You completed this part in Homework #5 A:

A company that provides a ride sharing service is expanding into another territory. A market analysis of the viability of this expansion has been completed and shows that there is indeed a viable market for this service. So, the company now wants to forecast a fuel budget for the new territory. Therefore, the company has purchased data that consists of a few hundred records of Uber and Lyft trips and several meteorological conditions at the time of the customer requests for service. The company wishes to use this data to forecast customer trip distances which will in turn influence the projected budget for fuel.

1. *Download the file named rideshare.csv that contains data reflecting Uber and Lyft rides. The features comprising this data consist of the following in order:*

   > *hour*
   > *day*
   > *month*
   > *distance*
   > *temperature*
   > *precipProbability*
   > *humidity*
   > *windSpeed*
   > *windGust*
   > *dewPoint*
   > *pressure*
   > *cloudCover*
   > *ozone*
   > *moonPhase*

2. *Read this file into a Pandas DataFrame and print all rows and columns.*
   Remember that the following Python statements can be used to guarantee all columns and rows will be displayed:

   > *pd.options.display.max_columns = None*
   > *pd.options.display.max_rows = None*

   The data should be "clean", but use the pandas DataFrame info method to confirm that there are no null values.

3. *Using correlation tests, identify which variable is the best predictor of distance. Print the results of all correlation tests.*

4. *Create a scatterplot of the best predictor that you identified in #3 and distance. This is to confirm the relationship between the predictor and distance.   Print the scatterplot and submit for credit.*

5. Create a Linear Regression *statsmodel* OLS model for the predictor (we will call X now) as the independent variable and the **distance** as the dependent variable. Make sure that you understand WHY the predictor is the independent variable! Print the OLS summary information. Separately print the equation of the line for this LR model. Print the equation by supplying values for 'm' and 'b' in y = m * x + b.

6. Print the min and max values for X, clearly labeled.

7. Use the linear regression model obtained in #5 to predict the **_distance_** for three new valid values of the independent variable. Be sure to very clearly label your output.

8. Plot the prediction points printed in #7 along with the regression line. Print the plot and submit for credit.

Print your Python code file and output and submit along with the plots for credit. In addition, upload your Python code in a zip file to the course Canvas site.