Problem Statement: Smoking & Cancer Risk Analysis

This dataset comprises detailed records of patients, focusing on the relationship between smoking behaviour and cancer risk. It integrates demographic, lifestyle, environmental, and medical information to support analytical research in Oncology, Epidemiology, and preventive healthcare.

The dataset has been normalized into three relational tables:

- Patients: General demographic and lifestyle attributes
- Habits: Smoking history and exposure levels
- Results: Cancer type and Cancer stage.

Smoking & Cancer: A Structured Query Analysis of Patient Data

1. List all patients who are older than 50 and have a poor diet.
2. Display the distinct physical activity levels recorded in the dataset.
3. Show the top 5 patients with the highest number of cigarettes smoked per day.
4. Find the average BMI grouped by physical activity level.
5. Count how many patients fall into each smoking status category.
6. List the number of cancer patients per cancer type ordered by patient count
7. Find the average years of smoking for patients who are current and former smokers.
8. Get the gender-wise average BMI of patients with cancer.
9. List patients who smoke more than the average number of cigarettes per day.
10. Create a new column named Risk_Level using CASE:
    - "High" if smoking status is 'Current'
    - "Medium" if smoking status is 'Former'
    - Else "Low"
11. Identify patients who have smoked for more than 20 years but have not been diagnosed with any cancer.
12. Find the top 3 cancer types with the highest number of female patients.