

Relatório CAGED (tabelas) automatizado

Fábio Rocha* | rochaviannafj@gmail.com

16 de outubro, 2023

Introdução

Mensalmente, o Ministério do Trabalho divulga um relatório contendo dados sobre as movimentações no mercado de trabalho formal regido pela Consolidação das Leis do Trabalho (CLT). Estes dados estão disponíveis em dois formatos: microdados acessíveis via FTP e tabelas em formato Excel disponibilizadas diretamente no portal do Ministério.

Trabalho com essas tabelas há quase dois anos e enfrentei desafios consideráveis em compreender a complexidade da formatação utilizada. As tabelas são frequentemente complicadas devido ao uso extensivo de células mescladas, o que pode prejudicar a eficiência do trabalho, especialmente para analistas iniciantes.

Assim tomei a iniciativa de desenvolver um script em R que automatiza integralmente o processo de obtenção, carregamento, tratamento, visualização de dados em gráficos, geração de tabelas e criação de um relatório em formato PDF para qualquer cidade do Brasil.

Embora tenha demandado um esforço considerável, o resultado me agradou. A criação deste código de automação proporcionou uma notável melhoria na minha produtividade. Aquilo que costumava me consumir cerca de 40 minutos no início, hoje é resolvido com um único clique.

Pacotes necessários

```
library(readxl)
library(dplyr)
library(tidyverse)
library(gt)
```

Download dos dados via link

```
url <- "http://pdet.mte.gov.br/images/Novo_CAGED/2023/202308/3-tabelas.xlsx"
destino <- "P:/DESENVOLVIMENTO ECONOMICO/02 - Publico/ESTUDOS ECONÔMICOS/Trabalhos/Estudos Ouro Preto/M"
download.file(url, destfile = destino, mode = "wb")
```

*Economista pela Universidade Federal de Ouro Preto

The screenshot shows an Excel spreadsheet with the following data structure:

UF	Código do Município	Município	Janeiro/2020				Fevereiro/2020				Variação Relativa (%)	Estoque
			Estoque	Admissões	Desligamentos	Saldo	Estoque	Admissões	Desligamentos	Saldo		
RO	110001	Ro-Alta Floresta D Oeste	1.506	62	74	-12	1.512	79	73	6	0,40	1,4
RO	110002	Ro-Ariquemes	15.009	675	672	3	15.091	727	645	82	0,55	15
RO	110003	Ro-Cabixi	355	45	8	37	357	18	16	2	0,56	3
RO	110004	Ro-Cacoi	15.387	787	839	-52	15.579	942	750	192	1,25	15
RO	110005	Ro-Cerejeiras	1.792	148	65	83	1.819	91	64	27	1,51	1,7
RO	110006	Ro-Colorado do Oeste	1.593	44	75	-31	1.592	71	72	-1	-0,06	1,5
RO	110007	Ro-Corumbiara	969	84	49	35	988	75	56	19	1,96	9
RO	110008	Ro-Costa Marques	414	14	8	6	410	8	12	-4	-0,97	4
RO	110009	Ro-Espigão D Oeste	3.793	158	186	-28	3.790	152	155	-3	-0,08	3,7
RO	110010	Ro-Guajará-Mirim	2.427	92	89	3	2.422	70	75	-5	-0,21	2,4
RO	110011	Ro-Jaru	8.929	348	391	-43	8.838	255	348	-93	-1,04	8,8
RO	110012	Ro-Ji-Paraná	25.835	1.179	1.202	-23	25.664	1.286	1.457	-171	-0,66	25
RO	110013	Ro-Machadinho D Oeste	2.147	105	119	-14	2.185	118	80	38	1,77	2
RO	110014	Ro-Nova Brasilândia D Oeste	1.075	37	39	-2	1.093	56	38	18	1,67	1
RO	110015	Ro-Ouro Preto do Oeste	4.527	181	195	-14	4.556	217	188	29	0,64	4,5
RO	110016	Ro-Pimenta Buena	8.397	474	567	-93	8.534	457	320	137	1,63	8,3

Figure 1: Tabela CAGED

Carregamento dos dados

```
caged <- read_xlsx("P:/DESENVOLVIMENTO ECONOMICO/02 - Publico/ESTUDOS ECONÔMICOS/Trabalhos/Estudos Ouro Preto/2023/01 - CAGED/2023/01 - CAGED.xlsx")
```

Alguns tratamentos

```
caged_filtrado <- caged[2:5580, 3:ncol(caged)]

# Colunas que você deseja converter para numéricas
colunas_para_converter <- 2:ncol(caged_filtrado)

# Use a função lapply para converter as colunas
caged_filtrado[colunas_para_converter] <- lapply(caged_filtrado[colunas_para_converter], as.numeric)
```

Pivotagem do dataset

```
dados_pivotado <- pivot_longer(caged_filtrado,
                               cols = colnames(caged_filtrado)[2:ncol(caged_filtrado)],
                               names_to = "variavel",
                               values_to = "valor")
```

Renomear colunas e filtragem para o município desejado

```
dados_pivotado <- dados_pivotado %>%  
  rename(municipio = ...3)  
  
caged_regiao = dados_pivotado %>%  
  filter(municipio %in% c("Mg-Ouro Preto")) # Escolha a cidade e escreva neste mesmo formato  
  
caged_regiao_final <- caged_regiao %>%  
  slice(1:(n() - 8))
```

Remoção de números da variável que contem “admissão, estoque, etc”

```
# remover numeros e pontos da variavel  
caged_regiao_final <- caged_regiao_final %>%  
  mutate(variavel = gsub("[0-9.]", "", variavel))
```

Filtragem da informação “Variação Relativa (%) e criação da variável data”

```
# remoção de variação relativa  
caged_regiao_final <- caged_regiao_final %>%  
  filter(variavel != "Variação Relativa (%)")  
  
# criação da variavel data  
caged_regiao_final <- caged_regiao_final %>%  
  mutate(data = rep(seq(ymd("2020-01-01"), by = "1 month", length.out = n() / 4), each = 4))  
  
caged_regiao_final <- caged_regiao_final %>%  
  filter(variavel != "Estoque****\r\n(Estoque de referência de )")
```

Gráfico de movimentações e estoque

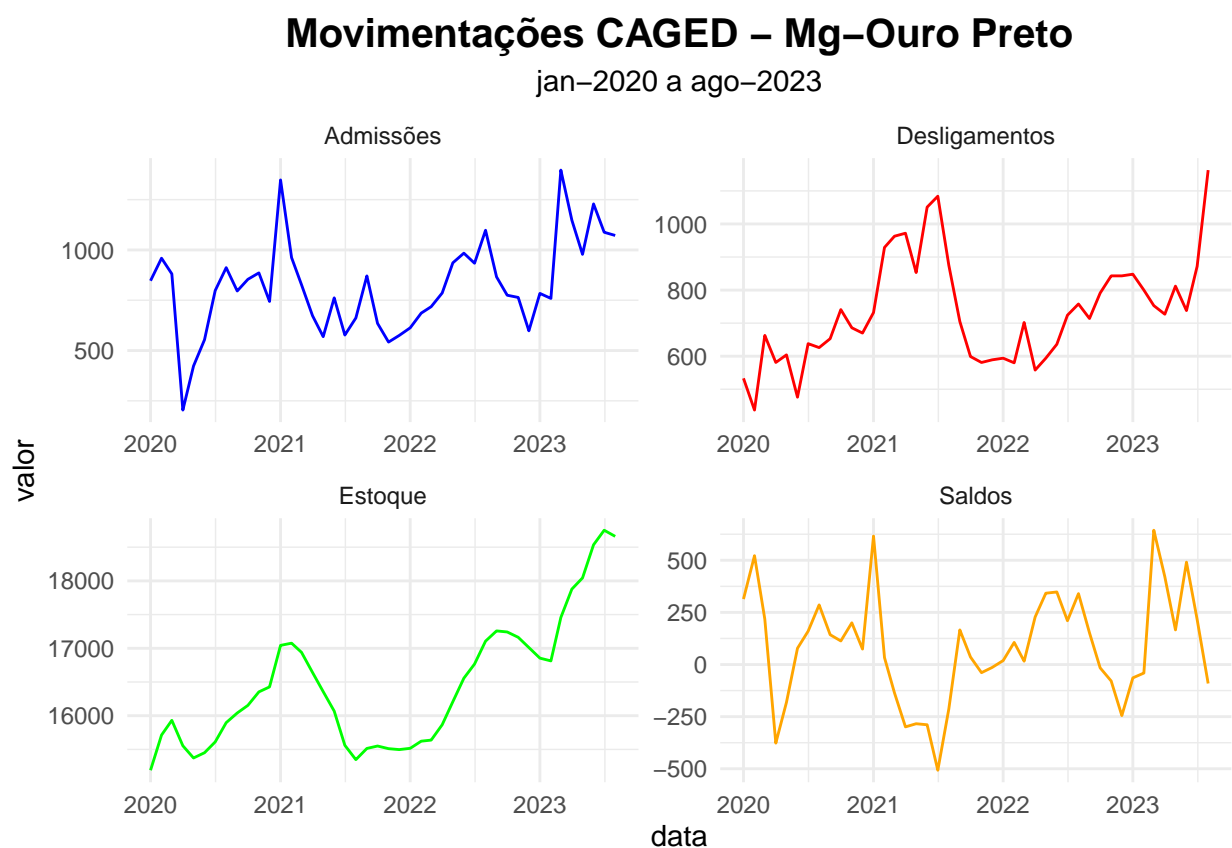
```
# Defina uma paleta de cores  
cores <- c("blue", "red", "green", "orange", "purple")  
  
# Crie o gráfico com cores diferentes para cada gráfico e destaque o título  
grafico <- ggplot(caged_regiao_final, aes(x = data, y = valor, color = variavel)) +
```

```

geom_line() +
theme_minimal() +
facet_wrap(~variavel, scales = "free") +
scale_color_manual(values = cores) + # Atribui cores da paleta
labs(subtitle = "jan-2020 a ago-2023") +
theme(legend.position = "none") # Ajusta a posição da legenda

# Adicione um título destacado
cidade <- caged_regiao_final[1, 1] # Acessa o nome da cidade
grafico +
ggtitle(paste("Movimentações CAGED -", cidade), subtitle = "jan-2020 a ago-2023") +
theme(plot.title = element_text(hjust = 0.5, size = 15, face = "bold")) +
theme(plot.subtitle = element_text(hjust = 0.5))

```



Tabela

```

tabela = tail(caged_regiao_final, 8)
tabela = tabela[,2:4]

tabela$data <- format(tabela$data, format = "%B-%Y")

```

```
tabela %>%
  gt(groupname_col = "data") %>%
  tab_header(title = md("**Movimentações CAGED**"),
             subtitle = paste(cidade, " - (julho-2023 agosto-2023)")) %>%
  tab_source_note(md("**Dados**: CAGED | **Elaboração**: Fábio Rocha")) %>%
  cols_label(variavel = "Movimentação/Estoque",
             valor = "Quantidade",
             data = "Período")
```

Movimentações CAGED

Mg-Ouro Preto - (julho-2023 agosto-2023)

Movimentação/Estoque	Quantidade
julho-2023	
Estoque	18751
Admissões	1088
Desligamentos	872
Saldos	216
agosto-2023	
Estoque	18660
Admissões	1072
Desligamentos	1163
Saldos	-91

Dados: CAGED | Elaboração: Fábio Rocha