Bike	Kilometers	Range
1 CD 70	50 to 60	1 liter
2. Express	40 to 50	1 liter
3. Jinan	40 to 50	1 liter
4. Hi Speed	40 to 50	1 liter
5. Impress	40 to 50	1 liter
6. Deluxe	40 to 50	1 liter
7. Unique	40 to 50	1 liter
8. Honda 100	30 to 40	1 liter
9. Suzuki 150	35 to 40	1 liter
10. Yamaha RX 115	30 to 40	1 liter
Total Bike :	**Range:**	**Total liter:**
10	345	10

“The Code Start Here”

```
# Import libraries
import pandas as pd
import numpy as np
import matplotlib as plt

import LinearRegression as Li_Re
```

```
import pandas as pd

# Define the file path using forward slashes
file_path = "C:/Users/HP
ZBOOK/OneDrive/Desktop/python/Assighment.csv"

# Read the CSV file into a DataFrame
```

```
data = pd.read_csv(file_path)

# Display the first few rows of the DataFrame
print(data.head())
```

```
data .describe()
```

```
import seaborn as sns
import matplotlib.pyplot as plt
```

```
# Import necessary libraries
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

# Define the file path using forward slashes or a raw
string
file_path = r'C:\Users\HP
ZBOOK\OneDrive\Desktop\python\Assighment.csv'

# Read the CSV file into a DataFrame
data = pd.read_csv(file_path)

# Convert 'Salary' column to numeric, setting
errors='coerce' to handle non-numeric values
data['Salary'] = pd.to_numeric(data['Salary'],
errors='coerce')

# Drop rows with NaN values in the 'Salary' column
data = data.dropna(subset=['Salary'])

# Set the title using matplotlib's pyplot
plt.title('Salary Distribution')

# Plot the distribution of the 'Salary' column using
seaborn's histplot
sns.histplot(data['Salary'])
```

```
# Show the plot
plt.show()
```

```
plt.scatter(data['EMPLOYE'], data['Salary'],
color='lightcoral')
plt.title('Salary vs Work/Houre')
plt.xlabel('Work Houre')
plt.ylabel('Salary')
plt.box(False)
plt.show()
```

```
# Splitting variables
X = data.iloc[:, :1]  # independent
y = data.iloc[:, 1:]  # dependent
```

```
import sklearn


from sklearn.model_selection import train_test_split
```

```
from sklearn.linear_model import LinearRegression
```

```
# Regressor model
regressor = LinearRegression()
regressor=(X_train, y_train)
```

```
# Splitting dataset into test/train
X_train, X_test, y_train, y_test = train_test_split(X,
y, test_size = 0.2, random_state = 0)
```

```
import numpy as np
```

```
# Importing necessary libraries
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression

# Define the file path using forward slashes or a raw
string
file_path = r'C:\Users\HP
ZBOOK\OneDrive\Desktop\python\Assighment.csv'

# Read the CSV file into a DataFrame
data = pd.read_csv(file_path)

# Convert 'Salary' column to numeric, setting
errors='coerce' to handle non-numeric values
data['Salary'] = pd.to_numeric(data['Salary'],
errors='coerce')
```

```
# Drop rows with NaN values in the 'Salary' column
data = data.dropna(subset=['Salary'])

# Splitting variables
X = data[['EMPLOYE']]  # independent variable
y = data['Salary']     # dependent variable

# Splitting dataset into test/train
X_train, X_test, y_train, y_test = train_test_split(X,
y, test_size=0.2, random_state=0)

# Initialize Linear Regression model
regressor = LinearRegression()

# Fit the model to the training data
regressor.fit(X_train, y_train)

# Prediction result
y_pred_test = regressor.predict(X_test)     # predicted
value of y_test
y_pred_train = regressor.predict(X_train)   # predicted
value of y_train
```

```
# Ensure that X_train and y_train are 1-dimensional
arrays


# Plot the data
plt.scatter(X_train, y_train, color='lightcoral')
plt.plot(X_train, y_pred_train, color='firebrick')
```

```
plt.title('Salary vs Experience (Training Set)')
plt.xlabel('Years of Experience')
plt.ylabel('Salary')
plt.legend(['X_train/Pred(y_test)', 'X_train/y_train'],
title='Sal/Exp', loc='best', facecolor='white')
plt.box(False)
plt.show()
```