Enrollment No:_____ **Machine ID: _____**

Faculty of Computer Applications and Information Technology
iMScIT Programme – Sem VIII
Mid Term Exam-2024

Subject Name: Machine Learning With Python
Date: 28/02/2024 **Time: 09:30-10:30 AM**

Set 1:

1. Load the CSV file named "carsales.csv" into a Pandas DataFrame. Display the first 5 rows.
2. Calculate the mode of the "Make" column using Pandas.
3. Group the DataFrame by the "Year" column and calculate the median price for each year.
4. Create a scatter plot using Matplotlib to visualize the relationship between the "Price" and "Mileage" columns.
5. Filter the DataFrame to only include rows where the "FuelType" column is 'Petrol'. Display the first 5 rows.
6. Calculate the total number of missing values in each column of the DataFrame using Pandas.
7. Create a bar chart using Matplotlib to show the count of each unique value in the "BodyType" column.
8. Calculate the standard deviation of the "EngineSize" column using NumPy.
9. Plot a histogram of the "Price" column using Pandas.
10. Filter the DataFrame to only include rows where the "Year" is greater than 2015. Display the last 5 rows.

Enrollment No:_____ **Machine ID: _____**

Faculty of Computer Applications and Information Technology
iMScIT Programme – Sem VIII
Mid Term Exam-2024

Subject Name: Machine Learning With Python
Date: 28/02/2024 **Time: 09:30-10:30 AM**

Set 2:

1. Load the "carsales.csv" file into a Pandas DataFrame. Display the column names.
2. Calculate the median of the "Price" column using NumPy.
3. Group the DataFrame by the "Make" column and calculate the mean mileage for each make.
4. Create a line plot using Matplotlib to visualize the trend of "Price" over time.
5. Filter the DataFrame to only include rows where the "FuelType" column is 'Diesel'. Display the first 5 rows.
6. Calculate the sum of the "Mileage" column using NumPy.
7. Create a box plot using Matplotlib to visualize the distribution of the "Price" column.
8. Calculate the mean of the "EngineSize" column using NumPy.
9. Plot a pie chart using Matplotlib to show the distribution of values in the "BodyType" column.
10. Filter the DataFrame to only include rows where the "Mileage" is less than 50000. Display the last 5 rows.

Enrollment No:_____ **Machine ID: _____**

Faculty of Computer Applications and Information Technology
iMScIT Programme – Sem VIII
Mid Term Exam-2024

Subject Name: Machine Learning With Python
Date: 28/02/2024 **Time: 09:30-10:30 AM**

Set 3:

1. Load the "carsales.csv" file into a Pandas DataFrame. Display the shape of the DataFrame.
2. Calculate the total number of missing values in the DataFrame using Pandas.
3. Group the DataFrame by the "Year" column and calculate the median mileage for each year.
4. Create a scatter plot using Matplotlib to visualize the relationship between the "Mileage" and "EngineSize" columns.
5. Filter the DataFrame to only include rows where the "FuelType" column is 'Hybrid'. Display the first 5 rows.
6. Calculate the mean of the "Price" column using NumPy.
7. Create a bar chart using Matplotlib to show the count of each unique value in the "Make" column.
8. Calculate the standard deviation of the "Mileage" column using NumPy.
9. Plot a histogram of the "EngineSize" column using Pandas.
10. Filter the DataFrame to only include rows where the "Price" is greater than 20000. Display the last 5 rows.

Enrollment No:_____ **Machine ID: _____**

Faculty of Computer Applications and Information Technology
iMScIT Programme – Sem VIII
Mid Term Exam-2024

Subject Name: Machine Learning With Python
Date: 28/02/2024 **Time: 09:30-10:30 AM**

Set 4:

1. Load the "carsales.csv" file into a Pandas DataFrame. Display the summary statistics of the DataFrame.
2. Calculate the median of the "Mileage" column using NumPy.
3. Group the DataFrame by the "Make" column and calculate the median price for each make.
4. Create a scatter plot using Matplotlib to visualize the relationship between the "Price" and "Year" columns.
5. Filter the DataFrame to only include rows where the "FuelType" column is 'Electric'. Display the first 5 rows.
6. Calculate the total number of missing values in each column of the DataFrame using Pandas.
7. Create a bar chart using Matplotlib to show the count of each unique value in the "BodyType" column.
8. Calculate the standard deviation of the "Price" column using NumPy.
9. Plot a histogram of the "Mileage" column using Pandas.
10. Filter the DataFrame to only include rows where the "Year" is greater than 2010. Display the last 5 rows.

Enrollment No:_____ **Machine ID: _____**

Faculty of Computer Applications and Information Technology
iMScIT Programme – Sem VIII
Mid Term Exam-2024

Subject Name: Machine Learning With Python
Date: 28/02/2024 **Time: 09:30-10:30 AM**

Set 5:

1. Load the "carsales.csv" file into a Pandas DataFrame. Display the data types of each column.
2. Calculate the mode of the "FuelType" column using Pandas.
3. Group the DataFrame by the "Year" column and calculate the median price for each year.
4. Create a scatter plot using Matplotlib to visualize the relationship between the "Mileage" and "Price" columns.
5. Filter the DataFrame to only include rows where the "Make" column is 'Toyota'. Display the first 5 rows.
6. Calculate the total number of missing values in each column of the DataFrame using Pandas.
7. Create a bar chart using Matplotlib to show the count of each unique value in the "BodyType" column.
8. Calculate the standard deviation of the "Mileage" column using NumPy.
9. Plot a histogram of the "Price" column using Pandas.
10. Filter the DataFrame to only include rows where the "Mileage" is less than 60000. Display the last 5 rows.

Enrollment No:_____ **Machine ID: _____**

Faculty of Computer Applications and Information Technology
iMScIT Programme – Sem VIII
Mid Term Exam-2024

Subject Name: Machine Learning With Python
Date: 28/02/2024 **Time: 09:30-10:30 AM**

Set 6:

1. Load the "carsales.csv" file into a Pandas DataFrame. Display the shape of the DataFrame.
2. Calculate the mode of the "Make" column using Pandas.
3. Group the DataFrame by the "FuelType" column and calculate the mean price for each fuel type.
4. Create a scatter plot using Matplotlib to visualize the relationship between the "Price" and "EngineSize" columns.
5. Filter the DataFrame to only include rows where the "Year" column is 2020. Display the first 5 rows.
6. Calculate the total number of missing values in each column of the DataFrame using Pandas.
7. Create a bar chart using Matplotlib to show the count of each unique value in the "BodyType" column.
8. Calculate the standard deviation of the "Mileage" column using NumPy.
9. Plot a histogram of the "Price" column using Pandas.
10. Filter the DataFrame to only include rows where the "Year" is less than 2015. Display the last 5 rows.