

# Introdução à análise de dados

## Atividade dois

Felipe Carvalho<sup>1</sup>, Felipe Menino<sup>1</sup>

<sup>1</sup>FATEC "Jessen Vidal"

`felipe.carvalho@inpe.br,`

`felipe.carlos@fatec.sp.gov.br`

***Resumo.** A aplicação dos conceitos vistos no mini-curso de introdução a análise de dados é extremamente importante para a fixação dos conteúdos. Esta é uma lista de atividades para auxiliar o aluno na fixação de todo o conteúdo.*

### 1. Contexto

A Universidade Federal do Oeste da Bahia (UFOB), disponibiliza dados nominais de alunos que trancaram suas matrículas na universidade. Os dados disponibilizados são referentes aos anos de 2014 à 2017.

Ao verificar estes dados, algumas dúvidas surgiram, como por exemplo, qual o curso tem mais registros, ou qual o ano.

Os são disponibilizados aqui: <http://dados.gov.br/dataset/estudantes-com-matricula-trancada>

Veja que, neste caso, não sabemos se todos os dados estão completamente atualizados, vamos apenas supor que eles estão completos.

### 2. Atividade

Como demonstrado ao longo das atividades do curso, a análise de dados é basicamente uma atividade utilizada para responder perguntas através de diferentes métodos de tratamento de manipulação dos dados.

Com base nisto, faça a análise do conjunto de dados **MIAD002**, que está no repositório do minicurso, dentro do diretório **atividades**.

Com este conjunto de dados, responda as seguintes questões:

- Qual o curso com maior quantidade de matrículas trancadas no ano de 2015 ?
- Faça um **plot** com a variação entre os anos do número de matrículas trancadas por curso.
- Qual semestre (de todos os anos) contém a maior taxa de matrículas trancadas ?