Interpreting Plots

Interpret scatter and bar plots in the context of the project.

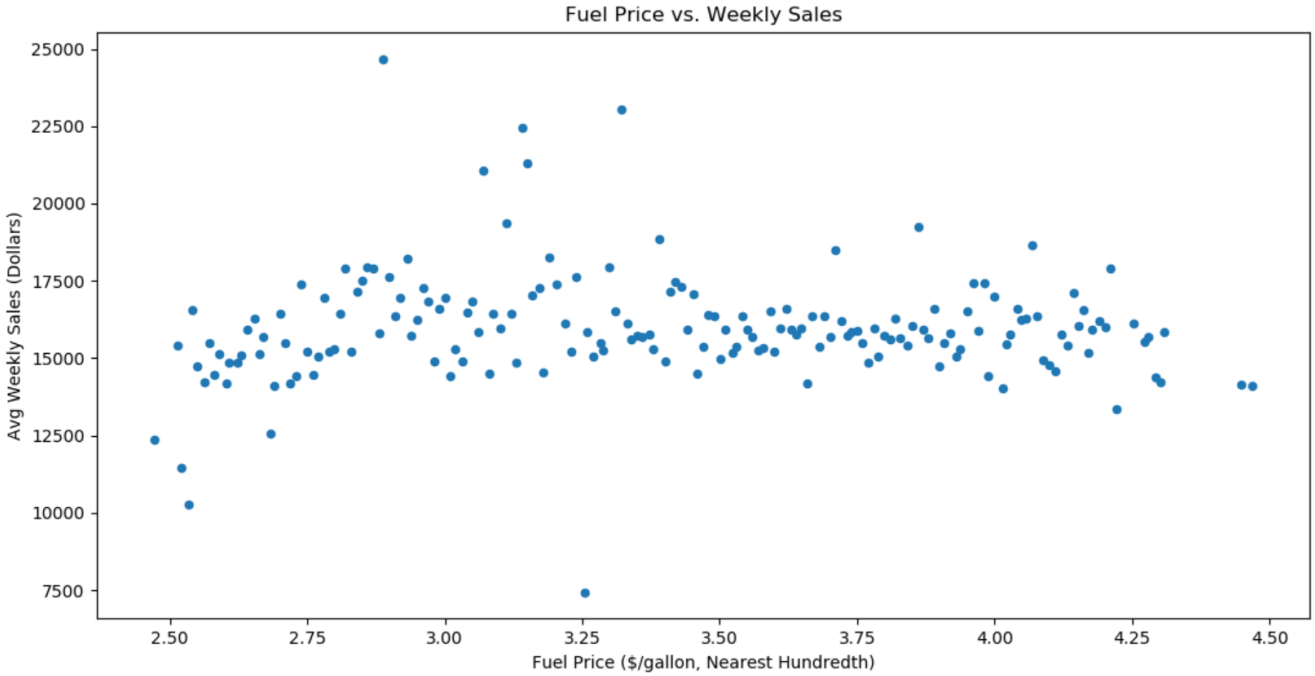
Chapter Goals:

- Analyze plots of the dataset's features with respect to weekly sales and usage in a machine learning model

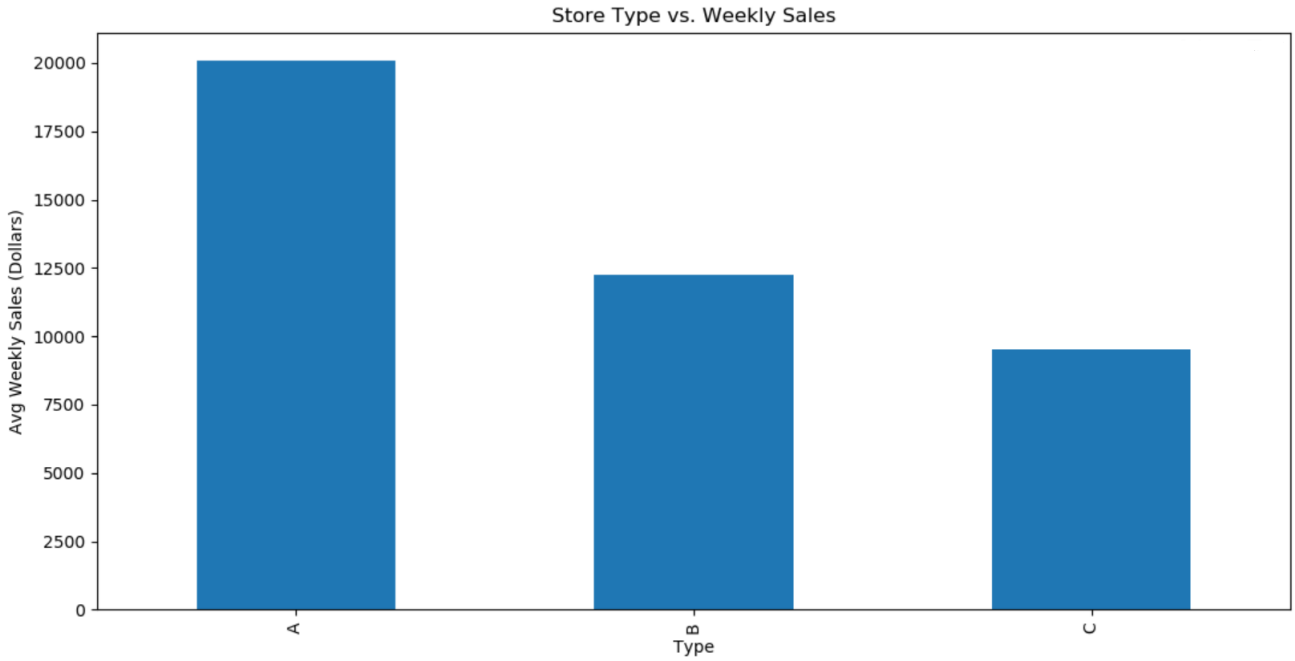
A. Analyzing within context

After creating the dataset plots, we should analyze them to determine whether they're useful in the context of our problem. For us, this means deciding whether there's enough correlation between the dataset features and weekly sales to train a machine learning model.

The main thing we're looking for in the plots is non-uniform distributions. A uniform distribution means that the weekly sales are identical regardless of the data feature's value. A non-uniform distribution, such as a normal distribution or a multi-modal distribution, shows that the dataset feature can potentially be used by a machine learning model to predict the sales.



The scatter plot for the 'Fuel' feature shows a relatively uniform distribution.



The bar plot for the 'Type' feature shows a non-uniform distribution.

It can also be beneficial to plot multiple features at once against the weekly sales. Sometimes, when plotting multiple features, we can find trends that would not have been found from the single feature's plot.