

# Ciência de Dados

**Licenciatura Engenharia Informática**  
**2º Semestre – 2021/2022**

Ricardo Jesus Ferreira  
[ricardojesus.ferreira@my.istec.pt](mailto:ricardojesus.ferreira@my.istec.pt)

# Objetivos

- Saber instalar e administrar o sistema de bases de dados relacionais SQL Server;
- Saber instalar e administrar o sistema de gestão de bases de dados relacionais SQL Server para grandes volumes de informação;
- Criar e aceder a bases de dados de grande dimensão e conteúdo em SQL Server;
- Saber instalar e administrar o sistema de bases de dados documentais MongoDB para Big Data;
- Saber utilizar o MongoDB num contexto de grande volume de dados;
- Saber instalar e administrar o sistema de bases de dados documentais Neo4j para dados de grande volume e dimensão;
- Compreender o modelo de dados e saber criar e aceder a bases de dados em Neo4j que armazenem dimensões volumosas de informação;
- Compreender Técnicas de Data Mining associadas a Big Data.

# Conteúdos Programáticos

- **SGBDR SQL Server**
  - Análise do sistema SQL Server
  - Implementação de bases de dados em SQL Server
  - Administração de bases de dados em SQL Server
  - Projeto prático usando SQL Server

# Conteúdos Programáticos

- **MongoDB**
  - Análise do sistema MongoDB
  - Implementação de bases de dados em MongoDB
  - Administração de bases de dados em MongoDB
  - Projeto prático usando MongoDB

# Conteúdos Programáticos

- **Neo4j**
  - Análise do sistema Neo4j
  - Implementação de bases em Neo4j
  - Administração de bases de dados em Neo4j
  - Projeto prático usando Neo4j

# Conteúdos Programáticos

- **Data Mining**
  - O processo de data Mining
  - CRISP-DM (Cross-Industry Standard Processs for Data Mining)
  - Outras tecnologias de Data Mining
  - Tarefas de Data Mining
  - Classificação
  - Segmentação ou Agrupamento
  - Estimação ou Regressão
  - Associação ou Modelação de dependências
  - Técnicas de Data Mining
    - Árvores de decisão
    - Regressão Linear
  - Projeto prático

# Componente Laboratorial

- 1. Trabalho SQL Server**
- 2. Trabalho MongoDB**
- 3. Trabalho Neo4j**
- 4. Trabalho Data Mining**

# Requisitos

- **BD Relacionais (MySQL)**
  - Modelo EA
  - Modelo Relacional
  - Modelo Fisico
- **Python**



# Avaliação

- Quatro laboratórios (40%)
- Exame Final (60%)

# Bibliografia

- B. Gomez,(2020) “Resolviendo problemas de Big Data”, Alfaomega.
- D. Insua, (2019)“Big data: Conceptos, tecnologías y aplicaciones”, CSIC.
- H. Jones, (2019)“Analítica de datos”, HJ,.
- J. Somed, (2020)“Big Data Analytics”, JLC.
- D. Petković (2020)“Microsoft® SQL Server® 2019 A Beginner’s Guide - Seventh Edition”, McGraw Hill.