The Carlson Department Store suffered heavy damage when a hurricane struck on August 31. The store was closed for four months (September through December), and Carlson is now involved in a dispute with its insurance company about the amount of lost sales during the time the store was closed. Two key issues must be resolved: (1) the amount of sales Carlson would have made if the hurricane had not struck, and (2) whether Carlson is entitled to an additional compensation for excess sales due to increased business activity after the storm.

File `hurricane.csv` gives Carlson's sales data for the 48 months preceding the storm. It also reports the total sales for the 48 months preceding the storm for all department stores in the county, as well as the total sales in the county for the four months the Carlson Department Store was closed. All sales in millions.

Carlson's managers ask you (1) to analyze the data and develop estimates of the lost sales at the Carlson Department Store from Sep 2016 to Dec 2016. They also ask you (2) to determine whether a case can be made for excess storm-related sales during the same period. If such a case can be made, Carlson is entitled to compensation for excess sales it would have earned in addition to ordinary sales. Use python (not MS Excel) and library `statsmodels.formula.api` to prepare a report as follows.

1. (20 pts) Fit a regression model (Month as categorical, no interactions) to estimate the loss of regular sales (had there been no hurricane) for Carlson Department Store from Sep 2016 to Dec 2016. Plot actual and estimated store sales (in the same chart).
2. (20 pts) Split the data into two sets (set 1: 2012 to Aug 2016, and set 2: from Sep 2016 to Dec 2016). Use data set 1 to fit a model for the county sales. Use this model to predict the county sales from Sep 2016 to Dec 2016. Report these four predictions.
3. (20 pts) Compare your monthly predictions from question 2 with the actual county sales from Sep 2016 to Dec 2016. Find monthly ratios of actual vs. predicted county sales. Ratios larger than 1.0 indicate that actual sales were larger than regular sales, most likely, due to hurricane properties damages.
4. (10 pts) Multiply these ratios by the extimated store sales from question 1. These quantities are the monthly total sales (regular and hurricane induced) lost by Carlson Store. Report them along with their sum. It is the amount that Carlson can demand from the insurance company.

Another approach to estimate the amount to be claimed is as follows.

5. (20 pts) Use data set 1 to compute the fraction of Carlson monthly sales to the county-wide department stores monthly sales. This is Carlson's market share. Build a regression model to predict Carlson's market share from Sep 2016 to Dec 2016.
6. (10 pts) Multiply Carlson's predicted market shares by the actual county department store sales month by month. These are the monthly total sales (regular and hurricane induced) lost by Carlson Store. Report them along with their sum. It is the amount that Carlson can demand from the insurance company.

Submit your report with your name and USC ID as a pdf file online (no screen captures).