

Especialização em Ciência de Dados Aplicadas a Políticas Públicas

D.9 - Introdução a Ciência de Dados

Professor(a):	Alex Lopes Pereira
Período:	08/11 a 07/12
Horário:	2a-feira: 14h30 às 17h30, 3a-feira: 9h30 às 12h, 3a-feira: 14h30 às 17h30,
Sala:	Zoom
Carga Horária:	42 horas

Objetivo de Aprendizagem

Ao final da disciplina, o estudante deverá ser capaz de: utilizar dados para criar painéis e relatórios dinâmicos com foco em narrativas (contar uma história). O estudante terá domínio de ferramentas que possibilitam tanto a produção de gráficos estáticos como gráficos dinâmicos.

Ementa

Tópicos da disciplina	
	Teoria sobre Story Telling com dados. Introdução a biblioteca de gráficos Seaborn.
	Teoria sobre Paineis (dashboards) e Introdução a ferramenta de painéis do Google, o Data Studio.
	Criação de tabelas dinâmicas (pivot table). Webscraping e Xpath. Validação de dados com schemas.
	Introdução à biblioteca de gráficos baseada em JavaScript: Observable e D3.

Metodologia

Serão realizados encontros síncronos. Nesses encontros serão apresentados conceitos teóricos ocupando por volta de 50% do tempo da aula, e serão aplicados exercícios práticos a serem desenvolvidos pelos alunos com supervisão e assessoria de monitores. Haverá alternância entre conteúdo expositivo curto, exercícios com codificação para consolidar o domínio das técnicas apresentadas e aplicações em novos conjuntos de dados, de forma guiada, para facilitar a experimentação das técnicas sobre dados reais e o ganho de autonomia do aluno. Todos os dias serão propostos exercícios práticos para proporcionar um domínio incrementalmente maior das habilidades de apresentação de dados. Os exercícios terão diversos níveis de dificuldade, o que pode proporcionar oportunidades de aprendizagem para diversos perfis de alunos numa turma heterogênea em relação ao domínio das ferramentas.

Avaliação da Aprendizagem

A avaliação será realizada por meio da resolução de exercícios práticos. Os exercícios serão disponibilizados ao final da aula de teoria em todas as 3^{as} Feiras a tarde. O prazo de entrega é até o sábado da respectiva semana.

PLANO DE AULA	
8/11/2021 - Segunda-feira	
14h30 às 17h30	<ul style="list-style-type: none"> • Introdução a Story Telling com dados – Parte 1 • Apresentação do Seaborn • Exercícios práticos com Seaborn
9/11/2021 - Terça-feira	
9h às 12h	<ul style="list-style-type: none"> • Introdução a Story Telling com dados – Parte 2 • Criação de tabelas dinâmicas • Exercícios práticos com Seaborn
9/11/2021 - Terça-feira	
14h30 às 17h30	<ul style="list-style-type: none"> • Introdução a Story Telling com dados – Parte 3 • Exercícios práticos com Seaborn • Apresentação de metodologia de trabalho com IDE Python
16/11/2021 - Terça-feira	
9h às 12h	<ul style="list-style-type: none"> • Webscraping com Python/Selenium e Xpath • Exercícios Práticos de webscraping
16/11/2021 - Terça-feira	
14h30 às 17h30	<ul style="list-style-type: none"> • Apresentação do Google Data Studio – Parte 1 • Pivot Table (Tabelas Dinâmicas) • Exercícios práticos com Google Data Studio
22/11/2021 - Segunda-feira	
14h30 às 17h30	<ul style="list-style-type: none"> • Apresentação do Google Data Studio – Parte 2 • Exercícios práticos com Google Data Studio
23/11/2021 - Terça-feira	
9h às 12h	<ul style="list-style-type: none"> • Apresentação do Google Data Studio – Parte 3 • Validação de dados com schemas • Exercícios práticos com Google Data Studio
23/11/2021 - Terça-feira	
14h30 às 17h30	<ul style="list-style-type: none"> • Apresentação do Google Data Studio – Parte 4 • Exercícios práticos com Google Data Studio
29/11/2021 - Segunda-feira	

14h30 às 17h30	<ul style="list-style-type: none"> • Apresentação dos Notebooks no Observable – Parte 1 • Apresentação de Operações com Tidyverse • Exercícios práticos com Observable
30/11/2021 - Terça-feira	
9h às 12h	<ul style="list-style-type: none"> • Apresentação dos Notebooks no Observable – Parte 2 • Conexão com fontes de dados no Observable • Exercícios práticos com Observable
30/11/2021 - Terça-feira	
14h30 às 17h30	<ul style="list-style-type: none"> • Introdução a biblioteca gráfica D3.js - Parte 1 • Exercícios práticos com Observable • Exercício Temático de Portifólio
6/12/2021 - Segunda-feira	
14h30 às 17h30	<ul style="list-style-type: none"> • Introdução a biblioteca gráfica D3.js - Parte 2 • Apresentação de trade-offs de ferramentas de visualização e de fontes de dados • Exercícios práticos de Comparação das Ferramentas
7/12/2021 - Terça-feira	
9h às 12h	<ul style="list-style-type: none"> • Apresentação dos Notebooks no Observable – Parte 3 • Exercícios práticos com Observable •
7/12/2021 - Terça-feira	
14h30 às 17h30	<ul style="list-style-type: none"> • Apresentação dos Notebooks no Observable – Parte 4 • Exercícios práticos com Observable • Exercício Temático de Portifólio

Bibliografia Básica

KNAFLIC, C. N. (2018). <i>Storytelling with data: a data visualization guide for business professionals</i> .
McKinney, W. (2018). Python for data analysis: Data wrangling with pandas, NumPy, and IPython.
HURST, L. (2020). <i>Hands on with Google Data Studio: a data citizen's survival guide</i> . https://onlinelibrary.wiley.com/doi/book/10.1002/9781119616238 .
WEXLER, S., SHAFFER, J., & COTGREAVE, A. (2017). <i>The big book of dashboards: visualizing your data using real-world business scenarios</i> . https://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=4850323 .

	https://seaborn.pydata.org/
	https://pandas.pydata.org/docs/
	https://d3js.org/
	https://numpy.org/doc/stable/reference/index.html

Bibliografia Complementar

	Downey, Allen. <i>Python for Software Design: How to Think Like a Computer Scientist</i> . Leiden: Cambridge University Press, 2009. Internet resource.
--	---

Docente (minicurrículo)



Alex Lopes Pereira

Doutor em Engenharia Eletrônica e de Computação pelo Instituto Tecnológico de Aeronáutica – ITA. É integrante da carreira de Analista de Planejamento e Orçamento. Atualmente, trabalha no Ministério da Economia. Foi Analista de Ciência e Tecnologia do Centro Gestor e Operacional do Sistema de Proteção da Amazônia – CENSIPAM e Empreendedor em startups de tecnologia da informação.