



PLANO DE ENSINO SEMESTRAL

DISCIPLINA: Análise, construção e visualização de dados (0053)

PRÉ-REQUISITOS: Pensamento computacional e lógicas de programação (0005)

CURSO: Jornalismo | Publicidade e Propaganda **SEMESTRE:** 2º
TURNO: Matutino

DOCENTE RESPONSÁVEL: Álvaro Campos Ferreira

E-MAIL: alvaro.ferreira@idp.edu.br

CARGA HORÁRIA: 80 h/a **ANO E SEMESTRE DE**
REFERÊNCIA: 2022/1

HORÁRIO DAS AULAS: Segundas, das 8h às 9h40 | Sextas, das 10h às 11h40

LOCAL: Sala de aula e Laboratório de iMACs

OBJETIVOS

Geral

- Compreender o que são dados na comunicação, para que servem, porque usá-los, e como trabalhar com eles.

Específicos

- Reconhecer as boas práticas do fluxo de trabalho com dados;
- Produzir análises com ferramentas como Jupyter e Open Refine;
- Elaborar visualizações de dados com Python, Flourish e Datawrapper.

METODOLOGIA DE ENSINO

A disciplina terá uma ênfase na resolução de problemas para que o estudante tenha um papel ativo no aprendizado. As aulas de laboratório serão realizadas nos laboratórios de iMacs ou de Windows, sendo que em ambos os softwares Anaconda Python, Open Refine, Flourish e Datawrapper estarão instalados.

O estudante irá desenvolver soluções para os problemas propostos utilizando as ferramentas indicadas para problemas que deverão ser enviados pelo Canvas, seja para os testes ou para o projeto final.



60% das aulas serão ministradas na sala de aula tradicional com quadro branco e projetor. 40% serão realizadas no laboratório, onde cada estudante terá acesso a um computador com as ferramentas necessárias.

RECURSOS DIDÁTICOS

Quadro branco, pinceis (azul, preto e vermelho), apagador, datashow. Ambiente Virtual de Aprendizagem, Anaconda Python, Open Refine, Flourish, Datawrapper.

EMENTA

Metodologia e boas práticas no fluxo de trabalho com dados; Formulação de hipóteses; Obtenção de dados; Limpeza e checagem de dados; Estatística básica; Vieses na análise de dados; Interpretação de dados com ferramentas de planilhas; produção e interpretação de infográficos.

PROGRAMA

- 1. O que são dados?**
 - 1. Tipos de dados**
 - 1. Primórdios dos registros de informações**
 - 2. Sistemas baseados na manipulação de arquivos**
 - 3. Sistemas de gerenciamento de banco de dados**
 - 2. Fontes de dados**
 - 3. Ética de dados**
 - 4. Fluxo de trabalho com dados**
- 2. Onde encontrar dados?**
 - 1. WebScrapping**
 - 2. Dados abertos**
 - 3. CSV, JSON, XML**
 - 4. Acessando dados pelo Python**
 - 5. Acessando bancos de dados pelo Python (ORM)**
- 3. Verificação e limpeza**
 - 1. Open Refine**
- 4. Análise de dados**
 - 1. Descrevendo um conjunto de dados**
 - 2. Estatística descritiva usando Pandas e Numpy**
- 5. Visualização de dados - Dataviz**
 - 1. Data storytelling**
 - 2. Qual gráfico usar?**



3. Flourish e Datawrapper
4. Pandas e Matplotlib
6. PowerBI

AValiação

A nota A1 será composta por uma avaliação (10,0 pontos) e um teste (5,0 pontos), dividido por 3/2. Já a nota A2 será composta por um projeto (10,0 pontos) e um teste (5,0 pontos) também dividido por 3/2.

- Além disso, será verificada a presença dos alunos em todas as aulas e atividades programadas da disciplina. De acordo com o Regimento Geral da Instituição, será exigida frequência mínima de 75% para aprovação.

Portanto, será considerado aprovado o estudante que obtiver nota final $[(A1 + A2)/2]$ maior ou igual a 6 (seis) e frequência mínima de 75%.

Observações:

- Sob justificativa plausível, o discente poderá solicitar a realização da prova substitutiva, seguindo os prazos estabelecidos pelo Calendário Acadêmico. Nos casos das avaliações divididas em diversas atividades (fichamentos, produtos, estudos de caso), não é possível solicitar a prova substitutiva para a atividade, exceto quando ela for composta por uma prova.

- É assegurado ao discente o direito de exigir revisão dos resultados obtidos nas avaliações 1 e 2, (A1) e (A2), assim como daqueles alcançados nas provas substitutivas, desde que respeite o prazo para solicitação especificado no Calendário Acadêmico e atenda as normas aprovadas pelo Consaepe. Isso significa que, ao requerer revisão de nota da Avaliação 2 (AV2), por exemplo, o(a) professor(a) irá reavaliar todas as atividades propostas para compor essa nota (fichamentos, projetos, produtos etc.).

- Além dos mecanismos avaliativos descritos anteriormente, indicamos a realização de uma avaliação diagnóstica no início do semestre letivo e duas autoavaliações a serem aplicadas, respectivamente nesta ordem, no término do primeiro e do segundo bimestre. É importante frisar que a estas avaliações não serão atribuídas notas, pois têm como objetivos aferir o conhecimento prévio dos alunos em relação aos conteúdos da disciplina, bem como identificar as deficiências na forma de ensinar, auxiliando, portanto, na reformulação e no aperfeiçoamento das estratégias de ensino adotadas.



CRONOGRAMA DE ATIVIDADES

Nº	Aula	Data	Conteúdo	Atividade
1		07/03/22	O que são dados?	
2		11/03/22	Fluxo de trabalho com dados	
3		14/03/22	Onde encontrar dados?	
4		18/03/22	Python e suas Bibliotecas	Laboratório
5		21/03/22	Acessando dados pelo Python	
6		25/03/22	Acessando dados pelo Python	Laboratório
7		28/03/22	Arquivos, planilhas e Dataframes	
8		01/04/22	Arquivos, planilhas e Dataframes	Laboratório
9		04/04/22	Dataframes e estatística	
10		08/04/22	Dataframes e estatística	Laboratório
11		11/04/22	Primeira Avaliação	Laboratório
12		15/04/22	Sexta-feira Santa	
13		18/04/22	Histórias de dados	Entrega teste
14		22/04/22	Webscrapping	
15		25/04/22	Webscrapping	Laboratório
16		29/04/22	Verificação e limpeza	
17		02/05/22	Verificação e limpeza	Laboratório
18		06/05/22	Open Refine	
19		09/05/22	Open Refine	Laboratório
20		13/05/22	Projetos - introdução	Entrega teste
21		16/05/22	Dataviz	
22		20/05/22	Qual gráfico usar?	
23		23/05/22	Flourish e Datawrapper	
24		27/05/22	Flourish e Datawrapper	Laboratório
25		30/05/22	Pandas e Matplotlib	
26		03/06/22	Pandas e Matplotlib	Laboratório
27		06/06/22	PowerBI	
28		10/06/22	Jupyter	
29		13/06/22	Jupyter	Laboratório
30		17/06/22	Projetos - Seguimento	Laboratório
31		20/06/22	Revisão	
32		24/06/22	Projetos - Seguimento	Laboratório
33		27/06/22	Rayanne Portugal - estratégia criatividade mkt digital escrita criativa	
34		01/07/22	Projetos - Seguimento	Laboratório
35		04/07/22	Apresentação Projetos	
36		08/07/22	Apresentação Projetos	
37		11/07/22	Welberth Douglas Nascimento Cientista de Dados	
38		15/07/22	Último dia de aula	



BIBLIOGRAFIA BÁSICA

MILANI, Alessandra Maciel Paz (coautor) et al. **Visualização de dados**. 1. ed. Porto Alegre: SAGAH, 2020. E-book. (1 recurso online). ISBN 9786556900278. Disponível em: <https://integrada.minhabiblioteca.com.br/books/9786556900278>. Acesso em: 9 fev. 2022.

ANDERSON, David R. (coautor) et al. **Estatística aplicada a administração e economia**. 5. ed. ed. completa São Paulo: Cengage Learning Brasil, 2020. E-book. (1 recurso online). ISBN 9786555583991. Disponível em: <https://integrada.minhabiblioteca.com.br/books/9786555583991>. Acesso em: 9 fev. 2022.

GRUS, Joel. **Data science do zero: noções fundamentais com Python**. 2. ed. Rio de Janeiro: Alta Books, 2016. E-book. (1 recurso online). ISBN 9788550816463. Disponível em: <https://integrada.minhabiblioteca.com.br/books/9788550816463>. Acesso em: 9 fev. 2022.

FERREIRA, Rafael Gastão Coimbra (coautor) et al. **Preparação e análise exploratória de dados**. Porto Alegre: SAGAH, 2021. E-book. (1 recurso online). ISBN 9786556902890. Disponível em: <https://integrada.minhabiblioteca.com.br/books/9786556902890>. Acesso em: 9 fev. 2022.

BIBLIOGRAFIA COMPLEMENTAR

AMARAL, Fernando. **Aprenda mineração de dados: teoria e prática**. Rio de Janeiro: Alta Books, 2019. E-book. (1 recurso online). ISBN 9786555206852. Disponível em: <https://integrada.minhabiblioteca.com.br/books/9786555206852>. Acesso em: 9 fev. 2022.

ALVES, William Pereira. **Banco de dados: teoria e desenvolvimento**. 2. ed. São Paulo: Erica, 2020. E-book. (1 recurso online). ISBN 9788536533759. Disponível em: <https://integrada.minhabiblioteca.com.br/books/9788536533759>. Acesso em: 9 fev. 2022.

PASSOS, Emmanuel; BEZERRA, Eduardo (coautor). **Data mining: conceitos, técnicas, algoritmos, orientações e aplicações**. 2. ed. Rio de Janeiro: GEN LTC, 2015. E-book. (1 recurso online). ISBN 9788595156395. Disponível em: <https://integrada.minhabiblioteca.com.br/books/9788595156395>. Acesso em: 9 fev. 2022.



VIDA, Edinilson da Silva (coautor) et al. **Data warehouse**. Porto Alegre: SAGAH, 2021. E-book. (1 recurso online). ISBN 9786556901916. Disponível em: <https://integrada.minhabiblioteca.com.br/books/9786556901916>. Acesso em: 9 fev. 2022.

D'IGNAZIO, Catherine. **Putting data back into context**. Why context is hard. Disponível em: <https://datajournalism.com/read/longreads/putting-data-back-into-context>. Acesso em: 27 dez. 2021.

COOK, Louise; HOLMES, Mike. **Econometrics Explained**. Disponível em: <https://ipa.co.uk/media/4779/econometricsexplained.pdf>. Acesso em: 27 dez. 2021.

SPAGNUOLO, Sérgio. **De um feed de Twitter para a farra dos verbetes**. Disponível em: <https://medium.com/volt-data-lab/de-um-feed-de-twitter-para-a-farra-dos-verbetes-869061a17bad>. Acesso em: 27 dez. 2021.

INDICAÇÕES EXTRAS

- Inserir aqui indicações de livros, artigos, séries, filmes, podcasts, documentários, sites etc.

Brasília-DF, 18 de fevereiro de 2022.

Prof. Álvaro Campos Ferreira