Data Mining to Predict Operational Outcome

Wen Xu, Celeste, Zhong Hern, Chavi, Hong Bo(advisor)

INTRODUCTION

Collin's Aerospace is an avionics service center situated in Singapore that provides avionics MRO services to commercial and military aviation companies in the Asia-Pacific region.

Problem Statement

OTD is a key metric to measure delivery performance and supply chain efficiency.

			
OTD	**OTD**	**Today**	**Goals**
Critical for retaining customers in the long run	Sales revenue is affected by missed OTD requests	Only 70% of units are meeting their OTD goal	Analyze and Improve OTD performance

METHODOLOGY

01

DATA RELEVANCE

Finding relevant features using the datasets given to establish a relationship between the content and areas of interest

02

PREPARATION

Focusing on the number of units that have missed the OTD and their reasons

03

MODELING

SQL was mainly used for analyzing product OTD, while R was used for modelling the relationship between employee clocked hours and capacity.

04

DEPLOYMENT

Implementing model into excel for ease of use, to make predictions and evaluate it with real data

Cleaned Data

Fiscal Wee	Fiscal year/pe	Product Type	Equip Type	Reason1	Reason2	CheckOTD for ABC	CheckOTD for non-ABC
14/2019	2019	A	FXX-XXXX	Not assigned	Not assigned	21	
14/2019	2019	A	FXX-XXXX	Not assigned	Not assigned	13	
14/2019	2019	B	EXX-XXXX	Not assigned	Not assigned	62	1
14/2019	2019	B	EXX-XXXX	Not assigned	Not assigned	35	1
14/2019	2019	B	DXX-XXXX	Not assigned	Not assigned	28	1

```
select Year, Section, Type, Equipment, Material,
case when Reason1 = 'Awaiting Parts' then 'Parts' when Reason1 = 'In Transit'
then 'Parts' when Reason1 = 'Engineering Eval' then 'Engineering' when Reason1 = 'Not assigned'
then 'NA' when Reason1 = 'Awaiting Subcontractor'then 'Subcon' else 'Others'
end as Reason, OTD, Count from Q3GTAT where OTD = 0
```

The cleaned data above was taken from the massive data files given to us. We required the reasons for missed OTD and the equipment type. The code above was part of the data cleaning process and produced the required clean data as seen.

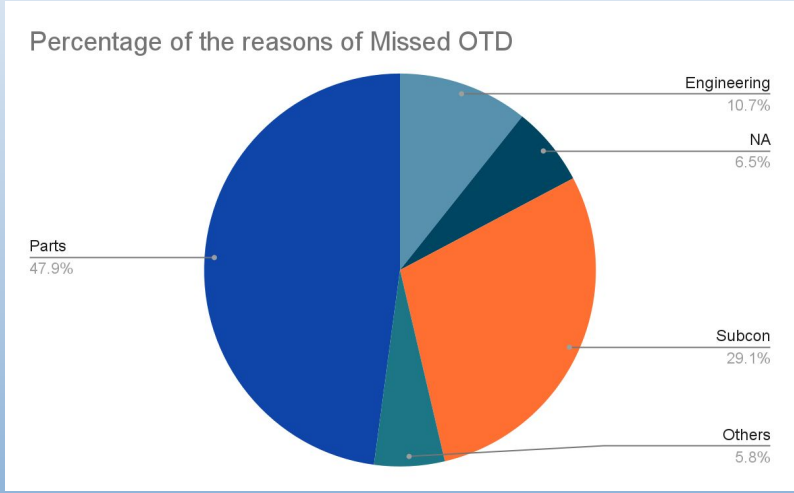
Cleaned Data

	Employee	PROD HR	LEAVE HR	OVERTIME HR	TOTAL HR	AVAIL HR	SHIFT NORMAL HR	PDTY W/ OT	PDTY W/O OT
ATE	Name								
32XXXX	Chan XXX	70	4		1XX	1XX	1XX	XX	XX

```
Select Station_name, AVG(Prod_hr) AS FEB_AVG
From FEB2019
Group by Station_name
```

From the original data shown in the first picture, we narrowed the employee data down to permanent staff and used productivity hours to create a trend line by station. Through the data, we also made correlation to OTD requests that were met to determine the capacity per station. This data also showed us the seasonality in requests.

Main Results



With our focus on Product ABC, we identified the main reasons for the Product ABC not meeting OTD which were these five factors.

Parts contributed most to units missing the OTD making it a major significant factor in the company not meeting the OTD target.

Main Results

Type	Total no. of units that missed OTD	Missed OTD due to missing parts
RXX	90(15%)	60
FXX	39(7%)	25
WXX	27(5%)	13

The **Pareto Analysis**, which is a decision-making strategy that is based on the "80-20 Rule" states that, for many events, **roughly 80% of the effects come from 20% of the causes in major parts.** The table was shows top 3 equipment types that contributed most to the total units that missed OTD and the number of units of those equipment types that missed OTD due to missing parts.

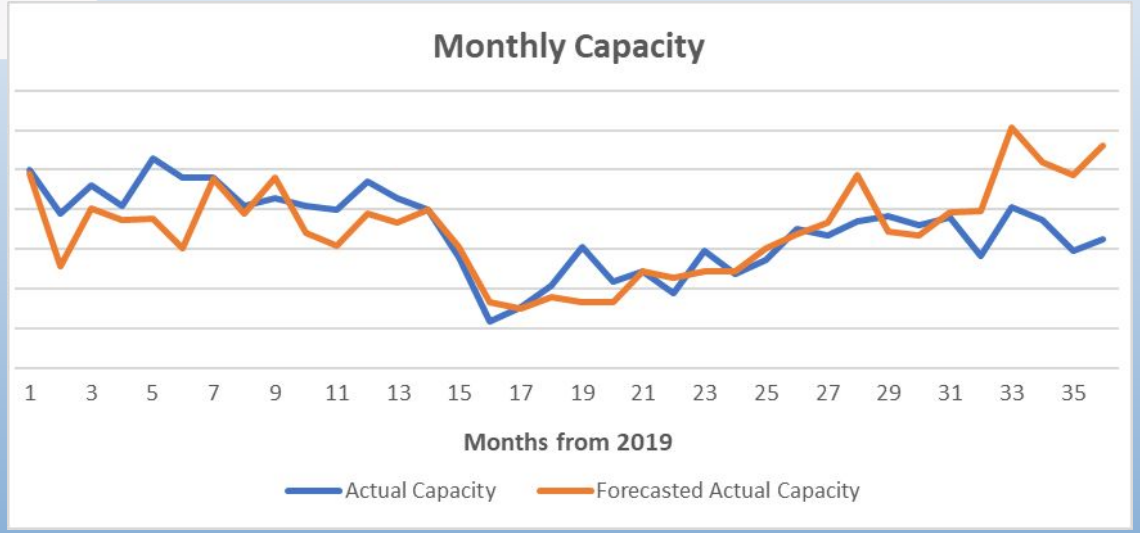
Main Results

Count of Material Description	Column Labels				Inventory Qty	Threshold Value
Row Labels	2020	2021	2022	Grand Total		
FXX	**15**	**24**	**7**	**46**		
xxx-xxxx-xxx	2	1	4		0	1
xxx-xxxx-xxx	7	12	2		11	3
xxx-xxxx-xxx	6	11	1		13	3
RXX	**56**	**52**	**13**	**121**		
xxx-xxxx-xxx	2	1	1		0	1
xxx-xxxx-xxx	2	1	2		6	1
xxx-xxxx-xxx	3	1	1		0	1
xxx-xxxx-xxx	2	1	3		0	1
xxx-xxxx-xxx	1	1	2		0	1
xxx-xxxx-xxx	25	30	1		0	7
xxx-xxxx-xxx	6	7	1		11	2
xxx-xxxx-xxx	10	8	1		7	3
xxx-xxxx-xxx	5	2	1		9	1
WXX	**97**	**105**	**27**	**229**		
xxx-xxxx-xxx	30	35	9		2	9
xxx-xxxx-xxx	20	39	4		5	8
xxx-xxxx-xxx	14	13	4		3	4
xxx-xxxx-xxx	1	7	1		0	1
xxx-xxxx-xxx	10	7	3		6	3
xxx-xxxx-xxx	14	1	2		4	2
xxx-xxxx-xxx	1	1	1		5	1
xxx-xxxx-xxx	7	2	3		1	2

To derive the threshold value, demands from historical data were divided to get the threshold value.

If the inventory quantity was lower than the calculated threshold value, the cell would be highlighted in red. This system allows the staff to conveniently note down the materials that can lead to a shortage issue and missing OTD.

Main Results



By predicting capacity, it gives the company an insight into the future and to better prioritize their array of requests to meet OTD.

Furthermore, in times when they can afford to entertain extra requests, the predictions will enable them to determine the quantity of these requests they can take without affecting day to day operations.

Conclusion

Recommendation

Ensure staff are aware of their inventory at all times

Limitations

Not able to fully resolve the problem due to reasons outside of the company's control

Outcomes

MRP Program
Predicted Capacity