

CENTRO UNIVERSITÁRIO CTC
DEPARTAMENTO DE INFORMÁTICA



INF2420	Introdução à Ciência de Dados	
	Carga horária total: 45h	Créditos: 3
	Pré-requisito: n/a	

OBJETIVOS	Apresentar conceitos fundamentais de Ciência de Dados, com ênfase em pesquisa reprodutível. No curso os alunos tomam contato com as diversas técnicas para a execução de um projeto em Ciência de Dados. Exemplos de implementação são feitos em R e em Python.
JUSTIFICATIVA	A Ciência de Dados visa transformar os dados puros em informação e a informação em conhecimento. Essa área está com uma alta demanda. A presente disciplina foi criada com o fim de apresentar aos alunos técnicas para implementar um projeto em Ciências de Dados de uma forma reprodutível com o uso de relatórios automáticos, por exemplo. Já foi oferecida por três vezes e em cada edição contou com mais de 20 alunos de diferentes departamentos.
EMENTA	Introdução a pesquisa reprodutível. Aquisição automática de dados. Preparação de dados para análise. Exploração de dados. Métodos de modelagem. Elaboração de relatórios automáticos de resultados através de documentos com código embutido, incluindo gráficos e tabelas.
PROGRAMA	<ol style="list-style-type: none">1. Introdução a pesquisa reprodutível.2. Aquisição automática de dados.3. Preparação de dados para análise.4. Exploração de dados.5. Métodos de modelagem.6. Elaboração de relatórios automáticos de resultados através de documentos com código embutido, incluindo gráficos e tabelas.
AValiação	A avaliação será baseada em: <ul style="list-style-type: none">▪ protótipo desenvolvido (em qualquer linguagem ou ambiente)▪ monografia▪ participação e apresentações em aula
BIBLIOGRAFIA PRINCIPAL	<ul style="list-style-type: none">▪ Zumel, Nina, and John Mount. Practical data science with R, 2014.▪ Artigos selecionados
BIBLIOGRAFIA COMPLEMENTAR	<ul style="list-style-type: none">▪ Gandrud, Christopher. Reproducible research with R and RStudio. CRC Press, 2013.▪ Kuhn, Max, and Kjell Johnson. Applied predictive modeling. New York: Springer, 2013.▪ Lamigueiro, Oscar Perpinan. Displaying Time Series, Spatial, and Space-time Data with R. CRC Press, 2014.▪ Xie, Yihui. Dynamic Documents with R and knitr. CRC Press, 2013.