

# Exploratory Data Analysis for e-commerce, retail and wholesale

The aim of this project is to solve business problems using a real Dataset from a US-based retail network, generating insights through this process.

A simplified .CSV file will be read, and we will address 10 business questions using this Dataset.

Python packages such as [Pandas](#), [Numpy](#), [Matplotlib](#) and [Seaborn](#) will be employed to generate insights from our Database.

Dataset Link:

<https://community.tableau.com/s/question/0D54T00000CW eX8SAL/sample-superstore-sales-excelxls>

Desired Insights:

1. Which city has the highest sales value for the 'Office Supplies' category?
2. What is the Total Sales by Order Date?
3. What is the Total Sales by State?
4. What are the top 10 cities with the highest total sales?

5. Which segment had the highest total sales?
6. What is the Total Sales per Segment and per Year?
7. The company's managers are considering offering different discount tiers and would like to simulate based on the following rule: If the Sales\_Value is greater than 1000, a 15% discount is given. If the Sales\_Value is less than 1000, a 10% discount is given. How many sales would receive a 15% discount?
8. Assuming the company decides to grant the 15% discount from the previous item. What would be the average Sales\_Value before and after the discount?
9. What is the Average Sales per Segment, per Year, and per Month?
10. What is the Total Sales by Category and Subcategory, considering only the top 12 Subcategories?