



CENTRO DE CIÊNCIAS JURÍDICAS E ECONÔMICAS
DEPARTAMENTO DE ECONOMIA

Plano de Ensino					
Universidade Federal do Espírito Santo			Campus:	Goiabeiras	
Curso:	CIÊNCIAS ECONÔMICAS				
Departamento Responsável:	ECONOMIA				
Data de Aprovação (Art. nº 91):	26/09/2024				
Docente Responsável:	Everlam Elias Montibeler (everlam.montibeler@ufes.br) Rodrigo Franklin (rodrigo.franklin@ufes.br)				
Qualificação/link para o Currículo Lattes:	http://lattes.cnpq.br/4076104093284079 http://lattes.cnpq.br/7268139028891824				
Disciplina:	Tópicos Especiais em Economia II			Código:	ECO-02234
Pré-requisito:	-			Carga Horária Semestral:	60
Créditos:	Distribuição da Carga Horária Semestral				
	04	Teoria	Exercício	Laboratório	
		60	---	---	---
Ementa: Introdução às ferramentas e técnicas fundamentais da Ciência de Dados, com ênfase na linguagem de programação R. Métodos de coleta de dados; técnicas de pré-processamento de dados; estrutura de dados; princípios de bancos de dados; armazenamento; conceitos de geoprocessamento; visualização de dados e criação de <i>dashboards</i> utilizando o pacote Shiny. Pipeline de dados e controle de versão na construção de soluções de dados.					
Objetivos Específicos: O objetivo da disciplina é desenvolver de forma introdutória as competências relacionadas a etapas complementares na área de ciência de dados normalmente não abordados com profundidade em cursos voltados para análises de dados na área da economia. Com isso, espera-se que ao final da disciplina os alunos de economia tenham a capacidade de sozinhos ou em colaboração com outros profissionais, desenvolverem soluções de dados completas, envolvendo as etapas de coleta, processamento e visualização.					
Conteúdo Programático: 1. Introdução: 1.1. Fundamentos de Ciência de Dados 1.2. Visão Geral sobre algoritmos 1.3. Data vs. Big Data 1.4. Ética em Ciência de Dados 1.5. Aplicações e casos de uso em diferentes indústrias 1.6. Projetos Práticos e Estudos de Caso 2. Introdução à lógica de programação em R: 2.1. Conceitos básicos; 2.2. Tipos de variáveis; 2.3. Comandos de atribuição; 2.4. Expressões aritméticas, lógicas e literais; 2.5. Estruturas de seleção; 2.6. Estruturas de repetição; 2.7. Modularização. 3. Controle de versão usando Git:					

- 3.1. Introdução ao controle de versão;
- 3.2. Comandos fundamentais (init, add, commit, status e log);
- 3.3. Trabalhando com repositórios remotos (clone, push, pull);
- 3.4. Ramificando e mesclando versões (branch e merge).

4. Métodos de coleta de dados:

- 4.1. Leituras de arquivos tabulares e de texto;
- 4.2. APIs;
- 4.3. Webscraping;
- 4.4. Coleta por automatização (Selenium);
- 4.5. Fontes de dados abertos.

5. Armazenamento:

- 5.1. Estrutura de dados (vetores, matrizes, arrays, data frames e listas);
- 5.2. Princípios de bancos de dados relacionais;
- 5.3. Linguagem SQL.

6. Processamento de dados:

- 6.1. Análise exploratória;
- 6.2. Limpeza;
- 6.3. Transformação;
- 6.4. Enriquecimento;
- 6.5. Manipulação (seleção, filtragem, e junção);

7. Conceitos de geoprocessamento:

- 7.1. O pacote sf em R;
- 7.2. Dados espaciais (pontos, linhas, polígonos e grades);
- 7.3. Operações básicas (união, interseção, recorte, junção espacial).

8. Visualização de dados:

- 8.1. Gráficos;
- 8.2. Mapas.

9. Criação de dashboards usando Shiny:

- 9.1. Introdução ao Shiny;
- 9.2. Construção da interface do usuário;
- 9.3. Princípios de HTML;
- 9.4. Lógica do servidor;
- 9.5. Publicação e compartilhamento de dashboards.

Metodologia:

A metodologia da disciplina combina aulas teóricas, práticas laboratoriais e projetos aplicados para promover uma compreensão introdutória da ciência de dados.

Critérios/Processo de Avaliação da Aprendizagem:

O critério de avaliação da disciplina baseia-se na metodologia de Aprendizado Baseado em Problemas (PBL), onde os alunos são avaliados por sua capacidade de aplicar os conceitos e ferramentas de ciência de dados para resolver problemas reais. Os trabalhos poderão ser realizados de forma individual ou em grupo. Os alunos que obtiverem nota parcial igual ou superior a 7,0 estarão dispensados da prova final.

Bibliografia Básica:

CARVALHO, André C. P. L. F. de; MENEZES, Angelo Garangau; BONIDIA, Robson Parmezan. **Ciência de dados: fundamentos e aplicações**. 1. ed., 2ª reimp. Rio de Janeiro: LTC, 2024.

WICKHAM, Hadley; GROLEMUND, Garrett. **R para Data Science: importe, arrume, transforme, visualize e modele dados**. Rio de Janeiro: Alta Books, 2021.

PIERSON, Lillian. **Data science para leigos**. Rio de Janeiro: Alta Books, 2019.

PUGA, Sandra Gavioli; FRANÇA, Edson Tarcísio; GOYA, Milton Roberto. **Banco de dados: implementação em SQL, PL/SQL e Oracle 11g**. São Paulo: Pearson, 2013.

Bibliografia Complementar:

AGUILAR, Luis Joyanes. **Fundamentos de programação algoritmos, estruturas de dados e objetos**. 3. Porto Alegre: AMGH, 2008.

GUILHERME, Luiz Fernando do Vale de Almeida. **Manual de proteção de dados: LGPD comentada**. 1. ed. São Paulo: Almedina, 2021.

HUA, Chee. **Artificial Intelligence, Analytics and Data Science: Core Concepts and Models**. Tradução técnica: Alex Marino Gonçalves de Almeida. Revisão: Fábio Gonçalves e Joana Figueiredo. 1. ed. São Paulo: Cengage, 2024.

LACERDA, Paulo Sergio Pádua de et al. **Programação em Big Data com R**. Revisão técnica: Gustavo Siqueira Vinhal e Fábio Josende Paz. Porto Alegre: SAGAH, 2021.

OLSEN, Wendy. **Coleta de dados: debates e métodos fundamentais em pesquisa social**. Tradução: Daniel Bueno. Revisão técnica: Dirceu da Silva. Porto Alegre: Penso, 2015.

Cronograma:

Prazo para definição do problema: 16/12/2024;

Prazo final para a entrega do trabalho: 17/03/2025;

Prova Final: 31/03/2025.