Underwear Price Analysis

Cesar Espitia, Ilya Kats and Michael Muller

May 2017

Objective

A new underwear brand is being launched that caters to the larger, more stylish gentlemen. Currently, the startup has one type of underwear designed (boxer briefs), but there is interest in picking one more type of underwear to include in the line. Additionally, so far there has been little analysis in how to price the produst. Complete market research done by a dedicated research firm such as Mintel, IBISWorld or Gartner is cost prohibitive for the startup. The goal of this analysis is to provide the company with the starting point in how to price and brand their line in the marketplace using freely available data.

Data

Data used in analysis was collected from two sources - several online shops and Bureau of Labor Statistics.

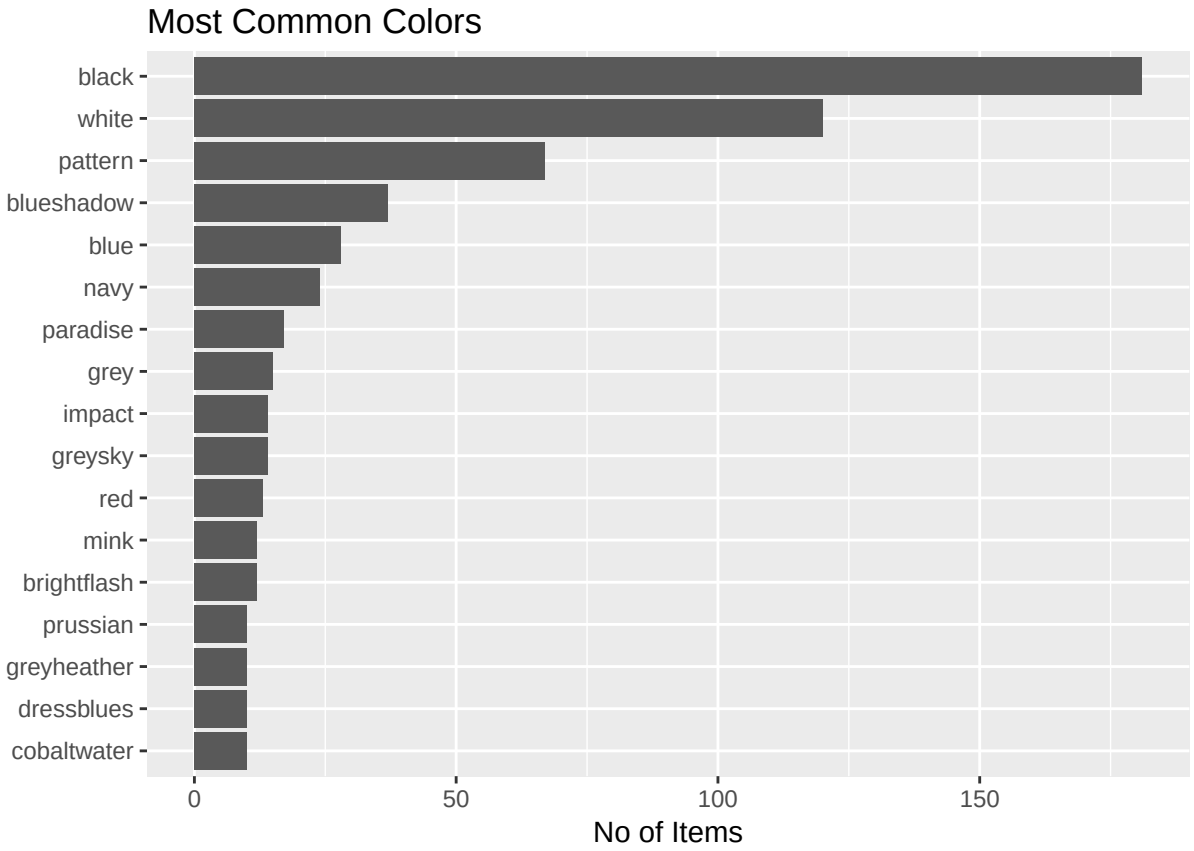
Information about related products was scraped from online shops of Bloomingdale's, Calvin Klein and Nordstrom. These three retailers were chosen because they provide a good sample of the type of marketplace the company is looking to enter (upscale and luxury). Other websites were considered, but were not included in the analysis because of resource constraints.

The scraping took place the week of May 7, 2017. The data was cleaned up and any observations missing product name, type or price were removed from the data set. The final set includes 511 observations - 134 from Bloomingdale's, 62 from Calvin Klein and 315 from Nordstrom.

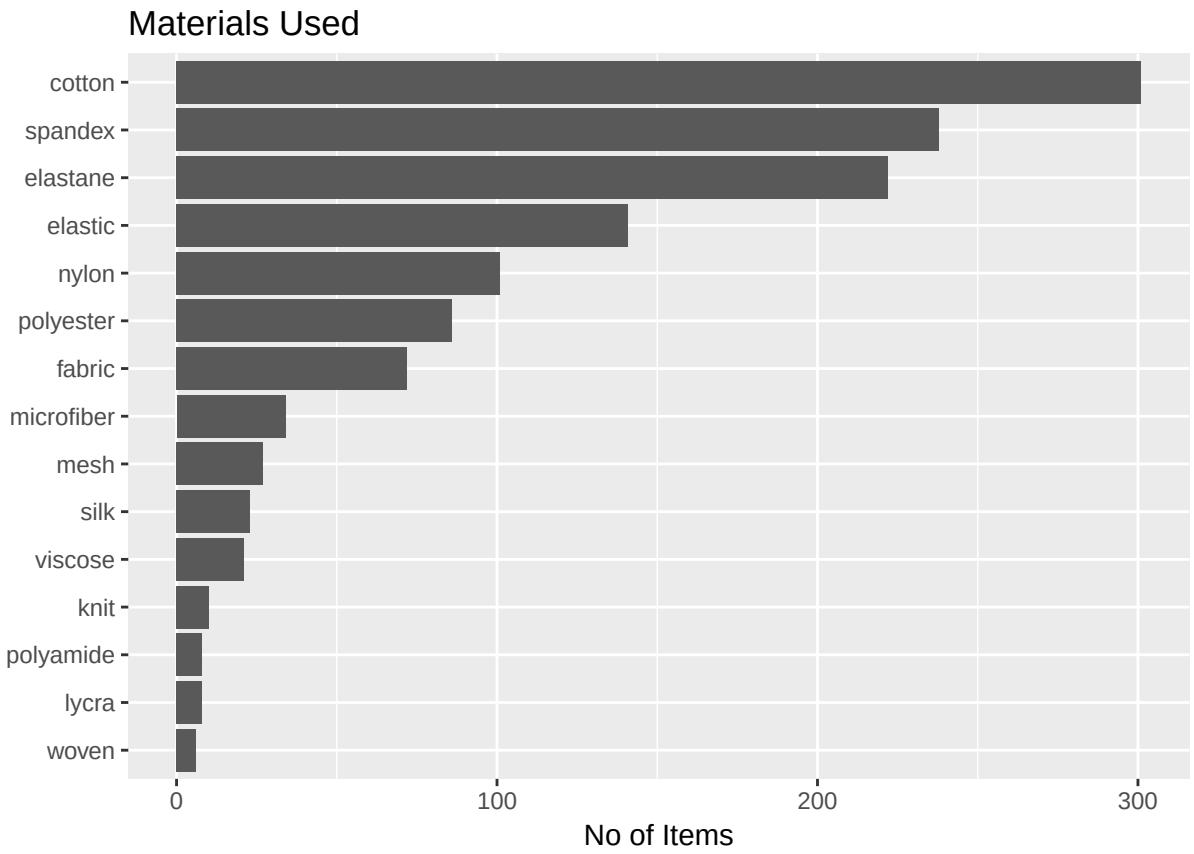
The following information was captured about each observation:

- **Product Name**
- **Product Type**: Includes one of four underwear types: *boxer briefs* (266 observations), *boxers* (30), *briefs* (73), *trunks* (142).
- **Color**: All colors the product is available in. Multi-color products were captured as patterned.
- **Material**: All materials used in the product. Not available for all products.
- **Original Price**
- **Sale Price**

Some data descriptive charts are included below.



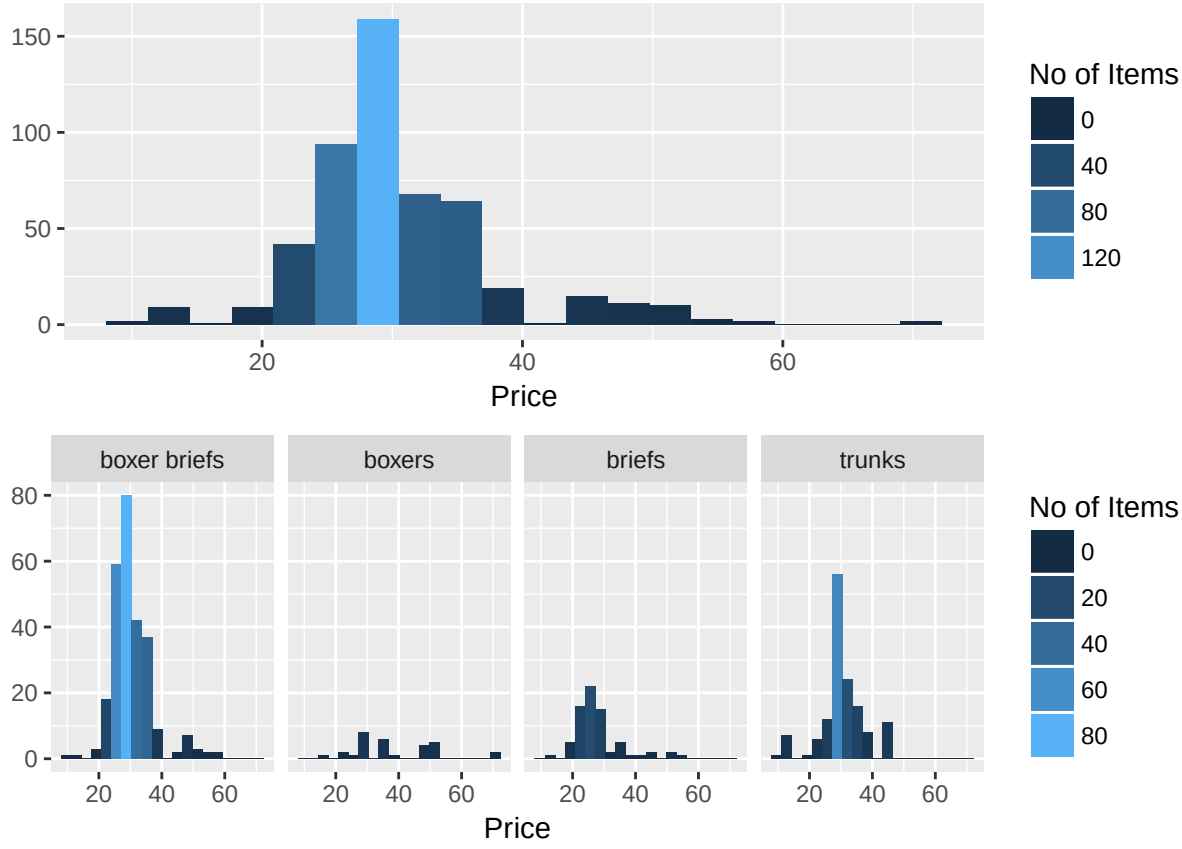
TIP: Please note that in this market segment there is a significant number of products with fancy color names. Research indicates that this naming strategy has an impact on consumer behavior. (*Skorinko JL, Kemmer S, Hebl MR, Lane DM. December 2006. A Rose by Any Other Name. . . : Color-Naming Influences on Decision Making. Psychology & Marketing, Vol. 23(12): 975-993*)



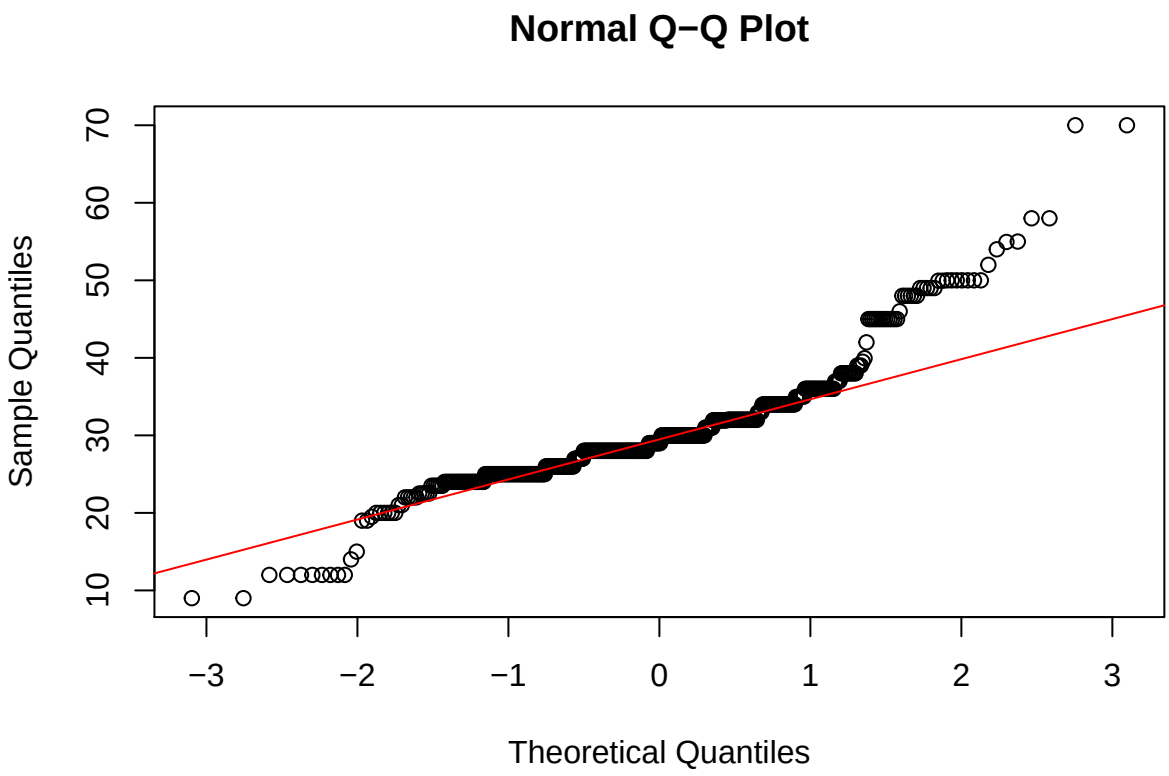
To further round out the data collection, data from a government source was used to determine household spend on particular goods. The Bureau of Labor Statistics (BLS) runs the Consumer Expenditure Surveys (CE) program that provides data on expenditures, income, and demographic characteristics of consumers in the United States. CE data is collected by the Census Bureau for BLS in two surveys, the Interview Survey for major and/or recurring items and the Diary Survey for more minor or frequently purchased items. The analysis uses data from 1st Quarter of 2015. The data was further filtered down only to category 200 (Undergarments).

Analysis

The graph below shows the distribution of original price from retail data. The distribution is unimodal with some right skew, but overall it appears to be nearly normal. **Mean price is $30.51** with standard deviation of $7.66. Median price is $29.00. Distribution of individual types are similar to the entire sample distribution.

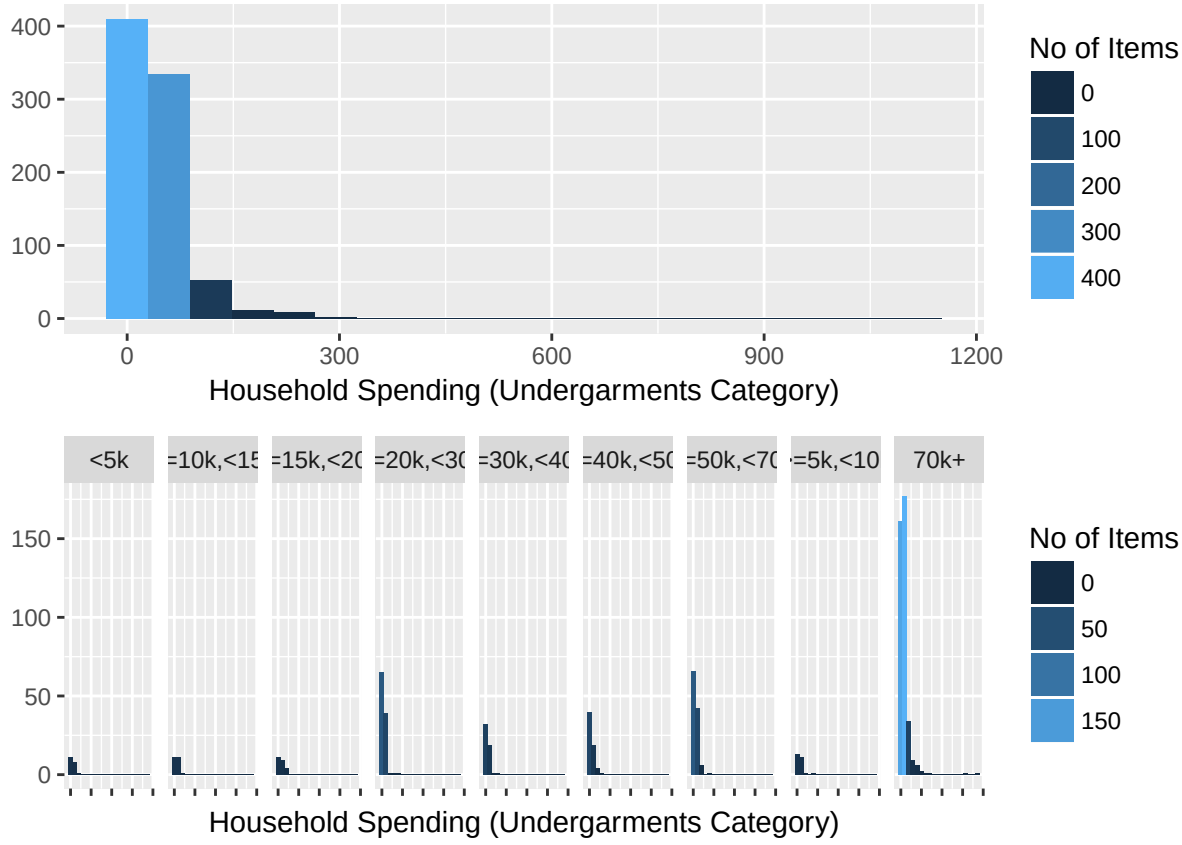


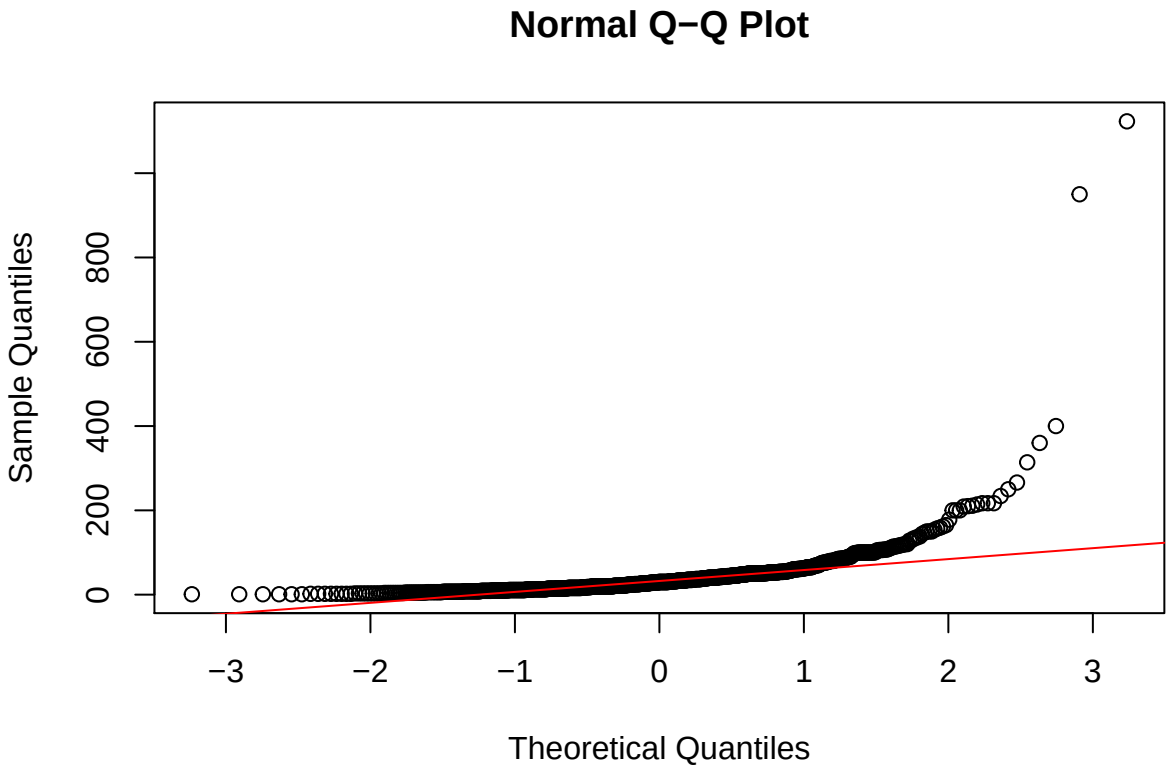
To test the normality of price distribution, below is the normal probability plot. The bulk of the data follows the line fairly close indicating normal distribution; however, there is some deviation especially in the right tail.



Considering the independence of observations it is important to note that there may be some dependence since the data comes from only three online shops and it is possible that pricing decesions made for one product influence prices of other products.

Below charts show the distribution of household spending per BLS data for the entire sample as well as by income level. Normal probability plot of the distribution is also included below.





IMPORTANT NOTE: Looking over the histogram and normal probability plot there are serious concerns whether the distribution can be considered nearly normal. It does not appear to be so. Continueing the analysis may not yield meaningul statistical results, but it may still be helpful for understanding the data for the startup.

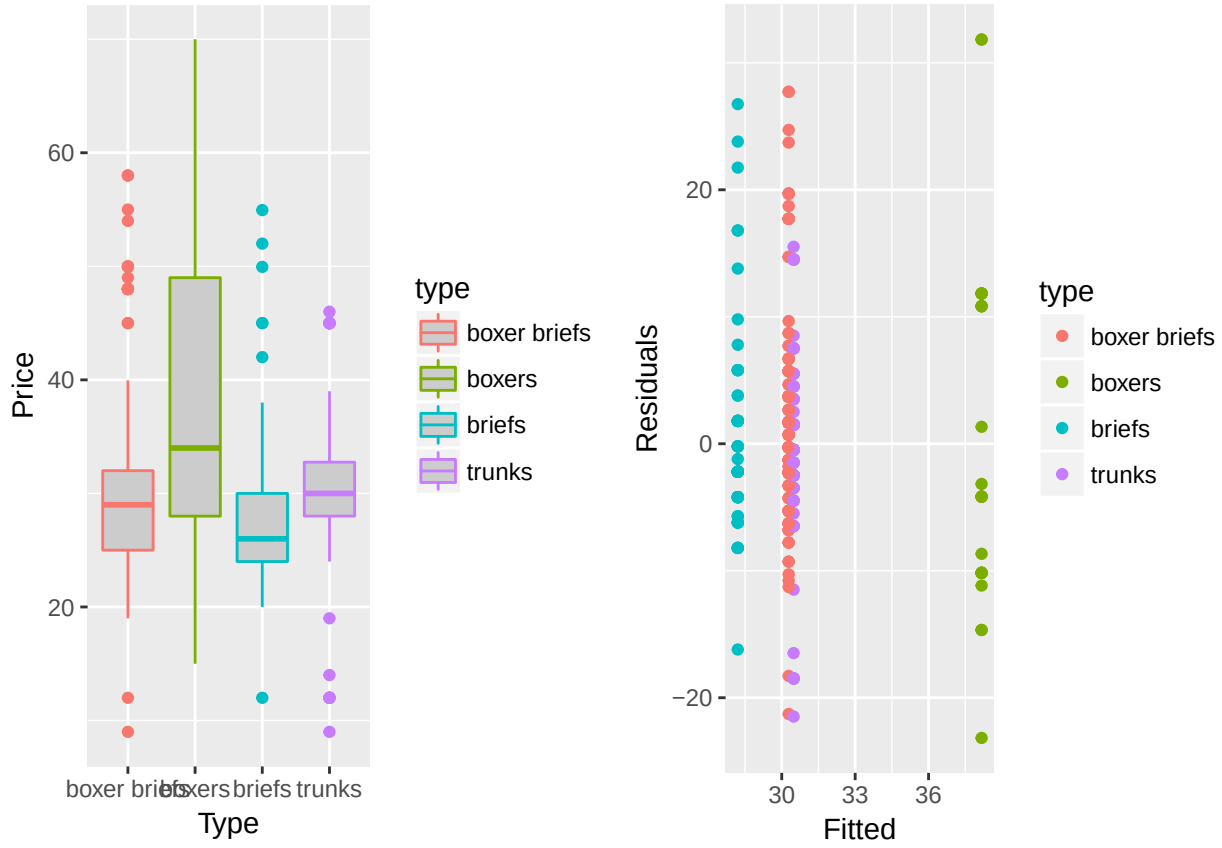
Since the goal is to determine a secondary underwear type to develop and the price point, the choice become apparent that an ANOVA analysis would be more appropriate across the undergarment types. Consider the following hypothesis.

H_0: There is no difference in pricing by underwear type. H_A: There is a difference in pricing by underwear type.

ANOVA summary table below shows results of analyzing product price by type categories.

```
## Analysis of Variance Table
##
## Response: price.original
##            Df  Sum Sq Mean Sq F value   Pr(>F)
## type        3  2160.5  720.18  13.137 2.84e-08 ***
## Residuals 507 27795.1   54.82
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Since the p-value is close to 0, we reject the null hypothesis. There is significant difference in pricing between at least some types. Box plots below illustrate each group.

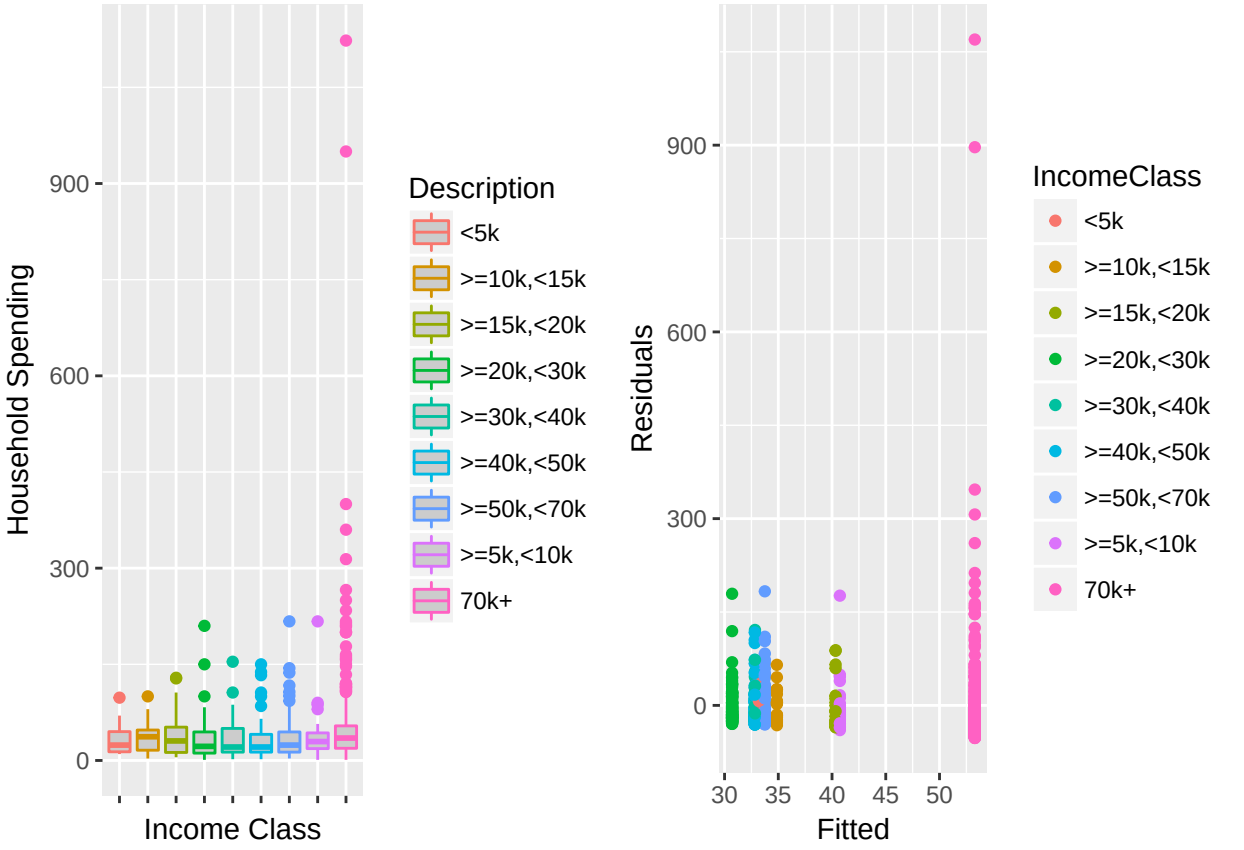


Now consider the household spending from the BLS data to see if there is difference between spending of various income groups. Set up similar hypothesis.

H_0: There is no difference in spending between income classes. H_A: There is a difference in spending between income classes.

```
## Analysis of Variance Table
##
## Response: CLOTHXA
##              Df  Sum Sq Mean Sq F value  Pr(>F)
## Description   8   83490 10436.3  2.5101 0.01068 *
## Residuals   816 3392690  4157.7
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

ANOVA test indicates that with a confidence level of 95%, we reject the null hypothesis. There is significant difference in household spending between at least some income groups.



From the discussion of assumptions above, this data is heavily right skewed which may invalidate this statistical test. However, the information and box plots have some value for descriptive purposes as a guiding factor for the company in terms of who to target.

Conclusion

The data collected provides strong starting guidance to the startup in determining direction of their underwear line. Trunks is a popular type of underwear with a similar pricing distribution of boxer briefs, so the startup should consider adding this type to their line. Approximate price point of a pair is $25-$35.

Not suprisingly, higher income increases spending in the undergarments category. The startup should consider targeting consumers based on their income level. Those with a household income of greater than $50,000 tended to have a higher spending.

Limitations of Analysis

Both price data and BLS income/spending data cover only US market. Factors governing other markets can differ significantly and this analysis should not be extended beyond the US market.

This analysis provides information only about a couple of factors relevant for market research. There are other factors that could be and should be researched, such as consumer buying habits, market size, market growth or decline, and any current trends. Managing the design and launch around the data driven process can generate confidence in business model assumptions.

The analysis covers price data from three major retailers. The retailers are typical for a certain market segment. However, results may not fully transfer to other market segments. Startup should take care in

applying results making sure that assumptions and data sources fit the business model.