🏥SQL Project: Hospital Intelligence Dashboard (100 Marks)

Project Theme:

You are hired as a **junior healthcare data analyst** at a leading hospital. The management team wants data-driven insights to improve services, manage costs, and understand patient behavior. You will query the hospital's database and submit your analysis to stakeholders.

🚃Database Coverage

You will be working across multiple tables:

- `encounters`
- `patients`
- `procedures`
- `payer_coverage` / `billing`
- `diagnoses`
- `medications`

🚡Submission Instructions:

- Submit a PDF report with **SQL queries**, **brief explanations**, and **results (screenshots or output tables)**.
- Write clean SQL with formatting and comments.
- Each question shows **concepts involved**, **hint(s)**, and **plain English formulae** for clarity.

🏫SECTION A: Encounter Trends (25 Marks)

💈 1. Total Encounters per Year (5 marks)

Q: How many encounters occurred in each calendar year?

- 💡 **Hint**: Extract the year from each encounter's start date and count how many rows belong to each year.

💈 2. Yearly % by Encounter Class (8 marks)

Q: For each year, show the percentage of encounters in each class (`inpatient`, `emergency`, etc.).

- 💡 **Hint**: For each year, calculate how many encounters fall in each class, and what percentage that is of the total for that year.

💈 3. Duration-Based Classification (12 marks)

Q: Categorize encounters as `Short Stay (<24h)` or `Long Stay (>=24h)` and show the yearly percentage split.

- 💡 **Hint**: Calculate the time difference in hours between the start and end of each encounter, classify it as short or long stay, then show what percent each type is by year.

SECTION B: Financial & Coverage Insights (25 Marks)

💈 4. Zero Payer Coverage (6 marks)

Q: How many encounters had no payer coverage? What % is that?

- 💡 **Hint**: Filter rows where the payer field is empty or null, then calculate what fraction of all encounters that is.
- 📐 **Formula**: (Number of encounters with no payer) ÷ (Total number of encounters) × 100

💈 5. Top 10 Frequent Procedures (7 marks)

Q: List the 10 most frequently performed procedures and the average base cost for each.

- 💡 **Hint**: Group by procedure name or ID, count how many times each appears, and average their base cost.

💈 6. Costliest Procedures (6 marks)

Q: Which 10 procedures have the **highest average base cost**, and how many times were they performed?

- 💡 **Hint**: Group by procedure, calculate average cost, and sort by it in descending order.

💈 7. Claim Cost by Payer (6 marks)

Q: Show the average total claim cost for each insurance provider.

- 💡 **Hint**: Join payer table and billing or encounters to group by payer and calculate average total billed amount.
- 📐 **Formula**: For each payer, sum up the total claims and divide by the number of claims to get average.

SECTION C: Patient Behavior & Risk Analysis (30 Marks)

💈 8. Unique Patients per Quarter (6 marks)

Q: How many **distinct patients** were admitted per quarter, per year?

- 💡 **Hint**: Extract the year and quarter from each encounter date, and count distinct patient IDs.

💈 9. Readmissions within 30 Days (10 marks)

Q: Count how many patients were readmitted within 30 days.

- 💡 **Hint**: For each patient, compare the start date of their current encounter with the end date of their previous one.
- 📐 **Formula**: If the difference between the current and previous encounter (for the same patient) is $\leq$ 30 days, it's a readmission.

💈 10. Top 5 Most Readmitted Patients (6 marks)

Q: Which 5 patients had the highest number of readmissions?

- 💡 **Hint**: From previous step, count how many readmissions per patient, then sort and take top 5.
- 📐 **Formula**: Count how many readmissions each patient has and list the top 5 by this count.

💈 11. First vs. Latest Encounter Analysis (8 marks)

Q: For each patient, show their first and latest encounter date, along with days between.

- 💡 **Hint**: Use the `MIN()` and `MAX()` of encounter start dates per patient.
- 📐 **Formula**: For each patient, calculate (latest encounter date - first encounter date) in days.

SECTION D: Advanced Logic (20 Marks)

💈 12. CTE + CASE Pivot Table (10 marks)

Q: Use a CTE to show a **pivot-style table**: for each patient, count how many of their encounters fall into each `encounter_class`.

- 💡 **Hint**: Use CASE WHEN logic to check each type and add them up with SUM.
- 📐 **Formula**: For each patient, count how many of their visits were 'Emergency', 'Inpatient', etc., using conditional logic.

💈 13. Most Recent Encounter per Patient (5 marks)

Q: Using a **CTE**, find the latest encounter per patient, including date and procedure.

- 💡 **Hint**: Use MAX on encounter date per patient, join to get details.
- 📐 **Formula**: For each patient, find the latest date and link it back to the original record to get full details.

💈 14. Top Diagnoses Per Age Group (5 marks)

Q: Show the most common diagnosis per **age group** (e.g., 0–20, 21–40, etc.)

- 💡 **Hint**: Use CASE to make age brackets, group by them and diagnosis.
- 📐 **Formula**: Assign age groups using CASE logic and count diagnoses within each.

Total Marks: 100

Let me know if you'd like this exported to PDF or formatted into a clean student handout!