Capstone project on smoking survey in UK.

**Executive Summary:**

* Examines smoking habits and demographic characteristics in the United Kingdom.
* Focuses on demographic attributes, regional disparities, and types of tobacco consumed.

**Problem Statement:**

* Smoking is a significant public health concern.
* Need for a detailed analysis of smoking habits in the UK for targeted interventions and policy development.

**Background and Objective:**

* Persistent health challenges associated with smoking.
* Objective: Identify key demographic factors, quantify prevalence, and analyze tobacco product consumption.

**Data Source:**

* Dataset: 1691 observations, 12 variables (gender, age, marital status, education, nationality, ethnicity, income, region, smoking status, cigarette details).
* Data collected through surveys in the United Kingdom.
* Link to dataset: [Smoking survey data](https://www.kaggle.com/datasets/utkarshx27/smoking-dataset-from-uk).

**Methodology:**

* Exploratory Data Analysis (EDA): Understand variable distribution, identify outliers, and explore relationships.
* Descriptive Statistics: Calculate summary statistics for key features.
* Data Visualization: Matplotlib, Seaborn.

**Expected Outcomes:**

1. Identification of Demographic Patterns: Uncover characteristics associated with smoking habits.
2. Quantification of Smoking Prevalence: Estimate prevalence in different groups.
3. Insights into Tobacco Consumption: Analyze popularity of different tobacco products among demographics.

**Tools and Technologies:**

* Dashboard Development: Tableau/Power BI.
* Data Cleaning and Pre-processing: Excel.
* Programming Languages: Python (Pandas, NumPy, Matplotlib, Seaborn).

**Risks and Challenges:**

1. Data Quality: Ensure accuracy and completeness.
2. Ethical Considerations: Handle sensitive information and ensure privacy.
3. Generalizability: Acknowledge limitations and potential biases.

**Conclusion:**

* Aims to provide critical insights for public health initiatives in the UK.
* Understanding demographic patterns and tobacco consumption to inform targeted interventions and policies.
* Contributes to reducing smoking rates and improving overall population health.

Insights from the dataset:  
  
  
  
1) The minimum age of male and female who smokes in 16 years old.  
2) As per the survey data, 27.23% of people smokes and 72.77% doesn’t smoke.  
3) Female smokes more compare to males.   
4) As per the data, the major people who said no for cigarettes are of no qualification people.

5) Based on martial status, we can see that single people smokes more.  
6) Based on region The north people smoke more.  
7) Based on nationality, we can see that the English and Britain people smoke more.   
8) The maximum income of smokers is 51K.   
  
  
  
  
  
  
  
  
  
  
  
  
  
  
1 What’s the minimum age of male and female who smokes  
2 Qualification of males and females who smoke  
3 Which nationality and ethnicity people smoke  
4 Average income of male and female who smokes  
5 What the average age of the male and female who smokes  
6 Average age of male and males who smokes  
7 Which ethnicity and nationality smokes mor  
8 Amount of cigarette, males and females smoke on weekdays and weekends.  
9 Which type of cigarette people mostly preferred on weekdays and weekends  
10 What’s the martial status of male’s and female’s who smokes.   
11 Region wise of male and female smokers  
12 Amount and type of cigarettes on weekends and weekdays. (Max and Min)