

Análise de dados do Enem 2015

Evandro Carlos B. dos Santos
Giuliard Cosmo Rodrigues



Introdução

O dataset escolhido para realização do trabalho foi extraído de um arquivo CSV do Enem 2015 que foi disponibilizado no Kaggle pelo usuário Gustavo Bonesso. Os dados fornecem informações detalhadas dos alunos dos estados do centro-oeste (Mato Grosso do Sul, Mato Grosso, Goiás e Distrito Federal).

Colunas do arquivo

- ▶ NU_INSCRICAO
- ▶ NU_ANO
- ▶ CO_MUNICIPIO_RESIDENCIA
- ▶ CO_UF_RESIDENCIA
- ▶ SG_UF_RESIDENCIA
- ▶ IN_ESTUDA_CLASSE_HOSPITALAR
- ▶ IN_TREINEIRO
- ▶ CO_ESCOLA
- ▶ CO_MUNICIPIO_ESC
- ▶ CO_UF_ESC
- ▶ SG_UF_ESC
- ▶ TP_DEPENDENCIA_ADM_ESC
- ▶ TP_LOCALIZACAO_ESC
- ▶ TP_SIT_FUNC_ESC
- ▶ NU_IDADE
- ▶ TP_SEXO
- ▶ TP_NACIONALIDADE
- ▶ CO_MUNICIPIO_NASCIMENTO
- ▶ CO_UF_NASCIMENTO
- ▶ SG_UF_NASCIMENTO
- ▶ TP_ST_CONCLUSAO
- ▶ TP_ANO_CONCLUIU
- ▶ TP_ESCOLA
- ▶ TP_ENSINO
- ▶ TP_ESTADO_CIVIL
- ▶ TP_COR_RACA
- ▶ IN_MATERIAL_ESPECIFICO
- ▶ IN_CERTIFICADO
- ▶ CO_UF_ENTIDADE_CERTIFICACAO
- ▶ SG_UF_ENTIDADE_CERTIFICACAO
- ▶ TP_PRESENCA_CN
- ▶ TP_PRESENCA_CH
- ▶ TP_PRESENCA_LC
- ▶ TP_PRESENCA_MT
- ▶ CO_PROVA_CN
- ▶ CO_PROVA_CH
- ▶ CO_PROVA_LC
- ▶ CO_PROVA_MT
- ▶ NU_NOTA_CN
- ▶ NU_NOTA_CH
- ▶ NU_NOTA_LC
- ▶ NU_NOTA_MT
- ▶ TP_LINGUA
- ▶ TP_STATUS_REDACAO
- ▶ NU_NOTA_COMP1
- ▶ NU_NOTA_COMP2
- ▶ NU_NOTA_COMP3
- ▶ NU_NOTA_COMP4
- ▶ NU_NOTA_COMP5
- ▶ NU_NOTA_REDACAO

Temas Abordados

- ▶ Porcentagem de homens e mulheres que participaram do exame;
- ▶ Número de candidatos por estado;
- ▶ Idade média dos candidatos;
- ▶ Nota média em cada área de conhecimento.

Infraestructura


Spark

 **mongoDB**®

