1. Data preparation:

Changing the formal date from **dd/mm/yyyy** to **yyyy-mm-dd** to be consistent to the PostgreSQL database.

A number of dates and numbers

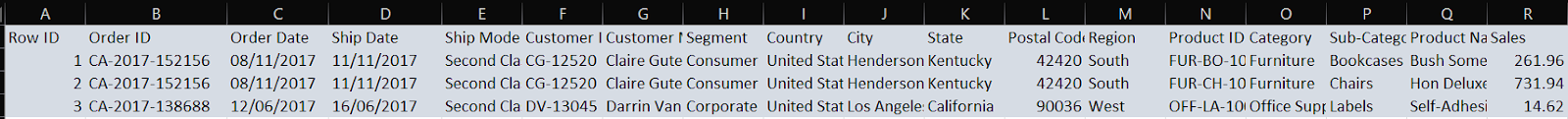
Description automatically generated with medium confidenceA black arrow pointing to the right

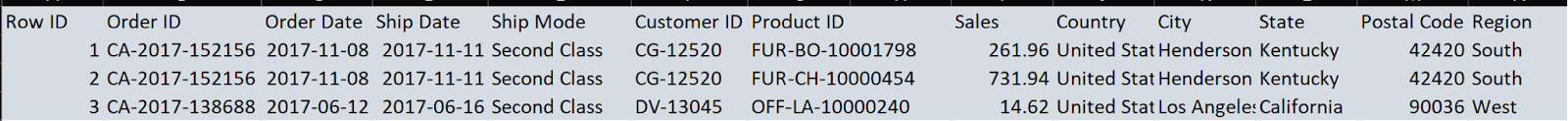
Description automatically generatedA close up of numbers

Description automatically generated

Dividing the train file into three files (tables) including: **Orders**, **Customers** and **Products** to make it easier for data analysis.

A black arrow pointing to the right

Description automatically generated



A close up of a name

Description automatically generatedA close-up of a sign

Description automatically generated

Filtering and removing **duplicated data** to improve accuracy and eliminate statistical errors.

A screenshot of a computer

Description automatically generatedA black arrow pointing to the right

Description automatically generatedA screenshot of a computer

Description automatically generated

1. Entity-Relationship Diagram (ERD)

Creating Entity-Relationship Diagram (ERD) to design a dataset and determine data type, primary and foreign keys.

A screenshot of a computer

Description automatically generated