INITIAL FINDINGS AND ANALYSIS PLAN

Phase 1: Setup and Planning

Project Name: **CareNova Health Patient Care Analysis**

Datasets Provided:

1. DoctorDetails.csv
2. Patientinfo.csv
3. TreatmentRecord.csv

Summary

1. **DoctorDetails.csv:**

- This dataset consists of a total of 300 records which includes details about Doctors like Doctor's ID, their Names, area of specialization and names of the hospitals where their duties are assigned.

- There are a total of 20 hospitals and 300 distinct doctors in the dataset.

- 5 different specializations which includes (Cardiologist, Endocrinologist, General Physician, Neurologist and Pulmonologist)

- The number of doctors in each specialization are:

  (a) Cardiologist – 69
  (b) Endocrinologist – 62
  (c) General Physician – 61
  (d) Neurologist – 58
  (e) Pulmonologist – 50

- The maximum years of experience of a doctor among the record is 29 and there are 12 doctors with same years of experience.

2. **Patientinfo.csv:**

- This dataset consists of 4000 unique records.
- The records consist of details about the patients which includes (Patient ID, Name, Gender, Age, Region and Diseases)
- Majority of the patients are from age group 50 and above (2228).
- Majority of the patients are suffering from Cardiac Issues.
- The patients suffering from various diseases are as follows:

  (1) Cardiac Issue - 833
  (2) Diabetes - 761
  (3) COVID-19 - 842
  (4) Asthma - 762
  (5) Hypertension - 802

- The Male and Female count of patients are almost similar.

  (a) Male - 1385
  (b) Female -1318
  (c) Others – 1297

- Majority of the patients are from North region.

  (1) North - 1059
  (2) East - 1000
  (3) South - 981
  (4) West – 960

3. **TreatmentRecord.csv:**

- This dataset consists of a total of 10000 records.
- It consists of records like (Patient ID, Doctor ID, Treatment Date, Outcome, Treatment Duration Days, Treatment Cost and Satisfaction Score).
- This particular historical data will play a major role in our descriptive analysis. From these records we can analyse the patients who received treatments from the hospital, Efficiency of the doctors, identify the trends and readmission rates, success rates of the treatments, cost analysis and satisfaction level of the patients.

OBJECTIVES

Support CareNova's leadership with insights that improve patient outcomes, reduce readmissions, optimize departmental performance, and enhance revenue efficiency across 20 hospitals.

Key KPIs to Focus On:

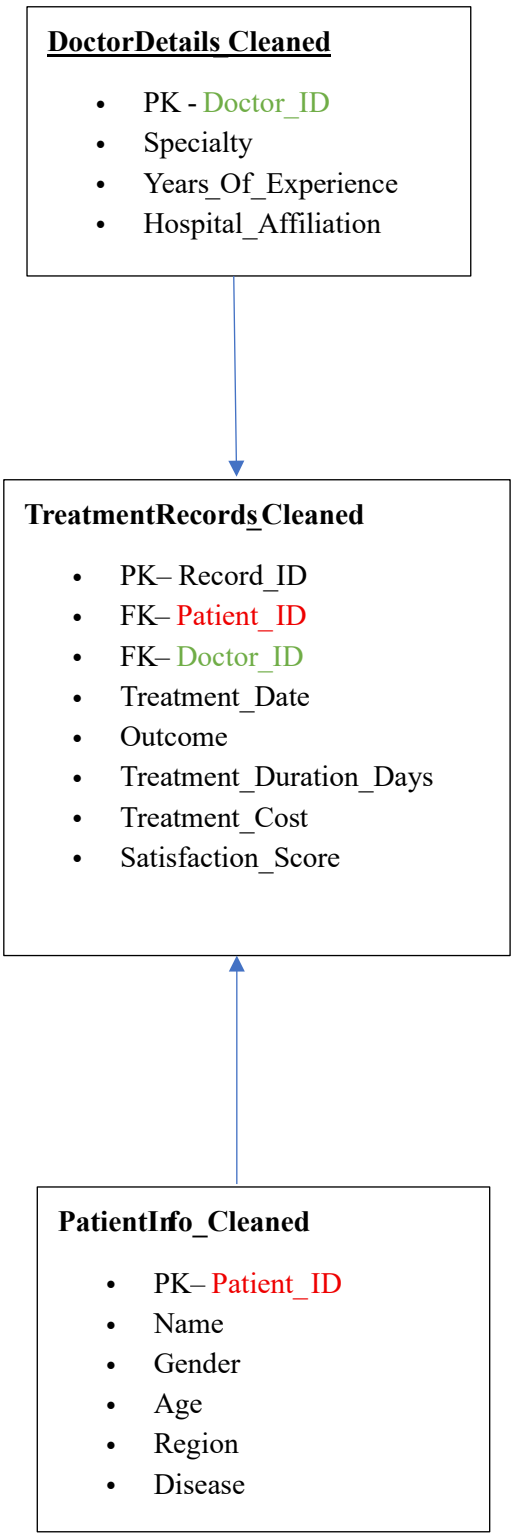
Based on the initial analysis, the following KPIs will be important for evaluating the performance or outcomes related to the data.

- KPI 1: Average Treatment Cost
- KPI 2: Treatment Success Rate
- KPI 3: Readmission Rate
- KPI 4: Average Treatment Duration
- KPI 5: Patient Satisfaction Score
- KPI 6: Cost per Recovery

Tools & Methods

- **Excel:** Initial EDA, pivot tables, validations
- **SQL Server:** Primary source for joins, aggregations, readmission logic
- **Python:** Cleaning, merging, correlation, visualizations ➢ **Power BI:** Interactive dashboard with KPIs and drill-through
- **AI Tools Used:**
  - ChatGPT for SQL/Python drafting ▪ Power BI Copilot for DAX insights
  - Excel Copilot for formula assistance

EI DIAGRAM


DoctorDetails_Cleaned
PK - Doctor_ID
Specialty
Years_Of_Experience
Hospital_Affiliation
TreatmentRecords_Cleaned
PK– Record_ID
FK– Patient_ID
FK– Doctor_ID
Treatment_Date
Outcome
Treatment_Duration_Days
Treatment_Cost
Satisfaction_Score
PatientInfo_Cleaned
PK– Patient_ID
Name
Gender
Age
Region
Disease