Project Goal:

Based on the company needs, our project is set to provide Nordstrom's buying teams with predictive models that inform both the selection breadth and depth at a category and supplier level for their online business. This will be executed through a series of steps that consist of:

1. Initial exploratory data analysis. We will get insight into the data and generate summary statistics.
2. Formulation of demand model. This model is used to forecast the one-period ahead demand and this will be the input into in the final selection model
3. Formulation of optimization model. We will balance the demand the supply and find out the optimal buying assortments and replenishment level.

The project will be limited to online service of women's handbags, and the analysis will be segmented into 2 departments: Bridge and Contemporary.

Each student will be assigned and work with only one department, and the assignment will be decided later. And we may be required to find the difference between these two department.

The project requires us working with historical sales and use different statistical tools and machine learning techniques to do the analysis and build the optimization selection model. Finally, the final model will be applied in other products and assist the stakeholders to take action in the future business decisions. This model is helpful in the supply chain to avoid delay, out of stock, unsold and to keep the right inventory level.

Approaches:

For the Data Exploration:

Firstly we will wrangle the data and filter out what part we could use. Do some aggregates and merging

For time series analysis requirement, we may aggregate the transaction data on different time scales.

We will merge some datasets, such as merging product hierarchy dataset to transaction datasets, to find the relationship between the product attributes and the sales/.

And we will visualize those observations and make the trends or prominent features been seen easily. And we will create the initial takeaways, like the summary statistics in R notebook, for better understanding our objectives,

Demand Model:

This is the fundamental part of the inventory plan and it will be used as an in our final selection model. And the accuracy of the demand model directly affects our final result. It is an important task.

But compared to other retailing service industries, sales forecasting is hard in fashion retailing because

* the fashion product’s life cycle is very short
* fashion product’s demand is highly volatile with ever-changing taste of the consumers
* sales of fashion products are affected” by seasonal factors, fashion trend factors [6], and many tricky variables (e.g., weather, marketing strategy, political climate, item features, and macroeconomic trend).
* It is more challenging when Nordstrom also has problem that it is carrying a large number of stock-keeping-units (SKUs) with limited historical sales data, a

**Time series models:**Nordstrom online transaction data is a *time series* of data. techniques which are based on this model are:

* Weighted moving average
* Exponential smoothing
* Autoregression integrated moving average model

Advantages: simplicity and fast speed, but it has shortcoming in neglecting the other exogenous factors that may affect sales and cannot detect the irregular patterns

Casual models: It is based on the assumption that future demand being forecast are related to other variables in the environment.

Longchamp is the most popular item in the online service,

Black size

* Multiple Linear Regression,
* The classification and the regression tree (suitable for new product with no historical data)
* Artificial Neural Network
* ANN model is very useful for forecasting demand in fashion industry for short term forecasting
* hybrid techniques
* Hybrid forecasting methods are usually developed based on the fact that they can utilize the strengths of different models together to form a new forecasting method. As such, many of them are considered to be more efficient than the pure statistical models and pure AI models.
* NN model with extreme learning machine for fashion sales forecasting with a short term horizon is proposed. Their model enables to quantify the relationship between sales amount and some significant fashion product attributes such as color, size and price.

ll make sales forecasting challenging and call for more sophisticated and versatile analytical tools. On the other hand, it is known that the fashion apparel supply chain is a relatively long one which includes upstream cotton plants, fiber manufacturers, apparel factories, distributors, wholesalers, and retailers. As a consequence, the notorious bullwhip effect [7] will have a particularly strong influence on the fashion supply chain. Since forecasting is a critical factor relating to the presence and significance of the bullwhip effect, improving forecasting can help reduce the bullwhip

replenishment and inventory costs

budget constraint

minimum dollar amount or units to sell to carry new