James Lee

Jet Case Assignment

05/09/2017

Preface:

1. Data Adjustments
   1. These data points are unlikely to be due to variability and thus removed from the data set.
      * 0 dollar prices
      * (-) prices
      * 99,999.00 prices
   2. For calculation on weighted average, NA data have been removed. i.e. some products didn’t have corresponding sales.

Case:

1. ***Overall, how are the three merchants priced relative to each other? Which merchant has the best pricing?***

For each product and merchant, I looked at the percentage deviation of its price from the product’s mean price. I used R for the calculation because of its robust ability to do column and aggregation calculations. Code is attached in the appendix.

Then to look at the deviations on a merchant level, I took an average of all the percentage deviations by merchant.

Result:

| **Merchant** | | **Average\_Price\_Deviation** |
| --- | --- | --- |
|  |  |  |
| **1** | Alexs Store | -0.02444468 (-2.44%) |
| **2** | Jasmines Shop | -0.04359199 (-4.36%) |
| **3** | Leos Bodega | 0.06842134 (+6.84%) |

Result essentially tells us that on average, Jasmine’s shop’s prices are 4.36% below product average. By this measure, Jasmine’s Shop has the best pricing.

On the other hand, Leo’s Bodega’s prices are 6.84% above the average, and has the worst pricing.

However this analysis may contain some bias in a sense that we want to identify vendors that provide the cheapest prices in products that drive our revenue the most. So it makes sense to weight these averages by how much impact each product has on Jet’s bottom line.

I assumed that average price of all vendors \* sales is a reasonable benchmark of how much the product impacts the revenue since sales allocation numbers were not provided. This is an obviously not a good measure in real life.

Merchant mean wmean

1 Alexs Store -0.02419016 -0.0004054448

2 Jasmines Shop -0.04393084 0.0585137272

3 Leos Bodega 0.06852133 -0.0608791728

What the result is showing is that while Leos Bodega has the highest prices in pure average sense, they are actually providing the cheapest prices in the product categories that has the most impact on the revenue. In this sense, Leos Bodega provides the best prices among all the vendors. This is a completely different conclusion than an analysis based purely on percentage deviations.

1. *What products have the most competitive pricing? Which ones have the least competitive pricing?*
2. *What else could we learn from this data?*

Appendix:

1. Removed Data Points
2. R Code used for the case

library(readxl)

**#Import Data**

Merchant\_Data <- read\_excel("C:/Users/James/Desktop/Programming/R Environment/Analytics\_Interview\_Case\_-\_With\_Sales\_Data.xlsx",

sheet = "Merchant Data")

Category\_Data <- read\_excel("C:/Users/James/Desktop/Programming/R Environment/Analytics\_Interview\_Case\_-\_With\_Sales\_Data.xlsx",

sheet = "Category Look Up")

Sales\_Data <- read\_excel("C:/Users/James/Desktop/Programming/R Environment/Analytics\_Interview\_Case\_-\_With\_Sales\_Data.xlsx",

sheet = "Sales Data")

**#Get mean price for each JET SKU ID**

SKU\_ID\_Avg <- setNames(aggregate(Merchant\_Data[, 4],Merchant\_Data[,3],mean ),c("Jet SKU ID", "Average\_Price"))

**#Combine average prices into the original Table**

Merchant\_Data\_With\_Avg <- (merge(Merchant\_Data,SKU\_ID\_Avg, by = 'Jet SKU ID'))

**#Compute % deviations of prices from the means**

Merchant\_Data\_With\_Avg\_Ndeviations <- within(Merchant\_Data\_With\_Avg,

Perc\_Deviation <-

(Price - Average\_Price)/Average\_Price)

#ANSWER TO #1

**#Calculate Overall Average deviations by Merchant**

Merch\_Perc\_Deviation\_Avg <- setNames(aggregate(Merchant\_Data\_With\_Avg\_Ndeviations$Perc\_Deviation,

by= list(Merchant = Merchant\_Data\_With\_Avg\_Ndeviations$Merchant),

mean),c("Merchant", "Average\_Price\_Deviation"))

#ANSWER TO #2

**#Calculate Average Deviations by Product**

Product\_Perc\_Deviation\_Avg <- setNames(aggregate(Merchant\_Data\_With\_Avg\_Ndeviations$Perc\_Deviation,

by= list(Jet\_SKU\_ID = Merchant\_Data\_With\_Avg\_Ndeviations$`Jet SKU ID`),

mean),c("Jet\_SKU\_ID", "Average\_Price\_Deviation"))