

Data Analysis Specialist

Project Title: Real Estate Sales Data Analysis

**Presenter's Name: Adham Osama, Ahmed Adel, Ahmed Anwar,
Shawkat Gamal**

Date: 10/21/2024

Project Idea

- 1- Problem: Analyzing real estate sales performance over four years (2020-2024) to understand trends, targets, and commissions for each sales rep.
- 2- Solution: We processed the company's sales data using SQL, Python, and Excel to calculate key metrics like sales performance, commission, and target achievements.
- 3- Unique Value Proposition: Comprehensive analysis of four years' worth of real estate data, offering insights on sales performance by region, city, and individual rep, visualized in a dynamic Tableau dashboard.

Project Wireframe

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

End Users + Features

End Users: Real estate companies, sales managers, and analysts.

Key Features:

- 1- Breakdown of sales by region and individual rep.
- 2- Commission calculations per sale.
- 3- Target achievement tracking and performance analysis.

User Problem Solving: Helps businesses assess sales reps' performance, track commissions, and set future targets based on past performance.

Data Structure

Database Architecture:

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

Key Entities:

Companies, regions, sales reps, unit types, sales amounts, commissions, and sale dates.

Data Flow:

Steps Taken:

Data imported, cleaned, missing values handled, types converted, commissions calculated, and targets set .

Summary:

Final summaries and visualizations created in Tableau.

Programming Languages + Frameworks

1- Languages: SQL, Python, Excel.

2- Frameworks/Tools: Tableau for data visualization.

3-Supporting Technologies: CSV files for data input, Python for analysis, SQL for aggregation.

Live Application

Current State: Tableau dashboard displaying real-time sales performance, target achievements, and growth forecasts, with charts for each state and rep .

Key Visuals: Sales by state, sales by rep, commissions, and performance metrics.

Deliverables

Reports:

- 1- Sales summaries by region and rep.
- 2- Performance charts and commission breakdowns.
- 3- Detailed reports on sales targets and achievements.

Other Products: Source code for SQL and Python scripts, Tableau dashboard link.

Project Documentation

SQL

```
BULK INSERT SalesData
FROM 'G:\Real_State_Sales_Data\5.csv'
WITH (
    FIELDTERMINATOR = ',',
    ROWTERMINATOR = '\n',
    FIRSTROW = 2 -- Skip header row
);
```

```
-- how many sales per each Salesperson ?
select * from SalesPerPersonPerYear
Create view SalesPerPersonPerYear as
select Salesname, Company, Year, count(*) Sales_Count, sum(SaleAmount) Total_Sales, sum(SalesCommission) Total_Commission
from SalesData where Salesname != 'unkown sales'
group by salesname, Company, year;
```

```
--who is the top performer in each year ??
select * from Top_performers_salesinYear
Create view Top_performers_salesinYear as
WITH RankedSales AS (
    SELECT
        SalesName,
        Company,
        Year,
        Total_Sales,
        Total_commission,
        RANK() OVER (PARTITION BY Year ORDER BY Total_Sales DESC) AS SalesRank
    FROM
        SalesPerPersonPerYear
)
SELECT
    SalesName,
    Company,
    Year,
    Total_Sales,
    Total_commission
FROM
    RankedSales
```

Results Messages

SalesName	Year	Total_Sales	Total_commission
Charlie	Countrywide	74645846.1937	5225209.23355901
Oweis	Savills	69925037.1638	5594002.97310399
Oweis	Savills	112401768.9314	8992141.51451197
Oweis	Savills	68085032.8496	5446802.62796799
Charlie	Countrywide	69414954.5032	4859046.81522401

Project Documentation

SQL

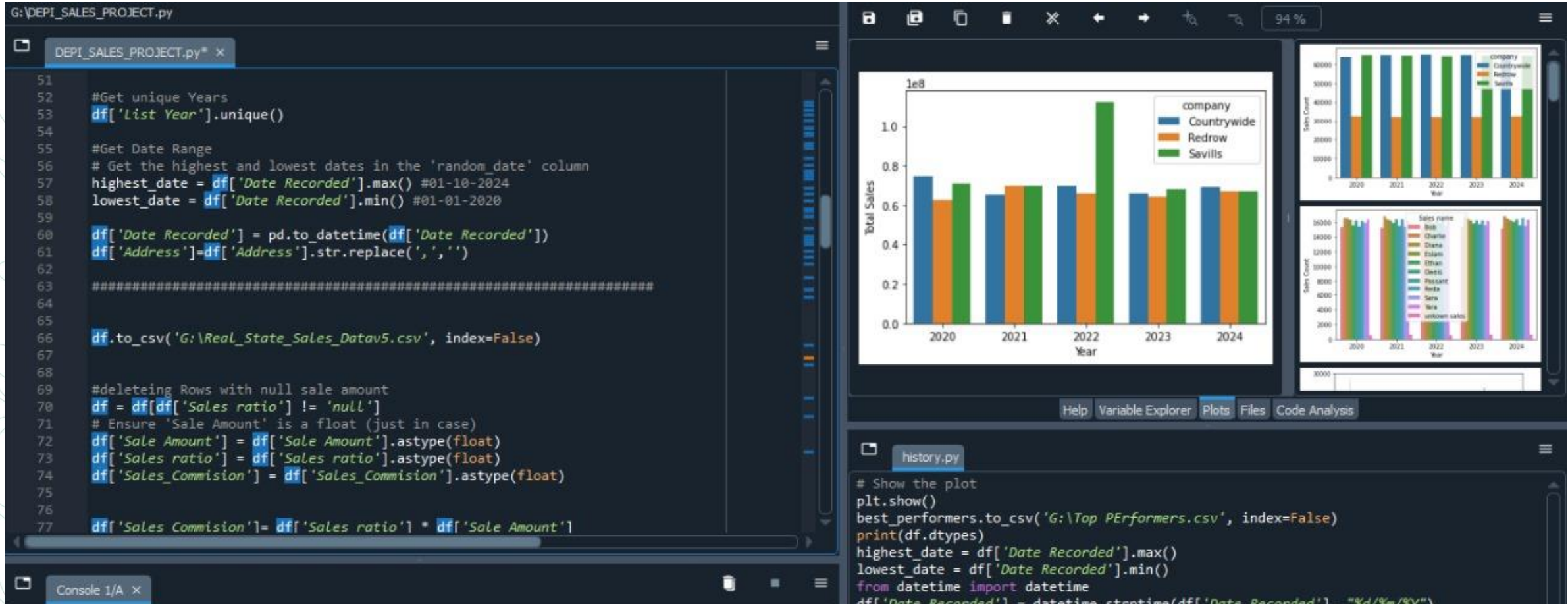
```
CREATE TABLE SalesData (
    SalesName VARCHAR(255),
    SerialNumber float,
    Year int,
    DateRecorded datetime,
    Town VARCHAR(100),
    Address VARCHAR(255),
    SaleAmount float,
    PropertyType VARCHAR(100),
    ResidentialType VARCHAR(100),
    Company VARCHAR(100),
    SalesRatio float,
    SalesCommission float
);
```

```
--View
-- how many sales per each Salesperson ?
select * from SalesPerPersonPERYear
Create view SalesPerPersonPERYear as
select Salesname,Company,Year,count(*) Sales_Count,sum(SaleAmount) Total_Sales ,sum(SalesCommis
from SalesData where Salesname != 'unkown sales'
group by salesname,Company,year;
```

Salesname	Company	Year	Sales_Count	Total_Sales	Total_Commission
Bob	Savills	2021	15273	57946625.57	4635730.0456
Bob	Savills	2024	15163	61093452.8112	4887476.224896
Charlie	Countrywide	2020	16548	74645846.1937	5225209.23355901
Diana	Savills	2020	16589	71083189.3348	5686655.14678399
Diana	Savills	2023	16472	63301582.2261	5064126.57808799
Eslam	Redrow	2023	16273	62941256.2684	3776475.37610399
Ethan	Countrywide	2021	15884	64097087.8435	4486796.14904501
Ethan	Countrywide	2024	16057	66252320.7442	4637662.45209401
Oweis	Savills	2022	16247	112401768.9314	8992141.51451197
Passant	Countrwide	2023	15790	62343284.2665	4364029.89865501

Project Documentation

Python



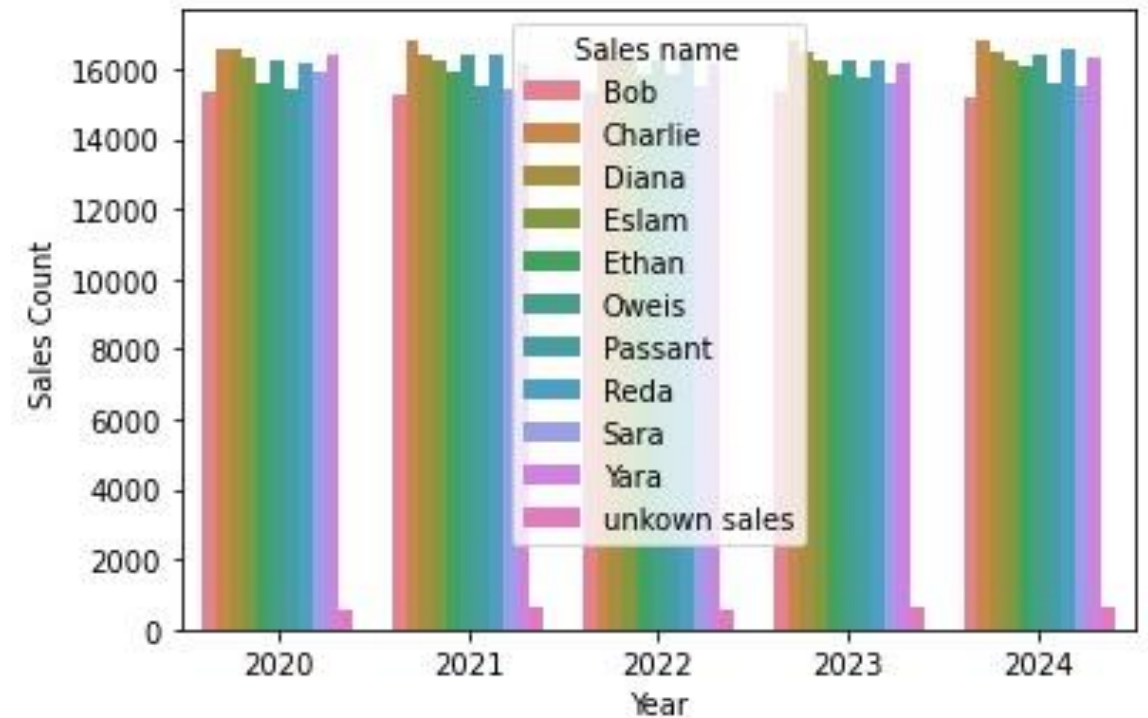
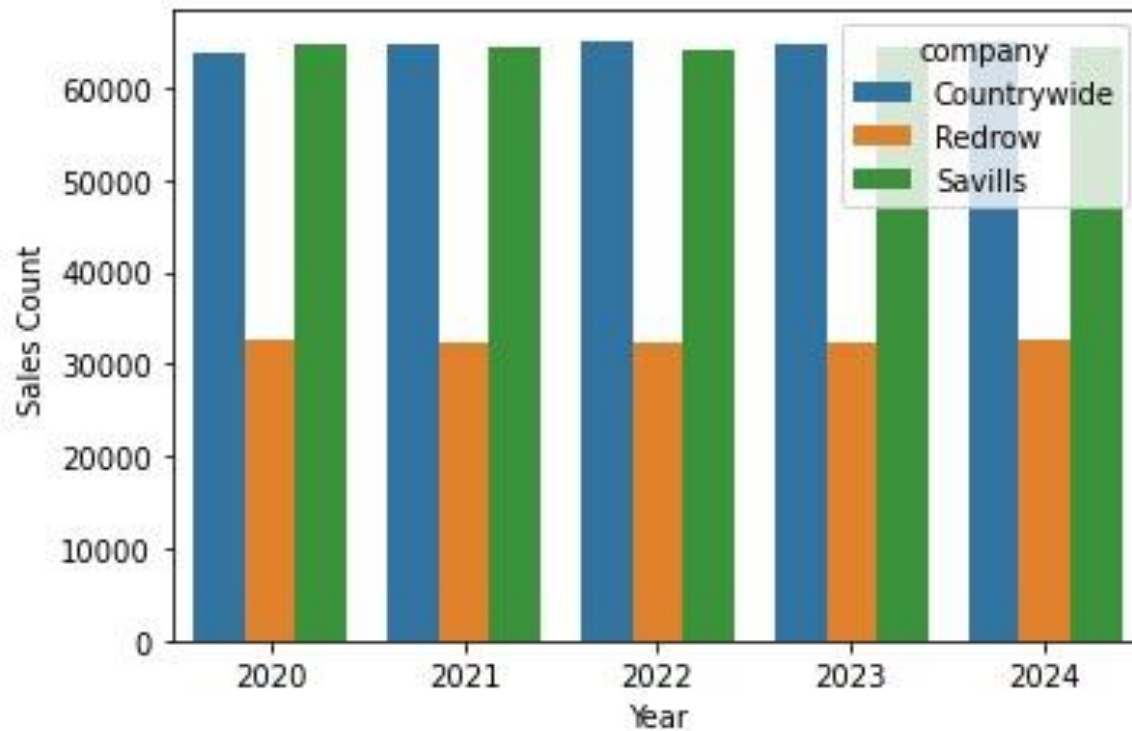
Project Documentation

Python

```
DEPI_SALES_PROJECT.py* X
1  #-----#
2  import pandas as pd
3  import seaborn as sns
4
5
6  df = pd.read_csv('G:\Real_Estate_Sales_2001-2022_GL.csv')
7
8  df.info()
9
10 df.dtypes
11 # View the first few rows
12 print(df.head())
13
14 # Get a summary of the dfset
15 print(df.info())
16
17 # Convert Serial Number and Year to object type
18 df['Serial Number'] = df['Serial Number'].astype(str)
19 df['Year'] = df['Year'].astype(str)
20 # Rename column
21 df.rename(columns={'Alice': 'Sales name'}, inplace=True)
22 df.rename(columns={'List Year': 'Year'}, inplace=True)
23
24
25 #Filling Null Data to avoid data loss
26 df.loc[df['Residential Type'] == 'Commercial', 'Property Type'] = 'Commercial'
```

Project Documentation

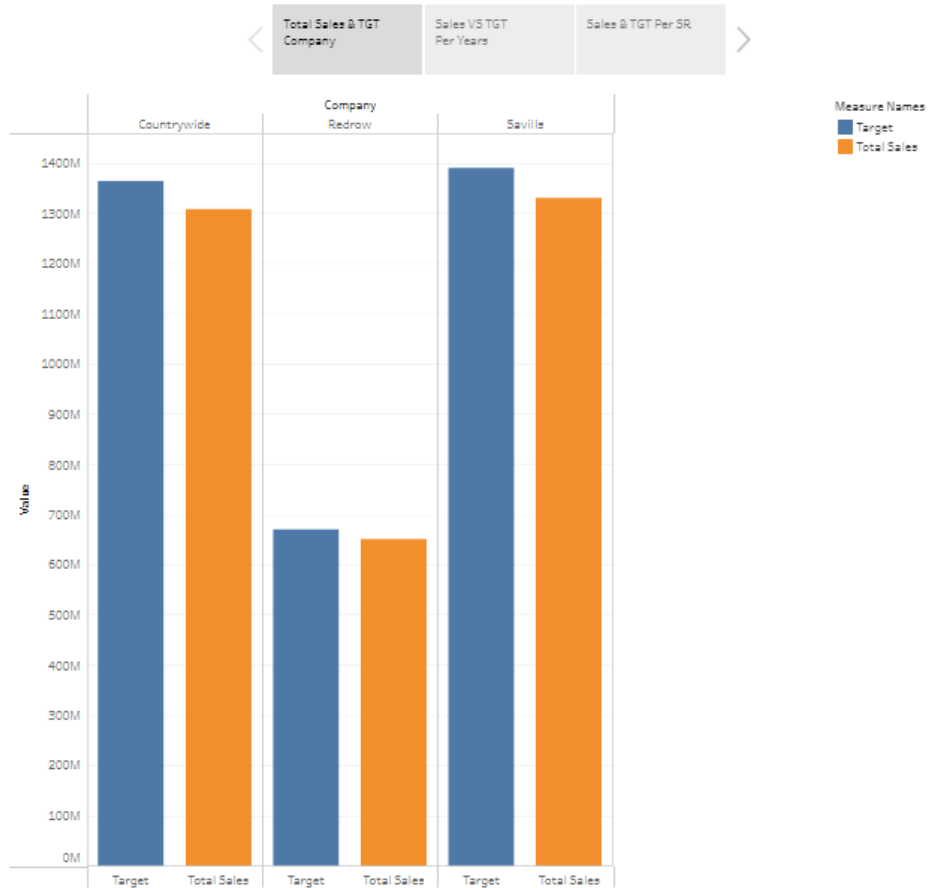
Python



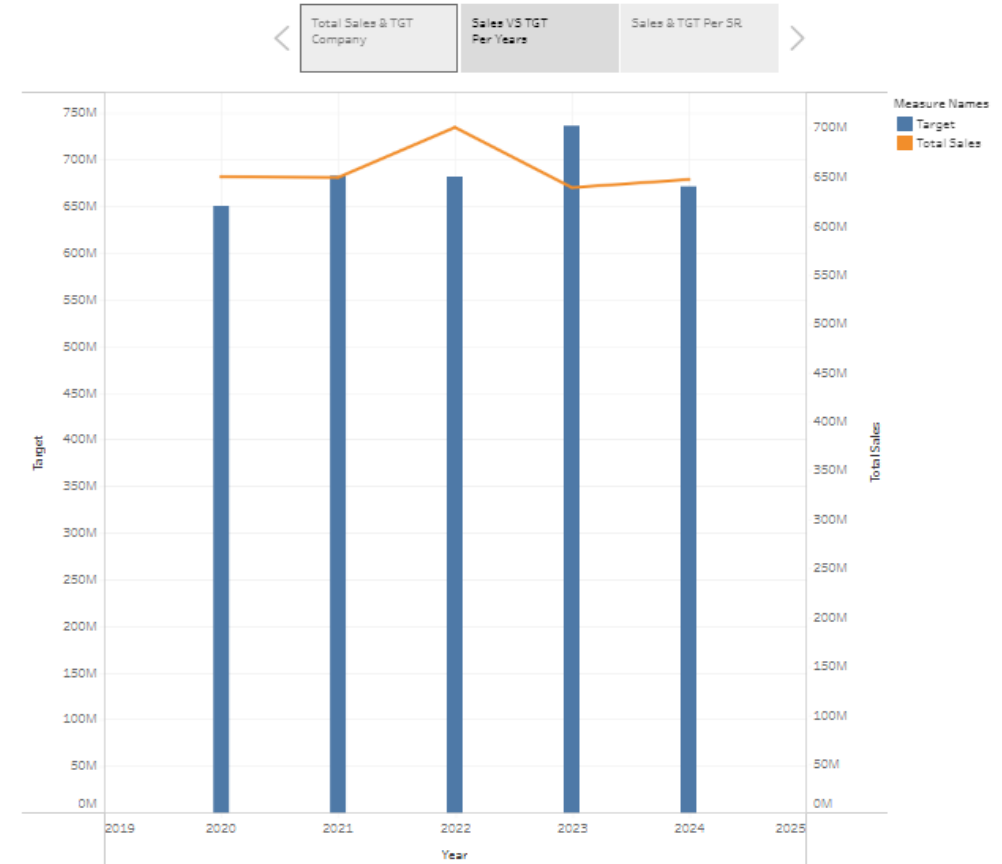
Project Documentation

Tableau

Project



Project



10/22/2024

Tableau

Project



Project Team + Roles

Team Members:

Adham Osama : Lead Data Analyst, Data Cleaning and Transformation (Python)

Ahmed Adel : Data Structuring and Aggregation (SQL).

Ahmed Anwar : SQL and Python Analysis.

Shawkat Gamal : Visualization and Reporting (Tableau).

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

Thank You!