

**PLANO DE ENSINO**

**DISCIPLINA:** Python para Ciências de Dados

**CURSO:** Ciência de Dados e Big Data Online

**PROF (A):** Leandro Lessa

**CARGA HORÁRIA TOTAL** (sala de aula + trabalho orientado): 24 h-a

**TRABALHO ORIENTADO** (descrição das atividades extracurriculares):

O Trabalho Orientado busca conduzir aos alunos por todas as etapas fundamentais do processo de projeto de ciência de dados, proporcionando uma experiência prática abrangente. Os alunos serão desafiados a aplicar conceitos teóricos e práticos aprendidos em sala de aula em um cenário fictício baseado no mundo real, permitindo a familiarização com ferramentas e técnicas essenciais para análise de dados.

**Etapas do Projeto:**

**1. Coleta de Dados:**

Os alunos deverão identificar fontes de dados relevantes para o problema proposto documentando a origem dos dados e o método de coleta para garantir transparência e replicabilidade.

**2. Limpeza e Tratamento de Dados:**

Após a coleta dos dados, os alunos irão realizar a limpeza e o tratamento necessário para lidar com problemas comuns, como valores ausentes, dados duplicados, inconsistências e outliers. Eles devem documentar todas as transformações realizadas nos dados para garantir a reprodutibilidade do processo.

**3. Visualização de Dados**

Nesta etapa, os alunos serão desafiados a criar visualizações eficazes que ajudem a entender os dados de forma clara e concisa. Eles podem utilizar uma variedade de gráficos, como histogramas, gráficos de dispersão, box plots, entre outros, para explorar diferentes aspectos dos dados e identificar padrões ou tendências.

**4. Análise Descritiva e Exploratória de Dados**

Com os dados preparados e visualizados, os alunos irão conduzir uma análise descritiva e exploratória para obter uma compreensão mais profunda das características dos dados. Isso pode envolver o cálculo de estatísticas resumidas, como média, mediana, desvio padrão, correlações, entre outras medidas relevantes.

**5. Busca por Insights:**

Por fim, os alunos serão incentivados a buscar insights significativos nos dados que possam informar decisões ou gerar novas perguntas de pesquisa. Eles devem apresentar suas descobertas de forma clara e justificar suas conclusões com base nas evidências encontradas nos dados.

**Entregáveis:**

- Relatório técnico detalhando todas as etapas do projeto, incluindo a descrição da fonte de dados, as transformações realizadas, as visualizações criadas e as conclusões da análise.
- Apresentação oral dos principais resultados e insights obtidos durante o processo de análise de dados.

**EMENTA** (conforme Projeto Pedagógico): Tipos de dados. Estruturas de controle: condicional e repetição. Estruturas de dados: listas, tuplas, conjuntos, dicionários, séries e dataframes. Funções. Vetorização e matrizes numéricas. Bibliotecas de manipulação de dados, de visualização de dados e vetorização de matrizes.

**UNIDADES DE ENSINO** (Conteúdo Programático):

1. Preparação do ambiente, Fundamentos de Python e Tipos de Dados
2. Estruturas de Dados, Estruturas de Controle e Funções
3. Coleta, Tratamento e Manipulação de Dados
4. Análise Descritiva e Exploratória de dados
5. Elaboração dos Trabalhos / Apoio aos grupos
6. Apresentação dos Trabalhos e entrega do relatório técnico

**OBJETIVOS/MÉTODOS DIDÁTICOS:**

**Objetivo geral**

Capacitar os alunos a conduzirem uma análise completa de dados utilizando python, desde a coleta até a descoberta de insights, utilizando técnicas e ferramentas de ciência de dados. Além disso, visa desenvolver habilidades de comunicação, pensamento crítico e resolução de problemas no contexto da análise de dados.

**Objetivos Específicos:**

- Compreender o Processo de Projeto de Ciência de Dados
- Aplicar Técnicas de Limpeza e Tratamento de Dados
- Criar Visualizações Efetivas
- Conduzir Análise Descritiva e Exploratória
- Buscar por Insights Significativos
- Desenvolver Habilidades de Comunicação
- Demonstrar Pensamento Crítico e Criatividade
- Fomentar Colaboração e Trabalho em Equipe:

**DISTRIBUIÇÃO DE PONTOS/CRITÉRIOS DE AVALIAÇÃO:**

Trabalho final em grupo valor de 100 pontos.

**Critérios de Avaliação**

1. Profundidade da análise realizada
2. Clareza e eficácia das visualizações criadas
3. Originalidade e relevância dos insights descobertos
4. Coerência e rigor técnico do relatório e da apresentação

**REFERÊNCIAS BIBLIOGRÁFICAS**

Grus, Joel. (2019). Data Science from Scratch: First Principles with Python. O'Reilly Media.  
VanderPlas, Jake. (2016). Python Data Science Handbook: Essential Tools for Working with Data. O'Reilly Media.  
McKinney, Wes. (2017). Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython. O'Reilly Media.  
Müller, Andreas C., & Guido, Sarah. (2016). Introduction to Machine Learning with Python: A Guide for Data Scientists. O'Reilly Media.

O'Reilly, Judea Pearl, & Mackenzie, Dana. (2018). The Book of Why: The New Science of Cause and Effect. O'Reilly

Media.

Pandas. (2024). Pandas Documentation. Disponível em: <<https://pandas.pydata.org/docs/>> Acesso em: 05 de abril de 2024

Matplotlib 3.8.4 documentation. (s.d.). Disponível em: < <https://matplotlib.org/stable/index.html> > Acesso em: 05 de abril de 2024

NumPy documentation. (s.d.). Disponível em: <https://numpy.org/doc/stable/> Acesso em: 05 de abril de 2024

#### Referências Complementares

KDD em 6 Etapas: Do Dado ao Insight. Disponível em: < <https://leandrolessa.com.br/dados/kdd-em-6-etapas-do-dado-ao-insight/> >. Acesso em: 05 de abril de 2024

3 passos para Instalar Bibliotecas em Python. Disponível em: < <https://leandrolessa.com.br/tutoriais/3-passos-para-instalar-bibliotecas-em-python/> > Acesso em: 05 de abril de 2024

Gráfico de Dispersão: Como Criar e Analisar na Prática. Disponível em: < <https://leandrolessa.com.br/tutoriais/grafico-de-dispersao-como-criar-e-analisar-na-pratica/> > Acesso em: 05 de abril de 2024

Integração de Dados: Descubra os 4 Tipos de Joins Essenciais. Disponível em: < <https://leandrolessa.com.br/tutoriais/integracao-de-dados-descubra-os-4-tipos-de-joins-essenciais/> >. Acesso em: 05 de abril de 2024