Government Travel Expense Analysis - 2019

Luiza Constâncio

2026-02-16

BUSINESS TRAVEL EXPENSE ANALYSIS - 2019

Data Source: Brazilian Transparency Portal

1. Problem Definition

Objective: Analyze government travel expenses to identify spending patterns and seasonal trends.

Business Questions: 1. What is the total amount spent per government department? 2. What is the total amount spent per city? 3. How many trips occurred per month?

2. Data Acquisition

```
travel <- read.csv(
  "C:/Users/Luiza/Downloads/2019_Viagem.csv",
  sep = ";",
  dec = ",",
  fileEncoding = "latin1"
)

dim(travel)
```

```
## [1] 756704     16
```

The dataset contains 756704 records and 16 variables.

3. Data Transformation

```
travel$data_inicio <- as.Date(
  travel$Período...Data.de.início,
  format = "%d/%m/%Y"
)

travel$year_month <- format(travel$data_inicio, "%Y-%m")
```

4. Exploratory Data Analysis (EDA)

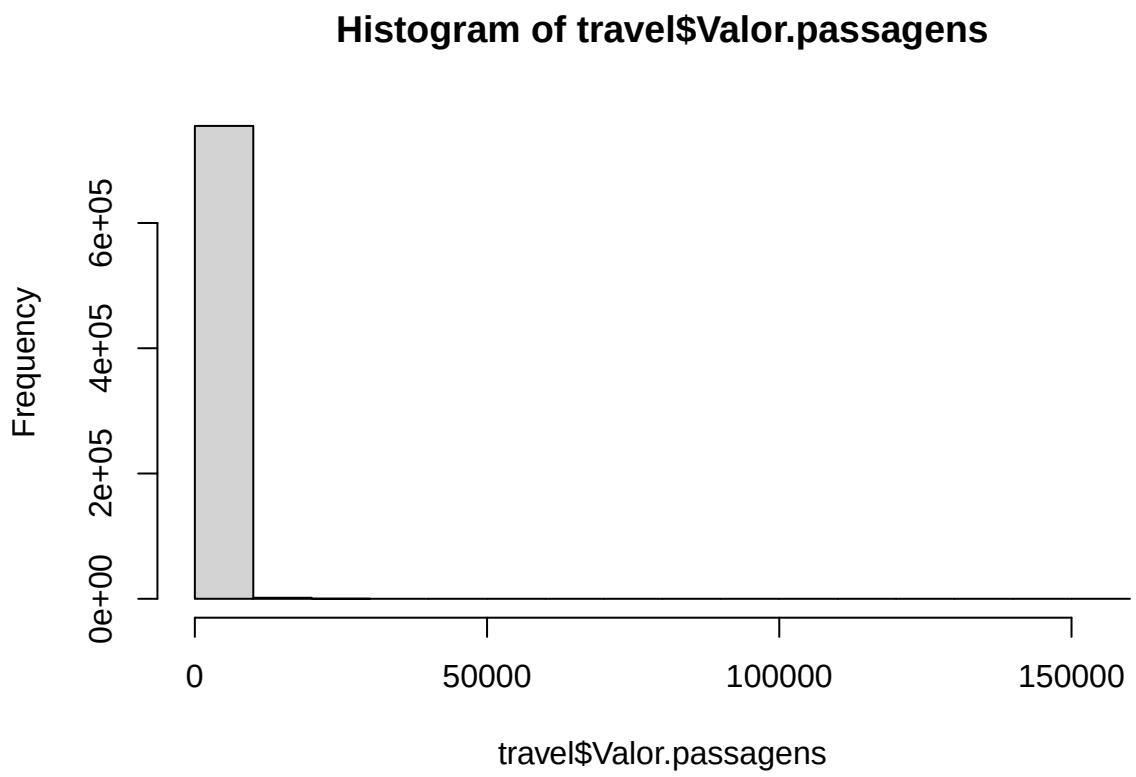
```
summary(travel$Valor.passagens)
```

```
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
##      0.0      0.0      0.0    590.1    907.0 155531.4
```

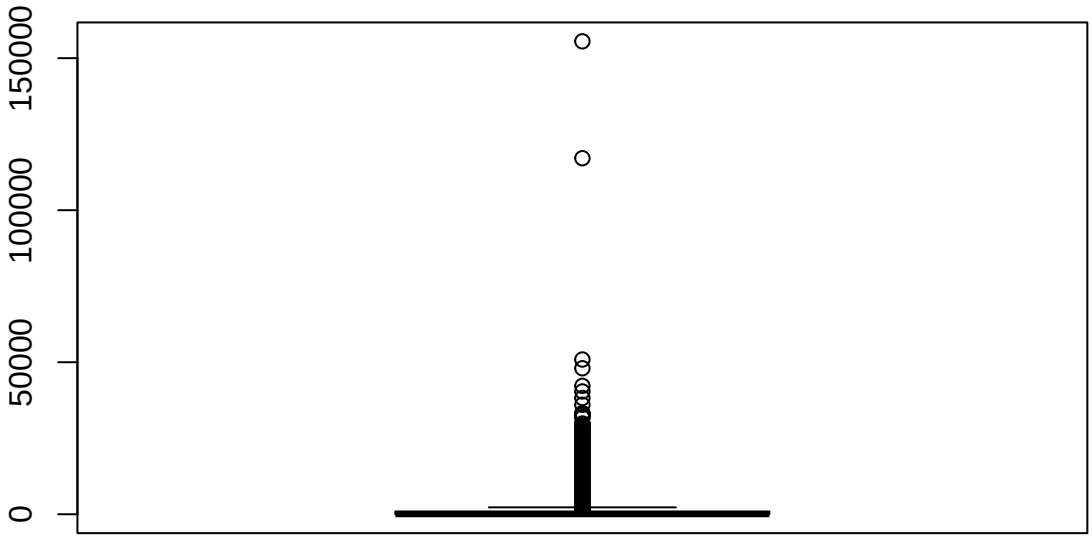
```
sd(travel$Valor.passagens)
```

```
## [1] 1278.517
```

```
hist(travel$Valor.passagens)
```




```
boxplot(travel$Valor.passagens)
```




```
colSums(is.na(travel))
```

```
## Identificador.do.processo.de.viagem                            Situação
##                                   0                                   0
##            Código.do.órgão.superior             Nome.do.órgão.superior
##                                   0                                   0
##            Código.órgão.solicitante             Nome.órgão.solicitante
##                                   0                                   0
##                       CPF.viajante                                Nome
##                                   0                                   0
##                              Cargo           Período...Data.de.início
##                                   0                                   0
##               Período...Data.de.fim                            Destinos
##                                   0                                   0
##                             Motivo                     Valor.diárias
##                                   0                                   0
##                     Valor.passagens                Valor.outros.gastos
##                                   0                                   0
##                         data_inicio                          year_month
##                                   0                                   0
```

```
travel$Situação <- factor(travel$Situação)
prop.table(table(travel$Situação)) * 100
```

```
##
```

```
## Não realizada     Realizada
##      2.293499     97.706501
```

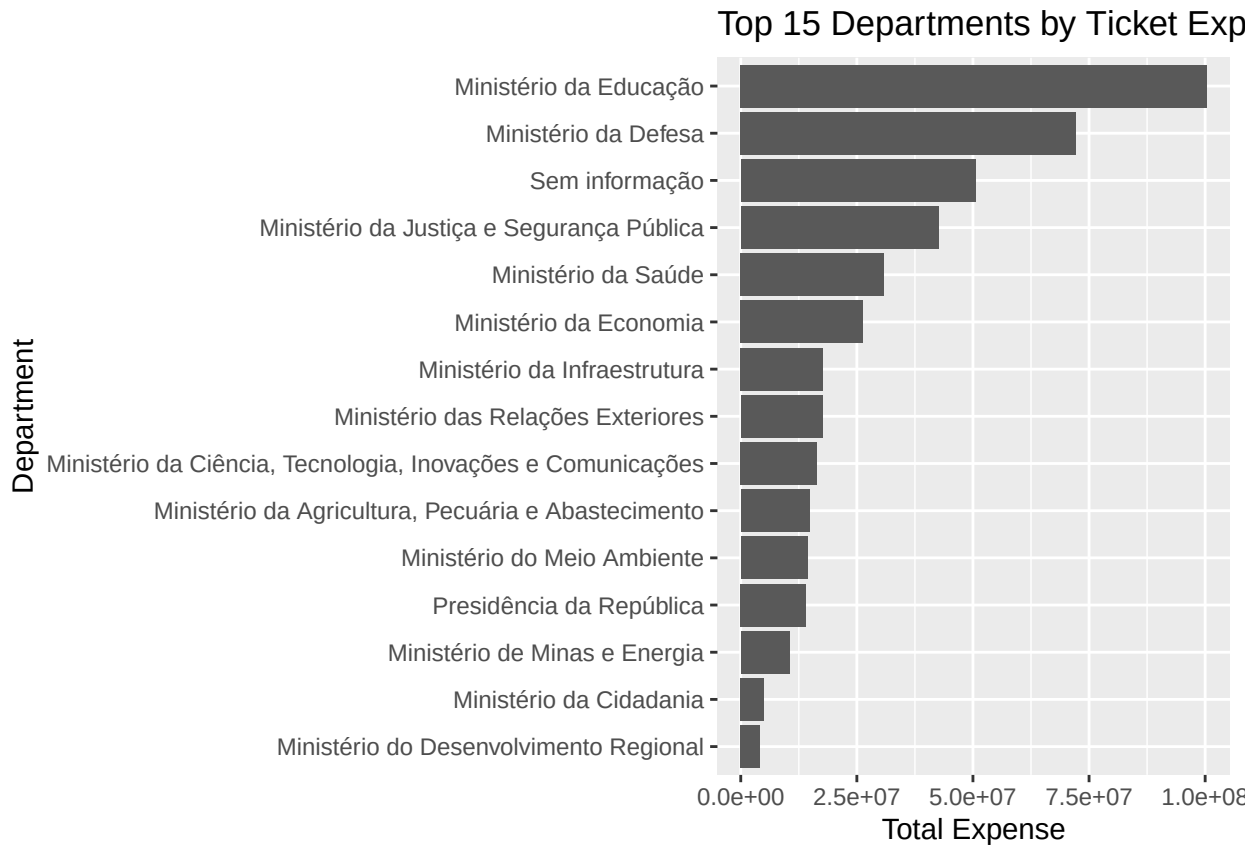
5. Results

5.1 Total Ticket Expenses by Government Department

```
expenses_by_department <- travel %>%
  group_by(Nome.do.órgão.superior) %>%
  summarise(total_expense = sum(Valor.passagens, na.rm = TRUE)) %>%
  arrange(desc(total_expense))

top15_departments <- expenses_by_department %>%
  slice_head(n = 15)

top15_departments
```

```
## # A tibble: 15 x 2
##    Nome.do.órgão.superior                                        total_expense
##    <chr>                                                                 <dbl>
##  1 Ministério da Educação                                          100414641.
##  2 Ministério da Defesa                                             72166086.
##  3 Sem informação                                                   50510733.
##  4 Ministério da Justiça e Segurança Pública                        42574181.
##  5 Ministério da Saúde                                              30867191.
##  6 Ministério da Economia                                           26361587.
##  7 Ministério da Infraestrutura                                     17746202.
##  8 Ministério das Relações Exteriores                               17630741.
##  9 Ministério da Ciência, Tecnologia, Inovações e Comunicações      16428722.
## 10 Ministério da Agricultura, Pecuária e Abastecimento              14840819.
## 11 Ministério do Meio Ambiente                                      14419791.
## 12 Presidência da República                                         14037021.
## 13 Ministério de Minas e Energia                                    10492274.
## 14 Ministério da Cidadania                                           4991945.
## 15 Ministério do Desenvolvimento Regional                            4104182.
```

```
ggplot(top15_departments,
       aes(x = reorder(Nome.do.órgão.superior, total_expense),
           y = total_expense)) +
  geom_bar(stat = "identity") +
  coord_flip() +
  labs(
    title = "Top 15 Departments by Ticket Expenses",
    x = "Department",
    y = "Total Expense"
  )
```



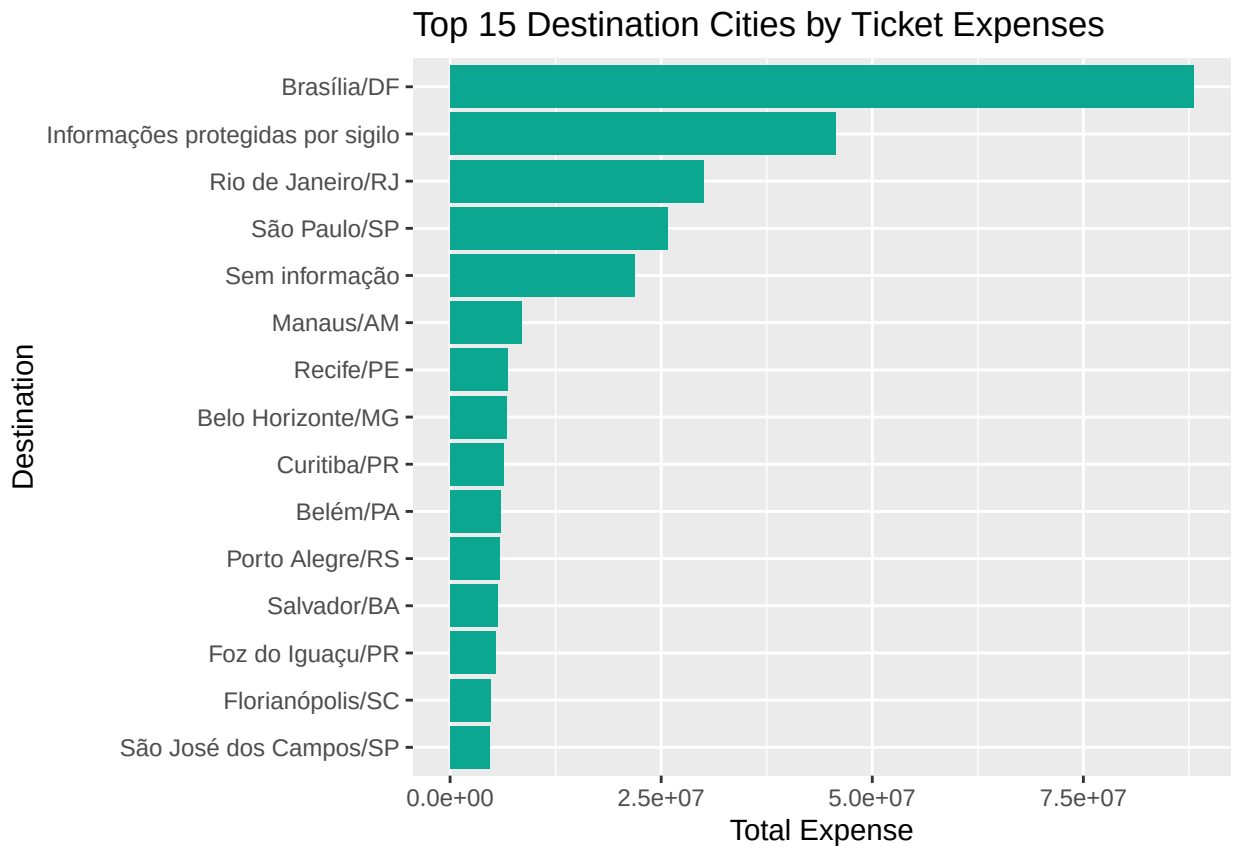
5.2 Total Ticket Expenses by Destination City

```
expenses_by_city <- travel %>%
  group_by(Destinos) %>%
  summarise(total_expense = sum(Valor.passagens, na.rm = TRUE)) %>%
  arrange(desc(total_expense))

top15_cities <- expenses_by_city %>%
  slice_head(n = 15)

top15_cities
```

```
## # A tibble: 15 x 2
##    Destinos                          total_expense
##    <chr>                                     <dbl>
##  1 Brasília/DF                            88009638.
##  2 Informações protegidas por sigilo      45675238.
##  3 Rio de Janeiro/RJ                      29999886.
##  4 São Paulo/SP                           25764054.
##  5 Sem informação                         21891215.
##  6 Manaus/AM                               8463212.
##  7 Recife/PE                               6833241.
##  8 Belo Horizonte/MG                       6691385.
```

```
##  9 Curitiba/PR                          6388793.
## 10 Belém/PA                             5982406.
## 11 Porto Alegre/RS                      5921046.
## 12 Salvador/BA                          5634670.
## 13 Foz do Iguaçu/PR                     5359965.
## 14 Florianópolis/SC                     4746875.
## 15 São José dos Campos/SP               4709705.
```

```r
ggplot(top15_cities,
       aes(x = reorder(Destinos, total_expense),
           y = total_expense)) +
  geom_bar(stat = "identity", fill = "#0ba791") +
  coord_flip() +
  labs(
    title = "Top 15 Destination Cities by Ticket Expenses",
    x = "Destination",
    y = "Total Expense"
  )
```



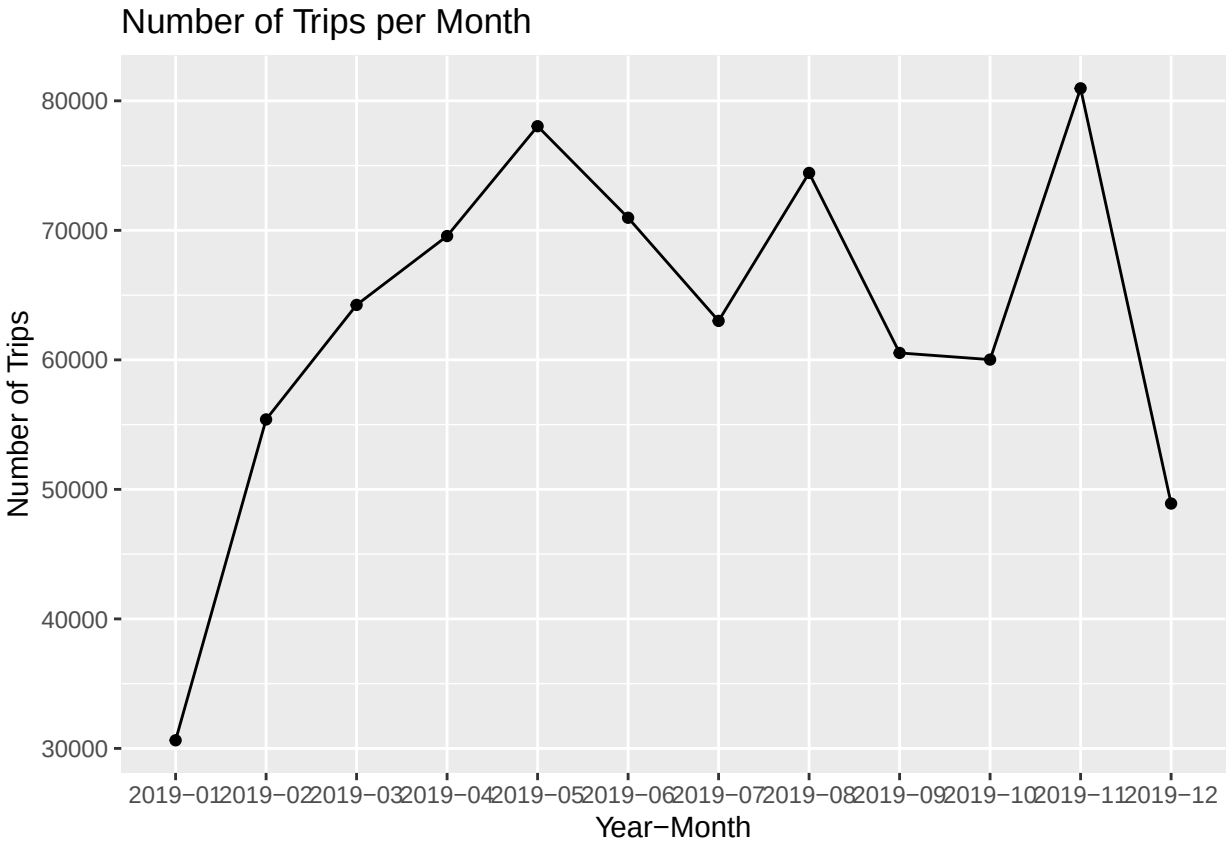
5.3 Number of Trips per Month

```r
trips_per_month <- travel %>%
  group_by(year_month) %>%
```

```r
  summarise(trip_count = n_distinct(Identificador.do.processo.de.viagem)) %>%
  arrange(year_month)

trips_per_month
```

```
## # A tibble: 12 x 2
##    year_month trip_count
##    <chr>           <int>
##  1 2019-01         30626
##  2 2019-02         55399
##  3 2019-03         64242
##  4 2019-04         69559
##  5 2019-05         78039
##  6 2019-06         70971
##  7 2019-07         63009
##  8 2019-08         74432
##  9 2019-09         60535
## 10 2019-10         60021
## 11 2019-11         80965
## 12 2019-12         48906
```

```r
ggplot(trips_per_month,
       aes(x = year_month,
           y = trip_count,
           group = 1)) +
  geom_line() +
  geom_point() +
  labs(
    title = "Number of Trips per Month",
    x = "Year-Month",
    y = "Number of Trips"
  )
```



6. Key Insights

- The dataset contains 756,704 travel records.
- 97.7% of trips were completed.
- The average ticket expense per trip was 590.1, with a high standard deviation of 1278.5.
- The Ministry of Education was the highest-spending department, exceeding 100 million in ticket expenses.
- Brasília/DF was the most expensive destination, totaling approximately 88 million.
- November recorded the highest number of trips, while January had the lowest.

Overall, the analysis reveals spending concentration patterns, strong geographic centralization, and seasonal travel variation.