

# Data Fundamentals



red.es

Centro de  
Referencia Nacional  
en Comercio Electrónico  
y Marketing

CRN  
Digital



"El FSE invierte en tu futuro"  
Fondo Social Europeo

# Data Fundamentals



# Data Fundamentals



red.es

Centro de  
Referencia Nacional  
en Comercio Electrónico  
y Marketing

CRN  
Digital



"El FSE invierte en tu futuro"  
Fondo Social Europeo



# Índice

## 1. La práctica definitiva



# La práctica definitiva





DATA FUNDAMENTALS

# INGREDIENTES

**GOOGLE (Herramientas Cloud)**

**PROGRAMACIÓN: Para manipular los datos**

**EDA mediante PYTHON**

**PROCESO REAL DE MANIPULACIÓN MASIVA MEDIANTE PANDAS**

**RESULTADOS**

# JUNTAMOS TODO

## NECESIDAD:

Datos  
Sector

Organización y Sector (**Business Insider**)  
Datos (Open Data SÓLO, sin Kaggle)



# JUNTAMOS TODO

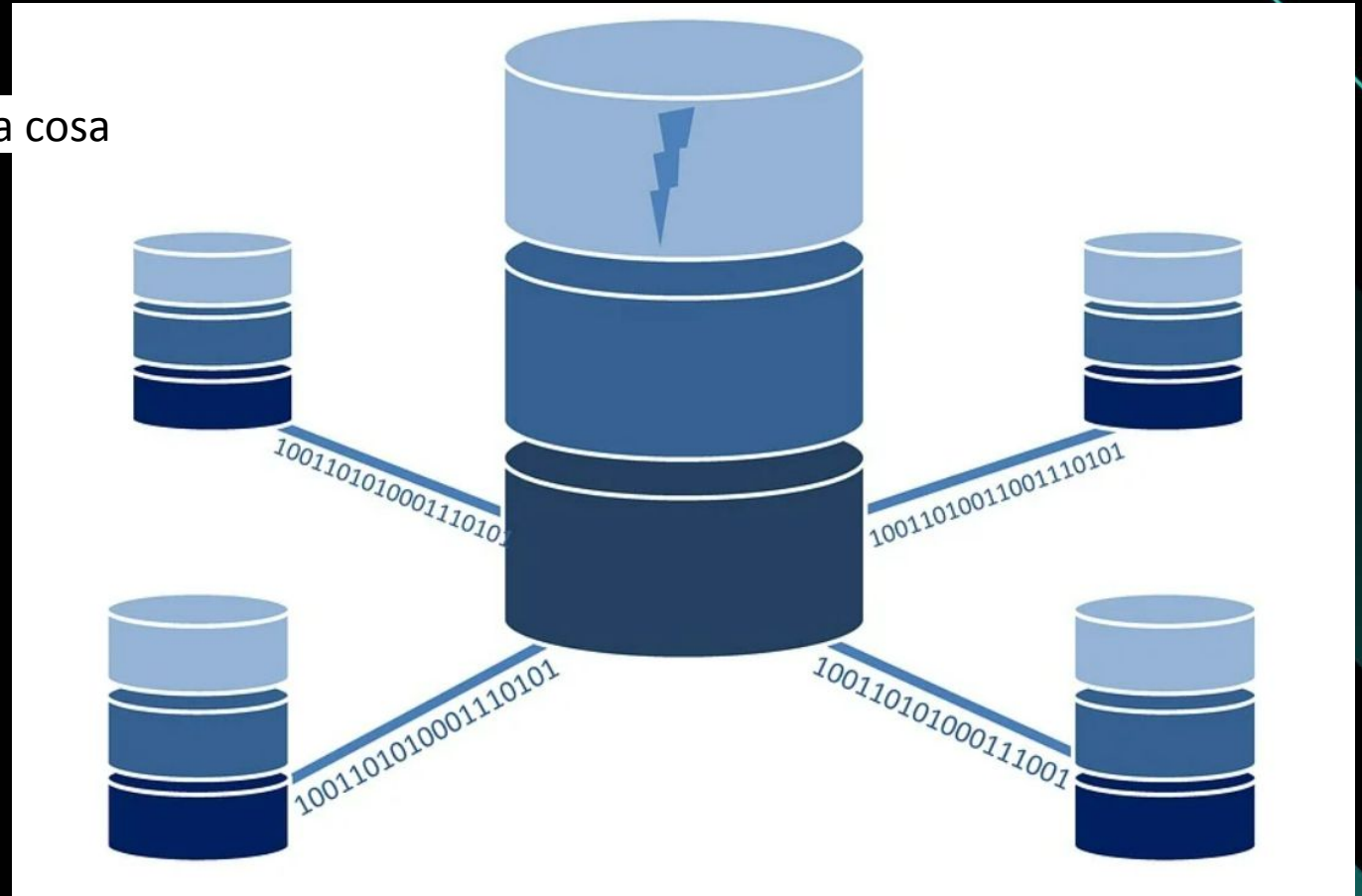
## Documento Inicial:

Criterio de Aceptación

Modelo de Datos: Qué significa cada cosa

Tipo de dato

Rango de valores





# JUNTAMOS TODO

## **EDA:**

Primera imagen

Primeros problemas de datos

Distribuciones (Pinta de variables)

## **Manipulación por dimensión de DATA WRANGLING:**

Incompletos

Incoherencias

Outliers

...



ECOSISTEMA DATA

# Ejercicios



## Ejercicio 2

Se hacen 3 o 4 equipos. Cada equipo escoge ser de...

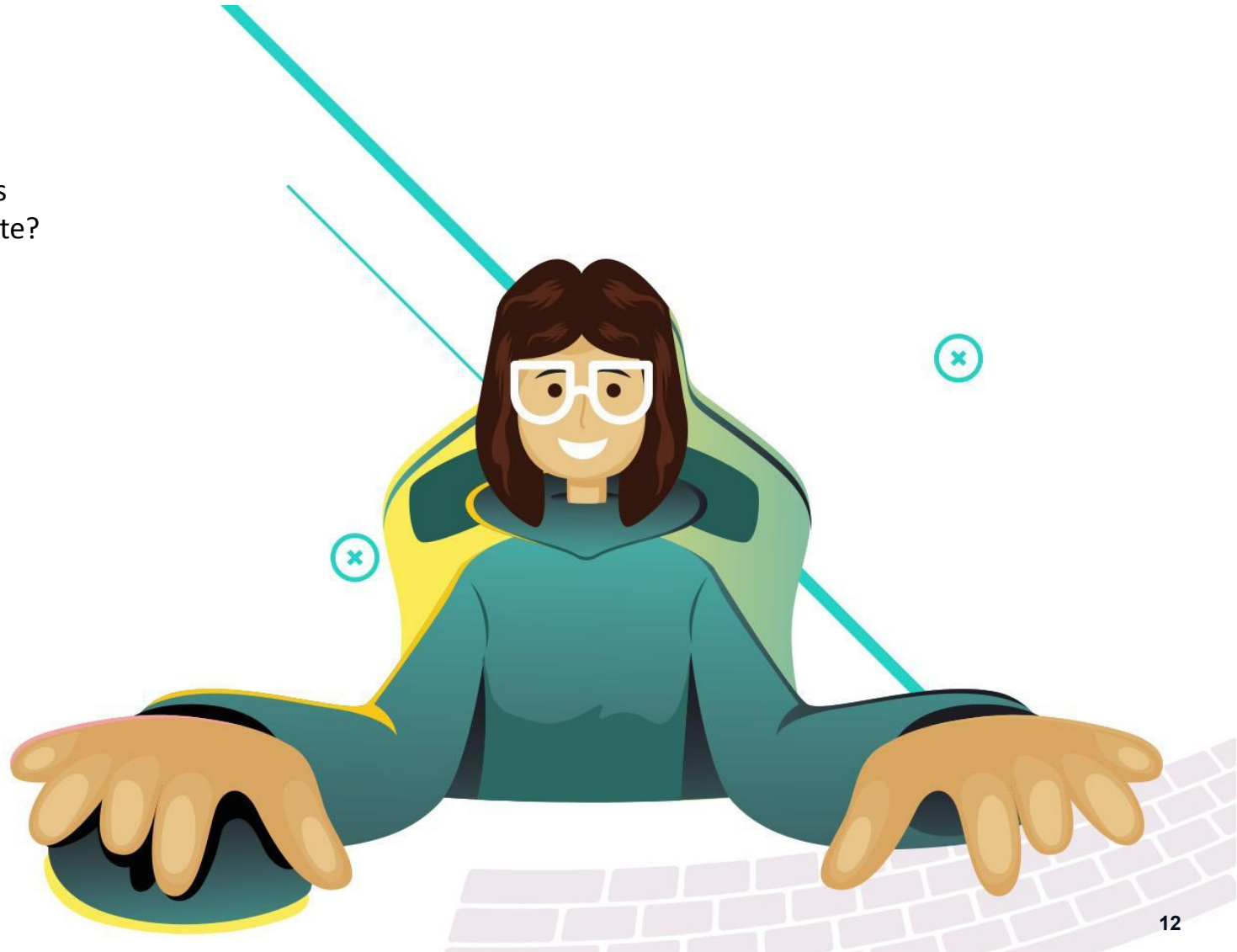
MercaMadrid  
Puerto de Santander  
Una panadería enorme  
Amazon



## Ejercicio 3

Partiendo de los mismos equipos...

- Definir un equipo de Ciencia de Datos
- ¿Qué rol le gustaría a cada participante?
- ¿Qué ha hecho?

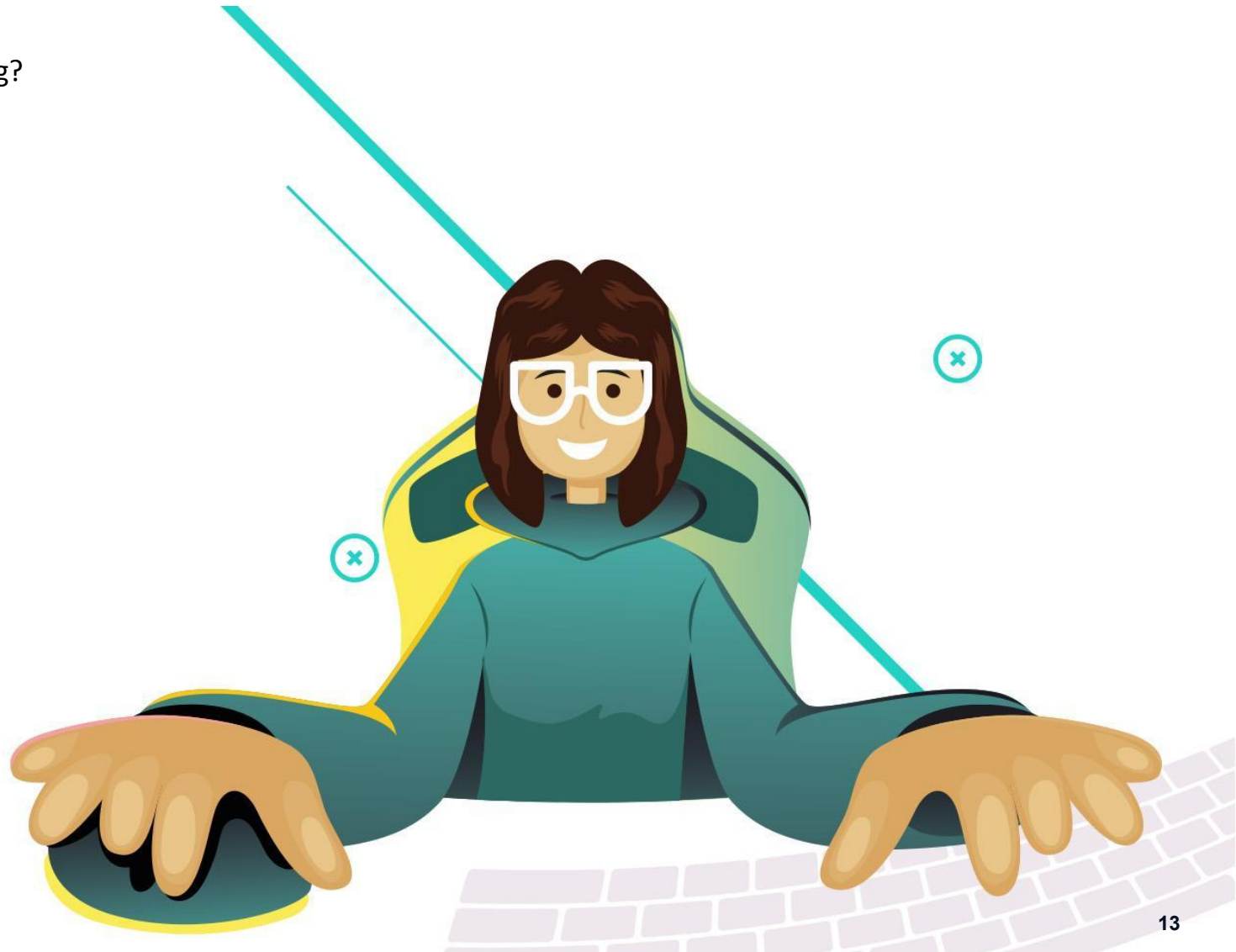




## Ejercicio 4

¿Qué es MapReduce? ¿Y Batch processing?

- ¿Cuándo nace MapReduce?
- ¿Qué es un clúster?
- ¿Qué es un NameNode?





**Contacto:**

[ibonreinoso@opendeusto.es](mailto:ibonreinoso@opendeusto.es)





red.es



UNIÓN EUROPEA

*"El FSE invierte en tu futuro"*

**Fondo Social Europeo**

