
Rutas Aéreas y Grafos con R

J. Javier Moralo

Presentación

Jesús Javier Moralo García

- Bioinformático por la UAM & CSIC
- Máster Big Data & Analytics de Datahack
- Gestión, Análisis e Integración de Datos
 - Global Biodiversity Information Facility (www.gbif.org)
 - Plataforma Europea Cibertaxonomía(<http://cybertaxonomy.eu>)
 - Proyecto Europeo para la Integración de Colecciones de Historia Natural (www.synthesys.info)
- Dirección de equipos Técnicos
- Data & IA Creative en datahack_
 - Diseño y Desarrollo de Bots
 - Proyectos en Big Data & Analytics

 <https://www.linkedin.com/in/jesusjaviermoralo/>

 @JJavierMoralo

 <https://github.com/javiermoralo>



Imagen de www.pinterest.com

Programa

- Introducción
- Conceptos básicos
 - I. Carga de datos
 - II. Georeferenciar aeropuertos
 - III. Mapas de rutas
 - IV. Diagrama de grafos



Imagen de Silvia Rabanete

Funcionalidades en R

- Clasificación y Regresión
- Minería de datos y análisis de textos
- Modelos de scoring y ranking
- Análisis de grafos
- Análisis de datos interactivo
- Cálculos Estadísticos
- Modelos predictivos avanzados:
 - Machine Learning
 - Deep Learning

Básico - Variables

- Permiten utilizar un nombre para almacenar un valor
- Se utiliza el operador de asignación `<-`
- Un nombre es mayúsculas no es lo mismo que en minúsculas (R es case sensitive)

```
> x <- 1  
> |
```

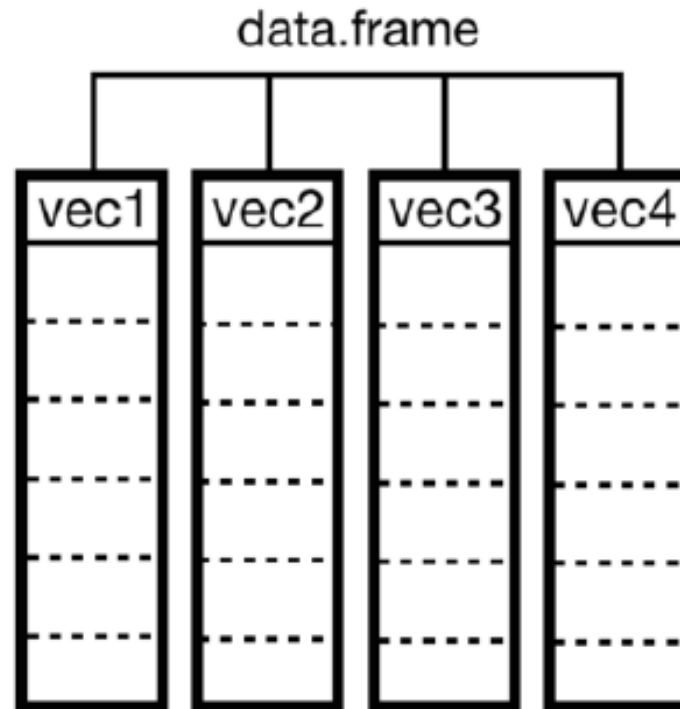
Básico - Vector

- La estructura de datos más básica con la que opera R es un **vector**
- Un vector sólo puede contener datos del mismo tipo



Básico - DataFrame

- Es una lista de vectores de igual tamaño y tipos distintos, que R visualiza como una tabla.



Básico – Obteniendo ayuda

- **help.start()**
 - Ayuda general
- **help(mean)**
 - Ayuda específica a una función
 - ?mean
- **help.search("mean")**
 - Encontrar una función en cualquier página de la ayuda
 - ??mean
- **example(mean)**
 - Mostrar un ejemplo de uso de una función

I.- Carga de datos

- Descargar el archivo flights.csv de:
 - <http://www.javiermoralo.es/datos/flights.csv>
- Guardarlo en el directorio de trabajo de R
 - `getwd()`
- Cargar el archivo en un dataframe
 - `flights <- read.csv("flights.csv", stringsAsFactors = FALSE)`
- Ver el contenido del dataframe
 - `head(flights)`

II.- Georeferenciar los aeropuertos

- Instalación de los paquetes ggplot2 y ggmap
 - `install.packages("ggplot2")`
 - `install.packages("ggmap")`
- Cargar la librería ggmap
 - `library(ggmap)`
- Obtener la coordenadas de los aeropuertos
 - `airports <- unique(c(flights$From, flights$To))`
 - `coords <- geocode(airports)`
- Crear un dataframe de aeropuertos georeferenciados
 - `airports <- data.frame(airport=airports, coords)`
 - `head(airports)`

II.- Crear el mapa de rutas

- Georeferenciar las lista de vuelos

- Origen:

- ```
flights <- merge(flights, airports, by.x="To", by.y="airport")
```

- Destino:

- ```
flights <- merge(flights, airports, by.x="From", by.y="airport")
```

- Instalar y cargar las librerías necesarias

- ```
install.packages("ggrepel")
```

- ```
library(ggplot2)
```

- ```
library(ggrepel)
```

## II.- Crear el mapa de rutas

---

- Crear el mapa base

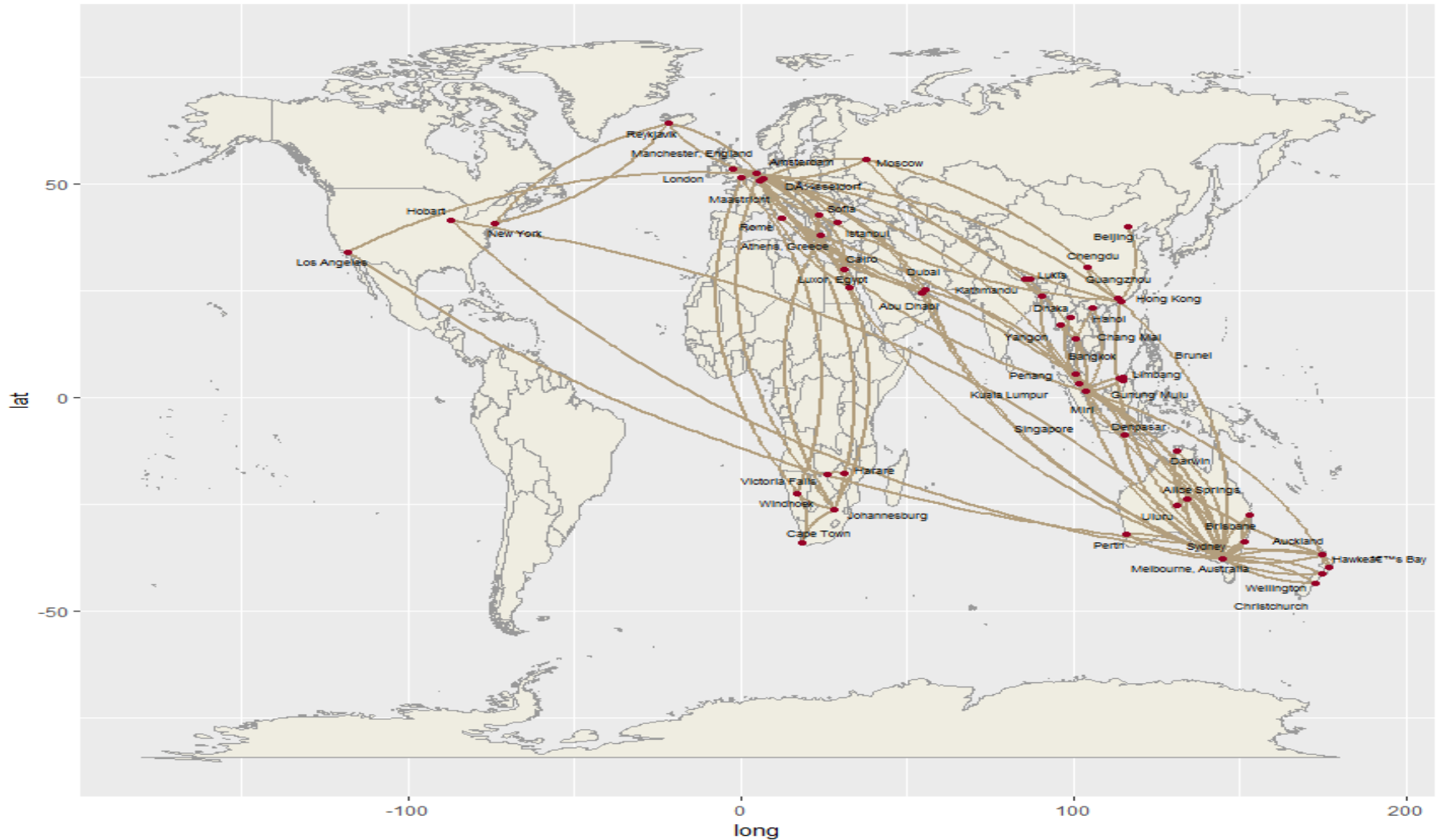
- `worldmap <- borders("world", colour="grey60", fill="#efede1")`

- Dibujar las rutas aéreas

- `ggplot() + worldmap + geom_curve(data=flights, aes(x = lon.x, y = lat.x, xend = lon.y, yend = lat.y), col = "#b29e7d", size = 1, curvature = .2) + geom_point(data=airports, aes(x = lon, y = lat), col = "#970027") + geom_text_repel(data=airports, aes(x = lon, y = lat, label = airport), col = "black", size = 2, segment.color = NA)`

1. Dibujar mapa base
  2. Pintar rutas
  3. Mostrar aeropuertos
  4. Nombre aeropuertos

## II.- Crear el mapa de rutas



# III.- Diagrama de grafos

---

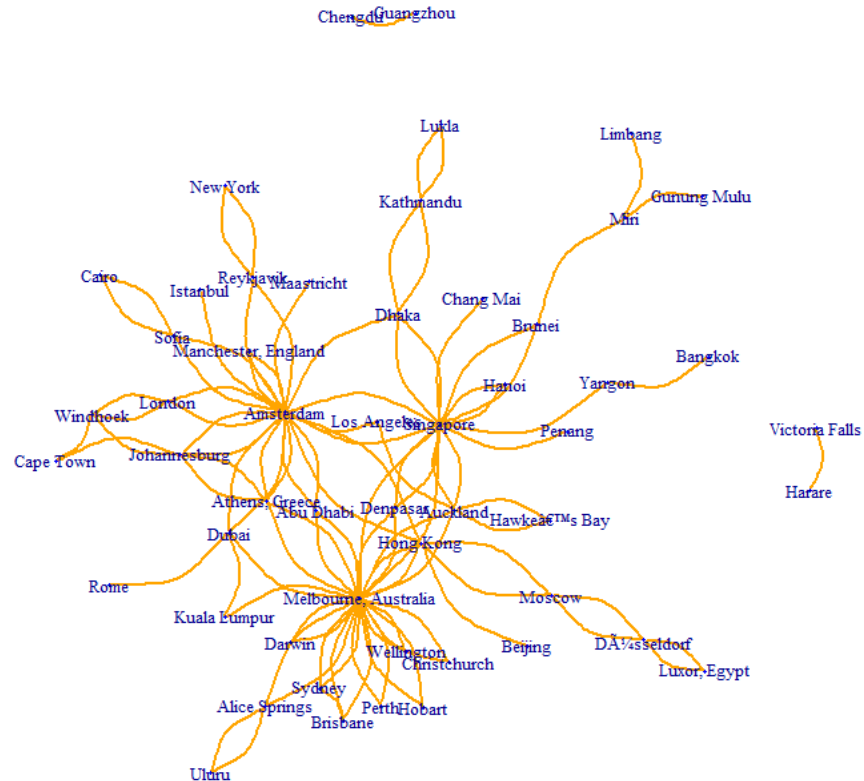
- Instalar y cargar la librería necesaria
  - `install.packages("igraph")`
  - `library(igraph)`
- Crear listado de grafos
  - `edgelist <- as.matrix(flights[c("From", "To")])`
  - `g <- graph_from_edgelist(edgelist, directed = TRUE)`
  - `g <- simplify(g)`

# III.- Diagrama de grafos

---

- Dibujar el Diagrama de Grafos
  - `plot.igraph(g, edge.arrow.size=0, edge.color="orange", edge.curved=TRUE, edge.width=2, vertex.size=1, vertex.color="blue", vertex.frame.color=NA, vertex.label=V(g)$name, vertex.label.cex=0.8, layout=layout.fruchterman.reingold)`

# III.- Diagrama de grafos





# ¡Gracias por vuestra atención!



*Imagen de Silvia Rabanete*



<https://www.linkedin.com/in/jesusjaviermoralo/>



[@JJavierMoralo](https://twitter.com/JJavierMoralo)



<https://github.com/javiermoralo>

# datahack