



# Spanish rail tickets sales

# Index

Project goals

---

Project phases

---

Workflow

---

Data sources

---

Data model

---

Database creation

---

Data preparation

---

Dashboard creation

---

# Project goals

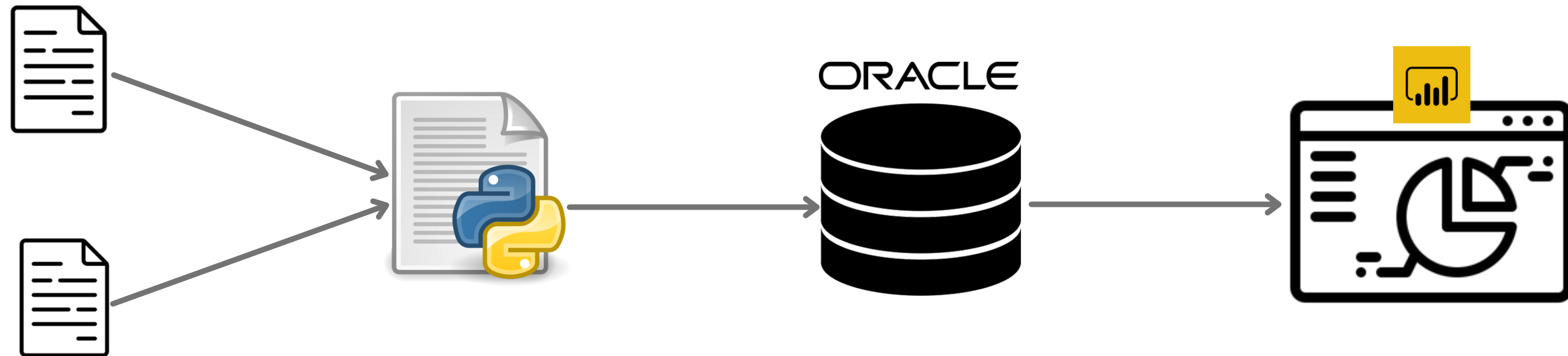
Using ticket sales data analyze:

- Sales: How are sales and revenue generated?
- Network: When, where and what kind of trains are offered?
- Pricing: What drives up and down the prices of the tickets?

# Project phases



# Workflow



# Data sources

Spanish Rail Tickets Pricing – Renfe dataset      source: [kaggle](#) – The Gurus

Web scrapped rail ticket sales data with price

Dataset information

- 38.8 million rows
- 14 columns

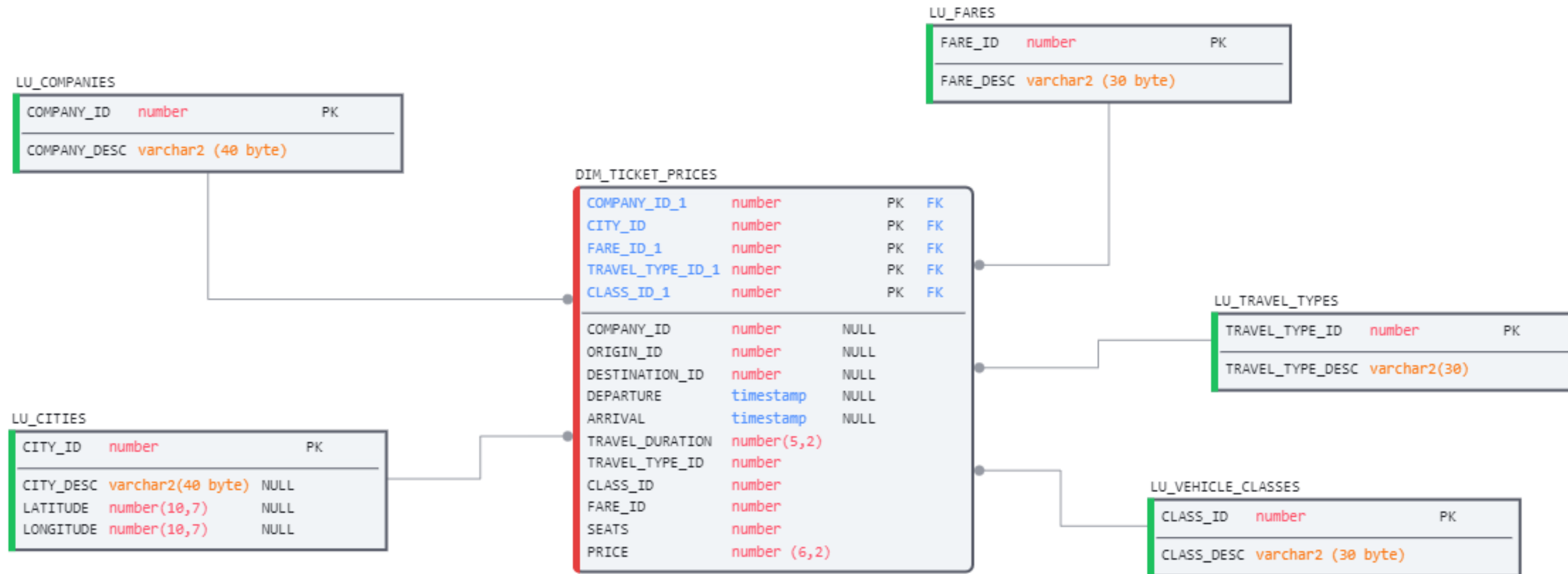
Spanish cities coordinates

Latitude and longitude data for every city in Spain

Dataset information

- 7893 rows
- 3 columns

# Data model



# Database creation

Oracle Database:

- Database installation
- Schema creation
- Table creation DDL script



# Data preparation

Python script with four main functionalities:

- Load files
- Clean and unify data
- Structure data to generate the data model
- Load data into the database

# Dashboard creation

## Power BI Dashboard

- Load tables from database
- Establish data model relations
- Create columns, measures
- Create visualizations