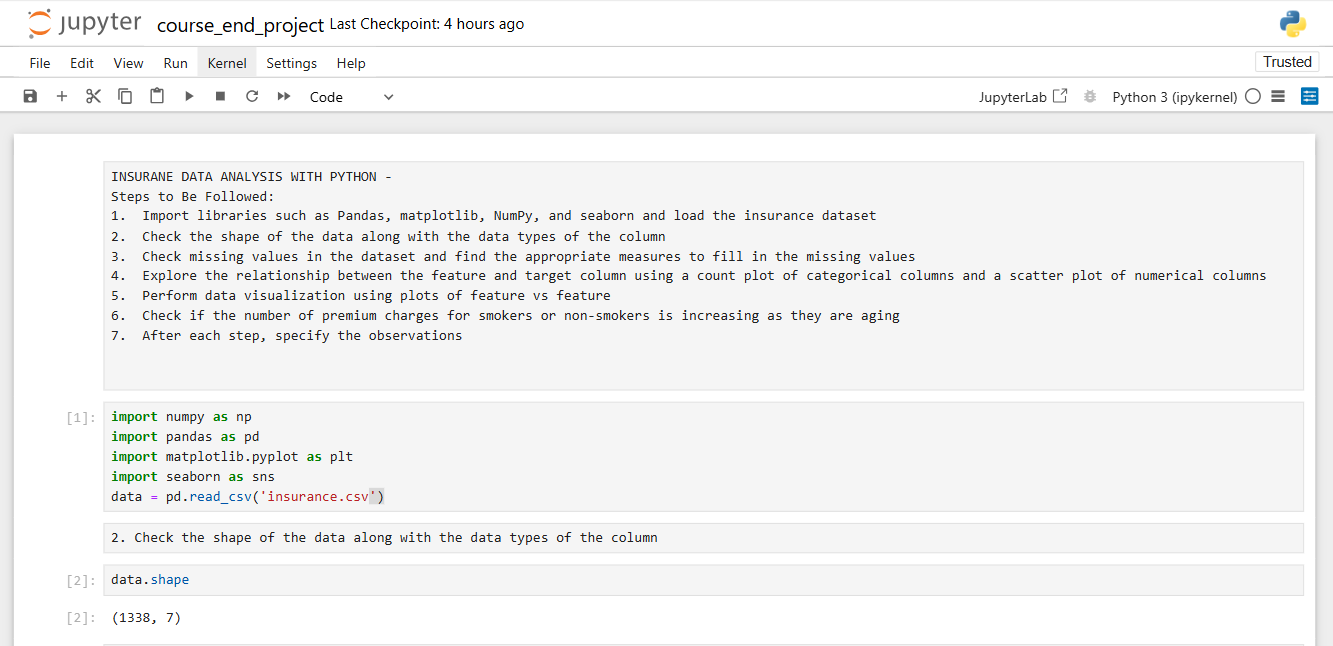
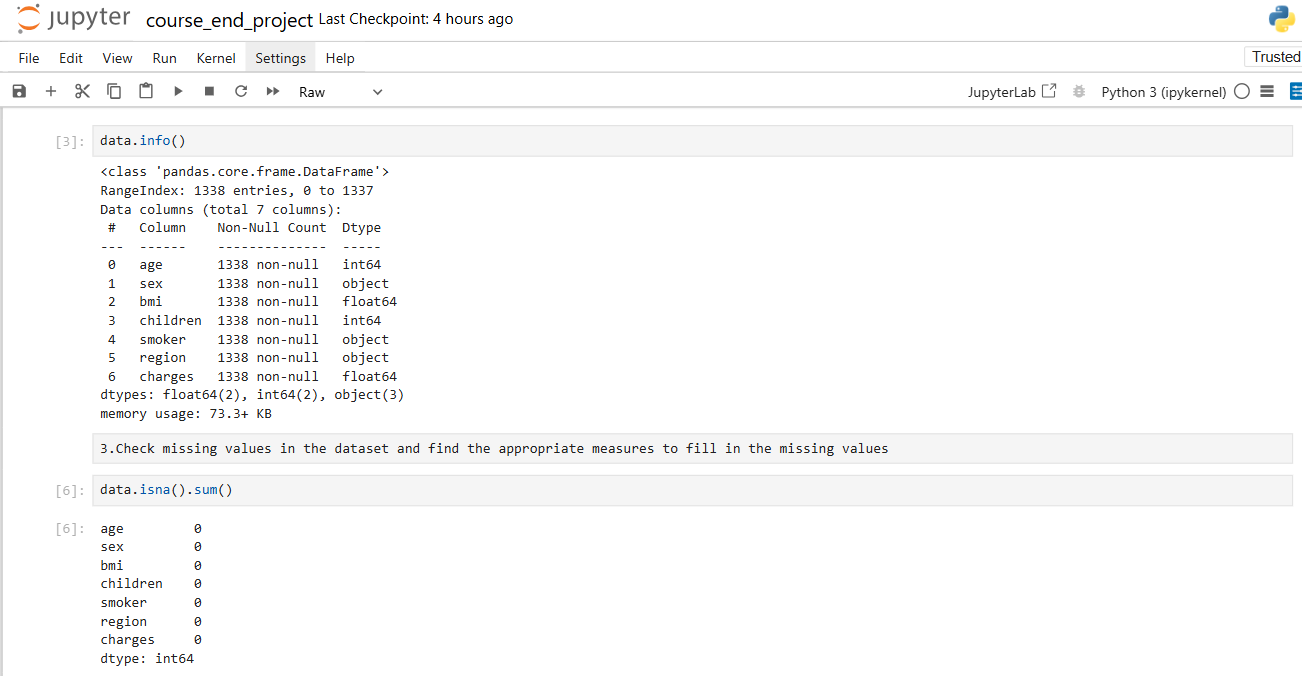
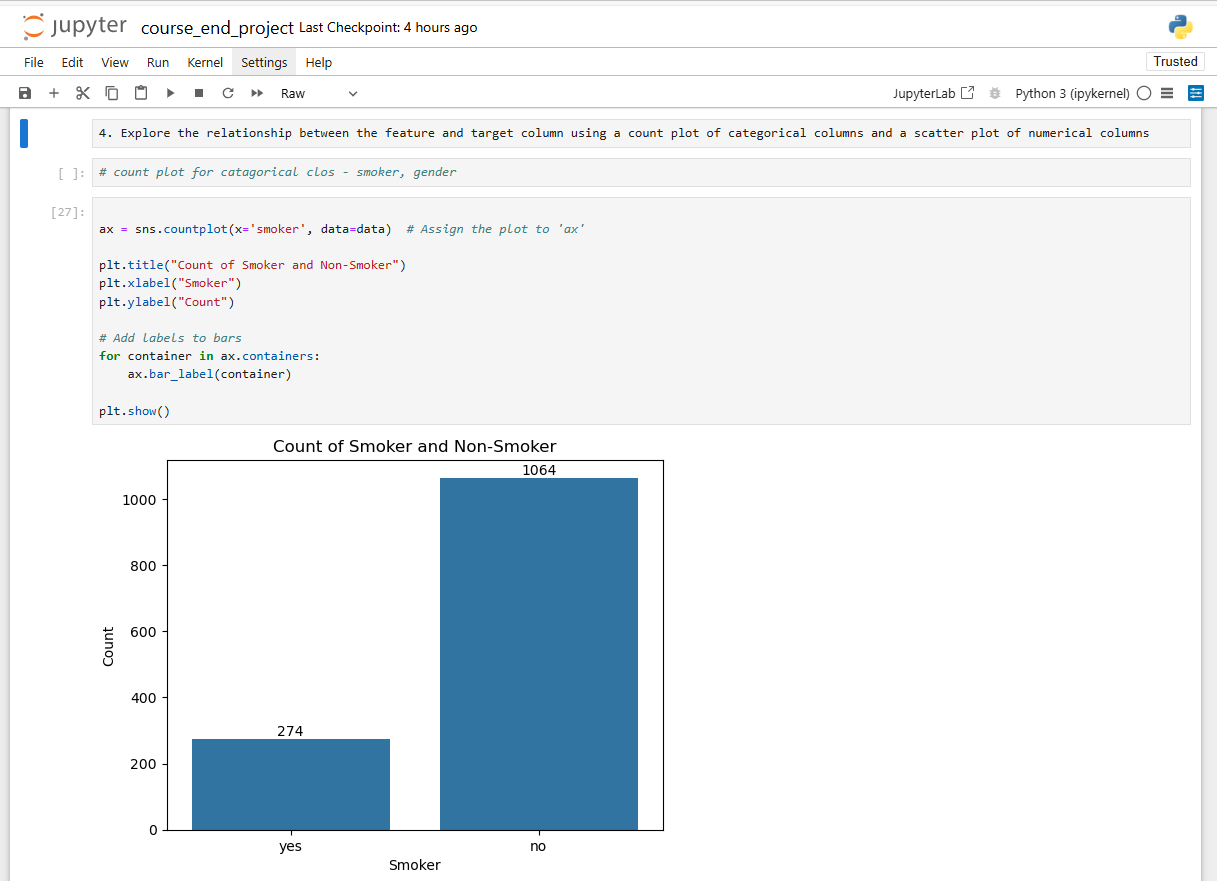
**Data Analytics with Python**  
**Insurance Data Analysis**

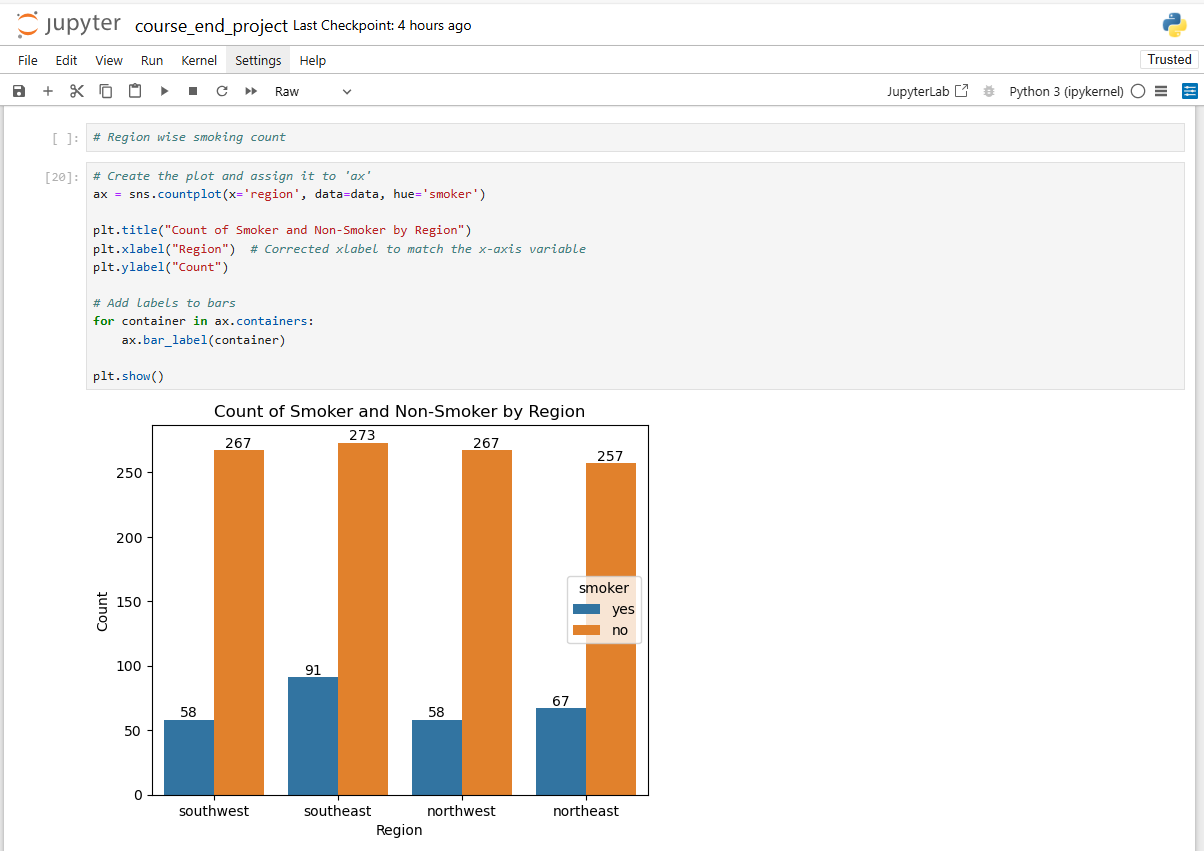
**Steps to Be Followed:**

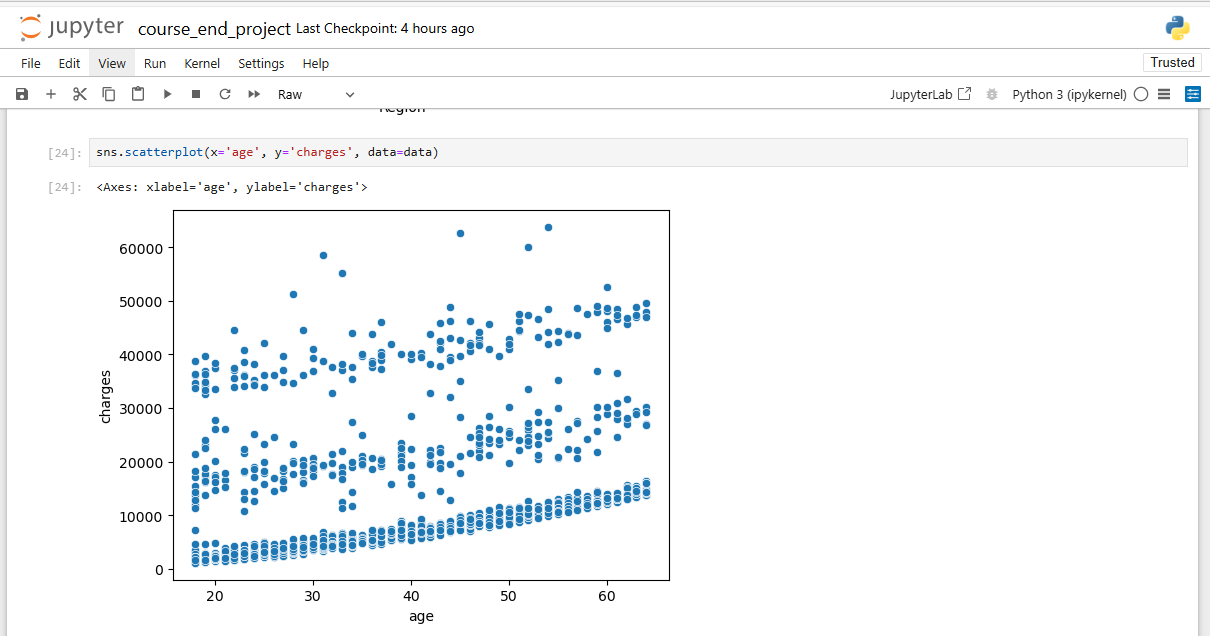
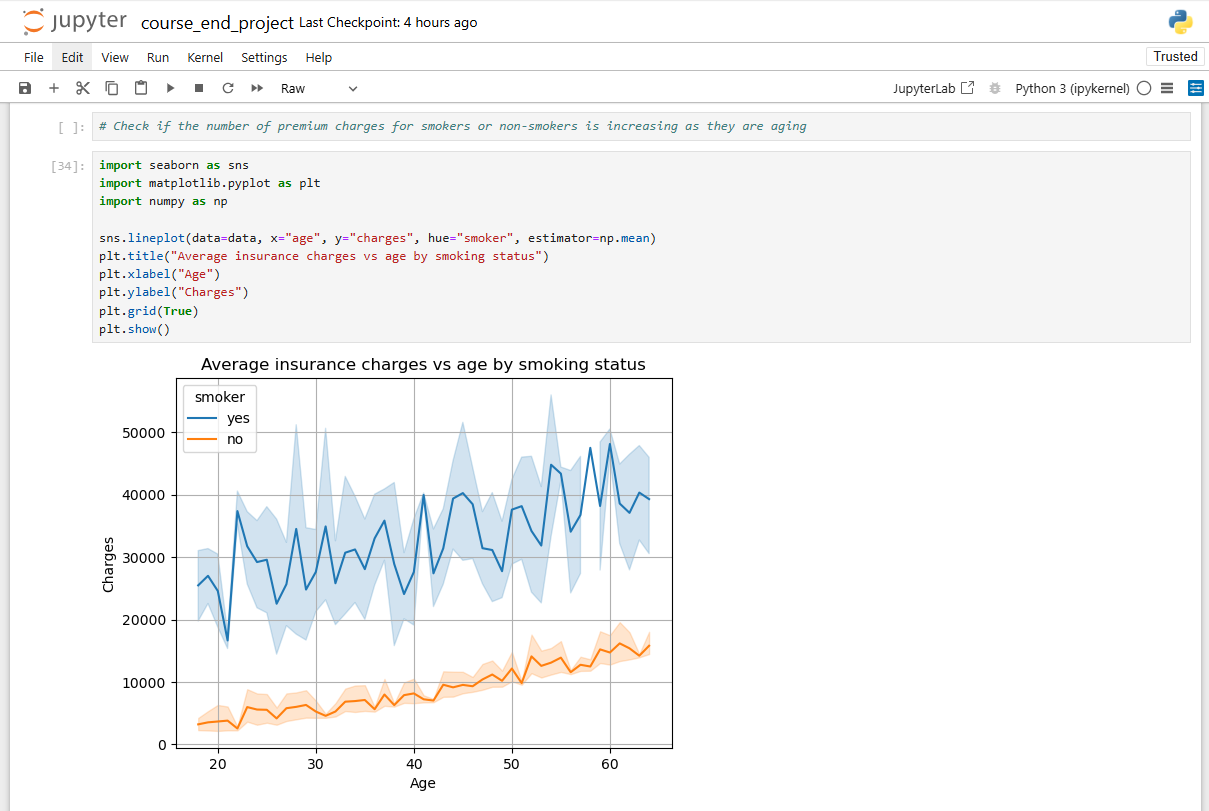
1. Import libraries such as Pandas, matplotlib, NumPy, and seaborn and load the insurance dataset
2. Check the shape of the data along with the data types of the column
3. Check missing values in the dataset and find the appropriate measures to fill in the missing values
4. Explore the relationship between the feature and target column using a count plot of categorical columns and a scatter plot of numerical columns
5. Perform data visualization using plots of feature vs feature
6. Check if the number of premium charges for smokers or non-smokers is increasing as they are aging
7. After each step, specify the observations

  
  
***Import libraries such as Pandas, matplotlib, NumPy, and seaborn and load the insurance dataset***

***Check the shape of the data along with the data types of the column***

There has total 7 columns in the given data set i.e. age, sex, bmi, children, smoker, regions, charges. Out of these 7 columns no missing value is present.  
  
***Check missing values in the dataset and find the appropriate measures to fill in the missing values***As per the analysis, out of this 1000 people smoker count is 274 and non-smoker count is 1064.  
  
***Explore the relationship between the feature and target column using a count plot of categorical columns and a scatter plot of numerical***



As per the analysis the smoker count in the southwest region is 58, in south west it’s 91, in north west it is 58 and in north east it is 67. In the non-smoker criteria the count for southwest is 267, in southeast it’s 273, in northwest it’s 267 and in the northeast, it is 257.   
  
Hence, out of total 274 smokers, southeast region has the largest number of smoker and out of total 1064 non-smoker people southeast has the largest quantity.  
  
  
 ***Perform data visualization using plots of feature vs feature***As per the BMI vs Age chart 🡪Non-smokers are present between the age group of 20 to 50.  
  
As per the Children vs age chart 🡪 Non-smokers are present between age group of 20 to 50.  
  
As per the Age vs Charges chart 🡪 Smokers are present between the age group of people 20 to 60  
  
As per the BMI vs Change chart 🡪 Charges are 20000 to 60000 between the BMI having 15 to 50.   
***Check if the number of premium charges for smokers or non-smokers is increasing as they are aging***

As per the insurance charges vs age by smoking status chart, people having age group having of 20 to 60 has the charges of 20000 to 50000.