Food Sales Predictions

Analysis and Machine Learning Recommendation
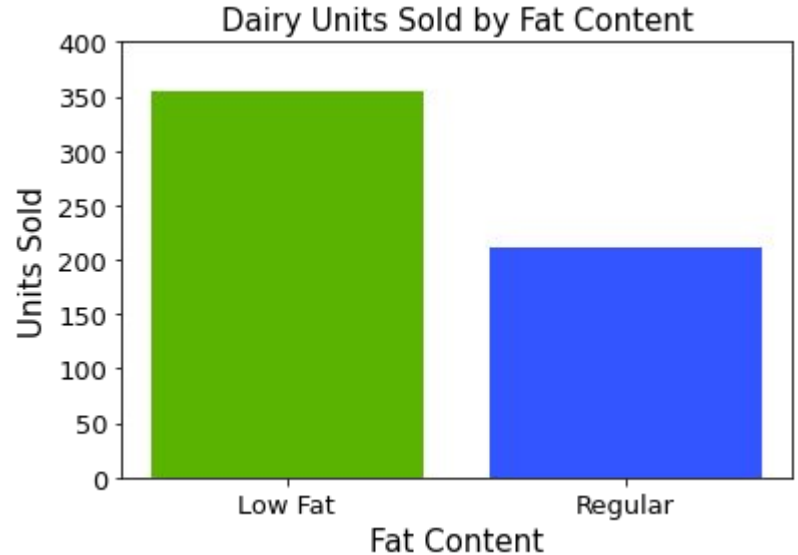By Nathan Billings

The Data

The data used in this project is a record of sales for groceries. Several things initially stood out for possible usefulness

- Item Maximum Retail Price
- Product Category
- Item Fat Content
- Visibility of the Product

Low Fat Dairy

Experimenting with subcategories of dairy products will help focus the right low fat offerings
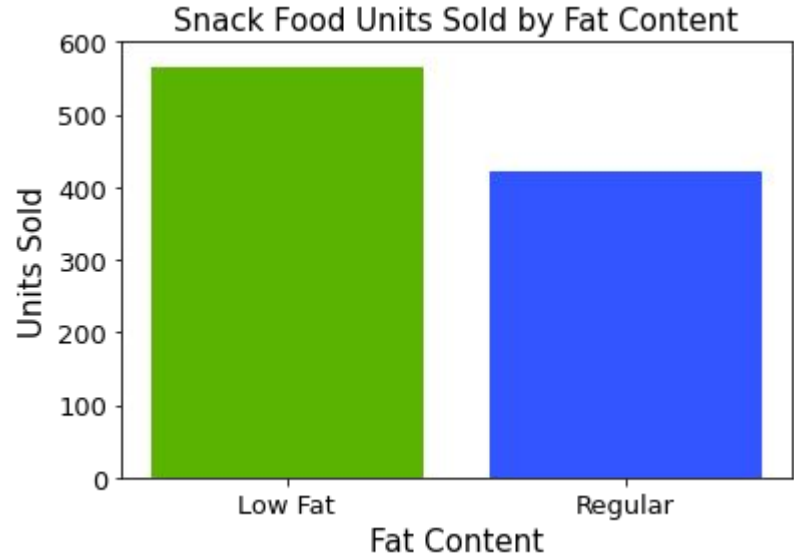
How do cheese or ice cream compare?

Dairy Units Sold by Fat Content

Low Fat Snack Foods

Correlation between fat content and snack sales not as strong compared to dairy

Splitting "healthy" snacks from regular snacks might be a better way of looking for more insight

Machine Learning Models Evaluated

A Decision Tree model was selected as it outperformed the Linear Regression model by roughly 1.3%

The data did not seem to have a strong linear correlation, allowing the Decision Tree to perform slightly better

Summary

- Low fat dairy products show promise
- Collect data on snack sub-categories such as "healthy" or "junk-food" to see if stronger relationship develops
- Machine Learning opportunities show moderate confidence with this data and further data and model tuning will increase performance

Questions? Feedback? Let me know!