
CURSO: Ciência de Dados e Inteligência Artificial – 2º semestre de 2023

DISCIPLINA: **Análise Exploratória de Dados e Visualização**

PROFESSOR(ES): **Asla Medeiros**

TUTOR(ES): **Antonio Neto**

CARGA HORÁRIA: 60h

PRÉ-REQUISITO: Não existe

HORÁRIO E SALA DE ATENDIMENTO: Sábado às 10:00H

SALA: via zoom.

PLANO DE ENSINO

1. Ementa

Essa é uma matéria expositiva que tem como objetivo apresentar os conceitos fundamentais da Análise Exploratória – permitindo assim a compreensão das principais ferramentas matemáticas e estatísticas utilizadas na compreensão e análise de conjuntos de dados – e os principais conceitos da Visualização de Dados – de modo a desenvolver a capacitação para a construção de bons gráficos, tanto exploratórios como explicativos/comunicativos. Nesse sentido, a matéria irá cobrir os seguintes tópicos:

Tipos de variáveis. Principais medidas de centralidade e dispersão. Limpeza de dados, dados faltantes, outliers. Estatísticas descritivas. Teste de hipótese. Clusterização: k-means e clustering hierárquico. Tipos de gráficos. Gráficos interativos. Princípios de design e apresentação de resultados.

Ferramentas:

R: ggplot2, ggthemes, esquisse

Python: matplotlib, plotly, seaborn, e bokeh

Tableau e Power BI

2. Objetivos da disciplina

Nesse sentido, esta disciplina tem por objetivo fornecer uma visão geral dos conhecimentos e ferramentas que permeiam a Análise Exploratória e a

Visualização de Dados como conjunto de técnicas necessárias à verificação de padrões e hipóteses, compreensão de conjuntos de dados e tomada de decisão. Para tal, serão perpassados os principais conceitos teóricos tanto matemáticos, como estatísticos e computacionais correlatos a tais temas. Além disso, o material foi cuidadosamente organizado para que de maneira objetiva auxilie os estudantes em questões profissionais reais, independente da área de atuação.

3. Procedimentos de ensino (metodologia)

As competências desenvolvidas ao longo desta disciplina:

- Utilizar ferramentas computacionais comumente utilizadas no desenvolvimento das atividades profissionais, e que são altamente valorizadas hoje;
- Esclarecer e ilustrar as dificuldades tipicamente enfrentadas durante o processo de análise de dados e tomada de decisão com base em dados;
- Capacitar o estudante a desenvolver soluções para problemas comuns às atuações profissionais e a avaliar diferentes conjuntos de soluções alternativos para o mesmo problema.
- Fomentar os aspectos éticos envolvidos no universo da Análise Exploratória e da Visualização de Dados.

4. Conteúdo programático detalhado

Esse é um curso online e assíncrono.

A proposta do curso é, em sua primeira metade, passar por todos os conteúdos relacionados à Análise Exploratória, indo desde os tipos de variáveis até teste hipótese.

Em sua segunda metade, serão lecionados os conceitos relativos à visualização de dados e a técnicas visuais de análises de dados.

5. Procedimentos de avaliação

As avaliações são definidas mediante aplicação de questionário com os alunos, sempre tendo como possibilidade a realização de provas, trabalhos individuais e

trabalhos em grupo. Para 2022, foi escolhida a realização de trabalho individual como avaliação tanto para a A1 como para a A2.

6. Bibliografia Obrigatória

- Tukey, J. Exploratory Data Analysis. Pearson. 1977
- Few, S. Information Dashboard Design: The Effective Visual Communication of Data. O'Reilly, 2006
- Kirk, A. Data Visualisation: A Handbook for Data Driven Design. Sage, 2016.
- Cole Nussbaumer Knaflic. Storytelling with Data: A Data Visualization Guide for Business Professionals. Wiley, 2015

7. Bibliografia Complementar

- Philipp K. Janert. Data Analysis with Open Source Tools: A Hands-On Guide for Programmers and Data Scientists. O'Reilly, 2011.
- Osvaldo Martin. Bayesian Analysis with Python. Packt. 2016
- Shai Vaingast. Beginning Python Visualization: Crafting Visual Transformation Scripts. Apress. 2014.
- Petrou, Theodore. Pandas Cookbook: Recipes for Scientific Computing, Time Series Analysis and Data Visualization
- Rossant, Cyrille. IPython Interactive Computing and Visualization Cookbook
- Wes McKinney. Python Data Analysis, O'Reilly. 2017

8. Mini currículo do(s) Professor(s)

Professora (Adjunto IV) da Escola de Matemática Aplicada EMap-FGV/RJ (<http://emap.fgv.br>). Pós Doutorado pelo Cultural Informatics Research Group da Universidade de Brighton na Inglaterra (2016 e 2012) financiado pela CAPES. Doutora em Ciências, com ênfase em Computação Gráfica, pelo IMPA- Instituto Nacional de Matemática Pura e Aplicada (2006). Mestre em Matemática Aplicada pela UFRJ (2001). Bacharel em Matemática pela UFRJ (1999). Áreas de Pesquisa: Computação Gráfica e Visualização de Informação em aplicações, com ênfase em herança cultural.

9. Link para o Currículo Lattes

<http://lattes.cnpq.br/1427153134693867>

<https://www.linkedin.com/in/amfneto/>