رواد مصر الرقمية

وزارة الاتصـــــالات
وتكنولوجيا المعلومات

E-Commerce Orders Analysis

Presented by :

- Ahmed Hassan Sayed
- Hisham Mohamed Mohamed
- Salma Samir Ziada
- Ehab Sayed Abd El Baky
- Eyad Abu Elhaggag Ahmed
- Mohamed Sayed Abdelhady

رواد مصر الرقمية

وزارة الاتصـــــالات
وتكنولوجيا المعلومات

Project Idea

1- Problem: Analyzing e-commerce orders performance over three years (1996-1997-1998) to understand trends and dealing with miss values and nulls , duplicated values and data taypes.

2- Solution: We processed by cleaning the data using Python to replacing the missing values , dealing with outliers and repairing the data types then we use SQL for overview the data and making validation and asking Q.

3- Unique Value Proposition: We use Tableau for analysis and Visualizing the data to understand Comprehensive analysis of Three years' worth of E-Commerce order data, offering insights on sales performance by Country’s, Customers segmentation and top products.

Project Wireframe

1- Visuals: Charts and dashboards showing sales by region, rep performance, commission breakdowns, and target achievements.

2- Cleaning Process: Data collection, cleaning, and analysis using SQL and Python, with final visualization in Tableau for easy monitoring of sales trends.

3- Focus: Usability and a clear user experience, allowing decision-makers to track real estate sales performance at a glance.

Project Wireframe

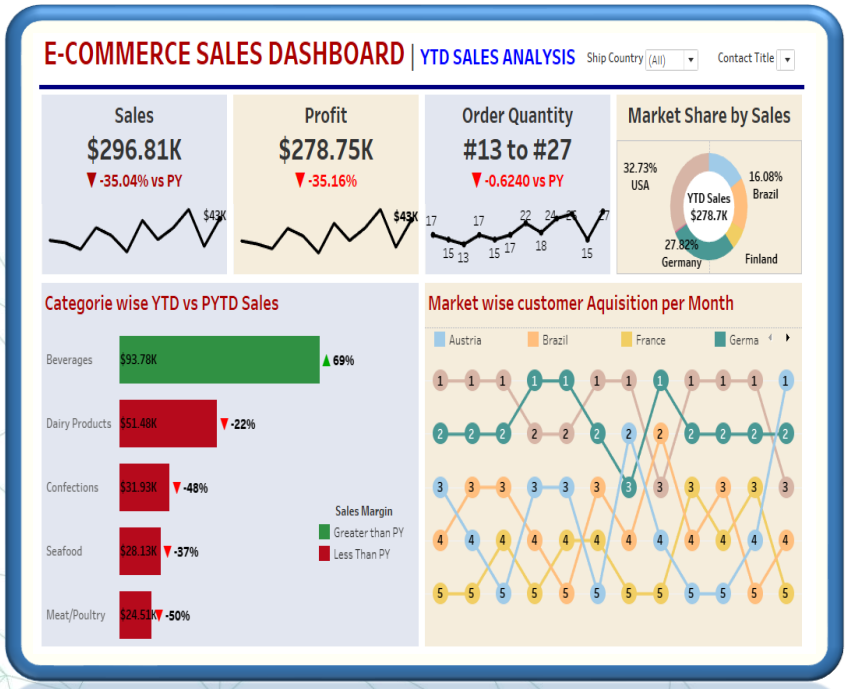
cleaning of Orders sheet

```
[33]: def data_Specs(Orders):
          print("Initial Specs on data: \n")
          print("Data Shape: ", Orders.shape)
          print("\n-----------------------------------------------\n")

          s = (Orders.dtypes == 'object')
          Cat_cols = list(s[s].index)
          print("Category Columns (",len(Cat_cols),") : \n",Cat_cols)
          print("\n-----------------------------------------------\n")

          s = (Orders.dtypes != 'object')
          Nums_cols = list(s[s].index)
          print("Numeric Columns (",len(Nums_cols),") : \n",Nums_cols)
          print("\n-----------------------------------------------")
      data_Specs(Orders)
```

```
Initial Specs on data:

Data Shape:  (830, 14)

-----------------------------------------------


Category Columns ( 8 ) :
 ['OrderID', 'Customer key ', 'ShipName', 'ShipAddress', 'ShipCity', 'ShipRegion', 'ShipPostalCode', 'ShipCountry']

-----------------------------------------------


Numeric Columns ( 6 ) :
 ['EmployeeID', 'OrderDate', 'RequiredDate', 'ShippedDate', 'ShipVia', 'Freight']

-----------------------------------------------
```

1- Cleaning Process: Data collection, cleaning, and analysis using Python and SQL dealing with missing values and nulls , duplicated values and data types.

Project Wireframe



2- Visuals: Charts and dashboards showing net sales , profit and total Quantity vs PY, Market Share by Sales , Categories wise Year to date vs PYTD and Market wise customer Acquisition per month.

3- Focus: Usability and a clear user experience, allowing decision-makers to track trends and market share.

End Users + Features

End Users: CEO,GM and sales managers.

Key Features:

1- Breakdown of Orders by Country and customers Segment .
2- Top N for all (Product's , Category , Country and Customers).
3- Most Shippers delivering orders .

User Problem Solving: Helps businesses to know which product or category with high profit , investing in which markets based on total orders and customers behavior.

Data Structure

Database Architecture:

CSV files used for data storage and processed through SQL and Python .

Key Entities:

Order, Customers, Employee , Order Details, Shippers, Suppliers, and Products dates.

Data Flow:

Steps Taken:

Data imported, cleaned, missing values handled, types converted, set measuring to know comparing sales , profit and total orders between current year vs previous year .

Summary:

Final summaries and visualizations created in Tableau.

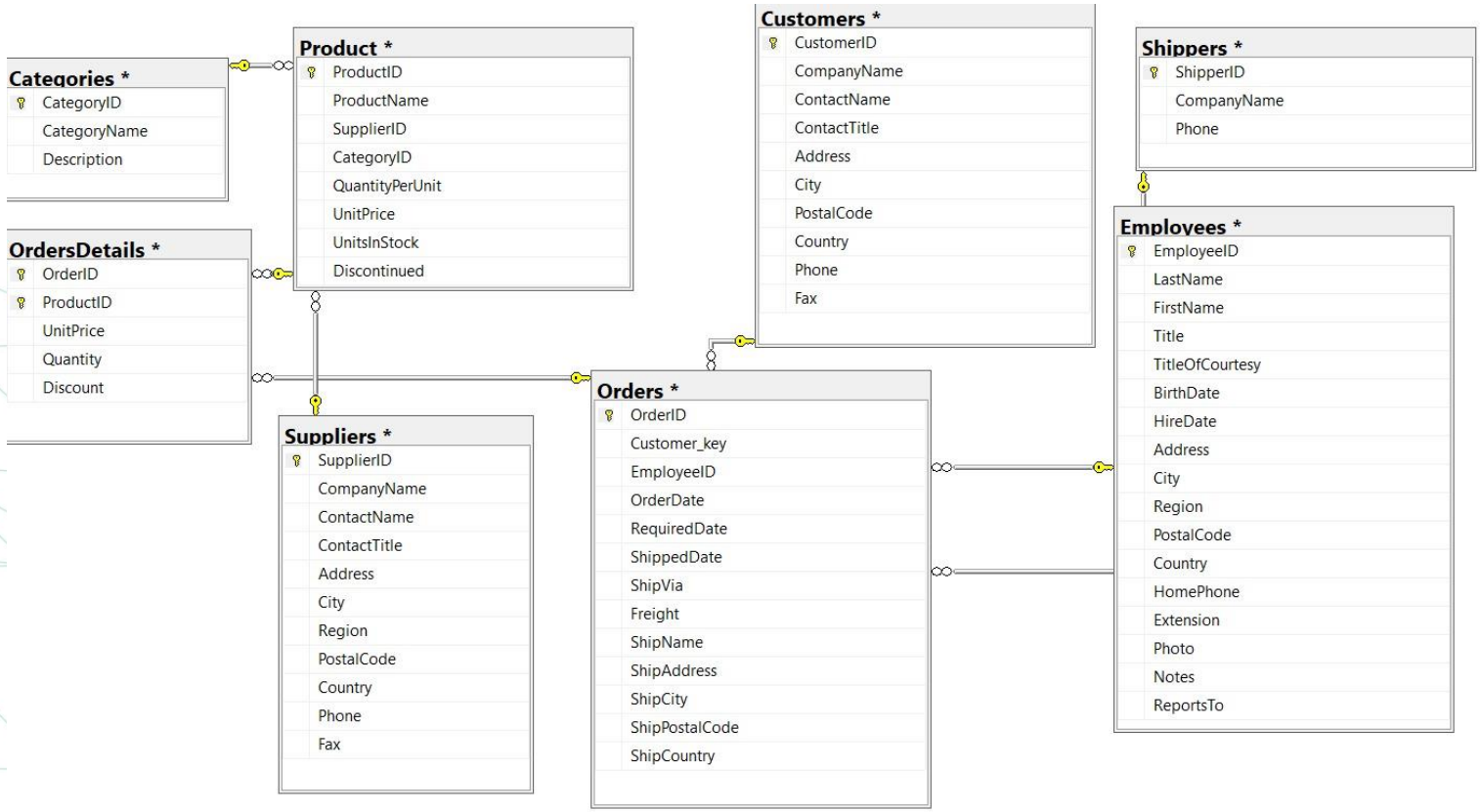
Programming Languages

1-Python Cleaning Process

```
-------------------------------------------------------------------------------
--total_quantity--

select sum([Sum_of_Quantity]) as Total_Quantity
from [dbo].[OrdersDetails]

-------------------------------------------------------------------------------
--top 10 products based on sales--

select top 10 [ProductName], sum([Sum_of_Sales]) as Total_sales
from [dbo].[vw_Ord_OrdDet_Prod_Cust_Emp_Shipp]
group by [ProductName]
order by sum([Sum_of_Sales]) desc

-------------------------------------------------------------------------------
--top 3 categories based on sales--

select top 3 [CategoryName], sum([Sum_of_Sales]) as Total_sales
from [dbo].[vw_Ord_OrdDet_Prod_Cust_Emp_Shipp] vw join [dbo].[Categories] c
on c.CategoryID=vw.CategoryID
group by [CategoryName]
order by sum([Sum_of_Sales]) desc

-------------------------------------------------------------------------------
--top 10 customers based on buying--

select top 10 [ContactName] , sum([Sum_of_Sales]) as Total_sales
from [dbo].[vw_Ord_OrdDet_Prod_Cust_Emp_Shipp]
group by [ContactName]
order by Total_sales desc

-------------------------------------------------------------------------------
```

Programming Languages

2-SQL diagram relationships



2- Part of SQL Aggregations

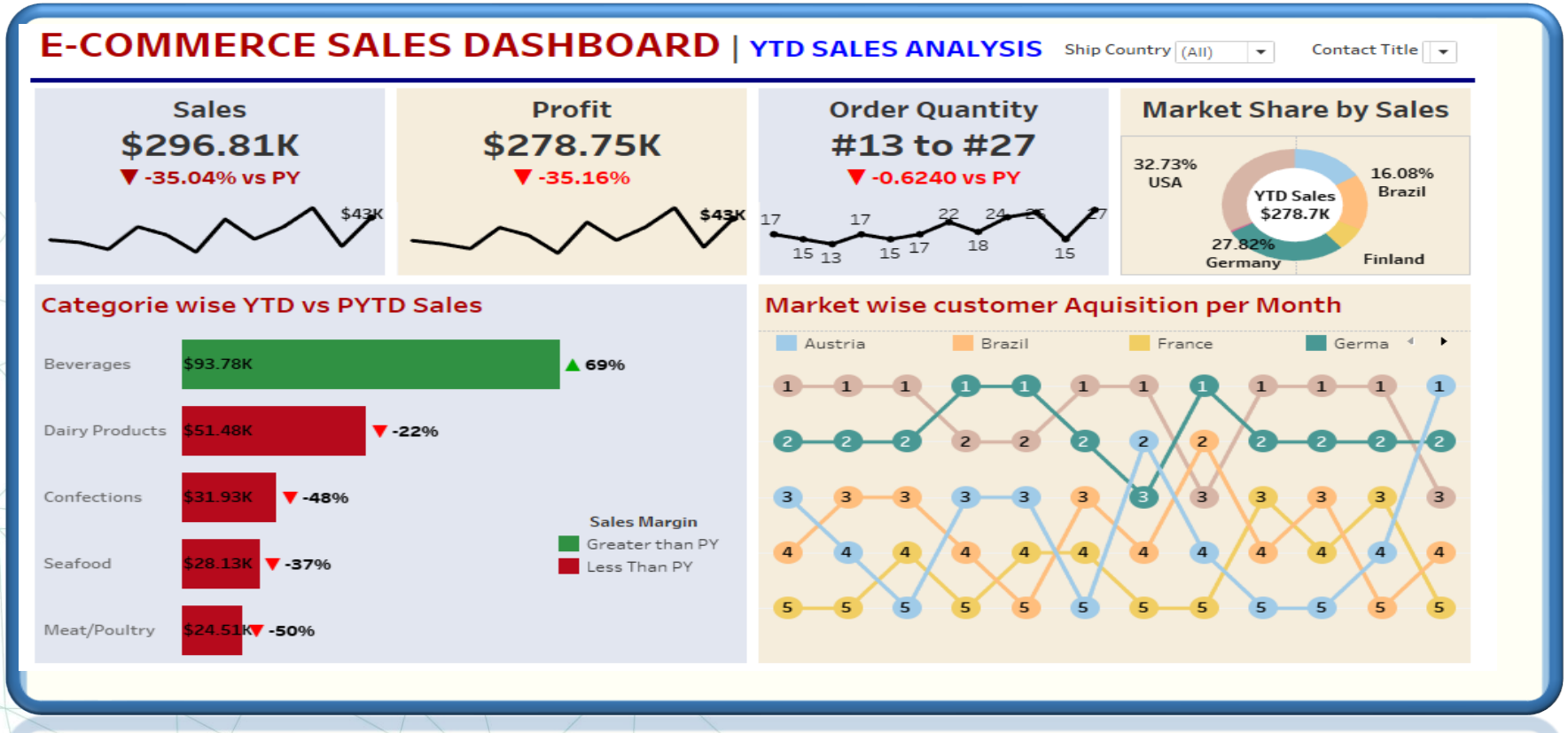
cleaning of OrdersDetails sheet

[29]:
```
def data_Specs(OrdersDetails):
    print("Initial Specs on data: \n")
    print("Data Shape: ", OrdersDetails.shape)
    print("\n-----------------------------------------------\n")

    s = (OrdersDetails.dtypes == 'object')
    Cat_cols = list(s[s].index)
    print("Category Columns (",len(Cat_cols),") : \n",Cat_cols)
    print("\n-----------------------------------------------\n")

    s = (OrdersDetails.dtypes != 'object')
    Nums_cols = list(s[s].index)
    print("Numeric Columns (",len(Nums_cols),") : \n",Nums_cols)
    print("\n-----------------------------------------------")
data_Specs(OrdersDetails)
```

```
Initial Specs on data:

Data Shape:  (2155, 5)

-----------------------------------------------

Category Columns ( 0 ) :
 []
```

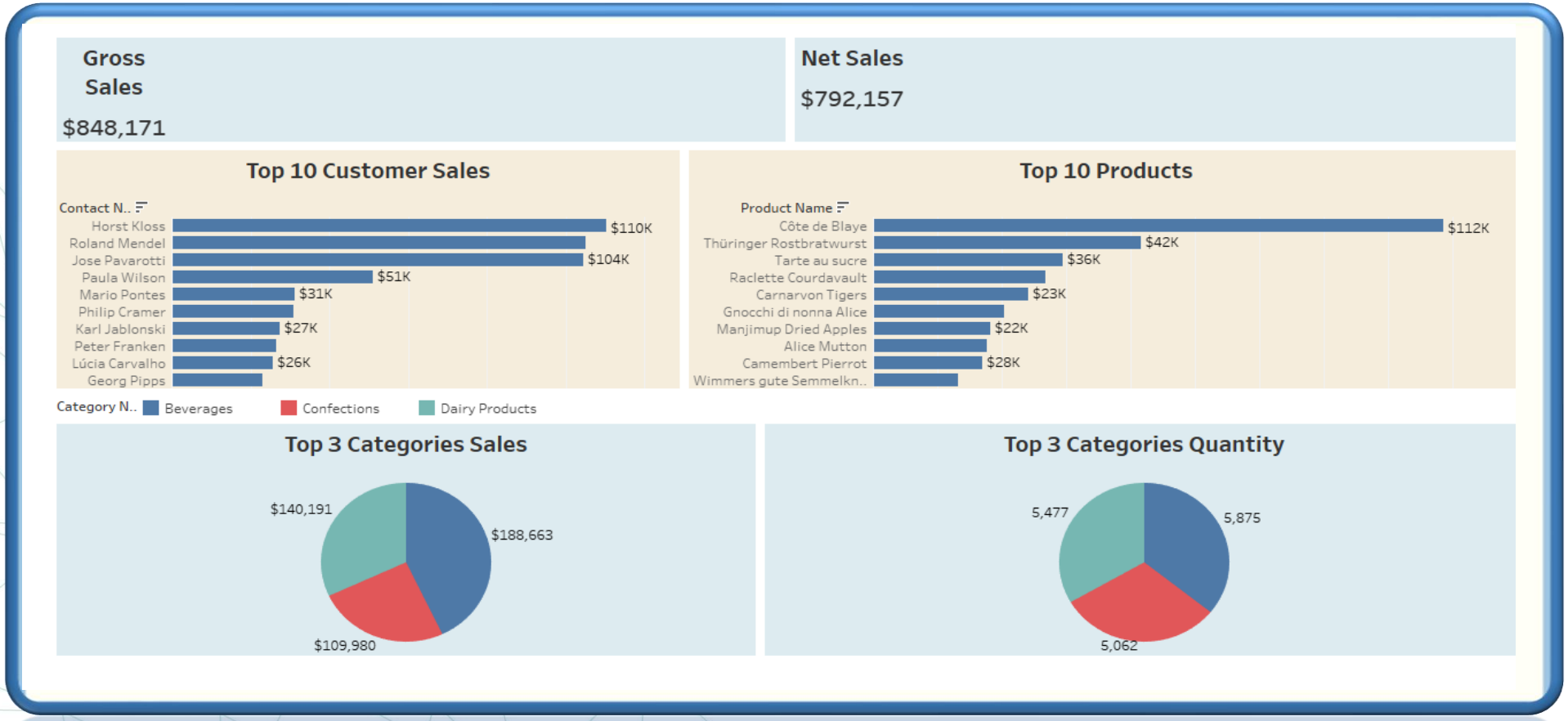
Programming Frameworks

3- Tableau Dashboard



Programming Frameworks

3- Tableau Dashboard



Live Application

Current State: Tableau dashboard displaying real-time Orders performance, Charts and dashboards showing sales by region, rep performance, commission breakdowns, and target achievements.

Key Visuals: Focusing on Market Share per Country.

Deliverables

Data Preparation and Cleaning:

1- cleaning data from all missing values for all rows that not contain any data .

2- replacing nulls by logically based on data view like in order ID there’s missing data and the id go like (10250 , 10251 , 10252 , Null, 10254,10255) so the null will be (10253) after viewing all Orders ID to be sure not be a duplicated number .

3- in Employee titles there’s missing value and we replacing it also based on logical like in Colum report all the Employee who's reporting to the number 3 is Sales Rep.

Deliverables

Data Analysis and Insights:

1- Performance charts and Profit breakdowns.

3- Detailed reports on Top Products and Categories.

4- focusing on Loyalty customers based on frequency making orders.

5- focusing on top customers based on total purchasing amount.

Project Team + Roles

Team Members and Roles

- Salma Samir Ziada
- Eyad Abu Elhaggag Ahmed

: Data Cleaning and Transformation (Python)

- Hisham Mohamed Mohamed
- Ehab Sayed Abd El Baky

: Data Structuring and Aggregation (SQL).

- Ahmed Hassan Sayed
- Mohamed Sayed Abdelhady

: Visualization and Reporting (Tableau).

Project Pipeline : Data collection → Cleaning → Analysis → Visualization.

Thank You!