June 17, 2018

Course: CIS570 – Business Intelligence

Name: Robert Palumbo

Assignment: Reading Discussions – Week4-Session1

Due Date: Sunday, June 17 @ 11:59pm

Question 1. For a data set with a minimum of 100 rows and two columns of numeric data, a) calculate descriptive statistics (i.e., mean, median, mode, variance, standard deviation, minimum, maximum, range), and b) create plots (e.g., box-and-whiskers, scatter, histogram).

For this assignment I chose the following data-set:

Website: https://www.quandl.com

URL: https://www.quandl.com/data/EOD/HD-Home-Depot-Inc-The-HD-Stock-Prices-Dividends-and-Splits

Content: Home Depot stock prices, splits, and dividends, 1981-2017 monthly, quarterly, and yearly.

Sample Data: 9527 records

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Date | Open | High | Low | Close | Volume | Dividend | Split | Adj\_Open | Adj\_High | Adj\_Low | Adj\_Close | Adj\_Volume |
| 6/5/2018 | 190.44 | 192.38 | 189.7 | 191.94 | 4231073 | 0 | 1 | 190.44 | 192.38 | 189.7 | 191.94 | 4231073 |
| 6/4/2018 | 188.22 | 191.54 | 187.97 | 191.36 | 4101118 | 0 | 1 | 188.22 | 191.54 | 187.97 | 191.36 | 4101118 |
| 6/1/2018 | 187.21 | 188.83 | 186.52 | 187.35 | 3706013 | 0 | 1 | 187.21 | 188.83 | 186.52 | 187.35 | 3706013 |
| 5/31/2018 | 187.16 | 187.36 | 185.25 | 186.55 | 4839928 | 0 | 1 | 187.16 | 187.36 | 185.25 | 186.55 | 4839928 |
| 5/30/2018 | 184.83 | 187.8 | 184.01 | 187.09 | 3976000 | 1.03 | 1 | 184.83 | 187.8 | 184.01 | 187.09 | 3976000 |

From this dataset I created the plots shown below which provide trending of the stock’s open, close, high, low prices over the full yearly span of the data. In general, the stock seems to hold fairly stable in terms of both open/close and high/low prices as shown by the means of each of these. The yearly average open and close prices are within 0.005 of each other. The stock would generally be considered a highly active stock as shown in Figure 5. While the market does not have a specific benchmark for this, it is accepted that trade volumes on the order of millions per days is considerered *active.* HD averaged ~5.5mil trade over this period and as of 2017 it still trades with a corresponding volume.

The most interesting aspect of what I have seen with this data is that shown for the period between ~2006-2010. This period is when the housing market crashed and the recession began. As shown in Figure 5, HD experienced it’s highest ever trade volume of ~94.5mil. There was huge panic in the market during this timeframe and many investors were cutting there losses. Huge sell-offs were occurring and HD being a home improvement supplies company incurred a major hit. Home improvement supplies are typically not needed when new home are not being built and that is just what happened during the housing collapse.

In 1999, HD average high was $67.18. It had a slow decline and in 2005 the average high was ~$39.7. However during the brunt of the recession the stock reached a yearly average low of ~$24.61 (actual daily low of $17.05 10/10/2008). However, during the recession recovery as the economy start to get back on track, the stock has since recovered quite nicely and as of 06/08/2018 the current trading prices is ~$196.86. As they say, hindsight is 20/20. UGH!

Figure 1 Descriptive Statistics



Figure 2 Yearly Average Open, Close Stock Price Barchart

Figure 3 Open, High, Low, Close Stock Price Box and Whisker Chart

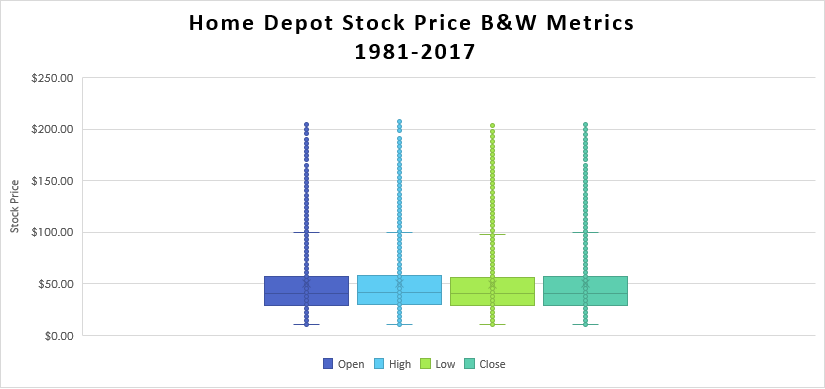


Figure 4 Yearly Average High/Low Line/Scatterplot Chart

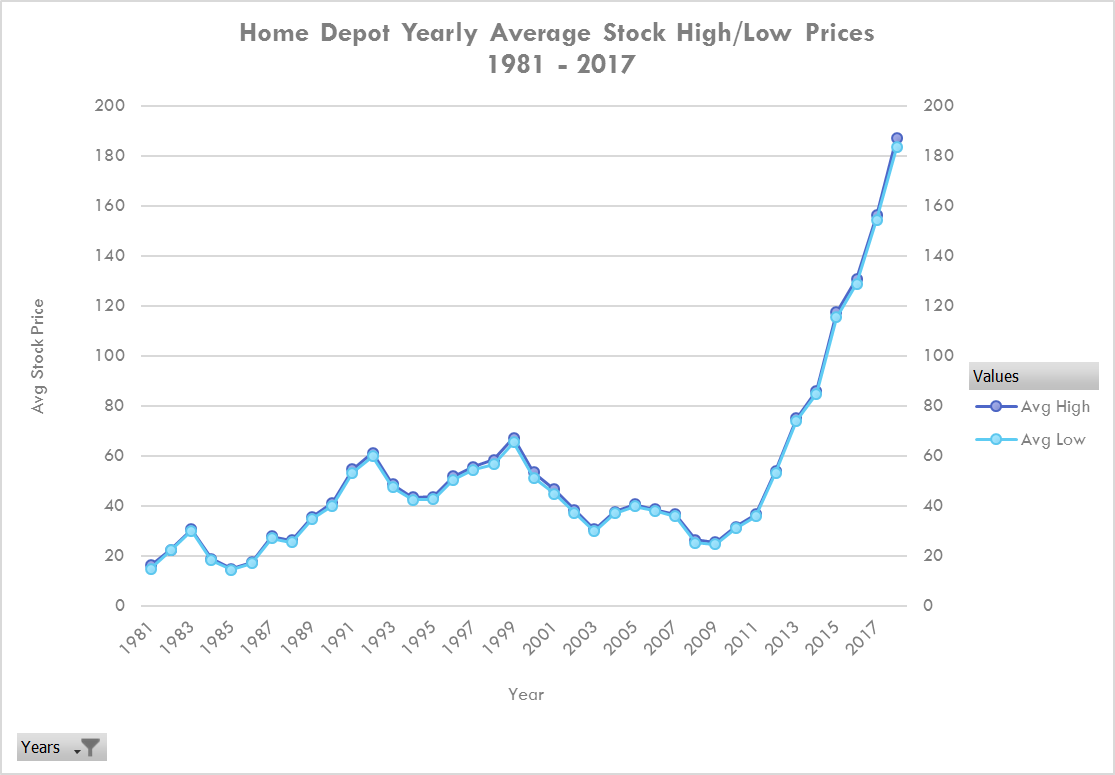


Figure 5 Yearly Volume Histogram