

ANLY482 Analytics Practicum

Project Proposal

*Optimizing operations productivity and developing an interactive*

*dashboard for local supply chain company*

Team 03

*Submitted by OPTEAMIZATION*

1. Tay Jing Ying (jytay.2014@sis.smu.edu.sg)

2. Russell Yap Song Chen (russell.yap.2014@sis.smu.edu.sg)

3. Chua Weilun (weilun.chua.2014@sis.smu.edu.sg)

14 January 2018

Table of Contents

[Project Sponsor 3](#_Toc503724479)

[Motivation 3](#_Toc503724480)

[Objectives 3](#_Toc503724481)

[Data 4](#_Toc503724482)

[From Warehouse Management System 5](#_Toc503724483)

[Handling In Report 5](#_Toc503724484)

[Handling Out Report 6](#_Toc503724485)

[From Reporting System 7](#_Toc503724486)

[Inbound Order Status Summary Report 7](#_Toc503724487)

[Outbound Order Status Summary Report 8](#_Toc503724488)

[From Human Resources System 9](#_Toc503724489)

[Overtime Hours Report 9](#_Toc503724490)

[Methodology 9](#_Toc503724491)

[Data Preprocessing 10](#_Toc503724492)

[Visualizations 10](#_Toc503724493)

[Bar Chart 10](#_Toc503724494)

[Treemap 11](#_Toc503724495)

[Time Series Line Chart 12](#_Toc503724496)

[Technology 13](#_Toc503724497)

[Scope of Work 14](#_Toc503724498)

[Data Gathering and Scoping 14](#_Toc503724499)

[Data Cleaning, Wrangling and Restructuring 14](#_Toc503724500)

[Data Visualization and Reporting 14](#_Toc503724501)

[Work Plan 14](#_Toc503724502)

[Limitations 15](#_Toc503724503)

[Important Information about the Business Process 16](#_Toc503724504)

# Project Sponsor

Our sponsor is a home-grown supply chain solutions company and leading regional supply chain management partner to many of the world’s leading brands such as Dell, Motorola, Samsung, LG, Exxon Mobil, Unilever and Heinemann. The company has operations all over Asia Pacific, such as China, the whole of ASEAN, India and Korea and they service various industry segments: Chemicals & Healthcare, Consumers & Retail, Electronics & Technology, E-Commerce and Cold Chain.

A strong advocate for technological and operational innovation, the company advocates productivity and sustainability improvements through interactive solutions so that they can provide their diverse consumer-base from various industry clusters with world-class logistics services.

# Motivation

The company has stayed ahead of their competition by embracing technology in the early days and now, they are starting to dabble into analytics to help make use of the large amount of data that they have generated throughout the course of their business. Currently, these information is being stored in their proprietary Warehouse Management System (WMS) which captures transaction information whenever goods are moving in, out or within the warehouse. The amount of information being managed by the system is massive and it is very complicated, thus managers must spend time to use pivot tables in excel to better visualize the information being generated to get a better view of how the current operations is being run. This uses up valuable time which could have been used for other more important operational matters.

Also, man-hour management is also very important as the amount of volume that the material handlers can handle is limited to 7 containers for inbound and 7 containers for outbound. Any more than this and overtime will be required. Currently, material handlers are the ones that request for overtime hours whenever the volume of work exceeds the current baseline threshold. Thus, it is up to the judgement of the manager on the ground to make the final call on how long the material handlers need to overtime. As a result, there are cases of material handlers taking more overtime than necessary leading to increased cost which leads to the company having their margins beings reduced due to this inefficiency as the client would only pay for a certain number of overtime hours based on the agreed service level agreement.

# Objectives

The main objectives of the project would be to do the following:

* Create an Operations Dashboard to visualize the following KPIs (The dashboard will be populated using data from the currently generated reports.)
  1. Operations Productivity Chart
     1. The purpose of this chart is to visualize the operations performance at each hour of the day and see if the drop in the productivity level is in-line with the break timings and if there are additional timings which are unproductive.
     2. X-axis: Hour of the day
     3. Y-axis: Number of pallets successfully serviced
  2. Product Ranking Chart
     1. The purpose of this chart is to find out which are the best-selling items and which are the least popular items.
     2. Treemap view of the products filtered by month / year
  3. Product Seasonality Chart
     1. The purpose of this chart is to see the seasonality factors of the different products.
     2. X-axis: Month of the year
     3. Y-axis: Number of pallets
  4. Actual Overtime Hours Chart (Dual-axis Chart)
     1. These charts will be made up of 2 lines, one for the total number of overtime hours incurred and another being the total number of containers that were completed for the day. One chart will be used for the morning shift and another chart will be used for the night shift.
     2. The purpose of this chart is to visualize the number of overtime hours for the month in a timeline view so that managers can see the overtime performance of the operations in relation to the number of additional containers that need to be serviced for the day.
     3. X-axis: Date of the month
     4. Y-axis: Number of overtime hours
     5. Y-axis: Number of containers
  5. Overtime Performance Analysis Chart
     1. The purpose of this chart is to use the historical timings to come-up with a baseline overtime hour required for varying number of additional containers.
     2. X-axis: Number of containers over the threshold limit
     3. Y-axis: Average time taken to complete the additional order
* Find out if the current operation performance is in-line with the baseline performance of 7 containers of inbound and 7 containers of outbound per working shift.

# Data

The data provided by our sponsor are obtained from 3 of their in-house systems:

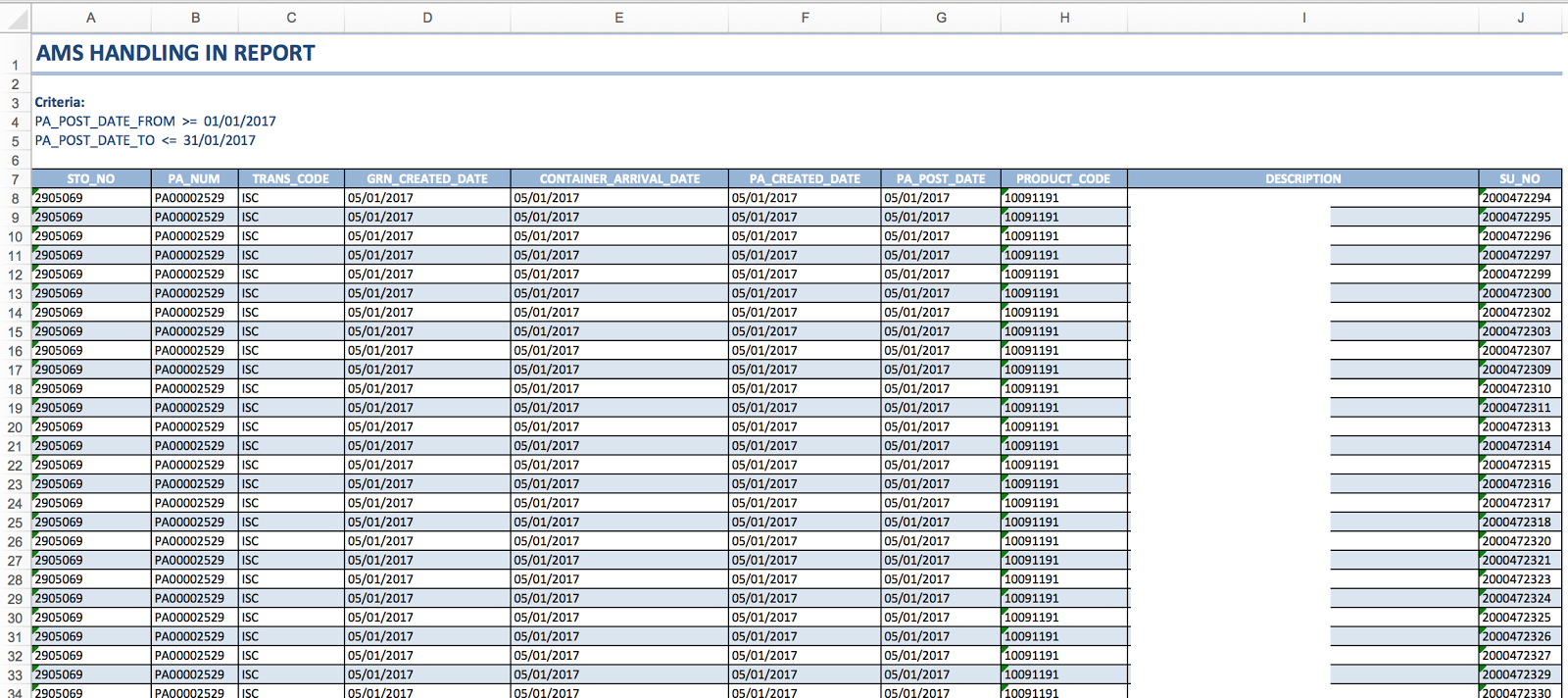
1. Warehouse Management System (WMS) - The warehouse management system is used to keep track of the movement of the goods coming in, going out and within the warehouse itself. The WMS is a proprietary system developed in-house.
2. Reporting System (RS) - This system provides the company with performance management systems which allow the company to manage their KPIs. Data from the WMS is being fed into the RS to generate summary reports.
3. Human Resources System (HRS) - This system provides information on the permanent employees’ working hours and overtime hours.

For this project, our sponsor has provided us with data for one of their client, which is a company selling powdered milk cans. This includes the inbound and outbound products for the year 2017. There are several excel workbooks provided which we will explain below.

## From Warehouse Management System

### Handling In Report

There is a total of around 140,000 rows of data spanning across January to December 2017 for the Inbound orders for this client. Each row represents a pallet.



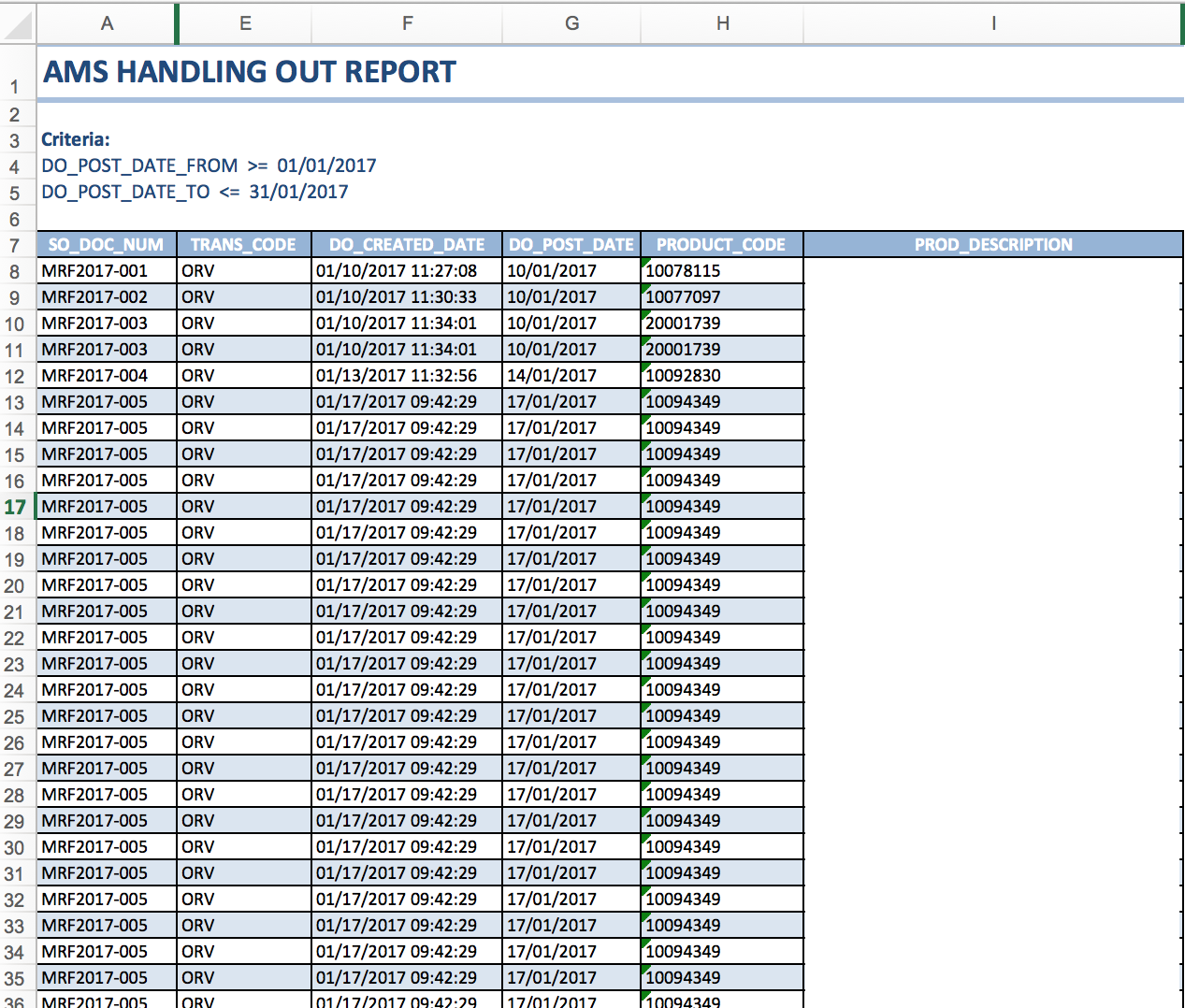
*\*The description is blanked out because it contains information about the client’s products*

#### Metadata Dictionary

|  |  |
| --- | --- |
| **Terminology** | **Description** |
| STO No | Unique identifier for an inbound order. It is the same as the Customer Ref No in Inbound Order Status Summary Report from RS. |
| Trans Code | There are several trans code, but we will only focus on:  IFP: Plant Transfer (Physical system change, movement of goods) |
| GRN Created Date | Goods Received Note. Once GRN is recorded, it means that inbound process is completed. |
| PA Created Date | The date when the putaway job is created. |
| PA Post Date | The date when the putaway job is completed. |
| Description | The product which will be stored in the warehouse e.g. Brand X 850g powdered milk cans or Brand X 400g powdered milk cans. |

### Handling Out Report

The number of Outbound orders for this client from January 2017 to December 2017 is approximately 81,000 rows of data. Each row represents a pallet.



*\*The prod\_description is blanked out because it contains information about the client’s products*

#### Metadata Dictionary

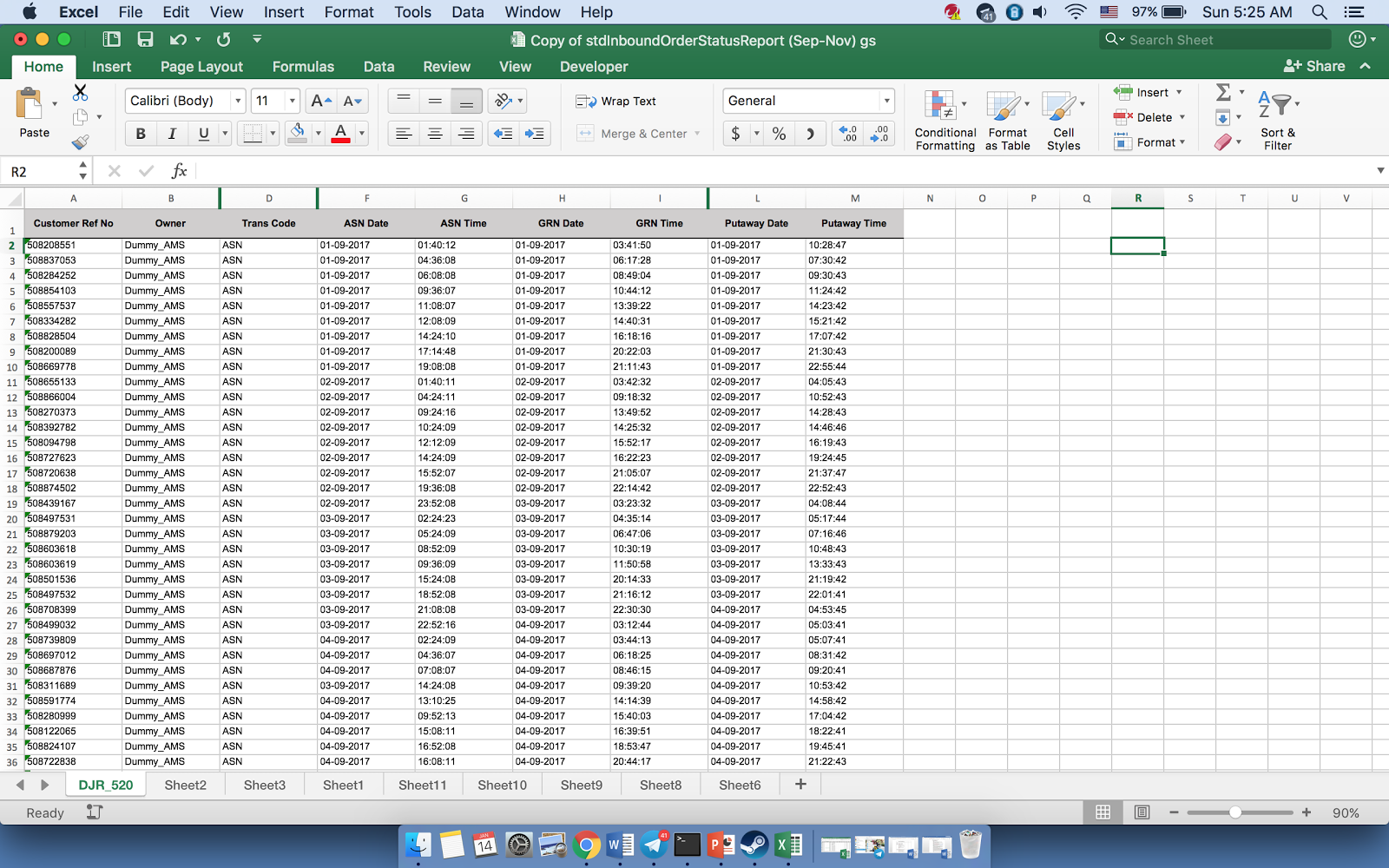
|  |  |
| --- | --- |
| **Terminology** | **Description** |
| SO Doc Num | Unique identifier for an outbound order. It is the same as the Order Num in Outbound Order Status Summary Report from RS. |
| Trans Code | There are several trans code, but we will only focus on:   * ORV - Return to vendor * OSD - Meant for disposal * OSO - For shipment outbound   These 3 trans codes are to be taken into consideration because there is a physical movement of goods. |
| DO Created Date | The date which the delivery order is created. |
| DO Post Date | The date which the delivery order is completed. |
| Product Code | Unique identifier of a product. |
| Prod Description | The product which will be stored in the warehouse e.g. Brand X 850g powdered milk cans or Brand X 400g powdered milk cans. |

## From Reporting System

### Inbound Order Status Summary Report

There is a total of around 3,000 rows of data spanning across January to December 2017 for the Inbound orders for this client from RS. This number is significant lesser than the Handling In Report as it is an aggregation of the data from WMS. Each row represents a container, which contains 22 pallets.

A snippet of the data for an Inbound Order Status Summary Report is shown below.



#### Metadata Dictionary

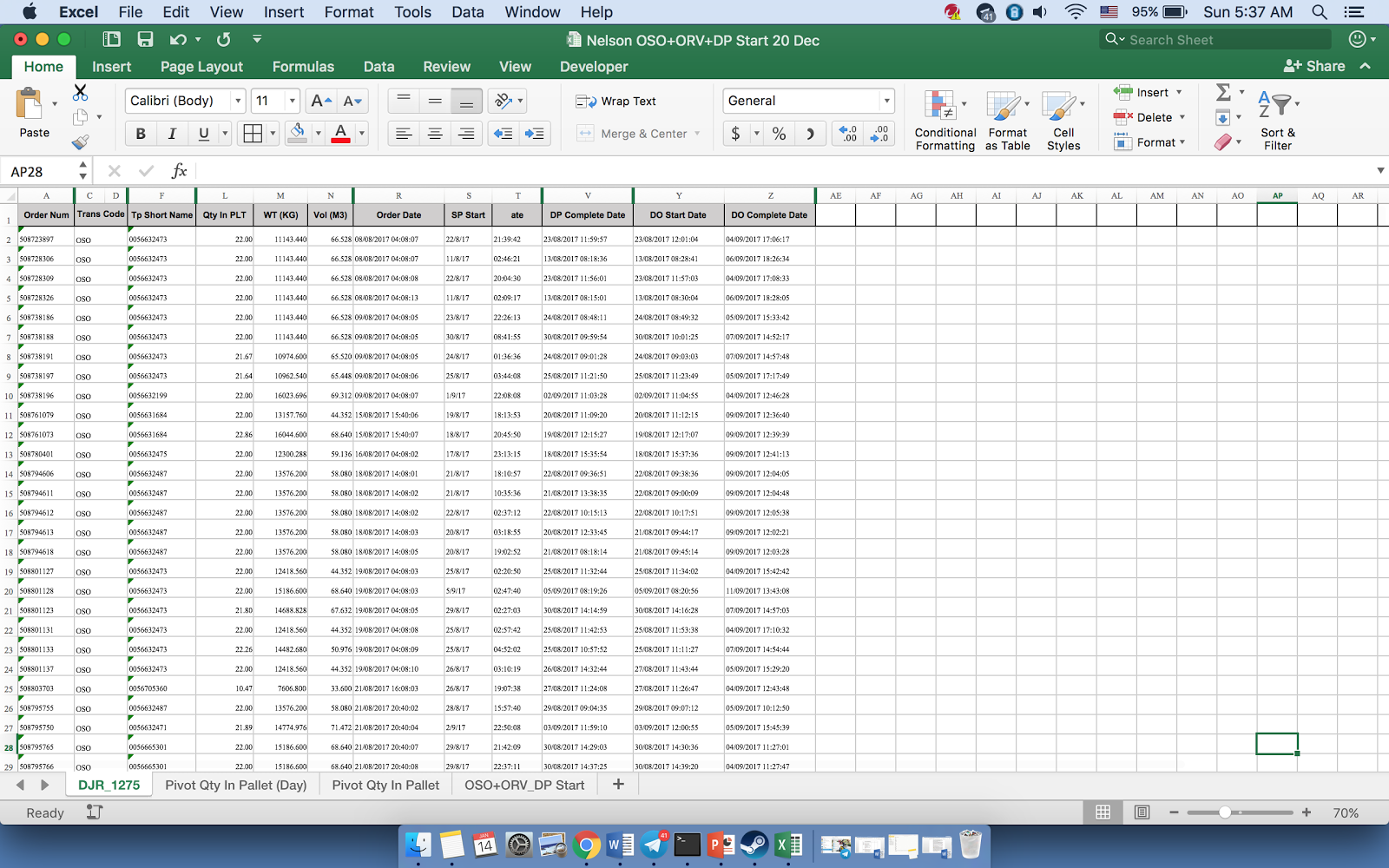
|  |  |
| --- | --- |
| **Terminology** | **Description** |
| Customer Ref No | Unique identifier for an inbound order. It is the same as the STO No in Handling In Report from WMS. |
| ASN Date/Time | Arrival Shipment Notice. It refers to the timing when the goods leave the production plant. |
| GRN Date/Time | Goods Received Note. Once GRN is recorded, it means that inbound process is completed. |
| Putaway Date/Time | The putaway timing refers to the time after which processing of a pallet is complete. It is usually 15 minutes after the reporting of GRN. |

The difference between ASN Date/Time and GRN Date/Time is the total time taken for goods to finish its inbound process after leaving the production plant. This includes the travelling time and the handling in time. However, we are only interested in the handling in time which is the time taken for the workers to complete the handling in. Yet, this timing is not available because the time which the truck arrives at the warehouse for handling in is not recorded. Therefore, we will have to assume on the average time taken to travel from the production plant to the warehouse, which is from Tampines to Tuas. To ensure a more accurate timing of the actual handling in time, we will recommend our sponsor to start collecting this information as well.

### Outbound Order Status Summary Report

There is a total of around 4,000 rows of data spanning across January to December 2017 for the Outbound orders for this client from RS. This number is significant lesser than the Handling Out Report as it is an aggregation of the data from WMS. Each row of data represents a container, which contains 22 pallets.

A snippet of the data for a typical Outbound Order Status Summary Report is shown below.



#### Metadata Dictionary

|  |  |
| --- | --- |
| **Terminology** | **Description** |
| Order Num | Unique identifier for an outbound order. It is the same as the SO Doc Num in Handling Out Report from WMS. |
| Trans Code | There are several trans code, but we will only focus on:   * ORV - Return to vendor * OSD - Meant for disposal * OSO - For shipment outbound   These 3 trans codes are to be taken into consideration because there is a physical movement of goods. |
| Qty in PLT | Quantity in Pallet, which will be rounded up to whole number. |
| DP Start Date | DP (direct picking) start refers to the operational trigger time when operation starts. |
| DP Complete Date | DP complete refers to the time when operation is completed. (actual timing is usually about 45 minutes after DP Start Date) |

## From Human Resources System

### Overtime Hours Report

The report on the number of working hours and overtime hours for the permanent staff can be retrieved from the HRS. This information helps the Operations Manager to get a sense of the estimated number of overtime hours required for a client storing this type of products.

Given the nature of a warehousing business, a client typically releases a Request For Quotation (RFQ) and a logistics provider will bid for the project. The contract will consist of a service level agreement which determines the number of workers that can be used per shift, the total number of overtime hours that the client agrees to pay, the amount of transactions that the client must provide each month, etc.

Thus, since the amount of overtime that the client is willing to pay is fixed, overtime must be properly planned so that the company will not exceed the agreed amount of overtime as this would lead to reduced margins from the client.

# Methodology

In this section, we will explain the methodology which our team will implement to perform analysis on the data provided by our sponsor.

We will be using JMP Software and Microsoft Excel for Exploratory Data Analysis (EDA) to better understand the dataset given and its characteristics. As part of data preprocessing, our team will be performing the following steps to obtain a clean dataset. The steps will eventually be converted into a script which will be used to clean the data that is uploaded into the dashboard which we will develop for our sponsor.

## Data Preprocessing

With every new dataset, we first must clean the data to remove irrelevant data that should not be included in our analysis. For data cleaning, the steps include:

1. Handling missing values. If there are missing values in a row of record, the entire row will be excluded because it will be inaccurate to include it.
2. Handling duplicate data. Duplicate data could occur when the employees double scan the barcode upon inbound of goods. Similarly, in the event of duplicate data, we will remove the entire row as well.
3. Resolving redundancies caused by data integration.

With the clean dataset, we will proceed to further explore the data and find out potential visualizations and analysis that can be done with the dataset to provide a more in-depth analysis and dashboard that will be useful for our sponsor.

## Visualizations

The final product of our project is to create an Operations Dashboard to visualize the following KPIs:

1. Operations Productivity Chart
2. Product Ranking Chart
3. Product Seasonality Chart
4. Actual Overtime Hours Chart
5. Overtime Performance Analysis Chart

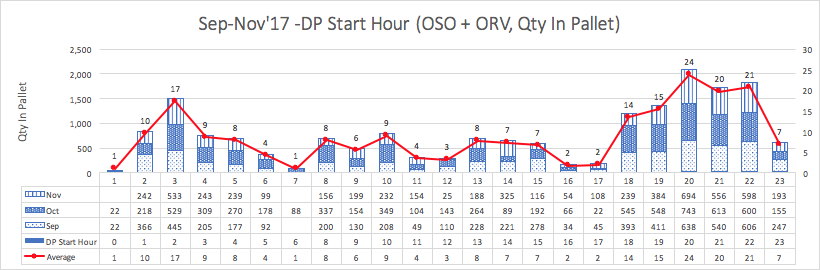
At the start of each day, the Operations Manager will upload 3 CSV files namely: Handling In Report, Handling Out Report and Overtime Hours Report for the day before. Upon uploading these files, the data cleaning script will be run and the relevant data will be stored into the database which will be used for data visualizations on the dashboard.

The visualizations to be used on our dashboard include Time Series Line Chart and Treemap. The Operations Manager will have the ability to select the period to view as well.

### Bar Chart

#### Operations Productivity Chart

Using a bar chart, we can visualize the various operations productivity and performance level during each hour of the day to see if there is a generic trend throughout the day, or whether the productivity levels are in line with the break timings allocated to the operators. From here we can further identify which time of the day could be better made use of to increase the overall productivity level, or whether there is an unexpected unproductive time in the day.

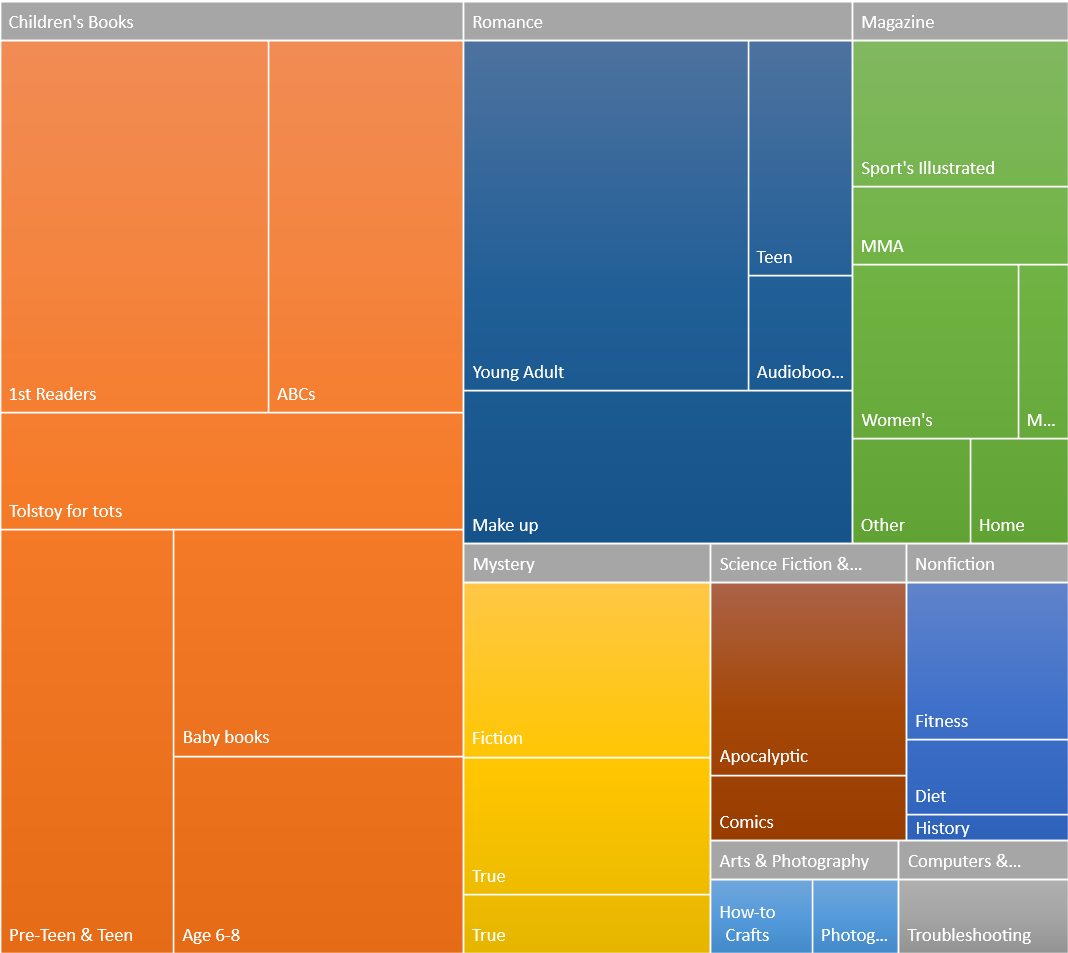


*Example of a Bar Chart that describes the average number of pallets per hour across Sep to Nov 2017.*

### Treemap

#### Product Ranking Chart

A treemap can be used to identify which product is the best-selling item and which is the least selling item. They are being ranked according to the size and proportion of their sales as compared to one another. As seen from the example below, we can identify the best-selling product by the proportion of its sales.

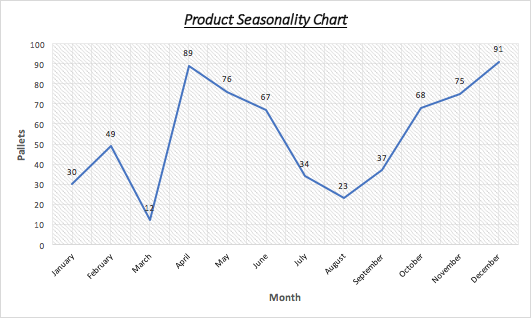


*Example of a Treemap that is used to understand which is the best-selling categories of books.*

### Time Series Line Chart

#### Product Seasonality Chart

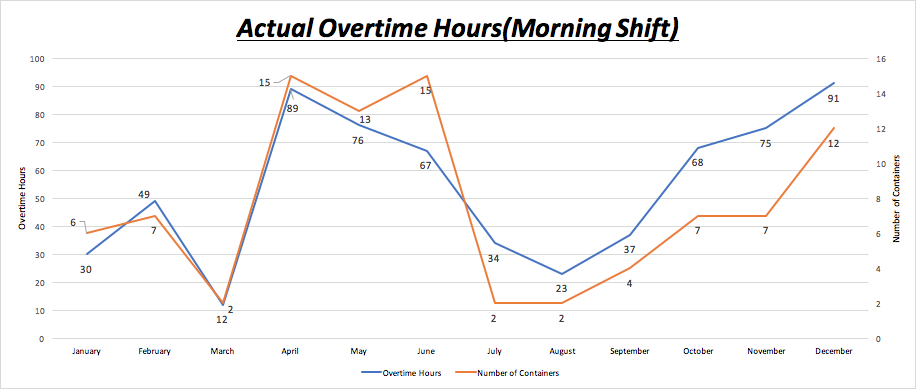
After identifying the ranking of the products by its sales volume through a Treemap, we can further drill down into the individual product seasonality performance through a Time Series Line Chart, where we can observe the sales of the product over the months or years.



*Example of a Time Series Line Chart that describes the number of orders over the year.*

#### Actual Overtime Hours Chart (Dual-axis Chart)

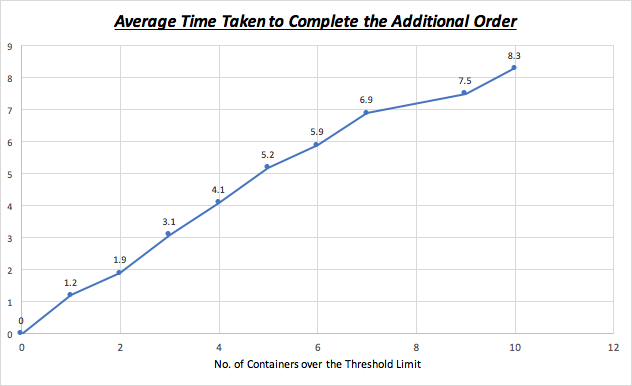
These charts will be made up of 2 lines, one for the number of overtime hours and one for the number of containers. The purpose of this chart is to visualize the number of overtime hours for the month in a timeline view so that managers can see the overtime performance of the operations in relation to the number of additional containers that need to be serviced for the day.



*Example of a Time Series Line Chart that describes the number of containers and overtime hours for Morning Shift.*

#### Overtime Performance Analysis Chart

With the chart, we can tell from historical data on how much time in hours is taken to fulfil and handle the number of specified containers that exceeds the threshold limit of a daily operation.



*Example of a Time Series Line Chart that describes the time taken to fulfil an additional order.*

# Technology

We will be utilising several tools to aid us in the development of our dashboard. This includes:

|  |  |
| --- | --- |
| **JMP Software** | JMP Software is developed by the JMP business unit of SAS institute and it is a powerful tool for analytics. Our team plans to use JMP Software for Exploratory Data Analysis. |
| **Highcharts and JavaScript** | Highcharts is a simple-to-use charting library written in JavaScript which allows us to add interactive charts to our web application for the dashboard. It will be used to develop our dashboard for the various visualizations mentioned above. |
| **Python** | Python is a programming language which we will use to create a script meant for cleaning the dataset when it is uploaded. |
| **PostgreSQL** | PostgreSQL is an object-relational database management system which will be used to store the data from the CSV files. |

# Scope of Work

The following describes the scope of work which the team will be working on across the span of the next few months till the final submission.

## Data Gathering and Scoping

Our sponsor has given us various excel spreadsheets with data for 2017, as mentioned above, for a client selling powdered milk cans. Our project scope includes analysing the products ranking, products seasonality, operations productivity and overtime hours used to meet the demand. With the data, we will perform Exploratory Data Analysis using JMP Software.

## Data Cleaning, Wrangling and Restructuring

Our team aim to combine the various excel spreadsheets into one for Inbound and one for Outbound by using their unique identifier to link the data together. Data cleaning will be done using Python to handling missing values and duplicate data, after integrating the data.

The team will also use Python to write a script which will combine the various excel spreadsheets and extract the relevant fields before consolidating them into a single spreadsheet that will be used as an input for the analysis and the Operations Dashboard. All the data that is uploaded will be stored into a PostgreSQL database, which serves as the database of information for the dashboard.

## Data Visualization and Reporting

Once the data is cleaned and combined, we will extract the data and create a dashboard with various visualizations as mentioned earlier. This will be done using Highcharts and JavaScript. The interactive dashboard will allow managers to have easy access to visualizations showing the performance of the current operations as compared to past operations.

We will also analyse the data provided so that we can evaluate if their current operations are operating within their current baseline threshold and is the current operations sustainable. Lastly, we will also research on the feasibility of doing a model to help the Operations Managers find out how much overtime hours are required based on the amount of outstanding work and workers on hand.

# Work Plan

The following Gantt chart shows the task allocations to each member for each iteration, as well as the important project milestones which we will prepare for.

****

# Limitations

As with any other real life projects, there are limitations which we face especially with regards to the availability of data.

|  |  |  |
| --- | --- | --- |
| **S/N** | **Limitations** | **Conclusion** |
| 1 | The current data provided only consists of information from 1 company, which sells powdered milk cans, that our sponsor company is handling. | We assume that the same analysis can be replicated for the logistic warehousing of other companies’ products which are similar in nature. |
| 2 | There is a lack of information on the exact time taken to complete the inbound storage as they do not record the timestamp when the truck arrives at the warehouse for handling in to take place. | With the lack of this information, we cannot analyse the time to complete an order and predict the time required to complete an inbound order. We will recommend our sponsor to record this information because we feel that it will be useful to analyse the operations productivity. For our project, we will come up with an estimate for the travelling time from the production plant to the warehouse. |
| 3 | Out of the 12 workers involved in the logistics work for this company, there are 3 contract staff. However, there is also a lack of information on the overtime hours of these contract workers. | The lack of information on overtime hours for contract workers will affect the accuracy of our findings and analysis. However, we will assume that the average overtime hours of the contract staff is the same as the average overtime hours of the permanent staff. |
| 4 | The workers allocated to work for Inbound can also help the staff allocated to work for Outbound and vice versa. For overtime, staff working on Inbound can help to do overtime for Outbound processes. However, there is no record on which worker is allocated for Inbound and which worker is allocated for Outbound process. | We will assume that all overtime work will be split based on the excess demand, which is determined by more than 7 trucks, for Inbound and Outbound. For example, there are 10 Inbounds and 9 Outbounds for a shift and the overtime hours are 10 hours in total, the overtime used for Inbound = 10 / ((10-7) + (9-7)) \* (10-7) = 6 hours. |
| 5 | The data provided from the WMS may not be totally representative of all the Inbound and Outbound transactions for this client as there may be human errors involved e.g. Worker did not scan the barcode | However, as the data size provided is sufficiently large (over 100k of data), the small percentage of potential human errors can be neglected. Therefore, the analysis conducted will still be considerably accurate. |

# Important Information about the Business Process

1. One container can hold up to 22 pallets.
2. Morning shift is from 08:00 to 17:30. Night shift is from 18:00 to 07:30. Refer to the table below for an exact schedule for both shifts.
3. Each shift is made up of 6 people, 4 for outbound, 2 for inbound but they can do the roles of the other function if they have nothing to do.
4. The morning shift does more inbound activity while the night shift does more outbound activities due to the nature of the drivers. Drivers work normal office hours, so there is more inbound activity for the morning shift.
5. There are 6 forklifts being used throughout the operations, there is a break to charge the forklifts from 00:00 - 01:00 every day.
6. The current baseline level of operations activity per day is 14 containers, 7 for inbound and 7 for outbound. Anything more than this and overtime will be required.
7. For the data coming from the WMS, each row represents one pallet while for data coming from the RS, each row represents one container (22 pallets).

**Daily Schedule**

|  |  |
| --- | --- |
| 08:00 to 17:30 | Morning Shift |
| 10:00 to 10:15 | Break |
| 12:00 to 13:00 | Lunch Break |
| 17:30 to 18:00 | Dinner Break |
| 18:00 to 07:30 | Night Shift |
| 22:00 to 22:15 | Break |
| 00:00 to 01:00 | Supper Break |
| 03:00 to 03:15 | Tea Break |