Coffee Sales Analysis

Objectives:

1. Determine which predictors (from hour of day, cash type, coffee type, time of day, weekday, and month) are the most influential on revenue.
2. Determine which drink purchase predicts the highest revenue.
3. Determine which day of the week and hour of day predicts the highest revenue.

Software Used: RStudio, Python, and Tableau.

Variable Information:

- Hour of day: continuous
- Cash type: factor with two levels (cash or card)
- Coffee type: factor with eight levels (americano, americano with milk, cappuccino, hot cocoa, cortado, espresso, hot chocolate, and latte)
- Time of day: factor with three levels (morning, afternoon, or evening)
- Weekday: factor with seven levels (each day of the week)
- Month: factor with 12 levels (each month)

Figures 1-4 show exploratory data analysis conducted prior to the analyses.

To address these objectives, stepwise regression was used.

The results of the stepwise regression showed a slight reduction in AIC and no change in adjusted R-squared indicating that the fit was slightly better and variance explained did not change.

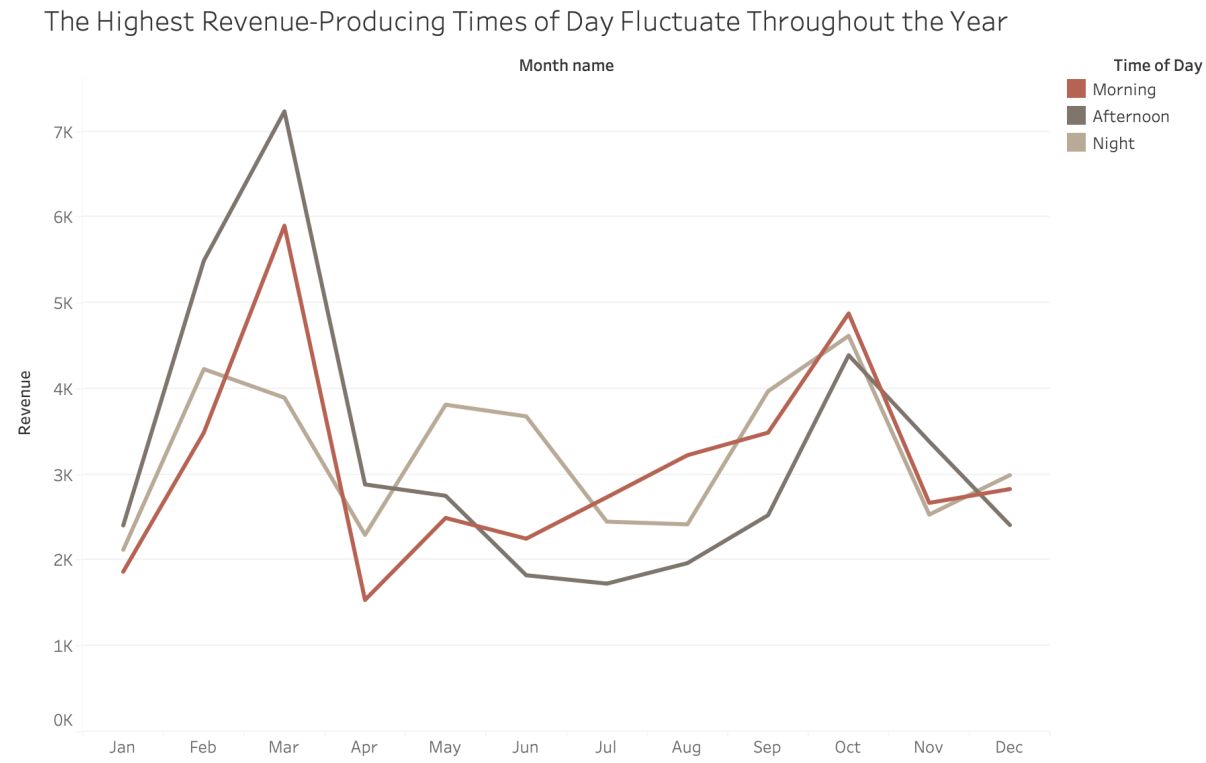
The results of the reduced model demonstrated that the best predictors of revenue were cash type, coffee type, time of day, and the month. The highlights include that all coffee types with the exception of espresso predicted higher revenues, morning was the time of day that predicted lower revenue.

As for Objective 2, we can see that hot chocolate predicted the highest revenue, followed by lattes and cappuccinos, when accounting for all other variables in the model.

Given that day of the week and the hour of the day were *not* found to be the most important predictors in price using stepwise regression, a separate multiple linear regression was used in order to see the impact of these variables (while not taking into account the others). The results, indicate that no day of the week or hour of the day were significant predictors. However, the adjusted R^2-squared was very low (0.052), indicating that little variance was explained by this model.

Figure 1.

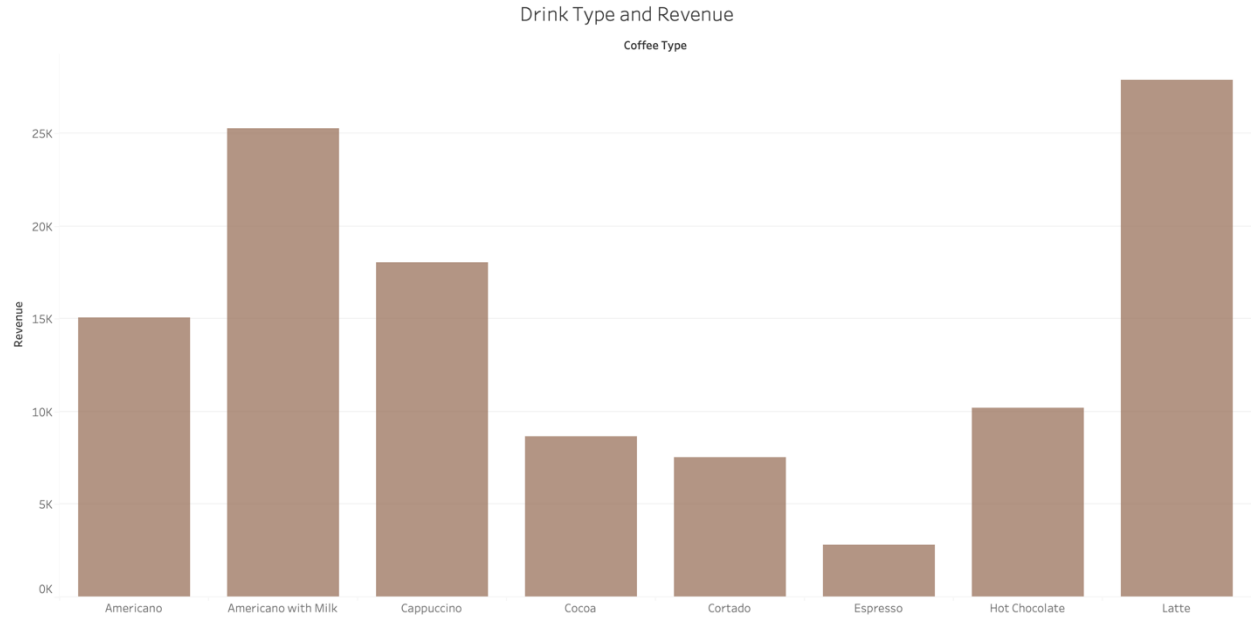
The highest revenue-producing times of day (morning versus afternoon versus evening) fluctuate throughout the year.



The trend of sum of revenue by month; the color shows details about the time of day.

Figure 2.

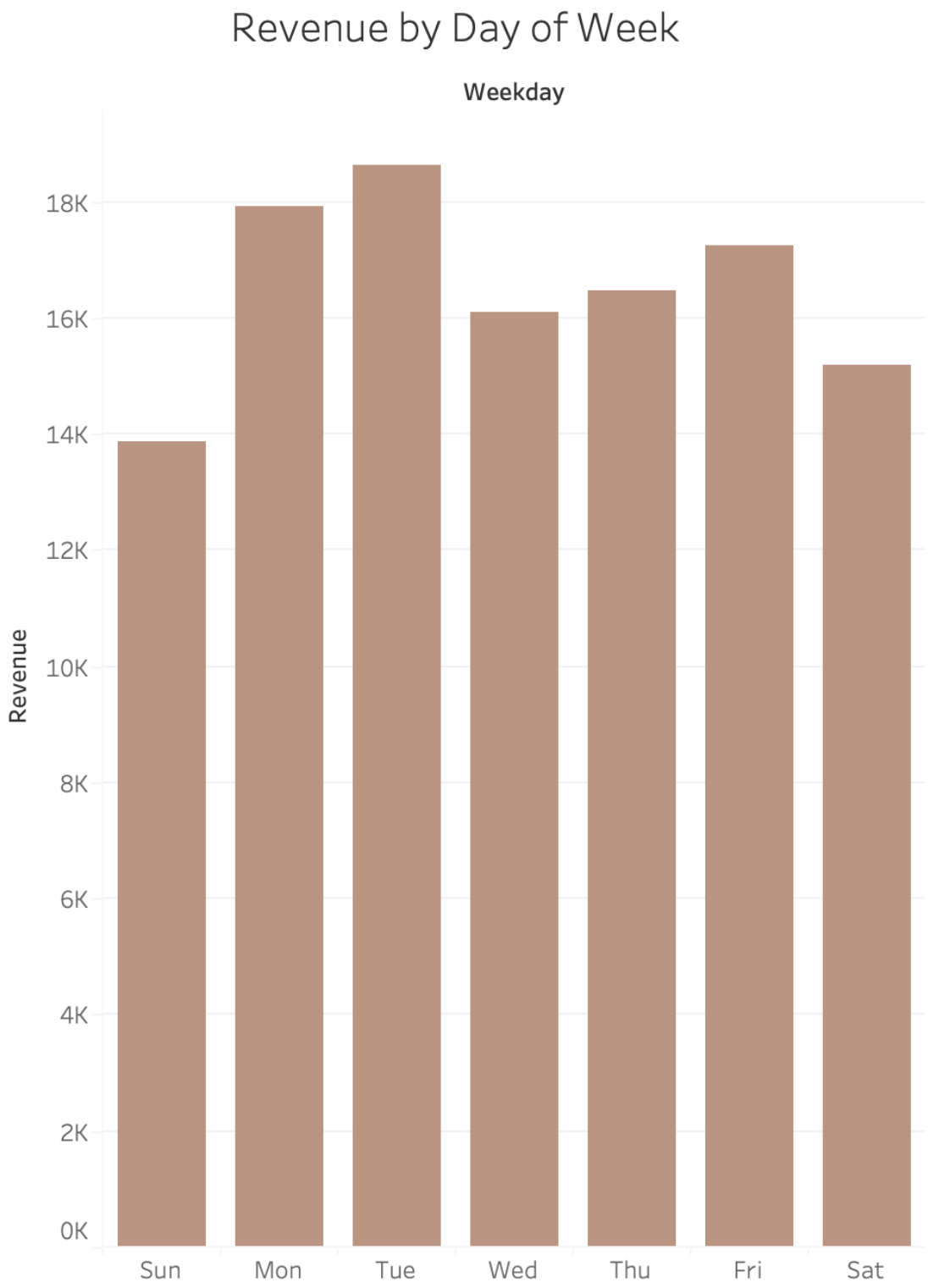
Drink type and revenue.



Sum of Revenue for each Coffee Type.

Figure 3.

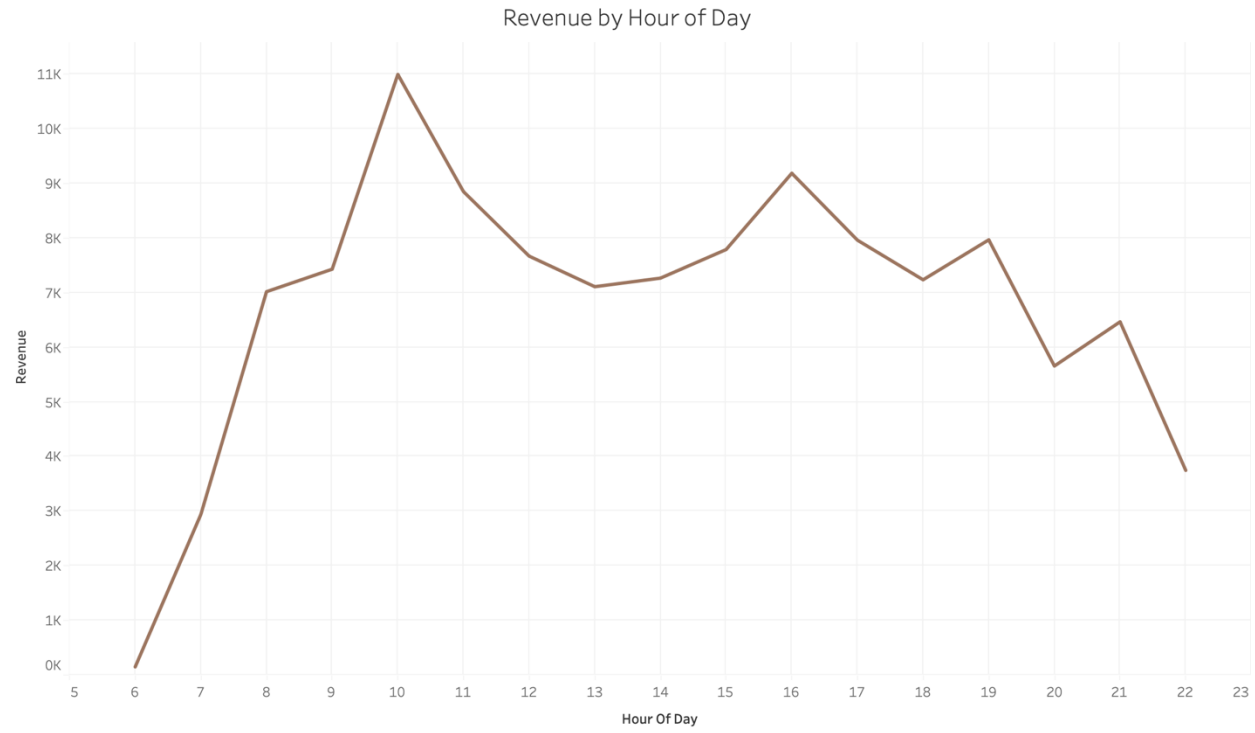
Revenue by day of the week.



Sum of Revenue for each Weekday.

Figure 4.

Revenue by hour of the day.



The trend of sum of Revenue for Hour Of Day.