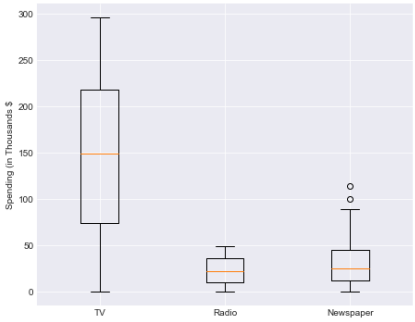
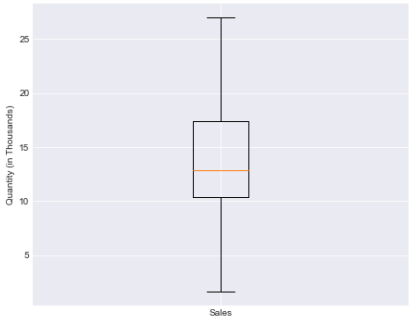
Mia Rodgers: https://github.com/miamrodgers/4310-ML/blob/main/ME1_simpleAnalysis/data_analysis.ipynb
Alex Larsen: https://github.com/alarsen123/ML-HW/blob/main/ME1_simpleAnalysis/data_analysis.ipynb
ME1

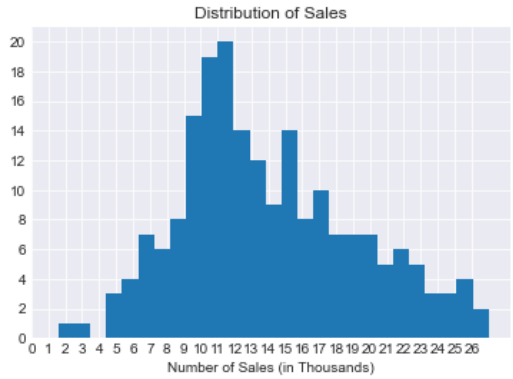
In this assignment, we discovered that the highest advertisement spending was for TV (mean: $147,042, median: $149,750), and much less was spent on Radio (mean: $23,264, median: $22,900) and Newspaper (mean: $30,554, median: $25,750). The average number of sales is 14,025 and the median number of sales is 12,900.

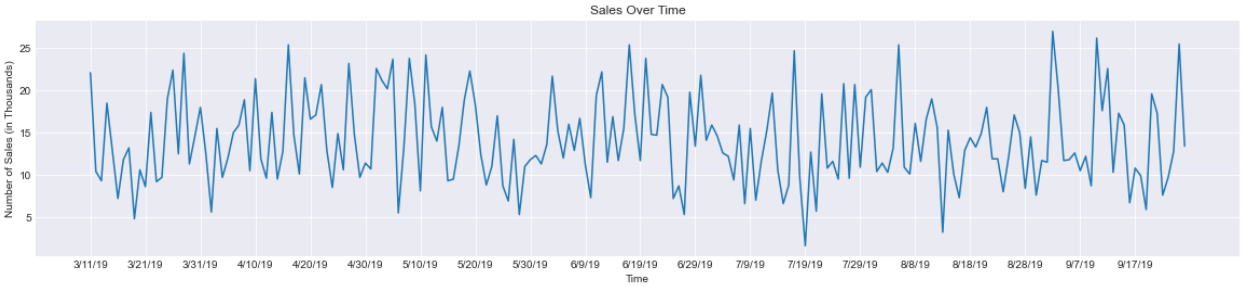
- The boxplots help us to see the differences in spending among the three different features and the overall amount of sales. The data is mostly evenly distributed, except the Newspaper attribute seems to be slightly skewed and contains a few outliers.



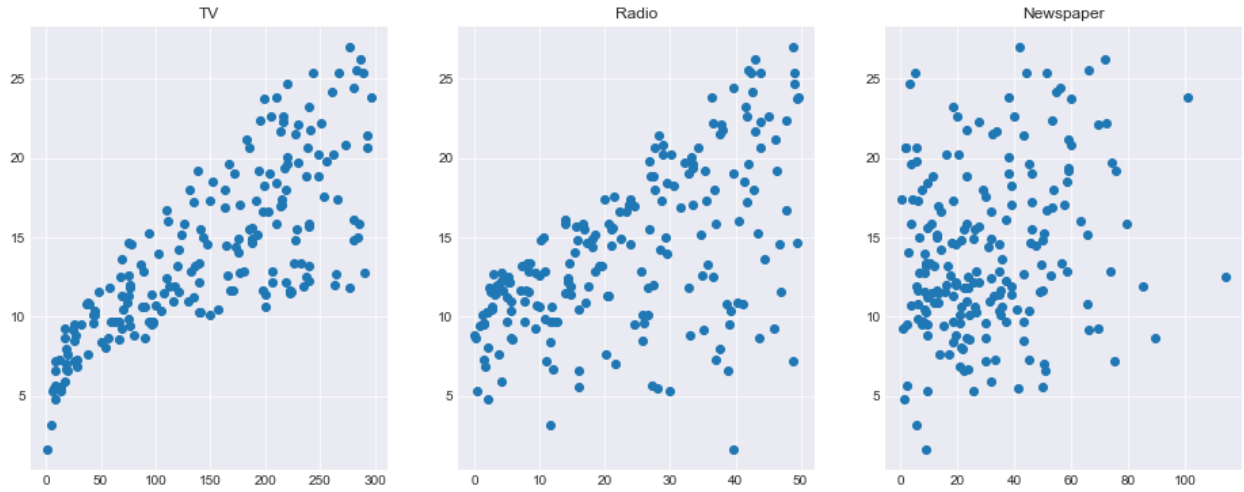


- The histogram visualizes the distribution of sales amounts across all of the sample points and shows a unimodal distribution, being slightly skewed toward the right.





- From the time series plot, there does not appear to be a pattern in sales for this product with respect to time.



- Looking at the scatterplots along with the correlation coefficient comparing each feature with Sales quantities, it can be seen that TV has the strongest correlation with sales (0.782), Radio is the next strongest (0.576), and Newspaper is the least correlated by a significant amount (0.228). In the plots, TV seems to have a logarithmic trend, where Radio seems to be a bit more linear, and it is difficult to see any relationship between Newspaper and Sales.

Overall, Television advertisements seem to be the most effective and have the greatest overall impact on sales.