



UNIVERSIDADE FEDERAL DE SANTA CATARINA – UFSC
CENTRO DE CIÊNCIAS DA EDUCAÇÃO – CED
DEPARTAMENTO DE CIÊNCIA DA INFORMAÇÃO – CIN
CAMPUS UNIVERSITÁRIO - TRINDADE
CEP: 88040-970 - FLORIANÓPOLIS - SC
Fone: (48) 3721-4075 E-mail: cin@contato.ufsc.br

PROGRAMA DE ENSINO

1 IDENTIFICAÇÃO

Disciplina: **CIN8000 – Introdução a Engenharia e Ciência de Dados**

Carga Horária: 36 h/a – 2 créditos (36 h/a teóricas)

Oferta: Obrigatória para a 1ª. Fase do Curso Superior de Tecnologia em Ciência de Dados

Pré-requisitos: **não há**

EMENTA

Fundamentos da área de Engenharia de Dados e Ciência de Dados. Aspectos sobre Integração de Dados e Análise de Dados.

2 OBJETIVOS

2.1 Objetivo Geral

Desenvolver habilidades introdutórias sobre a área de engenharia e ciência de dados, contextualizando problemas e soluções orientadas a dados.

2.2 Objetivos Específicos:

2.2.1 Compreender os conceitos básicos da área de Engenharia de Dados.

2.2.2 Compreender os conceitos básicos da área de Ciência de Dados.

2.2.3 Analisar problemas reais quanto a utilização de dados para sua solução.

2.2.4 Estudar a criação e utilização de conjuntos de dados.

3 CONTEÚDO PROGRAMÁTICO

3.1 Tipos de Dados

3.1.1 Dados estruturados

3.1.2 Dados semiestruturados

3.1.3 Dados não-estruturados

3.2 Engenharia de Dados

3.2.1 Persistência de dados

3.2.2 Criação e desenvolvimento de *pipelines*

3.2.3 Arquiteturas computacionais

3.2.4 Computação distribuída

3.3 Ciência de Dados

3.3.1 Coleta de dados

3.3.2 Preparação e limpeza de dados

3.3.3 Ferramentas para processamento de dados

3.3.4 Exploração de dados

3.3.5 Visualização de dados

4 BIBLIOGRAFIA

4.1 Bibliografia básica

DATE, C. J. Introdução a sistemas de bancos de dados. Rio de Janeiro: Campus, 2004. 865p. ISBN 8535212736.

GONZALEZ-AGUILAR, Audilio; PINTO, Adilson Luiz; SEMELER, Alexandre Ribas; SOARES, Ana Paula Alves. Visualização de dados, informação e conhecimento. Florianópolis: Ed. da UFSC, 2017. 210 p. ISBN 9788532807885.

PROVOST, Foster; FAWCETT, Tom. Data science for business. Sebastopol: O'Reilly, c2013. xxi, 386 p. ISBN 9781449361327.

4.2 Bibliografia complementar

O'NEIL, Cathy, and Rachel Schutt. Doing data science: Straight talk from the frontline. " O'Reilly Media, Inc.", 2013.

MEMBREY, Peter; PLUGGE, Eelco; HAWKINS, DUPTim. The Definitive Guide to MongoDB: The NoSQL Database for Cloud and Desktop Computing. 1st ed. 2010. Berkeley, CA: Apress: Imprint: Apress, 2010. xx, 328 p ISBN 9781430230526.

KIMBALL, Ralph; ROSS, Margy. The data warehouse toolkit: the complete guide to dimensional modeling. John Wiley & Sons, 2011.

KLEPPMANN, Martin. Designing data-intensive applications: The big ideas behind reliable, scalable, and maintainable systems. " O'Reilly Media, Inc.", 2017.

REIS, Joe, and Matt Housley. Fundamentals of Data Engineering. " O'Reilly Media, Inc.", 2022.