SEI Discount Program Profitability Analysis

Kan Ito

July 24, 2017

Overview

- Analysis based on transactional data from 2015-2016 on the following:
    - Bogota($75), South Face ($85), Pangea ($105) brands
    - Sold in, Jacksonville, Albany, Springfield, Eugene, Bend, and Tacoma
- Data includes
    - All transactions with quantities of jackets sold
    - Price of all jackets per day
    - National Discount Program schedule (seasonal & holidays)
- Objective: measure effectivity of discount program and recommend future discount programs for short and long term.

Initial Assumptions

▶ Define profitability as the sales of jackets

▶ Each transaction is limited to only one type of jacket, implied by the data

▶ Sales location and brands not correlated with each other.

　　▶ Location and brand variables are independent of each other.

▶ Empty data cells in daily prices were assumed to be at retail price with the respective discount.
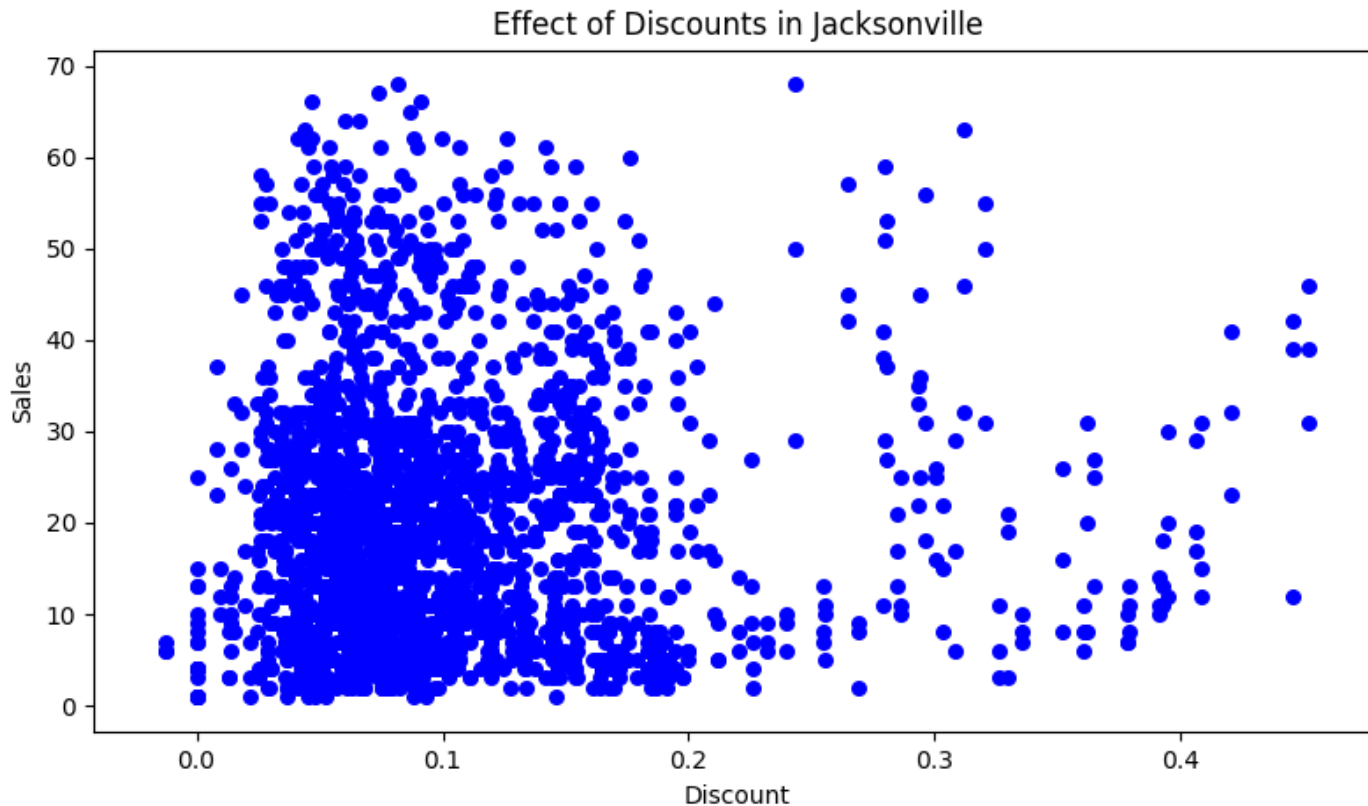
I. The Data

Seasonal variation

Other cities display the same sinusoidal pattern where sales are unsurprisingly correlated with time of year.
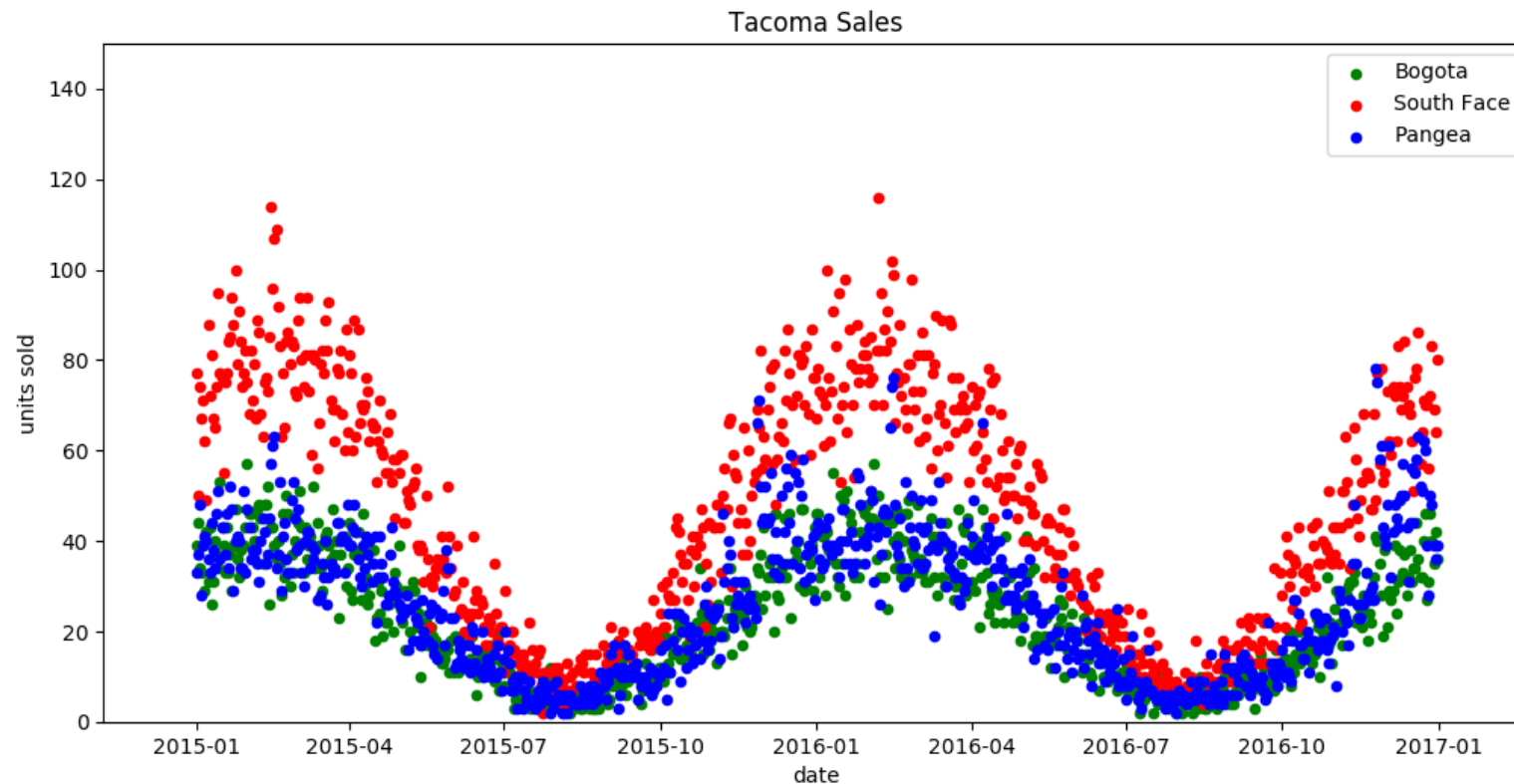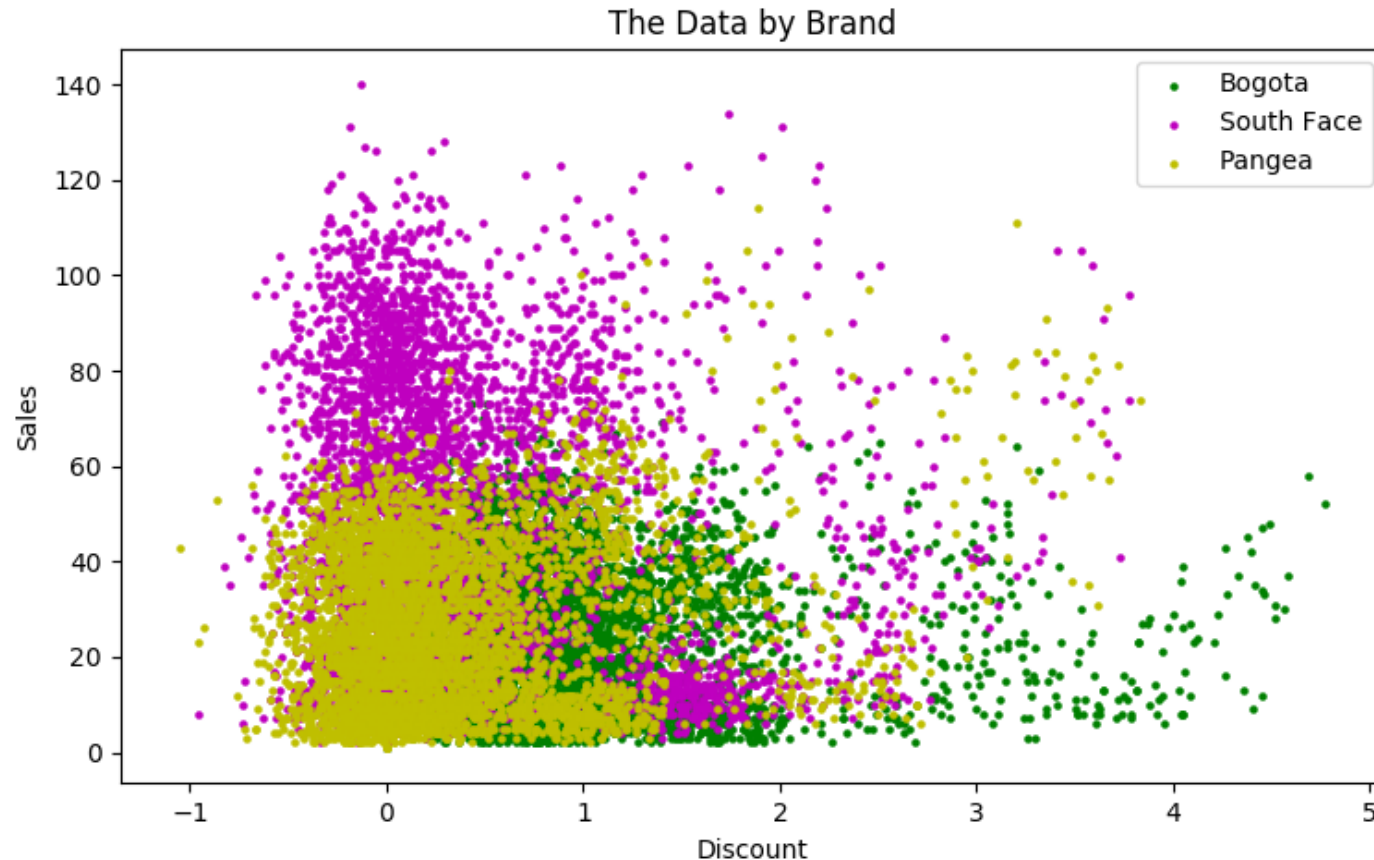
Discount Effect

Similarly, data extracted to find correlation for **sales** and **discounts.** Initially, this data shows weak signs of the correlation between the two.
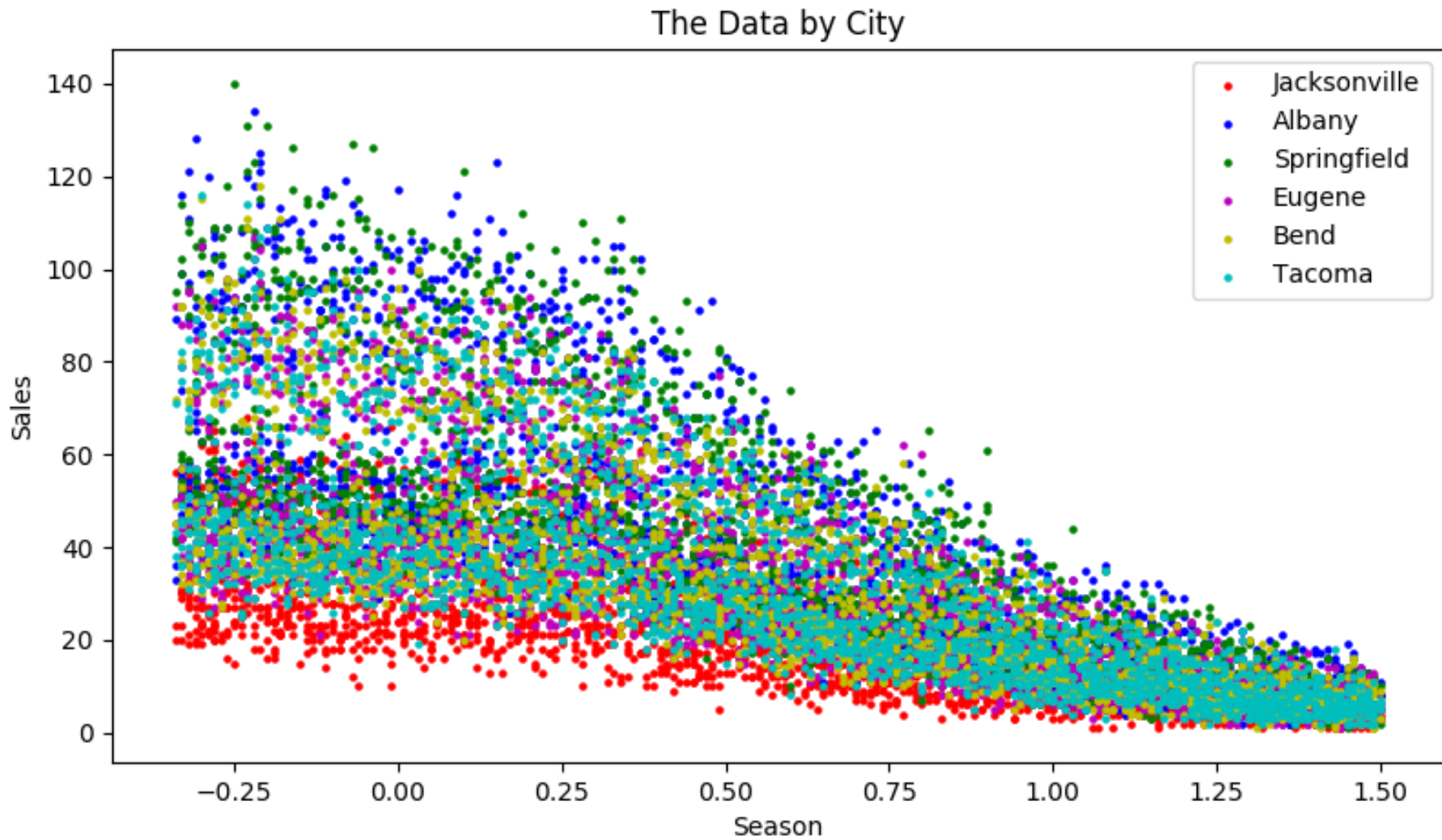
Brand Effects

The brand categories have also proved significant signs of influence in the sales of jackets. Below by brand,


The Data by Brand

by location...

The Data by City

II. Fitting a Model

Multiple Regression

Data suggests that seasons, cities, and brands are all significant variables that affects the sales. With several known variables, multiple regression is a suitable choice to model the sales. Furthermore, discount price will be added to the list of variables since this is our primary interest.

Initial Two Variable Regression

1. Mult-Reg Jacksonville Bogota

1. Mult-Reg Jacksonville Bogota
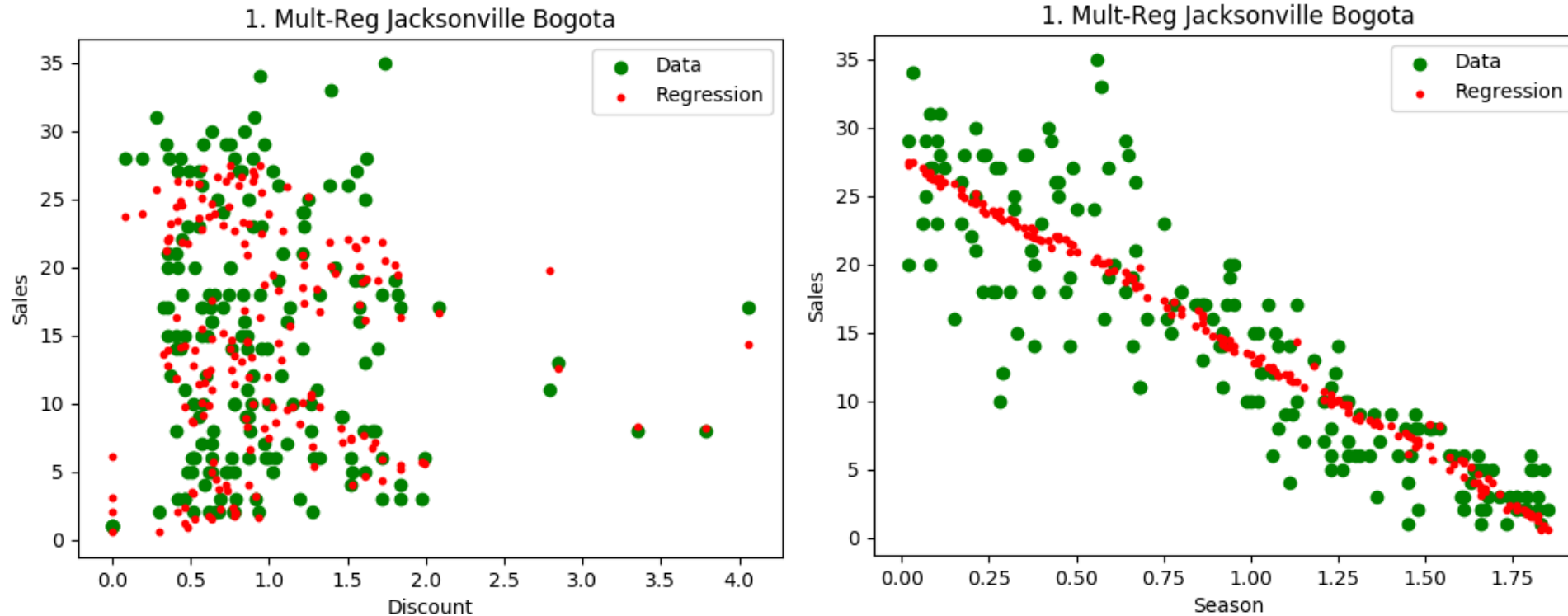
Sales = 27.1 + 0.91 * Discount + -14.5 * season
Coefficient of Determination: R-squared = 0.79
As expected, the coefficient for the discount is much smaller than that of seasons. This implies a much stronger correlation for seasonal changes in sales.
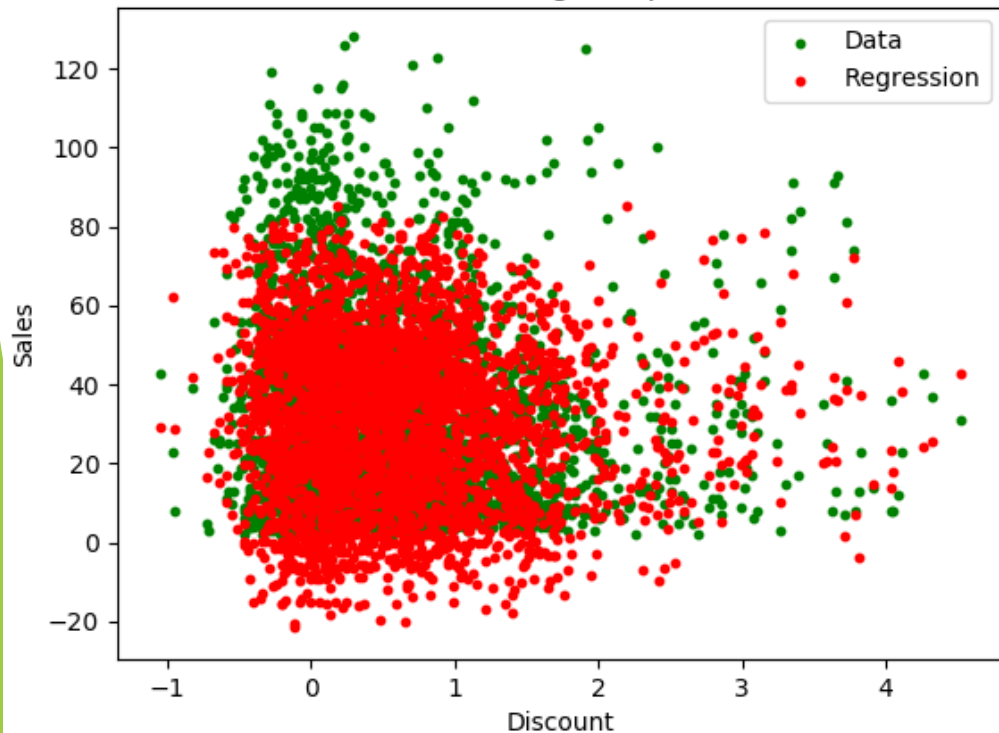
*Note that Discount is scaled by a factor of 10 and season is discretized into 365/2~183 slots for effective optimization

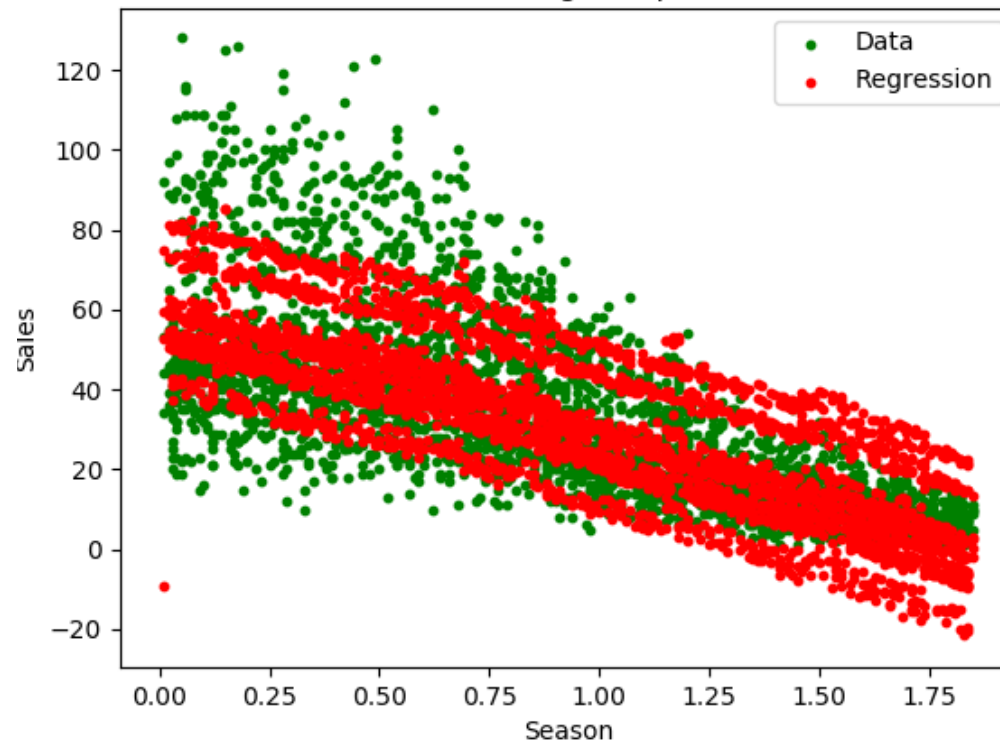# Complete Multiple Regression

Independent Variables: (15 total)
- constant
- discount
- season
- discount * season

- discount^2
- season^2
- bool constants for 6 cities
- bool constants for 3 brands

4. Mult-Reg Complete

4. Mult-Reg Complete

Complete Multiple Regression

Coefficients

Constant	32.42	Springfield	12.40
Discount	4.24	Eugene	4.92
Season	-30.50	Bend	5.90
Discount*Season	-1.15	Tacoma	5.67
Discount^2	-0.15	Bogota	0.92
Season^2	-1.87	South Face	26.2
Jacksonville	-5.63	Pangea	5.19
Albany	12.90		

Coefficient of Determination: R-squared = 0.81

III. Recommendation

Discounts incur larger sales variation in colder seasons. **Therefore, the discount program should be utilized more in colder seasons**, and less in warmer seasons. Separate simple regressions were conducted to verify such phenomena.

4. Mult-Reg Complete
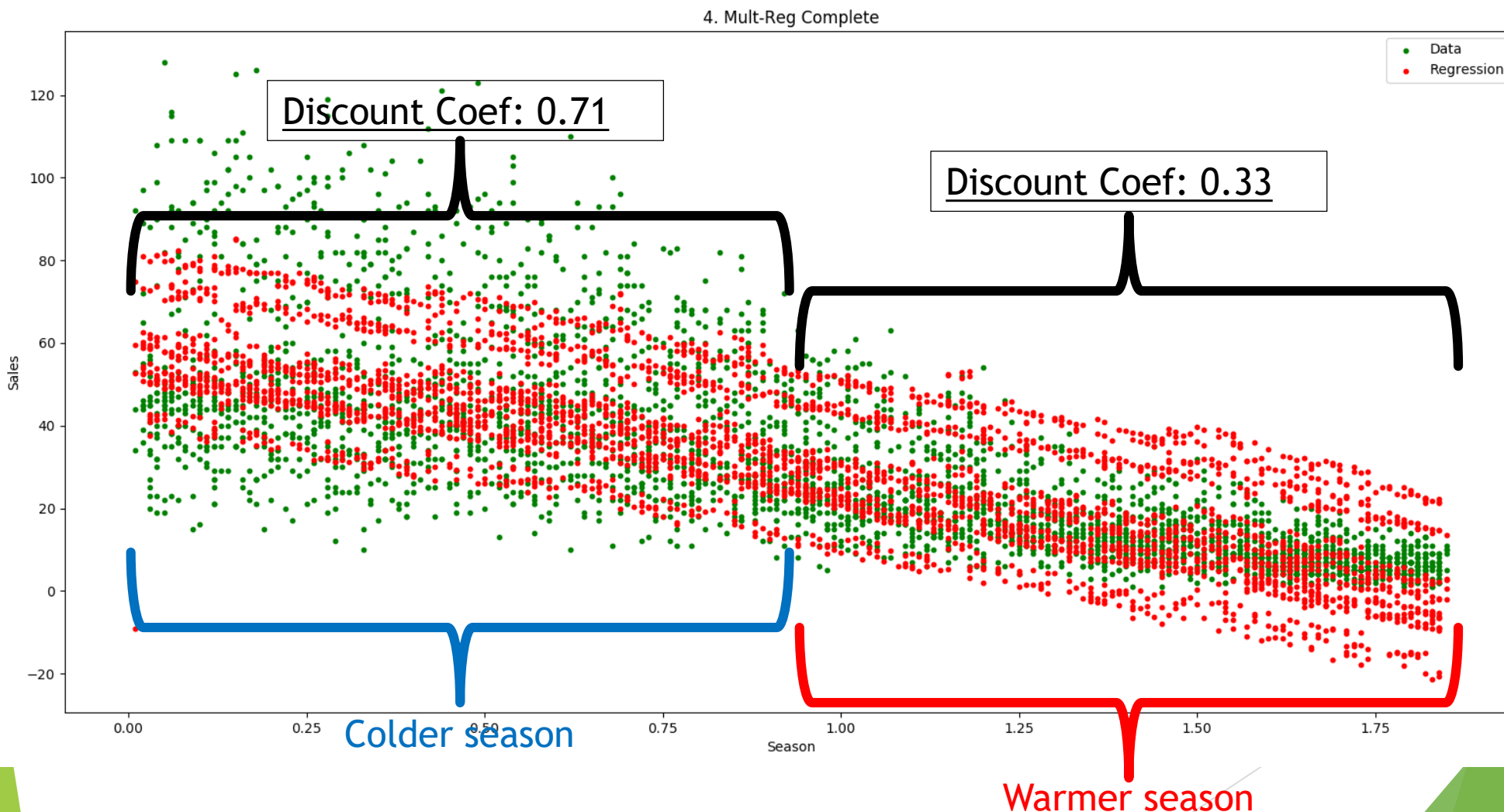
Similarly, the brands and locations can be utilized more from the discounts. Pangea displays the highest sensitivity to discounts. South Face is second, and Bogota shows very little sensitivity. **Therefore, the discount program should be utilized more for Pangea, then South Face in this order of preference.**

For location, the sensitivity to discounts was much less varied. **The locations of preference for discount programs are: Albany, Tacoma, Springfield, Eugene, Bend, Jacksonville.**

The coefficients for the separate regression results are below:

by brand:

Brand	Discount Coeff
Bogota	0.71
South Face	5.22
Pangea	8.99

by location:

Location	Discount Coeff
Jacksonville	4.38
Albany	8.44
Springfield	6.18
Eugene	5.90
Bend	4.41
Tacoma	6.27

IV. Conclusion

- Objective: provide recommendations for discount program to increase profit

- Data suggests the strongest correlation with seasons. Discounts have a relatively weak correlation.

- Model of choice: Multiple Regression with essential independent parameters discount, season, cities? Brands?

- The multiple regression model revealed variation in product of variables with discounts and several brands and cities in correlation with discounts. The variation in this correlation revealed that discounts should rather be implemented for winter seasons as opposed to summer.

- Similarly, the discounts should be emphasized for sales for the Pangea brand and locations in Tacoma and Albany.