

# Programa Analítico de Disciplina

## EST 362 - Mineração de Dados

Departamento de Estatística - Centro de Ciências Exatas e Tecnológicas

Catálogo: 2026

Número de créditos: 4

Carga horária semestral: 60h

Carga horária semanal teórica: 2h

Carga horária semanal prática: 2h

Carga horária de extensão: 0h

Semestres: I, II e Especial

### Objetivos

Apresentar os fundamentos estatísticos e metodológicos da Mineração de Dados, capacitando o estudante a compreender, aplicar e interpretar métodos de extração de padrões em bases de dados, com ênfase em análise exploratória, associação, classificação, regressão, agrupamento e detecção de outliers, destacando aspectos de qualidade, interpretabilidade, validação e relevância dos padrões, bem como suas implicações éticas e sociais.

### Ementa

Introdução à Mineração de Dados. Conhecendo os Dados: Estatística e Visualização. Pré-processamento de Dados. Mineração de Padrões Frequentes e Associações. Classificação e Regressão em Mineração de Dados. Agrupamento e Detecção de Outliers. Avaliação, Aplicações e Aspectos Éticos.

### Pré e correquisitos

INF 217 e EST 111

### Oferecimentos obrigatórios

Curso	Período
Ciência de Dados	6

### Oferecimentos optativos

*Não definidos*

## EST 362 - Mineração de Dados

Conteúdo					
Unidade	T	P	ED	Pj	To
<b>1. Introdução à Mineração de Dados</b> 1. O que é Mineração de Dados e por que utilizá-la, Mineração de Dados no contexto da Ciência de Dados, Tipos de dados e bases de dados, Tipos de padrões: descritivos e preditivos, Estatística, Aprendizado de Máquina e Mineração de Dados.	3h	3h	0h	0h	6h
<b>2. Conhecendo os Dados: Estatística e Visualização</b> 1. Tipos de atributos e escalas de mensuração, Estatísticas descritivas aplicadas à mineração, Visualização de dados para descoberta de padrões, Medidas de similaridade e dissimilaridade, Distâncias para dados numéricos, categóricos e mistos.	5h	5h	0h	0h	10h
<b>3. Pré-processamento de Dados</b> 1. Qualidade dos dados e impacto na mineração, Tratamento de dados ausentes e ruidosos, Integração e redundância, Redução de dimensionalidade (PCA – visão conceitual), Normalização, discretização e transformação.	6h	6h	0h	0h	12h
<b>4. Mineração de Padrões Frequentes e Associações</b> Análise de cestas de mercado, Itemsets frequentes e regras de associação, Medidas: suporte, confiança e lift, interpretação estatística das associações, Limitações e armadilhas na descoberta de padrões.	5h	5h	0h	0h	10h
<b>5. Classificação e Regressão em Mineração de Dados</b> Classificação como tarefa preditiva, Árvores de decisão: interpretação e poda, Classificação Bayesiana (Naive Bayes), Regressão como ferramenta de mineração, Avaliação de modelos: validação cruzada, bootstrap.	5h	5h	0h	0h	10h
<b>6. Agrupamento e Detecção de Outliers</b> 1. Análise de clusters: objetivos e desafios, Métodos de agrupamento (k-means e hierárquico), Avaliação de agrupamentos, Outliers: definição e métodos estatísticos. Aplicações práticas.	4h	4h	0h	0h	8h
<b>7. Avaliação, Aplicações e Aspectos Éticos</b> O que torna um padrão interessante?, Validação estatística dos resultados, Aplicações em negócios, ciência e setor público, Privacidade, viés e impacto social, Mineração de Dados e LGPD.	2h	2h	0h	0h	4h
<b>Total</b>	<b>30h</b>	<b>30h</b>	<b>0h</b>	<b>0h</b>	<b>60h</b>

Teórica (T); Prática (P); Estudo Dirigido (ED); Projeto (Pj); Total (To);

Planejamento pedagógico	
Carga horária	Itens

A autenticidade deste documento pode ser conferida no site <https://siadoc.ufv.br/validar-documento> com o código: 6TMC.4C4D.7ZDH

Teórica	Uso do PVANET pelos alunos (arquivo sobre a disciplina: objetivos, sistema de avaliação, recursos didáticos a serem usados, bibliografia e plano de aulas)/(arquivo com os slides a serem usados em sala de aula e vários artigos técnicos, dentre eles cartilhas, livros, dissertações e teses); Aulas Expositivas Dialogadas, com a apresentação dos conteúdos teóricos e práticos de forma interativa, incentivando a participação e a discussão dos estudantes; apresentação de conteúdo pelo professor e apresentação interativa do conteúdo.; e Apresentação de conteúdo oral e escrito com apoio de equipamentos multimídia, realização de trabalhos práticos em computador, atividades extras no PVANet/Moodle e apresentação de conteúdo pelo professor e apresentação interativa do conteúdo.
Prática	Aulas em laboratório de informática e execução prática das atividades pelos próprios estudantes.; e Aprendizagem Baseada em Problemas, Elaboração de relatórios, Aulas em laboratório de informática, Oficinas Didáticas e Prática executada pelos próprios estudantes
Estudo Dirigido	<i>Não definidos</i>
Projeto	<i>Não definidos</i>
Recursos auxiliares	<i>Não definidos</i>

## EST 362 - Mineração de Dados

### Bibliografias básicas

Descrição	Exemplares
J. Han, M. Kamber, and J. Pei, Data Mining: Concepts and Techniques. Morgan Kaufmann, 3rd ed. , 2011	0
Charu C. Aggarwal, Data Mining: The Textbook, Springer, 2015	0
U. Fayyad, G. Grinstein, and A. Wierse, Information Visualization in Data Mining and Knowledge Discovery, Morgan Kaufmann, 2001	0

### Bibliografias complementares

Descrição	Exemplares
T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd ed., Springer, 2009	0
P.-N. Tan, M. Steinbach and V. Kumar, Introduction to Data Mining, Wiley, 2005 (2nd ed. 2016)	0
Mohammed J. Zaki and Wagner Meira Jr., Data Mining and Analysis: Fundamental Concepts and Algorithms 2014	0
DA COSTA CÔRTEZ, Sérgio; PORCARO, Rosa Maria; LIFSCHITZ, Sérgio. Mineração de dados-funcionalidades, técnicas e abordagens. PUC, 2002.	0
GOLDSCHMIDT, Ronaldo; PASSOS, Emmanuel; BEZERRA, Eduardo. Data mining. Elsevier Brasil, 2015.	0