

# Quiz 1 Dataset Flights

David Moreno

2024-01-30

## Punto 1

- Explique cada una de las variables que están incluidas en el dataset. ¿Cuántas variables y cuántas observaciones tiene el dataset? ¿Qué tipo de datos tiene? Resuma la información del dataset y muestre una figura utilizando las primeras 100 observaciones. ¿Ve alguna tendencia o correlación relevante? Confirme este resultado usando una función de correlación de Pearson.

## Descripción del Dataset

- **year, month, day:** Fecha del vuelo.
- **dep\_time, arr\_time:** Hora de salida y llegada.
- **sched\_dep\_time, sched\_arr\_time:** Hora programada de salida y llegada.
- **dep\_delay, arr\_delay:** Retraso en salida y llegada.
- **carrier:** Código de la aerolínea.
- **flight:** Número de vuelo.
- **tailnum:** Número de registro del avión.
- **origin, dest:** Aeropuertos de origen y destino.
- **air\_time:** Tiempo en el aire.
- **distance:** Distancia del vuelo.
- **hour, minute:** Hora y minuto de salida programada.
- **time\_hour:** Fecha y hora de salida programada.

```
# Cargar paquetes necesarios
```

```
library(nycflights13)
```

```
library(dplyr)
```

```
##
```

```
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
```

```
##
```

```
##      filter, lag
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
##      intersect, setdiff, setequal, union
```

```
library(ggplot2)
```

```
# Cargar el dataset
```

```
data("flights", package = "nycflights13")
```

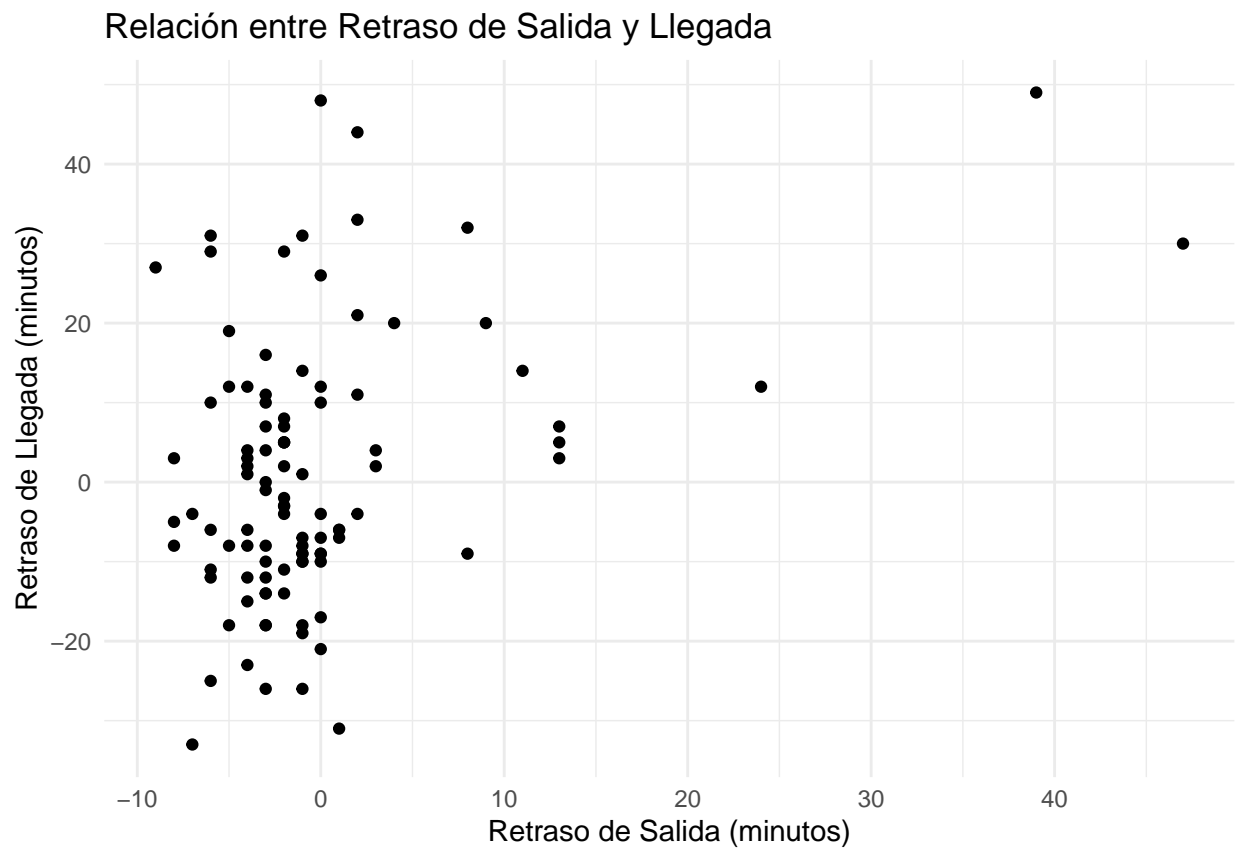
```
# Número de variables y observaciones
```

```
cat("Número de variables:", ncol(flights), "\n")
```

```
## Número de variables: 19
```

```
cat("Número de observaciones:", nrow(flights), "\n")

## Número de observaciones: 336776
# Análisis exploratorio con las primeras 100 observaciones
flights_100 <- head(flights, 100)
ggplot(flights_100, aes(x = dep_delay, y = arr_delay)) +
  geom_point() +
  theme_minimal() +
  labs(title = "Relación entre Retraso de Salida y Llegada",
       x = "Retraso de Salida (minutos)",
       y = "Retraso de Llegada (minutos)")
```



## Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.