas respectivas moedas locais.
- Moeda do Salário: Indica o código da moeda que representa o salário em questão.
- Salário em USD: Todos os salários foram convertidos para dólares americanos (USD) para permitir uma comparação uniforme.
- Localização da Empresa: Esta coluna especifica a localização das empresas, identificadas por códigos de país, como "US" para Estados Unidos e "NG" para Nigéria.
- Tamanho da Empresa: As empresas são classificadas em categorias de tamanho, que incluem grande, média e pequena.

Esses dados, cuidadosamente coletados e preparados, servirão como a base para nossas análises estatísticas avançadas, modelagem preditiva e visualizações de dados. Com esses insights, nossa equipe pode ajudar empresas a tomar decisões estratégicas informadas sobre remuneração, retenção de talentos e estratégias de aquisição de talentos em Ciência de Dados.

A integridade, qualidade e relevância desses dados são fundamentais para garantir que nossas análises sejam precisas e confiáveis, permitindo-nos fornecer serviços de consultoria e soluções personalizadas de alta qualidade para nossos clientes.

2) OBJETIVOS E METAS

Objetivo	Meta
Identificar a faixa salarial competitiva para Analistas de Dados Júnior em 2023	Determinar uma faixa salarial competitiva com base em dados coletados e análise

	estatística até o final do primeiro trimestre de 2023.
Criar um conjunto de recomendações para a empresa	Fornecer um conjunto de recomendações específicas para a empresa com base nas descobertas, incluindo ajustes salariais e estratégias de retenção, até o final do segundo trimestre de 2023.
Visualizar e interpretar tendências salariais em Ciência de Dados de 2021 a 2023	Desenvolver visualizações de dados interativas que mostrem as tendências salariais ao longo de três anos e interpretar os resultados até o final do terceiro trimestre de 2023.
Avaliar a eficácia das recomendações	Realizar uma análise de acompanhamento no quarto trimestre de 2023 para avaliar como as recomendações implementadas afetaram a retenção de Analistas de Dados Júnior e a competitividade salarial da empresa.
Publicar um relatório de tendências salariais	Lançar um relatório detalhado e informativo sobre as tendências salariais em Ciência de Dados em 2023 até o final do ano de 2023.
Oferecer treinamento sobre gestão de remuneração em Ciência de Dados	Desenvolver e lançar um programa de treinamento em gestão de remuneração voltado para profissionais de RH e gerentes de contratação no primeiro trimestre de 2024.
Expandir os serviços de consultoria em remuneração	Ampliar o portfólio de serviços de consultoria em remuneração para outras indústrias e setores até o final do segundo trimestre de 2024.

Esses objetivos e metas são formulados para orientar as atividades da empresa ao longo do projeto de análise de tendências salariais, garantindo que os resultados sejam alcançados de maneira eficaz e dentro do prazo estabelecido.

3) CRONOGRAMA DE ATIVIDADES

Fase 1: Preparação do Projeto	<ul style="list-style-type: none"> Definir os objetivos e escopo do projeto. Montar a equipe de projeto e designar responsabilidades. Estabelecer as metas de qualidade e critérios de sucesso. Preparar uma lista de verificação de coleta de dados.
Fase 2: Coleta de Dados 11/09	<ul style="list-style-type: none"> Identificar as fontes de dados relevantes. Realizar a coleta de dados das fontes identificadas.

	<ul style="list-style-type: none"> • Limpar e pré-processar os dados brutos para remover valores ausentes e erros. • Validar a qualidade dos dados coletados.
Fase 3: Manipulação de Dados com Python 15/10	<ul style="list-style-type: none"> • Importar os dados coletados para um ambiente de análise, como Jupyter Notebook. • Realizar a manipulação inicial de dados usando bibliotecas Python, como Pandas. • Executar transformações de dados, como filtragem, agregação e criação de novas variáveis. • Verificar a integridade e consistência dos dados manipulados.
Fase 4: Manipulação de Dados com R 15/10	<ul style="list-style-type: none"> • Importar os dados manipulados anteriormente para um ambiente de análise R, como RStudio. • Realizar manipulações de dados adicionais usando pacotes R, como dplyr e tidyr. • Aplicar análises estatísticas e modelagem de dados usando bibliotecas R relevantes. • Criar visualizações de dados usando ggplot2 ou outras bibliotecas de visualização.
Fase 5: Análise de Dados e Relatórios 01/11	<ul style="list-style-type: none"> • Realizar análises estatísticas avançadas usando Python e R, conforme necessário. • Interpretar os resultados das análises. • Preparar relatórios de análise de dados que incluam visualizações, gráficos e conclusões. • Revisar e validar as conclusões com a equipe e partes interessadas.
Fase 7: Documentação e Apresentação 01/11	<ul style="list-style-type: none"> • Documentar todo o processo de análise de dados, desde a coleta até a implementação de recomendações. • Preparar uma apresentação executiva para compartilhar com partes interessadas e a equipe de gerenciamento.
Fase 8: Encerramento do Projeto 29/11	<ul style="list-style-type: none"> • Avaliar o sucesso do projeto em relação às metas estabelecidas. • Realizar uma revisão pós-projeto para identificar lições aprendidas. • Arquivar todos os dados, códigos e documentação relevantes.

4) REFERÊNCIAS BIBLIOGRÁFICAS

- I) KUMARDATALAB, harish. Data Science Salary 2021 to 2023. Kaggle, 2023.
Disponível em: <https://www.kaggle.com/datasets/harishkumardatalab/data-science-salary-2021-to-2023>. Acesso em: 10 set. 2023.
- II) Mello, Oliveira Leornado. CIÊNCIA DE DADOS APLICADA A GESTÃO DE PROJETOS DE QUALITY ASSURANCE. Universidade Federal do Rio Grande do Sul, 2021.