# Phase 2: Data Exploration and Question Framing

## A. Excel: Exploratory Analysis and Validation

### 1. What is the monthly trend of patient admissions?

**Steps:**

1. Created 2 new columns(Admission_Month & Admission_Year) using formulas =TEXT([@[Treatment_Date]],"MMM") and =TEXT([@[Treatment_Date]],"YYY")

2. Created a Pivot table to show monthly admission counts and inserted a Line Chart based on the pivot table to show trends.

### 2. Highlight Departments with >80% Satisfaction

Steps:

a) Created a new column 'Speciality' in Treatment Records Table and used Vlookup to derive its values using Doctor_ID column.

Formula = =VLOOKUP([@[Doctor_ID]],DoctorDetails_Cleaned,3,FALSE)

b) Then added a pivot table with Specialty and Average of Satisfaction_Score. (Since Satisfaction_Score coumn contains decimal data, I had to use a helper column so that excel can treat the column values as numeric)

Formula(column) = =VALUE([@[Satisfaction_Score]])

c) Applied conditional formatting to highlight department with 80% satisfaction on pivot table.
(Since no department had 80% Satisfaction_score, it was not highlighted.)

### 3. Classify Patient Age Groups (Child, Adult, Senior)

Created new column and and applied
=IFS([@Age]<18, "Child", [@Age]<=50, "Adult", [@Age]>50, "Senior")

### 4. Average Treatment Cost by Department

Created a pivot table with 'Speciality' and 'Average of Treatment_Cost' and converted the number format to currency.

### 5. Identify Top 5 Doctors by Number of Patients Treated

Created a pivot table with 'Doctor_ID' and 'Count of Patient_ID' and used Top 10 filter.

6. **Lookup Patient Region Based on Patient ID**

   Formula : =VLOOKUP(B64,Patient_info!A:G,5,FALSE)

7. **Count Patients Admitted for "Cardiac" Conditions**

   Added new column 'Disease' to Traetment Records Table and derived values using Vlookup (=VLOOKUP([@[Patient_ID]],Patient_info!A:F,6,FALSE)).

   Then used 'CountIF' to find the count of Patients with Cardiac issues.
   =COUNTIF(TreatmentRecords_Cleaned[Disease], "Cardiac Issue")

8. **Restrict Invalid Department Entries (Data Validation)**

   Applied Data restriction to specialty columns using Data Validation. Any enteries other than (Cardiologist, Neurologist, Endocrinologist, General Physician, Pulmonologist) throws error message.

9. **Chart for Treatment Outcome Rates by Department**

   Added a pivot table,
   (Rows – Specialty. Columns – Outcome, Values – Count of Record_ID) and then inserted a Column chart to show distribution.

10. **How Many Patients Stay Beyond 7 Days?**

    Created a new column named 'Long Stay (>7 days)' using IF,
    =IF([@[Treatment_Duration_Days]]>7, "Yes", "No").

    Using COUNTIF, calculated the count of 'Yes' in 'Long Stay (>7 days)' column,
    =COUNTIF(TreatmentRecords_Cleaned[Long Stay (>7 days)],"Yes")