

# CAPSTONE PROJECT PROPOSAL

## Product Backorder Prediction

By: Hediye Ledbetter

### Introduction

Part backorders is a typical supply chain problem. Backordering happens when a customer places an order for a product that is temporarily out of stock with the vendor and order cannot be fulfilled. It is a dream for any business, but it is also a massive problem if we do not know how to handle it. In this project, the goal is to "identify parts at risk" of backorder before the backorder occurs, so the business has time to react. With the help of data analysis, a reasonable prediction on all products can go on backorders is expected. Such a prediction could immensely help the client to plan for a more efficient stocking and backorder handling. A company can manage its inventory more efficiently using a prediction on the backorder risk for the products.

### Problem Statement

The client is looking for ways to improve backorders handling. With the help of this analysis, a reasonable prediction on the products that can go on backorder expected. Such a prediction could immensely help the client to plan for a more efficient stocking and backorder handling.

### Dataset:

Dataset consists of the historical data around the backorders. It has 23 features and 1687860 entries, Dataset found at <https://www.kaggle.com/tiredgeek/predict-bo-trial>

### Feature details:

- sku - Random ID for the product
- national\_inv - Current inventory level for the part
- lead\_time - Transit time for product (if available)
- in\_transit\_qty - Amount of product in transit from source
- forecast\_3\_month - Forecast sales for the next 3 months
- forecast\_6\_month - Forecast sales for the next 6 months
- forecast\_9\_month - Forecast sales for the next 9 months

- sales\_1\_month - Sales quantity for the prior 1 month time period
- sales\_3\_month - Sales quantity for the prior 3 month time period
- sales\_6\_month - Sales quantity for the prior 6 month time period
- sales\_9\_month - Sales quantity for the prior 9 month time period
- min\_bank - Minimum recommend amount to stock
- potential\_issue - Source issue for part identified
- pieces\_past\_due - Parts overdue from source
- perf\_6\_month\_avg - Source performance for prior 6 month period
- perf\_12\_month\_avg - Source performance for prior 12 month period
- local\_bo\_qty - Amount of stock orders overdue
- deck\_risk - Part risk flag
- oe\_constraint - Part risk flag
- ppap\_risk - Part risk flag
- stop\_auto\_buy - Part risk flag
- rev\_stop - Part risk flag
- went\_on\_backorder - Product actually went on backorder. This is the target value.

### **Problem-solving steps:**

Problem understanding

Read data

Data wrangling

Data visualizations

EDA

Training The model

Model validation

### **Methodology**

The analysis will be done using Python, Pandas, Matplotlib, Scikit-Learn packages in IPython Notebook 3.6. Exploratory Data analysis will use to get a generalized prediction model. The following will analyze about the data

- Provide an overall insight from data using exploratory data analysis.

- Identifying what the main features are caused backorders the most?

- Predict product is going to backorder or not?

