INSURANCE DATASET

Beginner Level Questions:

1. Select all the columns from the dataset.
2. Retrieve the list of unique regions from the dataset.
3. Find the total number of patients from the dataset.
4. Get the total number of smokers.
5. Retrieve the patient IDs and claims of patients who are diabetic.
6. Count how many children the patients have in total.
7. Find all the male patients from the dataset.
8. Select all columns for patients whose BMI is greater than 25.
9. Get the average age of all patients.
10. Find the total number of patients from the 'Southwest' region.

Intermediate Level Questions:

1. Calculate the total insurance claim for patients who are smokers.
2. Retrieve the average blood pressure of patients who are diabetic.
3. Find the total number of female patients with more than two children.
4. Get the highest claim amount by any patient from the dataset.
5. List patients who are aged above 40, have a BMI greater than 30, and are smokers.
6. Retrieve the number of patients from each region.
7. Find the minimum claim amount for patients who are not diabetic.
8. Calculate the total insurance claim amount for patients from the 'Northeast' region.
9. List all patients where the claim amount is greater than 10,000 and who have more than two children.
10. Retrieve the number of diabetic patients in each region.

Advanced Level Questions:

1. Calculate the total insurance claims for patients grouped by gender and smoking status.
2. Find the top 3 regions with the highest average claims for smokers.
3. Identify the patients who are diabetic, above 50 years old, and have a blood pressure higher than the average.
4. Find the top 5 patients who have made the highest insurance claims.
5. Calculate the percentage of smokers in each region.
6. Retrieve the average claim of patients with at least two children, segmented by region.
7. Find the average claim amount for diabetic patients segmented by their smoking status and region.
8. List the top 5 patients in each region based on their claim amount.
9. Calculate the variance of insurance claims for each region.
10. Find the maximum claim amount for each region where the patient is a smoker and over 50 years old.