

Introdução à análise de dados

Felipe Carvalho *National Institute for Space Research (INPE), Brazil*

Felipe Menino *National Institute for Space Research (INPE), Brazil*

A aplicação dos conceitos vistos no minicurso de introdução à análise de dados é extremamente importante para a fixação dos conceitos. Desta forma, esta lista de exercícios busca criar situações que ajudem o aluno na aplicação e fixação de tais conceitos.

1. Contexto

Este conjunto de dados, representa uma serie temporal, do ano de 2001 até 2015, com dados relacionados ao número de mortes por transportes terrestres no Brasil. Os dados são disponibilizados pela [Sala de Apoio à Gestão Estratégica](#). Descrição dos dados:

Fazer a análise deste conjunto de dados, pode nos ajudar a entender os problemas de mortalidade relacionados a transportes terrestres.

2. Atividade

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

Com base nisto, faça a análise do conjunto de dados **MIAD001**, que está no repositório do minicurso, dentro do diretório **atividades**.

Com este conjunto de dados, responda as seguintes perguntas:

1. Qual o tipo de acidente que mais ocorreu independente do ano?
2. Qual ano apresentou a maior taxa geral de mortalidade por acidentes?
3. Qual ano apresentou a maior taxa de mortalidade em automóveis?
4. Qual o desvio padrão entre todos os anos, em acidentes sofridos por Motociclistas?

Dicas

Abaixo uma tabela com a relação de funções que podem ser úteis para a solução dos exercícios propostos

Descrição	R	Python
Leitura de dados	<code>readr::read_csv()</code>	<code>pd.read_csv()</code>
Agrupamento	<code>dplyr::group_by()</code>	<code>pd.DataFrame.groupby()</code>
Filtro	<code>dplyr::filter()</code>	<code>pd.DataFrame[Condição]</code>
Desvio padrão	<code>stats::sd()</code>	<code>pd.DataFrame.std()</code>
Média	<code>mean()</code>	<code>pd.DataFrame.mean()</code>
Mínimo	<code>min()</code>	<code>pd.DataFrame.min()</code>
Máximo	<code>max()</code>	<code>pd.DataFrame.max()</code>