**4. Exploratory Data Analysis**

Data should now have meaningful information to explore.

Prior to predictive analytics.

* Data is ready to be loaded
* Ensure any outliers or missing values before trends analysis.

Items with highest sales?

Items with contribution to maximum profit?

Tools: Power BI and R - Descriptive and Exploratory tools

Top 10 trending items

ROI and demand for items

Highest ROI this week not the same next week.

Seasonal items?

Or local events in sales regions or area?

Same trends don’t always continue.

Time series indices – items

Items contributing to ROI

Document thought process

Why you choose certain variables for your analysis?

**Explore and Clean Data**

Perform necessary data exploration to get the feel of the data in Power BI or R. Remove any outliers, normalize the data, etc. Discuss why and how you have explored specific aspects of the data. Are there any visual clues that demonstrate the outliers, distribution of the data? Are there any visible patterns or trends in the data? You could use either or both of the technologies effectively saving time. Here are some sample questions you may wish to use:

* What are the trends (highest, lowest, average) in demand, replenishment, sales, profit, ROI, etc.?
* What are the items with highest demand?
* What are the items with highest ROI?
* What are the replenishment rates of the items with highest demand?
* Identify whether the WWI is losing their business for under replenishment?
* Identify whether the WWI is losing their ROI for replenishment of incorrect inventory?
* What are the insights from customers segmentation based on size of the market basket?
* Are there any product/item affinities certain customers have?

The questions are supposed to point exactly what is expected. For example, what is the trend (with performing exploratory analytics) of demand? Identify the top 10 trending (highest demand) items. What is the predicted (with performing predictive analytics) demand over next one week for top trending items identified in the exploratory analytics? Similarly, what are the trends of ROI? Find top 10 items with highest ROI? What is the predicted ROI over next one week for the top 10 items with highest ROI?

These are examples of two exploratory analytics; however, students can choose to combine any other two factors such as Replenishment Rate and ROI. **Important Note:** *Be sure to answer at least two of the questions above, provide scripts, programs, queries, screenshots, results and documentation.*