Trabalho Individual 1: Mineração de Dados Descritiva Aplicada
GSI055 - Mineração de dados
Prof. Marcelo Keese Albertini

Neste trabalho devem ser usadas técnicas de pré-processamento, visualização e mineração de regras para estudar o dataset sobre "Cessões de Direitos Minerários" disponível em: `https://app.dnpm.gov.br/DadosAbertos/SCM/Cessoes_de_Direitos.csv`.

O objetivo deste trabalho é responder às seguintes perguntas utilizando tabelas, estatísticas-resumo, técnicas de visualização e mineração de regras de associação.

- Qual é a média de concessões de direitos de processos minerários por ano, por estado?
- Quais são os estados que mais e menos concedem direitos minerários?
- Quais as substâncias mais procuradas nos últimos 2, 5 e 10 anos?
- Quais são as empresas que recebem um número de concessões muito superior a outras?
- Quais são as empresas que têm monopólios/duopólios nacionais de extração de substâncias?
- Existe regiões de concentração geográfica de concessões?
  - Dica: usar o arquivo contendo as coordenadas geográficas das cidades brasileiras [link] para fazer mapa das concessões
- Quais são as associações mais fortes em relação à exploração de diferentes minérios para uma mesma empresa?
- Quais são as associações mais surpreendentes em relação à exploração de diferentes minérios para uma mesma empresa?

Ao usar esse arquivo deve-se trocar D"por D para não haver erros ao carregar os dados no R. Isso pode ser feito de seguinte forma:

- `wget https://app.dnpm.gov.br/DadosAbertos/SCM/Cessoes_de_Direitos.csv -O tmp.csv`
- `iconv tmp.csv -f iso8859-1 -t utf8//TRANSLIT > tmp2.csv`
- `sed "s/SERRA D\" /SERRA D/" tmp2.csv > direitos.csv`

O arquivo resultante pode ser carregado no R da seguinte forma:

```
require(dplyr)
direitos <- read.csv("direitos.csv")
str(direitos)
```

```
## 'data.frame': 27394 obs. of  10 variables:
##  $ Superintendência    : Factor w/ 25 levels "Superintendência / AL",..: 24 9 24 24 24 24 24 24 24
##  $ Processo            : Factor w/ 27363 levels "000022/1994",..: 7268 15569 8060 8172 8308 8309 8
##  $ Tipo.de.requerimento: Factor w/ 16 levels "Requerimento de Autorização de Pesquisa",..: 16 1 1
##  $ Fase.Atual          : Factor w/ 8 levels "Autorização de Pesquisa",..: 4 1 1 5 5 5 5 5 4 4 ...
##  $ CPF.CNPJ.do.titular : Factor w/ 10572 levels "00.004.212/0001-47",..: 1 1 3 3 3 3 3 3 4 4 ...
##  $ Titular             : Factor w/ 10552 levels "13janx Comércio de Águas Ltda Epp",..: 9541 9541
##  $ Municipio.s.        : Factor w/ 6066 levels "","ABADIA DOS DOURADOS - MG",..: 5160 3481 1520 15
##  $ Substância.s.       : Factor w/ 1748 levels "","AGALMATOLITO",..: 713 682 167 536 167 167 167 1
##  $ Tipo.s..de.Uso      : Factor w/ 513 levels "","Abrasivo",..: 483 409 483 483 195 195 195 195 19
##  $ Situação            : Factor w/ 1 level "Sim": 1 1 1 1 1 1 1 1 1 1 ...
```

O relatório deste trabalho deve ser feito usando o pacote `knitr` em conjunto com LaTeX em um arquivo ".Rnw". Arquivos Rnw são escritos em LaTeX com comandos em R dentro de uma sintaxe iniciando em uma linha com `<<>>=` e terminando em "@".

```
# comandos do R podem ser inseridos aqui
```

Usar os seguintes comandos de terminal para transformar um arquivo ".Rnw" em "PDF":

- `echo 'knitr::knit("trabalho1.Rnw")' | R --no-save`
- `pdflatex trabalho1.tex`

O relatório em Rnw e PDF deverá ser até 19/05.