



Optimizing Warehouse Processing

Project Proposal

ANLY482

Analytics Practicum

Submitted by:

Team Skulptors

1. Gabriel Leong Jun Kang
2. Tan Siying
3. Zhou Xuanyi

Professor:

Prof. Kam Tin Seong

Table of Contents

PROJECT SUMMARY	1
1 GROUP INTRODUCTION	2
2 PROJECT SPONSOR - COMPANY DESCRIPTION	2
3 SPONSORS & LIAISON INFORMATION	2
4 PROBLEM STATEMENT / MOTIVATION	2
5 PROJECT OVERVIEW	3
6 PROJECT OBJECTIVE / GOALS	3
7 DATA PROVIDED	4
8 METHODOLOGY	7
8.1 DATA PREPARATION	7
8.2 VISUALIZATIONS.....	7
8.2.1 <i>Control Chart</i>	7
8.2.2 <i>Time series line chart</i>	8
8.2.3 <i>Treemap</i>	9
9 TECHNOLOGY	9
9.1 D3.JS	9
9.2 JMP	10
9.3 SAS ENTERPRISE MINER.....	10
9.4 JAVA & BOOTSTRAP	10
9.5 OPENSIFT	10
10 TIMELINE & SCHEDULE	11
11 LIMITATIONS	11
12 RETURN OF INVESTMENTS (ROI)	12
13 RISK MANAGEMENT	12
13.1 RISK METRICS.....	12
13.2 RISK ASSESSMENT	13
14 REFERENCES	14

Project Summary

In recent years, data has evolved to provide answers and solutions to problems hidden from the surface. Analytics, the study of data, has evolved over the years to be an integral tool for many companies to further improve efficiency and sustainability. However, the ability to easily visualise data, quickly modify views and interrogate data can be difficult and time consuming due to the difficulty and effort required to update reports and dashboards. On top of that, in today's connected world there are many more data sources available to organisations, such as machinery sensors, mobile devices, wearables, web logs, etc. The Internet of Things (IoT) now allows us to collect and exchange data between systems, providing a rich set of data for analysis. Some of this data is unstructured, in siloes, making it difficult to analyse in its raw form let alone integrating it to gain valuable insights¹.

Today, a local multi-national corporation faces the same adversity as many others – the plight of unworked data. Visualizations of the data provided by their in-house warehousing solution has proven to be difficult and much of the information is not placed to better use. The company hopes to solve 3 main issues, namely how they should categorize their warehouse supplies into ABC categories (each category refers to how fast the goods move) to store their goods in the best location available, how can their data be visualized better so that it is easier for employees to see, and how can their warehouses be utilized to the fullest to ensure minimal wastage.

Our team, Skulptors, will be developing a dashboard application to solve the issues of this company. Control charts will be used to depict control of the movement of the SKUs, identifying which type of products have excessive or lack of movements from the warehouse. Treemaps will be used to identify products or SKUs that have minimal utilization of warehouses, and also to identify which warehouse location is not fully utilized. Java add-ons will be used to facilitate additional functions, such as allowing the upload of an excel file to update the database. This allows for our solution to remain sustainable. Time series line graphs will also be used for easy display of inflow and outflow rate of specified SKUs.

With our solution, we hope to assist the company and the employees, relieving them from the hassle of manually analyzing the given data. Through our visualizations and analysis, we also hope to improve upon their warehousing solution, allowing the company to better utilize their warehouses and thus improving the business value.

¹ Solving data warehouse problems with Cortana Analytics (18 December 2015). Received from <https://www.theta.co.nz/news-blogs/tech-blog/solving-data-warehouse-problems-with-cortana-analytics>

1 Group Introduction

We are Team Skulptors. The team consists of:

- a. Leong Junkang Gabriel. Year 4 student from School of Information Systems. 2nd Major in Analytics (Operational Track)
- b. Tan Siying. Year 4 student from School of Information Systems. 2nd Major in Analytics (Operational Track)
- c. Zhou Xuanyi. Year 4 student from School of Information Systems. 2nd Major in Analytics (Accounting Track)

2 Project Sponsor - Company description

Our project sponsor is a local multi-national corporation (MNC), specializing in end-to-end supply chain and logistics industry, with its headquarters located in Singapore. On top of providing transportation and warehouse services, our sponsor also provides specialized consulting services to various companies. Today, the company is one of Asia Pacific's leading supply chain company, partnering some of the largest MNCs in the world and aspiring growth companies, including Dell, Motorola, Samsung, LG, Exxon Mobil, Unilever, LVMH and Royal FrieslandCampina. A strong proponent of innovation, the company advocates productivity and sustainability improvements through interactive solutions to better cater to their diverse consumer-base and industry clusters, ranging from consumer and electronics to chemical and healthcare.

3 Sponsors & Liaison Information

The sponsors will be from the company's Operations Department, namely:

- a. Mr Khoo, Head of Operation Innovation & Development (Corporate)
- b. Miss Shen, Executive, Operations Innovation & Development (Corporate)
- c. Mr Sundar, Process Improvement Executive (Automation)

4 Problem Statement / Motivation

The complication of big data resonates strongly with the company, much like many others. The Warehouse Management System (WMS) captures data of logistics supplies in the warehouses of the company, such as quantity of movement, time of movement and identification codes of the packages stored. With extensive data provided by the WMS, analyzing of spreadsheets can be

cumbersome and inefficient due to visualization impediments. The company is unable to effectively analyze and sculpt solutions as a result.

In addition, the company is in the midst of implementing a vertical lift high-tech system² for its new warehouses in its upcoming Supply Chain City (SCC) project. In light of these recent developments, the company saw an opportunity to better analyze their WMS data to determine its Stock Keeping Unit (SKU)'s inbound rate, outbound rate, warehouse utilization trend, and ideally, by performing the aforementioned analysis, categorizing each SKUs into namely A, B, C categories. 'A' category refers to fast moving SKUs while 'C' category SKUs refers to slow moving SKUs.

With the analyzed results, the company hopes to determine the optimal warehouse location and vertical lift in which a particular SKU can be placed for picking. In addition, the company would also like to see the extent of productivity savings it can obtain with the adoption of a batch picking technique instead of an order picking technique. Order picking involves going into the warehouse to collect the supplies per order basis, while batch picking involves collecting for multiple orders in a batch.

5 Project Overview

This project will serve as the first of the many efforts which the company conducts regularly to constantly improve its entire logistics process via data analytics. Logistics process encompasses the entire flow of Stock Keeping Units (SKU), from the inflow of goods to the outflow of goods. The project is mainly aimed at providing in-depth SKU analysis based on different product brands which the company is handling. In this project, we will be looking into the products of 2 companies it is handling. The companies are:

- a. B. Braun Singapore - A company which manufactures and produces medical products.
- b. Heinemann Asia Pacific - A leading duty free company headquartered in Singapore.

6 Project Objective / Goals

The main objective of the project would be to develop the following:

- a. Creation of an application to sort SKUs by A, B and C categories
 - i. To provide employees with a high level view of the flow and demand changes for different SKUs. This allows higher level management employees to make

² More on vertical lift can be found here: <http://www.haenel.de/de/en/Products/Lean-Lift%C2%AE.html>

- important decisions based on it (e.g. eliminate SKUs which are extremely slow moving as they are taking up warehouse space which can be otherwise optimized.)
- b. Dashboard for quick visualization of inbound and outbound rate for different SKUs
 - i. Inbound rate: X-axis and y-axis to be date against number of inbound in pieces, carton and pallets per day.
 - ii. Outbound rate: X-axis and y-axis to be date against number of inbound in pieces per day.
 - iii. This reduces both the time and manpower needed for the manual analyzing of data as it provides a general trend and flow of SKUs. For instance, when an employee notices a particular product's supply running low he will be able to call for a refill immediately.
 - c. Warehouse Utilization Tool
 - i. To help employees understand fill and flow rate of the warehouse based on historical data.
 - ii. To determine pick rate of warehouse locations and identify most used and least used locations.

7 Data Provided

The data will be obtained from the in-house Warehouse Management System (WMS). The company updates the database whenever goods are being received (inbound) and released (outbound). For the purpose of the project's analysis, the team will be given one year worth of data for each of the companies - B. Braun Singapore and Heinemann Asia Pacific's SKUs. The size of the data is expected to be approximately 4.2 million rows of data (4 excel sheets) for each company. For each of the companies, there will be 3 sets of data provided, namely:

- a. Product master sheet
- b. Product inbound report
- c. Product outbound report

A sample of the **product master** data sheet (B. Braun Singapore) is as shown below:

	A	B	C	D
1	PROD_CODE	PROD_DESC	UNIT_P_PLT	UNIT_P_CTN
17	031250A	ALMARYS TWIN STANDARD BASE PLATES 50MM	4000	100
18	031260A	ALMARYS TWIN BASE PLATE 60/55MM	4000	100
19	031280A	ALMARYS TWIN BASE PLATE 80/75MM	4000	100
20	031365A	ALMARYS TWIN CONVEX 60 / 17-40MM	50	50
21	032250A	ALMARYS TWIN COLO BEIGE 50MM 30/BX	360	360
22	032260A	ALMARYS TWIN COLO BEIGE 60MM	360	360
23	032380A	ALMARYS TWIN COLO TRANSPARENT 80MM	8640	360
24	032450A	ALMARYS TWIN ILEO TRANSPARENT 50MM	360	360
25	032460A	ALMARYS TWIN ILEO TRANSPARENT 60MM	8640	360
26	032580A	ALMARYS TWIN ILEO BEIGE 80MM	360	360
27	032650A	ALMARYS TWIN ILEO BEIGE BAG W FILTER	8640	360
28	032660A	ALMARYS TWIN ILEO BEIGE FILTER 60MM	8640	360
29	032950A	ALMARYS TWIN URO TRANSPARENT 50MM	5760	240
30	032960A	ALMARYS TWIN URO TRANSPARENT 60MM	5760	240
31	036250A	ALMARYS TWIN+ CUT TO FIT 15-45	4000	100
32	036260A	ALMARYS TWIN+ CUT TO FIT 15-55	4000	100
33	036365A	ALMARYS TWIN+ CONVEX 60/17-40	2000	50
34	036780A	ALMARYS TWIN+ BELT, WHITE, 120 MM	100	100
35	037250A	ALMARYS TWIN+ COLO BEIGE 50 MM	8640	360
36	037350A	ALMARYS TWIN+ COLO TRANSPARENT 50 MM	8640	360
37	037750A	AT+ ILEO BEIGE WITH FILTER 50 MM	3000	30
38	037760A	AT+ ILEO BEIGE WITH FILTER 60 MM	8640	360
39	037850A	AT+ ILEO TRANSP WITH FILTER 50 MM	8640	360
40	037851A	AT+ ILEO TRANSP NO FILTER 50MM	360	360
41	037860A	AT+ ILEO TRANSP WITH FILTER 60 MM	8640	360

The interpretation for the **product master** data is as follows:

S/N	Column Number	Description
1	PROD_CODE	Unique ID assigned to a particular SKU, of a particular brand.
2	PROD_DESC	General description (Color, size etc.) of a particular SKU.
3	UNIT_P_PLT	Depicts the number of pieces which can be filled into a pallet for a particular SKU.
4	UNIT_P_CTN	Depicts the number of pieces which can be filled into a carton for a particular SKU.

A sample of the product **outbound** data sheet data (B. Braun Singapore) is as shown below:

	A	B	C	D	E	F
1	OWNER_SHORT_NAME	DOC_NUM	PRODUCT_CODE	ACTUAL_MOVEMENT_QTY	LOC_SHORT_NAME	ACTIVITY_COMPLETE_DATE
64613	SG01	7600274856	JK442	5	A02A18	04-JUN-15
64614	SG01	7600274856	JF120R	3	A02A18	04-JUN-15
64615	SG01	7600274856	JK442	4	A02A18	04-JUN-15
64616	SG01	7600274856	BC242R	6	MF01A	04-JUN-15
64617	SG01	7600274856	BC271R	30	MF01A	04-JUN-15
64618	SG01	7600274856	BH183R	5	MF01C	04-JUN-15
64619	SG01	7600274856	BH814R	5	MF01C	04-JUN-15
64620	SG01	7600274856	BV249R	3	MF01E	04-JUN-15
64621	SG01	7600274856	BV248R	2	MF01E	04-JUN-15
64622	SG01	7600274856	DL212R	1	MF02A	04-JUN-15
64623	SG01	7600274856	DB515R	5	MF02A	04-JUN-15
64624	SG01	7600274856	FK293R	10	MF02C	04-JUN-15
64625	SG01	7600274856	DO322R	3	MF02A	04-JUN-15
64626	SG01	7600277841	JK740	3	A01Z24	04-JUN-15
64627	SG01	7600277841	JK742	30	A01Z24	04-JUN-15
64628	SG01	7600277841	JK740	2	A01Z24	04-JUN-15
64629	SG01	7600277841	EB233R	5	MF02B	04-JUN-15
64630	SG01	7600277841	BT674R	2	A01Z20	04-JUN-15
64631	SG01	7600277841	JK740	3	A01Z24	04-JUN-15
64632	SG01	7600277841	JK744	2	A01Z24	04-JUN-15
64633	SG01	7600277841	BT674R	2	A01Z20	04-JUN-15
64634	SG01	7600277841	EB234R	5	MF02B	04-JUN-15
64635	SG01	7600277085	4056353	6800	A03A14	04-JUN-15
64636	SG01	7600277393	8713147	5	A04A08	04-JUN-15
64637	SG01	7600277393	8713147	2	A04A04	04-JUN-15

The interpretation for the product **outbound** data is as follows:

S/N	Column Number	Description
1	OWNER_SHORT_NAME	Country location of warehouse.
2	DOC_NUM	Unique order number.
3	PRODUCT_CODE	Unique ID assigned to a particular SKU, of a particular brand.
4	ACTUAL_MOVEMENT_QTY	Outbound quantity of a unique SKU in pieces.
5	LOC_SHORT_NAME	Location where a particular SKU has been placed at in the warehouse.
6	ACTIVITY_COMPLETE_DATE	The date where the outbound of a particular SKU is completed.

Note:

- The company did not provide the sample data for product inbound as it is similar to that of the outbound.
- The product master sheet is to be used in conjunction with both the product inbound and outbound report, via the common linkage of the product code.

8 Methodology

The following methodology will be implemented to perform analysis on the past year data provided by our sponsor for the two companies. Before we begin with our analysis, the data needs to be thoroughly examined and prepared to ensure there is no possibility of garbage-in-garbage-out.

8.1 Data Preparation

Before going into elementary explorations of the data, data cleaning is to be conducted to remove irrelevant data that should not be included in our analysis. Some types of reason for exclusion includes:

- **Repetition** – Double scanning happens many of a times in the company’s warehouse. As such, there can be repetition of rows in the past year data that may falsify our findings.
- **Missing data** – if there are any rows of incomplete data, the row will be excluded.
- **Over-spilling** – As our project focuses on inflow rate and outflow rate of the SKUs movement, we will exclude data there does not account for the start of the SKUs to the end of it.

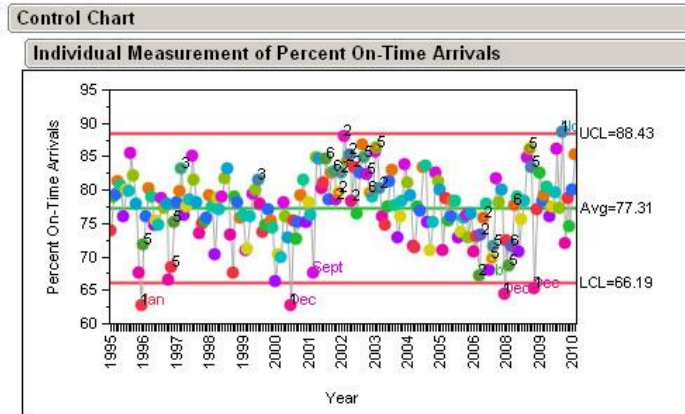
After the initial preparation, we will further develop and sort our data by investigating for potential groupings that can facilitate ease into performing our visualizations and analysis. This will also assist us in the development of our regression model to figure out what type of classification, A, B or C should we give to a specific SKU.

8.2 Visualizations

Our dashboard will consist of the following visualizations.

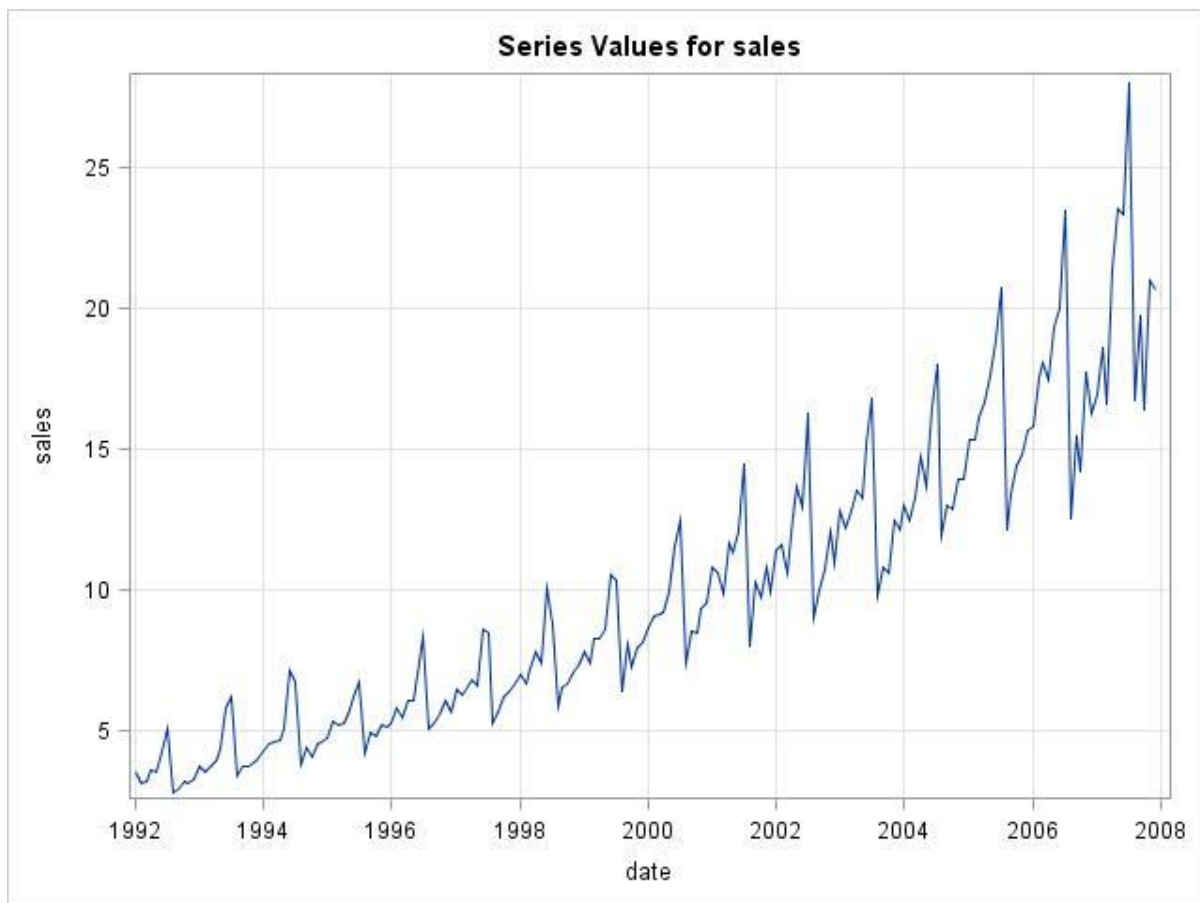
8.2.1 Control Chart

The control chart is a graph used to study how a process changes over time. Data are plotted in time order. A control chart always has a central line for the average, an upper line for the upper control limit and a lower line for the lower control limit. These lines are determined from historical data. By comparing current data to these lines, we can draw conclusions about whether the process variation is consistent (in control) or is unpredictable (out of control, affected by special causes of variation).



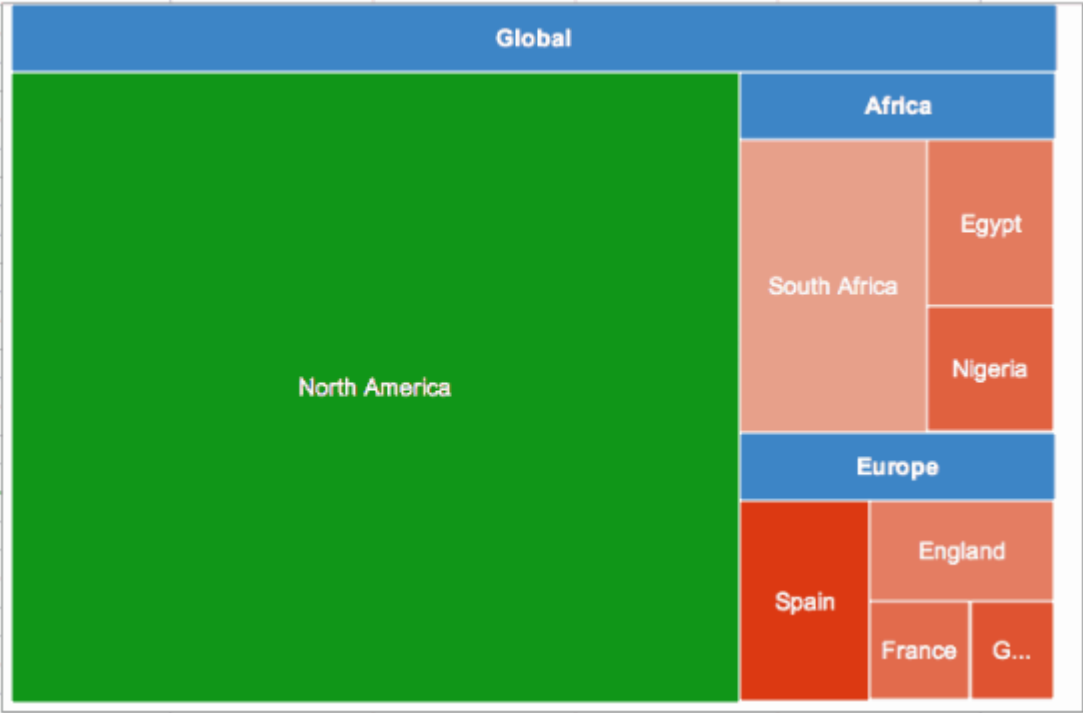
With respect to our project, these control charts will provide us with insights on whether certain movements of SKUs are out of control e.g. too little movements which results in wastage of space in the warehouse. It also simplifies the performance to be easily read by the human eye. Our sponsor emphasizes on sustainability of our solution, which will be further tackled by this in our dashboard.

8.2.2 Time series line chart



A time series is a sequence of data points, typically consisting of successive measurements made over a time interval. With relation to the inbound and outbound rate visualization for different SKUs, the time series line graph will cater to this purpose.

8.2.3 Treemap



A Treemap can convey our hierarchical data with 2 additional attributes via color and size, allowing us to dissect the relationship between the two. The size of each block represents the percentage of the warehouse utilized, and we can allow the user to fill in the colors with other indicators, such as the type of SKUs.

9 Technology

For our dashboard, we will be utilizing certain tools to help with our development of the application.

9.1 D3.js

D3.js is a Javascript library. As our client request for as low cost as possible, D3.js will be a good option. That is because it can work on websites. Thus, our client will be able to see the data visualization without paying and installing any software. Another benefit of D3.js is its flexibility.

It allows control over the final result. D3.js will be used for the visualizations that was mentioned above.

9.2 JMP

JMP is developed by the JMP business unit of SAS Institute. It is the tool of choice for data explorers in every industry. We will be using JMP to perform Control Chart Analysis for the inbound and outbound rate so that we can classify the SKUs into the ABC.

9.3 SAS Enterprise Miner

SAS Enterprise Miner is a software developed by SAS Institute. We may be using SAS Enterprise Miner to perform other needed analysis.

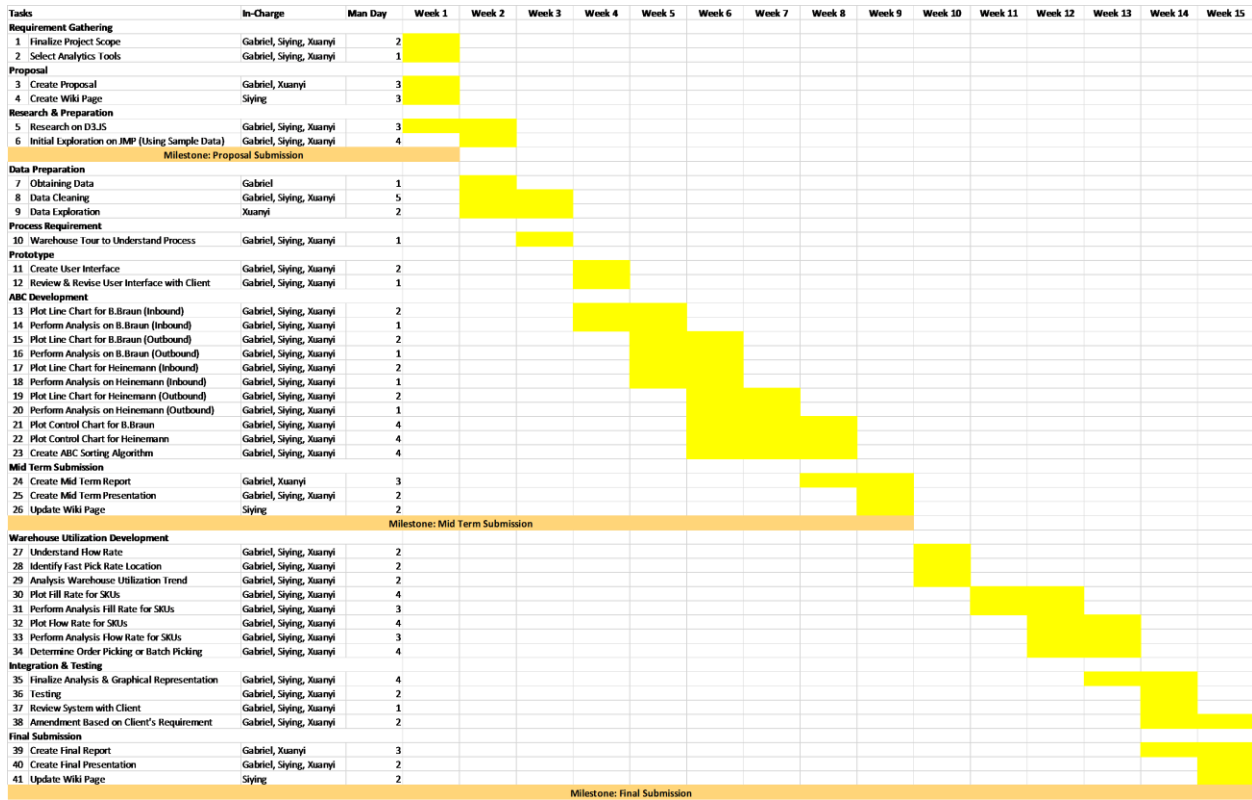
9.4 Java & Bootstrap

Java will be used to develop the skeleton of the application while Bootstrap will be used to beautify the application. As our client wanted a sustainable analysis, we will need to do up a simple application whereby they can upload their data to be processed for analysis.

9.5 OpenShift

OpenShift is a product from Red Hat. It is an open-source platform as a service. It serves as a platform to launch our application for live deployment.

10 Timeline & Schedule



11 Limitations

S/N	Limitations	Assumptions
1	Data provided from the WMS might not be accurately representative of all transactions due to human error (miss scanning of barcode), lost / misplaced SKUs etc.	As the data size provided is large, analysis conducted on it will still be sufficiently accurate.
2	Current data analysis will only be conducted on 2 company products that the logistics company is handling.	The same analysis can be replicated for other company products as they will share some similarities in the nature of warehouse processing.
3	The project scope analyses the nature of warehouse processing of SKUs in Singapore. However, there are also other warehouses of the logistics company which are located	Although the warehouses of the logistics company are located in different geographical locations, they share the same warehouse processing system. This

outside of Singapore. They may have a different approach in warehouse processing.	thus allows the possibility of the team's application to be replicated easily into them, should the logistics company choose to integrate their WMS with the team's project application.
---	--

12 Return of Investments (ROI)

The application provides a platform which will allow executives and managers in the operations department to have a high-level visualization of the happenings within the warehouses. This can be based in terms of unique SKUs, brand and date. Hence, not only will it allow strategic managerial decision to be planned and executed in an accurate and optimal way, it will also allow employees to pinpoint possible bottle necks (e.g. identify SKUs which have slow orders and thus, taking up warehouse spaces which could have been otherwise utilized), and make improvements to it. This will also allow both costs and time savings which will have been spend on manual analysing of data which can be tedious and takes up manpower. Furthermore, in terms of long time sustainability use, the company is also looking into the possibility of integrating the project's application into its in-house warehouse management system. This will serve as an additional add-on feature to further complement and enhance the existing WMS features.

13 Risk Management

13.1 Risk Metrics

Likelihood Impact	Low	Medium	High
High	B	A	A
Medium	C	B	A
Low	C	C	B

13.2 Risk Assessment

Risk Type & Description	Consequences	Likelihood	Impact	Level	Mitigation
<p>Technical Risk</p> <p>Steep learning curve as team are unfamiliar with D3.JS and JMP.</p>	<p>Project may be delayed due to:</p> <ol style="list-style-type: none"> 1. Incorrect time estimation. 2. Longer time needed to familiarize with D3.JS and JMP. 	High	High	A	<ul style="list-style-type: none"> • Start exploration of D3.JS and JMP. • Seek help from seniors who have experience with D3.JS and JMP to expedite learning.
<p>Technical Risk</p> <p>Team members have not taken IS428 - Visual Analytics for Business Intelligence (VA) module.</p>	<p>Team may be unsure of the best practices of information graphics to present data in the language of graphics effectively, clearly and eloquently.</p>	High	Medium	A	<ul style="list-style-type: none"> • Obtain and view past VA projects from its wiki page to have an overview of how data are displayed effectively. • Seek help from peers who have taken VA.
<p>Client Management Risk</p> <p>Client may be making frequent changes to the project scope.</p>	<p>Project may be delayed due to frequent readjustment of project scope and expectation may lead to an over-run in project.</p>	Medium	Medium	B	<ul style="list-style-type: none"> • Team to schedule regular meetings with clients to keep up to date with latest information.
<p>Project Management Risk</p> <p>There are certain features such as warehouse utilization tool, which the team are unsure of the time needed to complete it.</p>	<p>Project may be delayed due to incorrect estimation of time can lead to the feature not being completed on time.</p>	Medium	Medium	B	<ul style="list-style-type: none"> • Team lead to ensure regular meetings to gather feedback from all members on the project's progress. • Team to ask client's technical team on tips to approach such unfamiliar feature.

14 References

- A Guide to Control Charts. (n.d.). Retrieved January 9, 2016, from <http://www.isixsigma.com/tools-templates/control-charts/a-guide-to-control-charts/>
- How to Use a Time Series Graph. (n.d.). Retrieved January 9, 2016, from <http://statistics.about.com/od/Descriptive-Statistics/a/Time-Series-Graphs.htm>
- Overview of Time Series Characteristics. (n.d.). Retrieved January 9, 2016, from <https://onlinecourses.science.psu.edu/stat510/node/47>
- Skau, D. (2013, January 29). Why D3.js is so great for data visualization. *Visually*. Retrieved 9 January, 2016, from <http://blog.visual.ly/why-d3-js-is-so-great-for-data-visualization/>
- Solving data warehouse problems with Cortana Analytics (2015, December 18). Retrieved 9 January, 2016, from <https://www.theta.co.nz/news-blogs/tech-blog/solving-data-warehouse-problems-with-cortana-analytics>
- The Certified Six Sigma Black Belt Handbook Second Edition. (2015, November 14). Retrieved January 9, 2016, from <http://tpowercarpentry.com/the-certified-six-sigma-black-belt-handbook-second-edition.html>
- Treemaps for space-constrained visualization of hierarchies. (n.d.). Retrieved January 9, 2016, from <http://www.cs.umd.edu/hcil/treemap-history/>
- When to use a control chart (n.d). Excerpted from Nancy R. Tague's [The Quality Toolbox](#), Second Edition, ASQ Quality Press, 2005, page 155-158
- Wikipedia*. (2015, August 28). JMP (Statistical Software). Retrieved 9 January, 2016, from [https://en.wikipedia.org/wiki/JMP_\(statistical_software\)](https://en.wikipedia.org/wiki/JMP_(statistical_software))
- Wikipedia*. (2015, December 29). D3.js. Retrieved 9 January, 2016, from <https://en.wikipedia.org/wiki/D3.js>