**Data Columns:**

* age: age of primary beneficiary
* sex: insurance contractor gender, female, male
* bmi: Body mass index, providing an understanding of body, weights that are relatively high or low relative to height,  
  objective index of body weight (kg / m ^ 2) using the ratio of height to weight, ideally 18.5 to 24.9
* children: Number of children covered by health insurance / Number of dependents
* smoker: Smoking
* region: the beneficiary's residential area in the US, northeast, southeast, southwest, northwest.
* charges: Individual medical costs billed by health insurance

**Analysis Questions for Medical Cost Data(Each Question - 5 Marks) — Total : 50**

1. Basic Statistics and Data Understanding

- Question: Calculate the average, median, and standard deviation of medical charges, BMI, and age across the entire dataset.

- Relevant Excel Functions: AVERAGE, MEDIAN, STDEV

2. Gender Analysis

- Question: Compare the average medical charges and BMI between male and female beneficiaries. Which gender has higher average costs and BMI?

- Relevant Excel Functions: AVERAGEIF, PIVOT TABLE

3. Impact of Smoking on Costs

- Question: Analyze how smoking status influences medical costs. Calculate the average cost for smokers versus non-smokers.

- Relevant Excel Functions: AVERAGEIFS, PIVOT TABLE, CHARTS

4. Dependents Analysis

- Question: Evaluate the relationship between the number of children/dependents and medical charges. Do beneficiaries with more dependents incur higher medical costs?

- Relevant Excel Functions: CORREL, LINE CHART, SCATTER PLOT

5. Geographical Insights

- Question: Determine which US region has the highest average medical costs and BMI. Create a pivot table and chart that summarize this data by region.

- Relevant Excel Functions: PIVOT TABLE, AVERAGEIFS, BAR CHART

6. Cost Distribution Analysis

- Question: Create a histogram to visualize the distribution of medical charges. Identify the most common cost bracket.

- Relevant Excel Functions: HISTOGRAM, FREQUENCY, DATA ANALYSIS TOOLPAK

7. BMI Classification

- Question: Classify beneficiaries into BMI categories as Underweight (<18.5), Normal (18.5-24.9), Overweight (25-29.9), and Obese (30+). Count how many beneficiaries fall into each category.

- Relevant Excel Functions: COUNTIFS, VLOOKUP with range lookup

8. Advanced Correlation Analysis

- Question: Use a scatter plot to analyze the correlation between BMI and medical charges. Add a trendline to the scatter plot to assess the relationship.

- Relevant Excel Functions: SCATTER PLOT, TRENDLINE, CORREL

9. Smoking and Region Interaction

- Question: Investigate the interaction effect between smoking and region on medical charges. Which region has the highest cost for smokers?

- Relevant Excel Functions: PIVOT TABLE, SLICERS, AVERAGEIFS

10. Dashboard Creation for Stakeholders

- Question: Develop a comprehensive dashboard that provides stakeholders with insights into medical costs, including filters for gender, smoking status, region, and BMI categories.

- Relevant Excel Functions: DASHBOARD DESIGN, PIVOT CHARTS, SLICERS, INTERACTIVE ELEMENTS

Submission Format :

1. Solve all the questions and after finishing the project, create one presentation.
2. You need to record a video explaining your project and all main points in 10 minutes. Upload that video on youtube and add the link in your presentation.
3. Submit that ppt on the portal exactly below from where you downloaded this docx file.