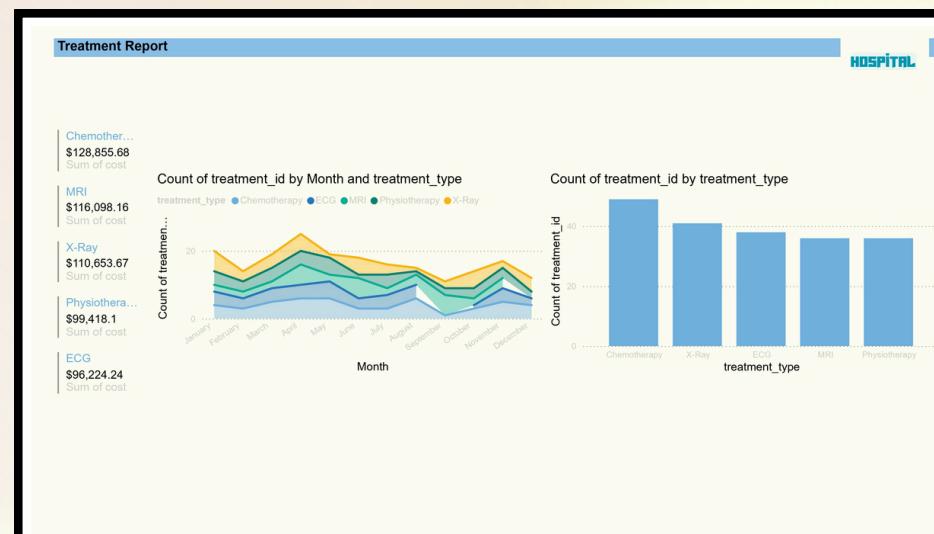
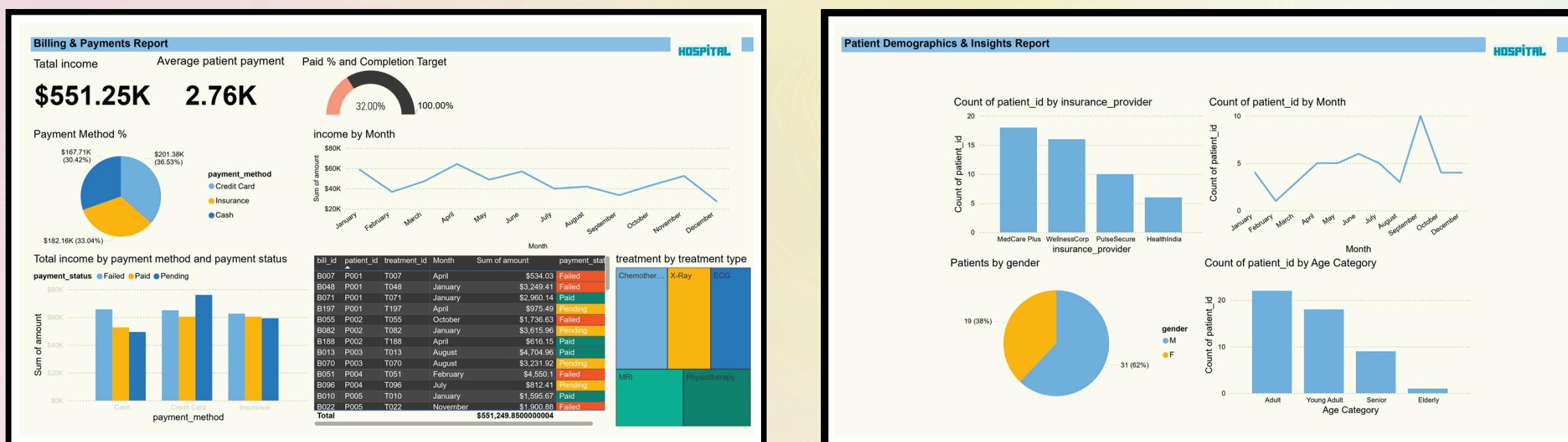
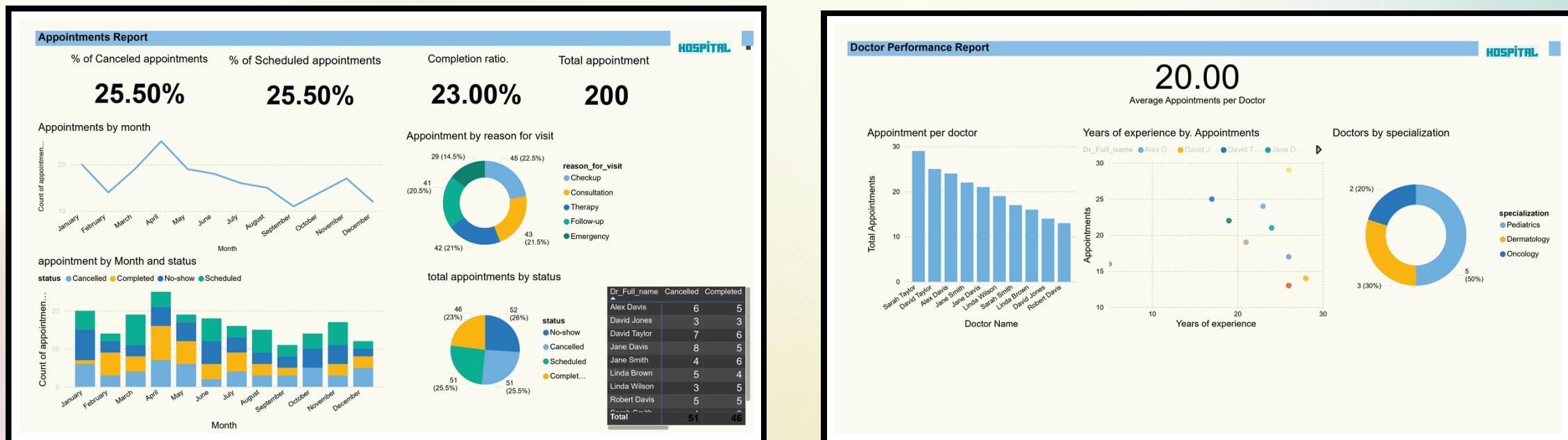


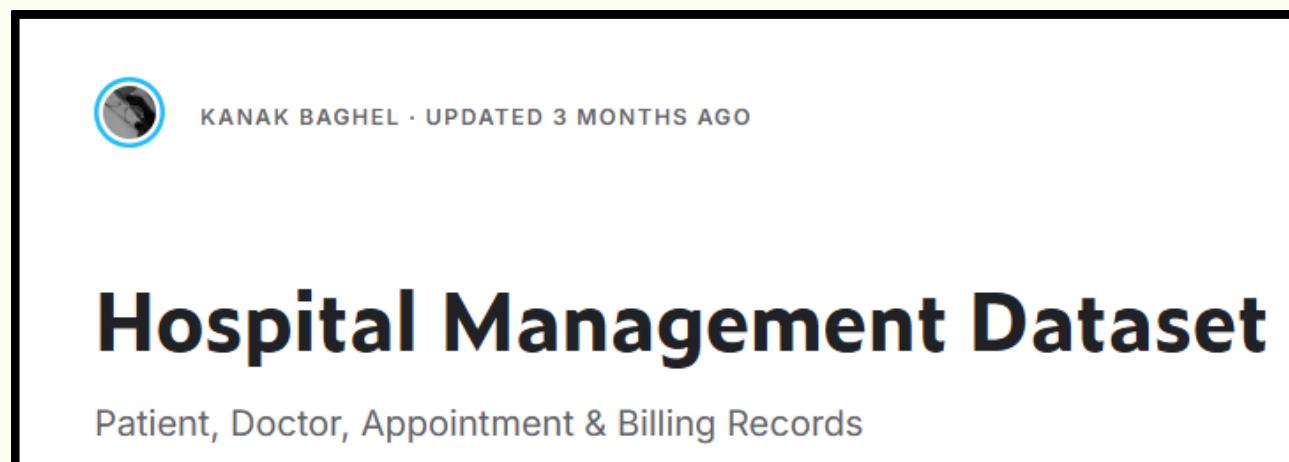
Enhancing Healthcare Data Insights with Python and Power BI

This project demonstrates how **Python** and **Power BI** can be leveraged to transform healthcare data into meaningful insights. By integrating ETL processes in Python and visualizations in Power BI, this project provides a comprehensive overview of the data analysis lifecycle.



Data Source and Import

- Data Source: The dataset used is sourced from Kaggle, containing comprehensive healthcare information including patient demographics, appointments, billing, and treatments.



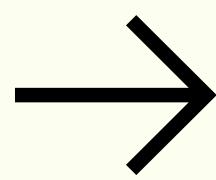
- Import Process: The data is imported into a Python environment using Pandas, allowing for efficient data manipulation and exploration.

```
# Define folder path where CSV files are located
folder_path = r'C:\Users\nawaf\Downloads\HospetalDB' # change if needed

# List CSV files in folder
csv_files = [f for f in os.listdir(folder_path) if f.endswith('.csv')]

# Load CSV files into a dictionary
dataframes = {f.split('.csv')[0].strip().lower(): pd.read_csv(os.path.join(folder_path, f))
              for f in csv_files}

# Assign variables
appointments = dataframes.get('appointments')
patients = dataframes.get('patients')
doctors = dataframes.get('doctors')
treatments = dataframes.get('treatments')
billing = dataframes.get('billing')
```



ETL Process in Python

- Explore: Utilized Pandas to examine the data with methods such as `.head()`, `.tail()`, and `.info()` to understand its structure.
- Transform: Although the data is already clean, this step illustrates the ETL process, including data transformation and preparation for analysis.

```
# 1. rename columns for consistency
patients = patients.rename(columns={"first_name": "patient_name", "Age": "patient_age"})
print(patients.head())
```

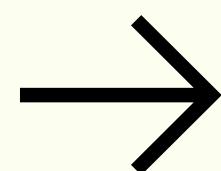
- Load: The processed data is then ready for visualization.

```
# Save cleaned tables for Power BI (CSV format)
export_path = os.path.join(folder_path, "cleaned")

os.makedirs(export_path, exist_ok=True)

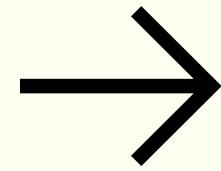
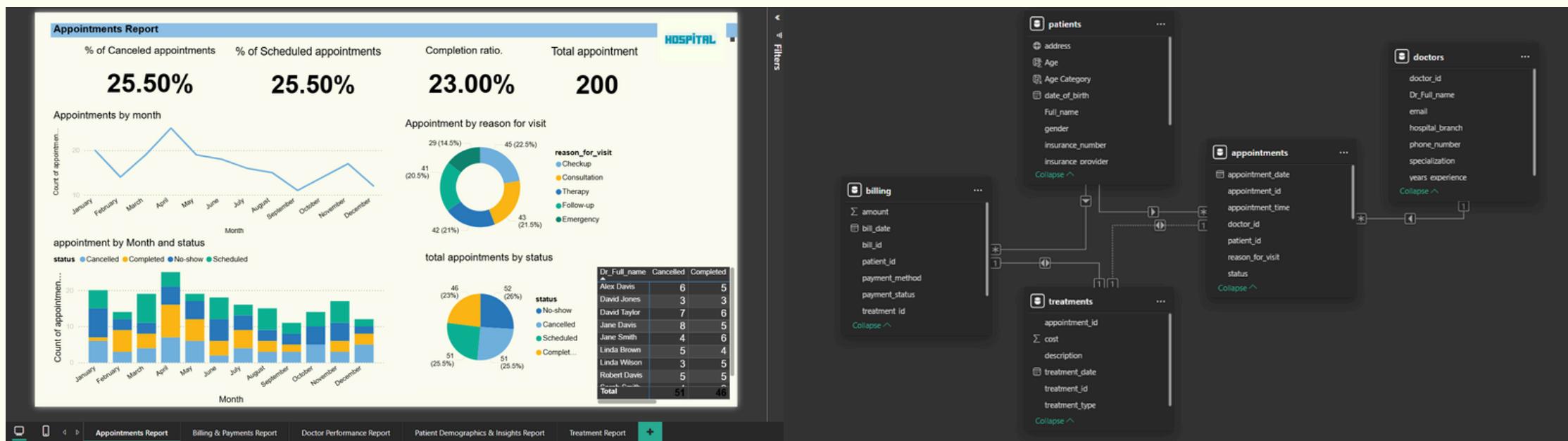
patients.to_csv(os.path.join(export_path, "patients_clean.csv"), index=False)
appointments.to_csv(os.path.join(export_path, "appointments_clean.csv"), index=False)
doctors.to_csv(os.path.join(export_path, "doctors_clean.csv"), index=False)
treatments.to_csv(os.path.join(export_path, "treatments_clean.csv"), index=False)
billing.to_csv(os.path.join(export_path, "billing_clean.csv"), index=False)

print("✅ Cleaned data exported successfully for Power BI.")
```



Visualization with Power BI

- Dashboard Components: The Power BI dashboard includes multiple visualizations:
 - **Patient Statistics:** Demographic breakdowns and visit patterns.
 - **Appointment Trends:** Analysis of appointment frequencies and no-shows.
 - **Billing Data:** Financial metrics and trends.
 - **Doctor Performance:** Insights into doctor-patient interactions
 - **Treatment Analysis:** Frequencies and outcomes of various treatments.



This project highlights the power of combining **Python** for data manipulation with **Power BI** for visualization, providing a robust framework for healthcare data analysis. It serves as a practical example of how data engineering and business intelligence can drive better decision-making in the healthcare sector.



By Nawaf
Alhomeed



Health Informatics,
Data Engineer

SAVE & SHARE