Final Report on Panasonic Kitting Capstone Project

Kefei Wu, Jong Hwi Park, Yingping Lu, Miao Zhou

Goergen Institute of Data Science
University of Rochester
{kwu19, jpark177, ylu63, mzhou31}@ur.rochester.edu

Introduction/Motivation

Due to the advancement of wireless system and IoT (internet of things) devices, various industries have started to incorporate IoT solutions in their manufacturing process. Notably, Rochester based IoT solutions company, ProVIEW, has been growing its influence in various industries, especially in providing material flow and tracking applications in large capital goods manufacturing enterprises. It was recently acquired by Panasonic, and trying to expand its customer list. In order to perfect the product, Panasonic strives to build an analysis pipeline for internal analysis and for customers. To do so, Panasonic generously provided their ProVIEW database gathered from Michigan General Motors Lansing Factory during product kitting process. For our capstone project, we analyzed the database by establishing Key Performance Indicators (KPIs) through creating interactive dashboards with Tableau and employing predictive analysis methods to optimize kitting efficiency and to minimize kitting errors.

Key words: ProVIEW, Kitting, Key Performance Indicators, Predictive Analysis

Data Set Description

A. Data Acquisition

When the job orders are requested by the customers through wireless connection, the process will be sent to the workers and gets logged on database. Once the worker has received the jobs, the process starts. The picking steps are that pick list is downloaded from the system to a pick book, the operator begins the job on the pick book and locates an available cart which is equipped with a view tag and automatically associates with the pick job. When the operator picks the part corresponds to a pick job, he or she pushes the button on tag to indicate the pick is completed and verifies that the job is completed or failed on the pick book. All the records will then be uploaded into the system, which is our database.

B. Data Accessing

At the very beginning, we have to access the backup files we received. Due to the limitation of operating-system of our devices, we cannot operate on the backup files directly. Thus, we utilize a virtual machine or container to make our device a suitable environment to use softwares which can open the backup files. In this project we choose Docker to build a container instead of VM since Docker is more lightweight than VM. After setting up the environment, we use Microsoft Azure to open the backup files.

C. Data Description

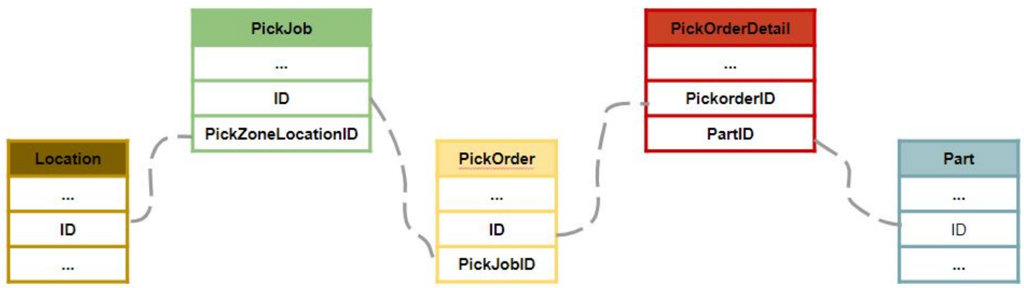


Figure 1. Overview of data tables

We are given five tables which are *Location, PickJob, PickOrder, PickOrderDetail* and *Part*. Each job may contain multiple orders and every order may have many parts included. Thus, the feature ID in the *PickJob* table could be used to identify each job and be a foreign key to link *PickJob* with *PickOrder* table. Each order also has its ID, which is a feature we use to join *PickOrder* table with *PickOrderDetail* table. In the *Part* table, we are provided with a feature named ID which is the identification of a single part, and this primary key is also a foreign key in the *PickOrderDetial* table. Lastly, the *Location* table has all the information for each pick zone where pick job takes place.

D. Data Cleaning

General Motors Lansing Factory updated their warehouse's data at the 15th of every month. Every time they update their data, a backup file will be generated. Since our data is from January to July, we have 7 files in total. Those 7 files have some duplicated data, since the job that was not completed in previous month but finished in current month will be recorded in both backup files. Also for those jobs that were failed in previous month but redone in this month will be recorded in two backup files too. In order to join all the backup files together without duplicates, we only keep the latest updated jobs. There are some dates and time data in our database, and our data are based on GMT time zone. We changed the time zone to be EST time zone and removed the daylight saving setting.

Exploratory Data Analysis

To develop intuition regarding the distribution of the data, we carried out our preliminary analysis of the raw dataset. First, an exploratory analysis was performed by examining tables and the attribute information in dataset. With the hierarchical structure defined in the previous section, the job information is organized into 5 tables as shown below, along with the dimensionality of each table accordingly:

Table 1. Table information

Table	Number of Records	Number of Attributes
Pick Job	105762	15
Pick Order	1859214	5
Pick Order Detail	3893469	7
Part	2322	7
Location	409	9

Table 2. Job status

State	Error Code	2019
Ready	Null	90
Active	Null	1
Picked	Null	7
Failed	Other	6,731
	General error	2,160
	Broken tag	1,322
	Missing parts	73
Completed	Null	95,310
	General error	37
	Other	31

To better serve our analysis in the next step, we also examined the rationality and outliers of our data that could have a potential influence on further investigation. For example, in table *pick job*, the status of each job is defined as either ready (order generated), active (picking started), picked (last part picked), failed (job failed) or completed (job succeeded) and indicated in an attribute named *State.* Since the data were collected from January to July, 2019 in the past, the status of each job should be either failed or completed by now. However, by summarizing the status information in Table 2, we found that there's still a small portion of the jobs that remains in incomplete status (i.e., ready, active or picked). To take a closer look, we explored the distribution of those active jobs with respect to location and time in days (Fig. 2).

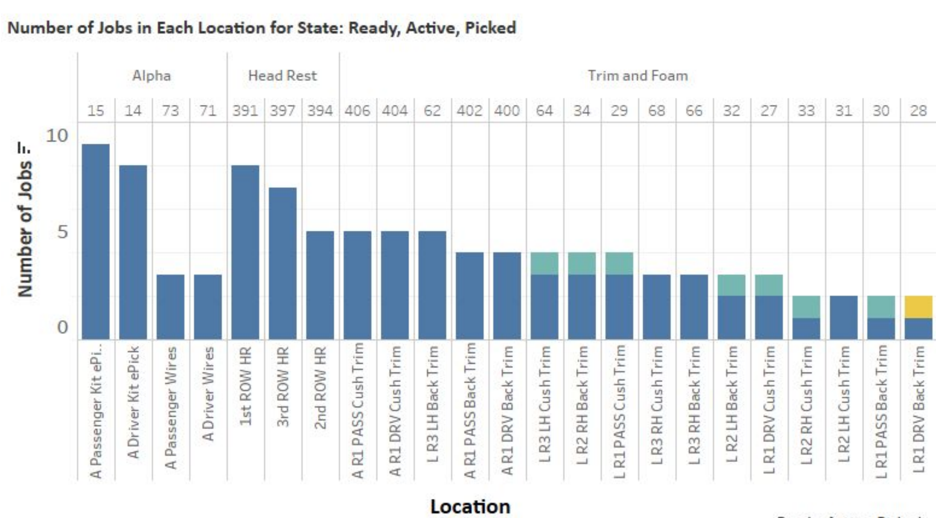


Figure 2(a). Distribution of unfinished jobs based on location.

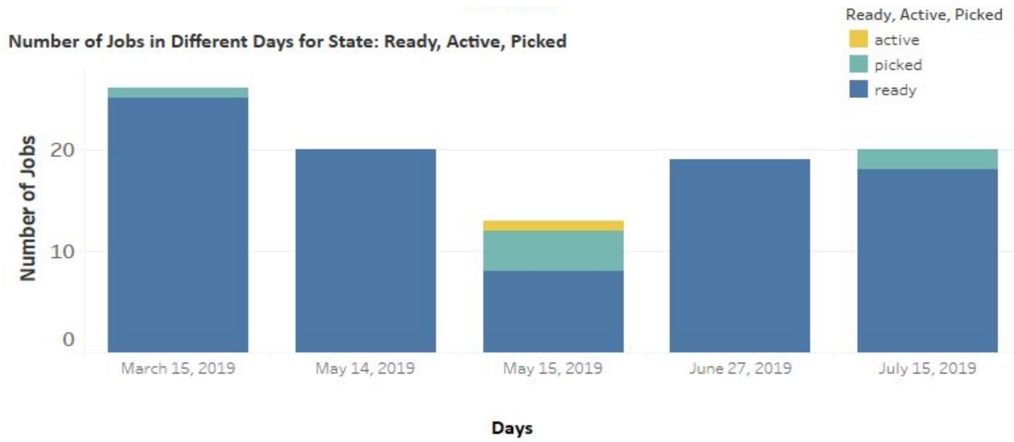


Figure 2(b). Distribution of unfinished jobs based on day.

The distribution implies that the abnormal records occurred only in 5 days. In particular, all jobs in active status occurred restrictively in one location on a specific day. As confirmed with our sponsor, we assume the problem listed above as a technical issue and remove those records from further analysis.

Furthermore, based on the structural design of the database, all failed jobs should be recorded as failed associated with an explicit error code as either a general error, broken tag, missing parts or other, while the completed jobs should be successfully completed without any error. However, we found that there is also a small number of completed jobs that has an error code associated with it without an actual picked date. In other words, it indicates that those jobs were actually failed but mistakenly categorized as completed jobs for some reason. Thus, we examined the issue by plotting the distribution of the abnormal records with respect to zone location and then grouped them by error code and end date, respectively (Fig. 3).

Figure 3(a). Distribution of completed jobs with error code. Number of Completed Jobs without Picked Date according to Error Code andPick Zone.

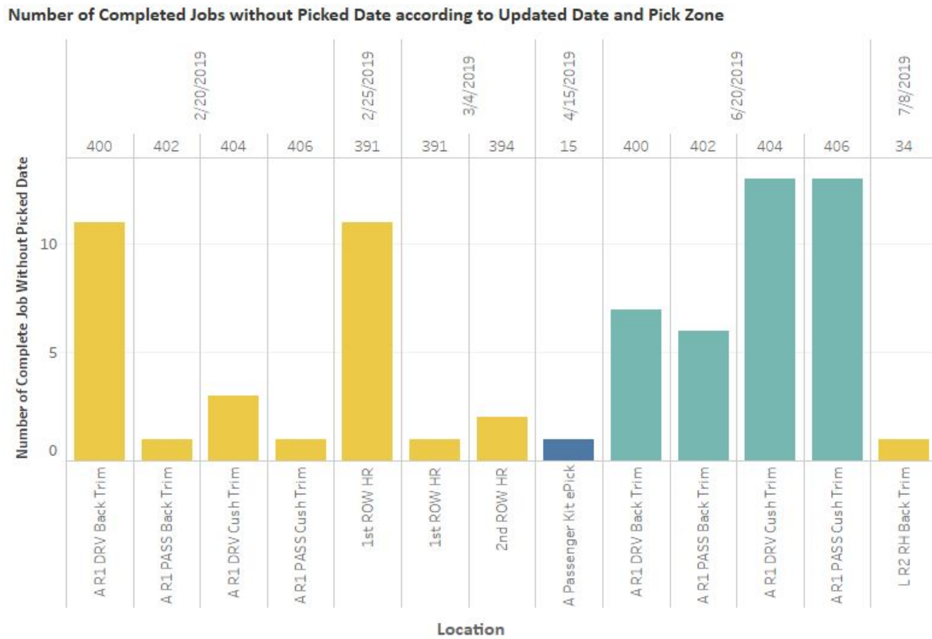


Figure 3(b). Distribution of completed jobs with error code. Number of Completed Jobs without Picked Date according to Updated Date and Pick Zone.

We see that most of the completed jobs with error codes occurred in a restricted number of locations (top). At the same time, there are over 20 failed jobs that were later on overwritten by Brooks, Zavier on February 20th, 25th, March 4th and July 8th as completed jobs, and all 39 failed jobs were overwritten by admin on June 20th. Especially, the plot shows that there's one record that was modified by a picker who shall not be authorized to do so (bottom).

By conducting a preliminary analysis, we were able to explore the dimensionality and detailed information of our data. We were also able to identify and handle some of the potential irrationality and issue in our data, as well as the need to further prepare the training dataset, rebalancing it for modularization.

Methods

Several key performance indicators were created to help improve General Motors Lansing Factory's efficiency and lower the picking error. For the purpose of this final report, selected(**bold**) KPIs are introduced in this report.

1. **Number of completed jobs per time interval (per day, week, month, etc)**
2. Number of failed jobs per time interval (per day, week, month, etc)
3. Number of components picked (Total per day per zone)?

4. Pick job completion time and reaction time per time interval
5. Compare number of completed picks and number of failed picks between zones
6. **Compare pick times between zones**
7. Average pick time per time interval (Pick Rate)
8. Most common parts picked – Part based
9. Most frequently picked parts – Time based
10. **Busiest pick zone / Slowest pick zone**
11. Job-In-Queue time
12. Number of jobs and time per shift

A. Key Performance Indicators (KPI)

KPI 1: Number of Completed Jobs per Time Interval

From Fig. 4(a), there are some days without completed jobs and also some days with low completed jobs. After checking those days without completed jobs, almost all of those days belongs to holidays or weekends. 28 days with low completed jobs. 24 days are from weekend, and 4 of them are from holidays, such as Martin Luther King Day, Easter Day, Memorial Day, and Independent Day.

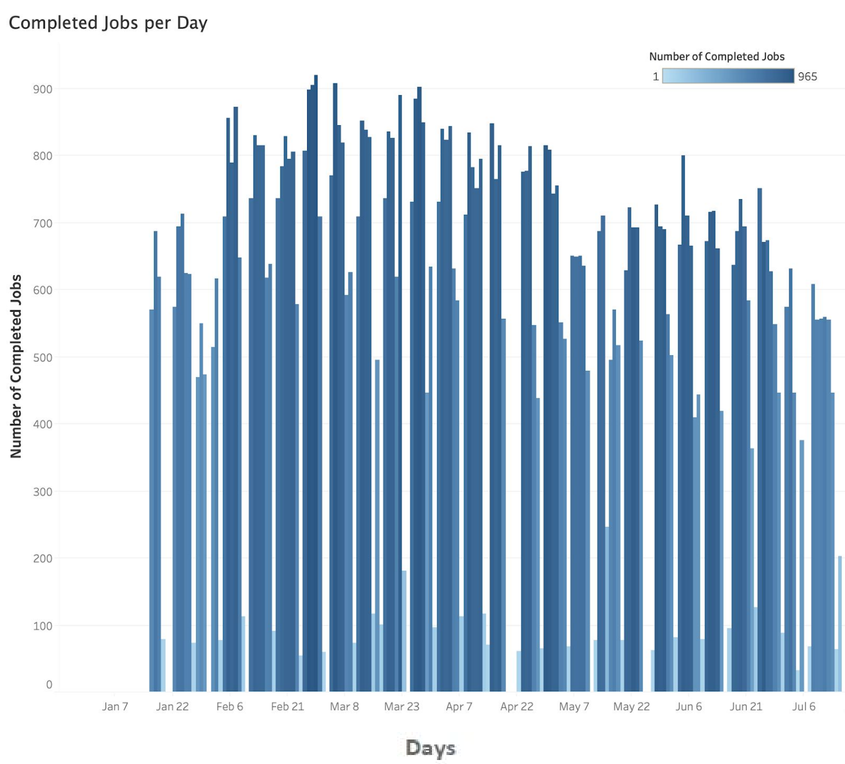


Figure 4(a): Number of Completed Jobs per Day. The graph shows the data from January to July. Lighter blue represents less completed jobs in that day, and darker blue represents more completed jobs in one day.

Fig. 4(b) shows the number of completed jobs per weekday. Sunday has really low completed jobs, and Saturday has the second least completed jobs. We can analyze the trend of the number of completed jobs within one week. At the beginning of each week - Sunday - the number of completed jobs are really low. However, on Monday, workers started to work a lot more than on Sunday. Tuesday achieve the peak of number of completed jobs. Starting from Wednesday, the number of completed jobs keep decreasing, since it's the end of the week.

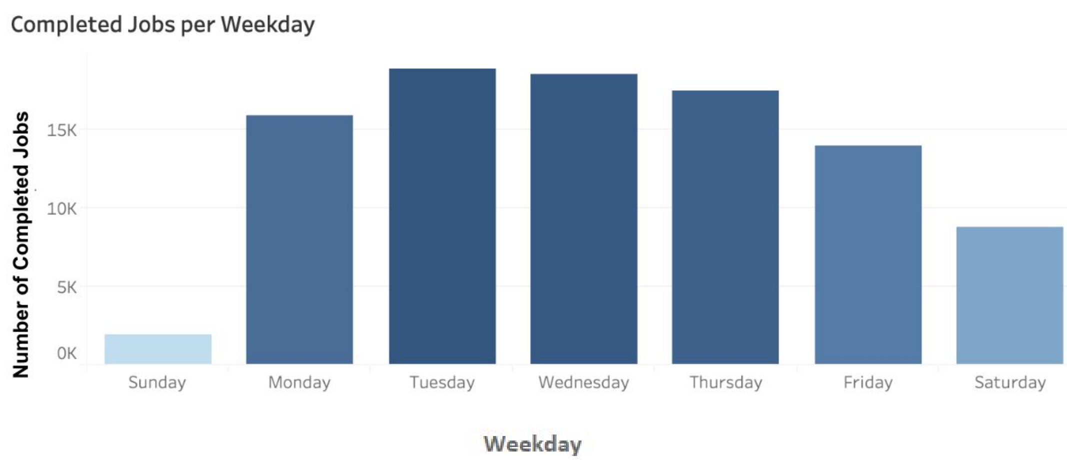
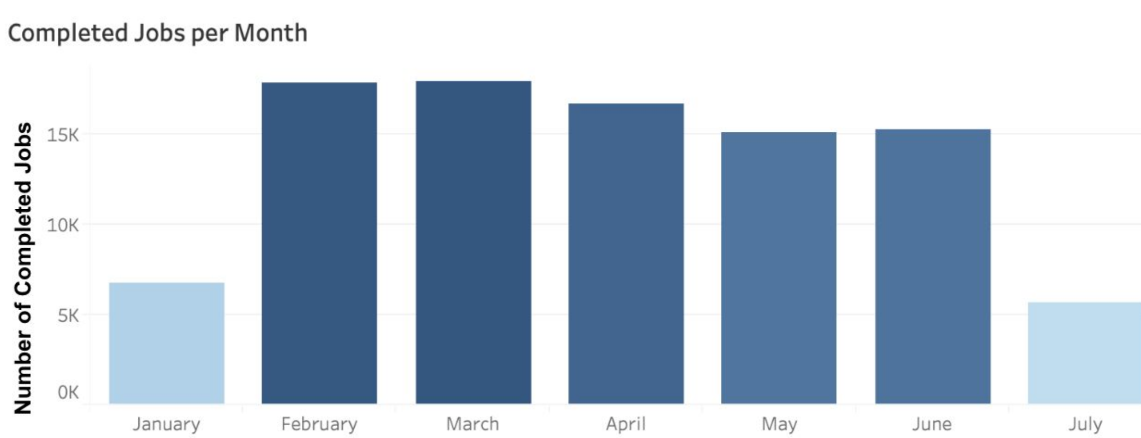


Figure 4(b): Number of Completed Jobs per Weekday. The graph shows the data from January to July. Lighter blue represents less completed jobs in that day, and darker blue represents more completed jobs in one day.

Fig. 4(c) shows the number of completed jobs per month. According to the graph, January and July have the lowest completed jobs. The first two week of January only had failed jobs, since the factory is testing their system. The reason that July has less complete jobs is that our data is till July 15th. So this graph only contains data from half month of July.



Month

Figure 4(c): Number of Completed Jobs per Month. The graph shows the data from January to July. Lighter blue represents less completed jobs in that day, and darker blue represents more completed jobs in one day.

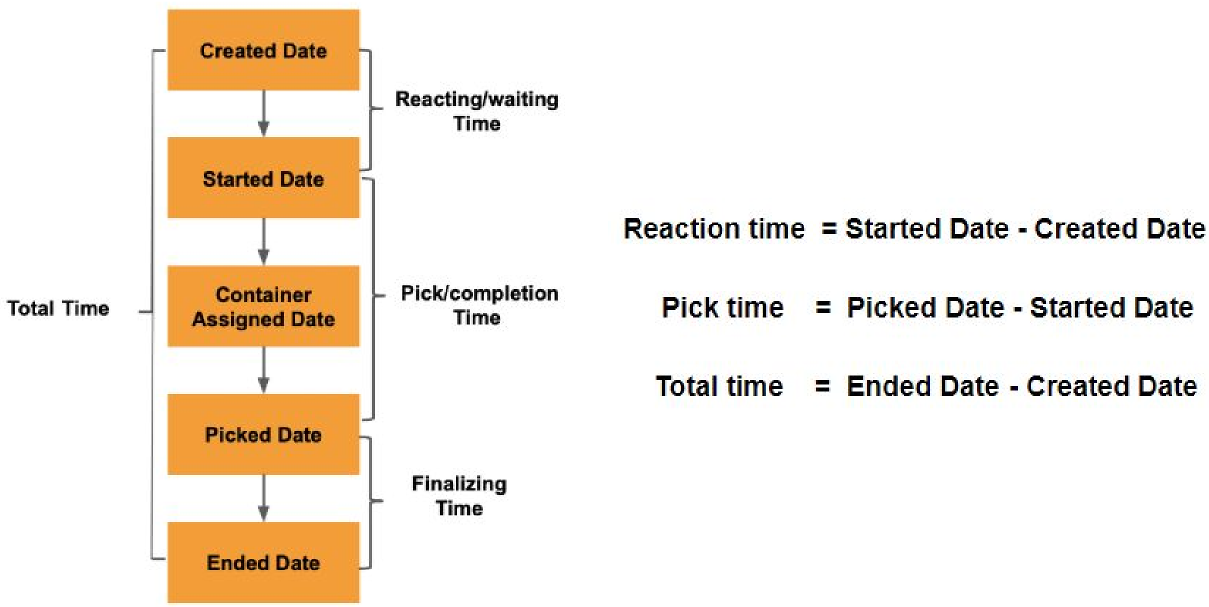


Figure 5: Time definition flow chart of reaction time, Pick(completion) time, finalizing time, and total time (Left). For better understanding, equations to compute the defined times are provided (Right).

KPI 6. Compare pick times between zones

In KPI 6, we compare pick times between zones, and scrutinize deeper based on newly established time definitions. Reaction time, difference between start date and created date, shows how much time it took for a worker to reacted to a job once a job is assigned. Pick (completion) time refers to actual time of kitting of a job. Moreover, finalizing time has information of a worker moving to an end of a zone and push a finished button after picking, and Total time is an entire duration of the process. With these definitions, we created new columns in each job with corresponding values, and utilized them for further analysis.

Since duration of time can vary based on each location or zone, we analyzed based on zones and visualized through tableau dashboard as shown in Fig. 6. Here, the figure shows average pick time / average reaction time with completion/failure rate. Especially, zone 14 and 15 take long time to react and pick, yet they have high completion rate compare to other zones. In fact, our sponsors confirmed that zone 14 and 15 have the most items to be picked and they are located far away from other zones, which explain relatively high reaction time and pick time.

With this analysis pipeline and visualization, customers can obtain a clear understanding in zone level.

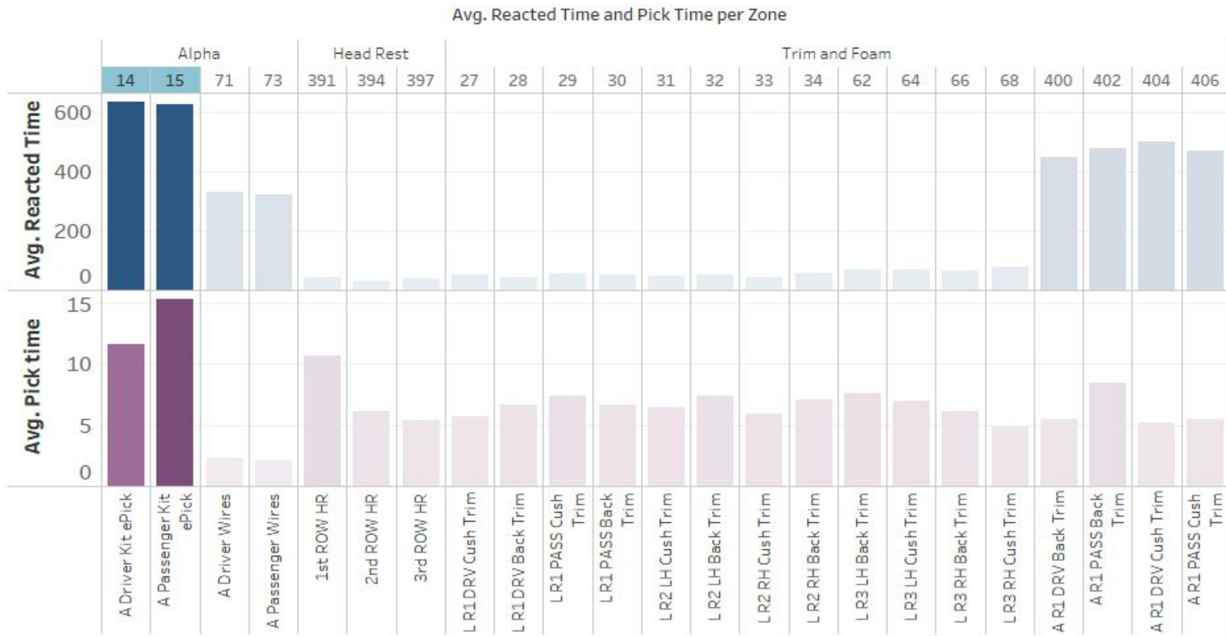


Figure 6(a). Visualization of average reaction time and average pick time of zone 14 and 15.

Failed Rate for Different Locations

Pick ..	Name	completed	failed
14	A Driver Kit ePick	98.57%	1.43%
15	A Passenger Kit ePick	97.05%	2.95%

Figure 6(b). Failed rate for jobs of zone 14 and 15.

KPI 10. Busiest pick zone / Slowest pick zone

In KPI 10, we evaluated the performance of the kitting process regarding the efficiency based on location. Specifically, the measurement is divided into two parts, in which we ranked the busiest (part A) and slowest zones (part B) with respect to the quantity of parts and time spent on jobs. In part A, we defined the busiest zone as the location with the most number of requested parts. Due to the variation in the total quantity of parts needed for each job, we evaluated part A on a part level for the sake of unbiasedness.

On the left (Fig. 7a), we see that the top 3 busiest zones are consistent throughout seven months, which are location 14, 15 and 62. We then cross referenced the number of requests with the pick rate of each location (Fig. 7b), which is defined as the proportion of the number of parts that were picked against the total number of parts that were initially requested by customers. Ideally, we want the pick rate to be as high as possible despite of the requested quantity. We see that location 14 and 15 have exceptionally high pick rates of over 95% given their busyness, while the pick rate of location 62 is relatively low with less than 85%.

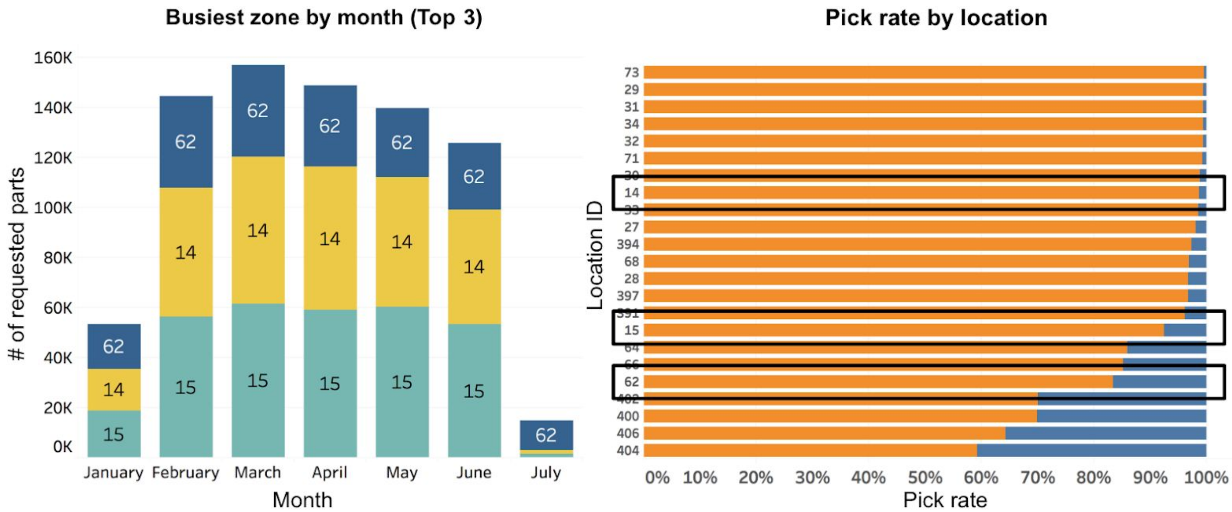


Figure 7. Busiest zones. a (left): Number of requested part in each month per location showing only the top 3 locations. The colors represent different locations associated with the location id explicitly written on each stacked bar. b (right): Pick rate by location. The picked and non-picked parts are displayed in orange and blue, respectively. The top 3 busiest locations are marked with black boxes.

Next, for part B, we define the slowest zone as the one with the longest average completion time, which is then calculated by the total completion time divided by the total number of parts picked in each location.

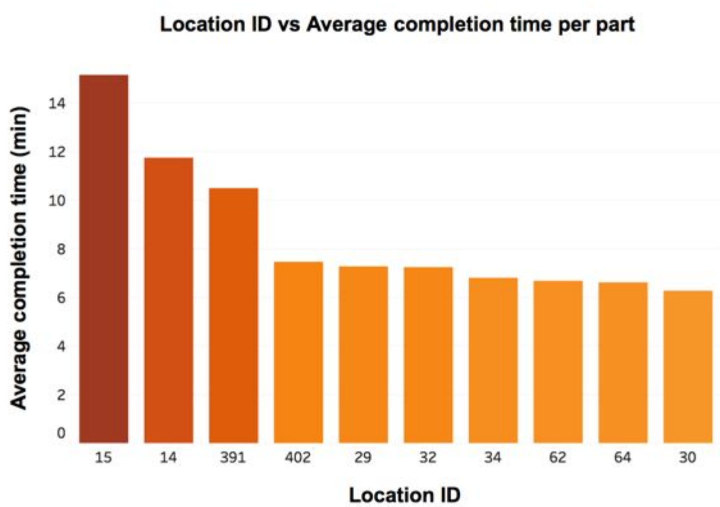


Figure 8. Average completion time for each part of each location. Showing only the top 10 locations.

Taken the number of parts requested into consideration, location 14 and 15 still have the longest average completion time after normalization. Especially, location 15 has an average pick time as twice as the of the majority of the rest zones, while location 62, which also appears to be one of the top busiest zones, has a relatively low pick time compared with the other two.

B. Predictive Analysis

Pipeline

After establishing Key Performance Indicators, we performed additional predictive analysis for two tasks: 1) predict failure or completion of a pickjob, 2) predict slow or fast of a completed pickjob. However, we focused on analyzing key features that drive model to make a binary classification to pinpoint which factor has the highest influence on prediction. Thus, 'white-box' models such as Random Forest, Logistic Regression, and XGBoost that have the capacity to show which features are considered to make a classification were chosen.

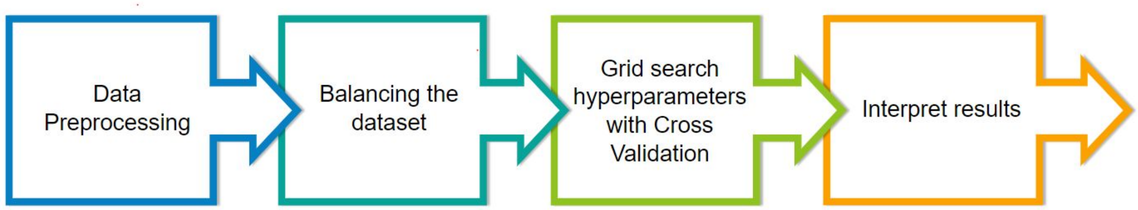


Figure 9: Procedure for predictive analysis. For both tasks, data preprocessing, data balancing, grid search with cross-validation, and interpretation were done in order.

To begin with predictive analysis, the following procedures were followed (Fig. 9): First, data was preprocessed through standardizing features with continuous values with z-score standardization. Other features like zones, locations, and shifts were kept as is since they are one-hot encoded. Also, we put completed job as label 0 and failed as lab 1 for easier interpretation. For slow and fast completion of a job, we discarded failed jobs and labeled jobs as 1 (slow) whose completion time was above mean + 1 standard deviation and rest as 0 (fast).

Then, data was split into training and testing set with 80/20 ratio and balanced through USWOR (undersampling without replacement). Since the dataset was heavily imbalanced for both tasks, undersampling was employed on training dataset to adjust model's tendency in having bias toward a dominant label. Yet, the testing data remained untouched to reflect distribution of the original dataset. After processing the data, models were employed with grid search on hyperparameters paired with 5-fold cross-validation. By doing so, we were able to find the best hyperparameters that yields the best performance with robustness. Finally, the results were analyzed via comparing models with two evaluation metrics, precision and recall, and scrutinizing feature importance of the models.

Table 3: Result of predictive analysis on binary classification of failure vs completion of a pickjob.

	Random Forest	**Logistic Regression**	**XGBoost**
Precision	0.19	**0.24**	**0.24**
Recall	***0.77**	0.69	0.69
Accuracy	0.63	**0.74**	**0.74**
Hyperparameters	N_estimatros: 150 Criterion: entropy Max_depth: 4	C: 0.5 Max_iter: 500	Gamma: 1.5 Subsample: 6 Max_depth: 5

Result Interpretation (Performance and Result)

1) Predict failure or completion of a pickjob

The result shows that all three models have moderate to high recall and accuracy, ranging from .63 to .77. Yet, since the data is imbalanced, we sought for a model with the highest recall value, and random forest had the highest value with 0.77 (Table 3). On the other hand, precision was extremely low for all three models, suggesting that there were lots of false positives (completed jobs are classified as failed).

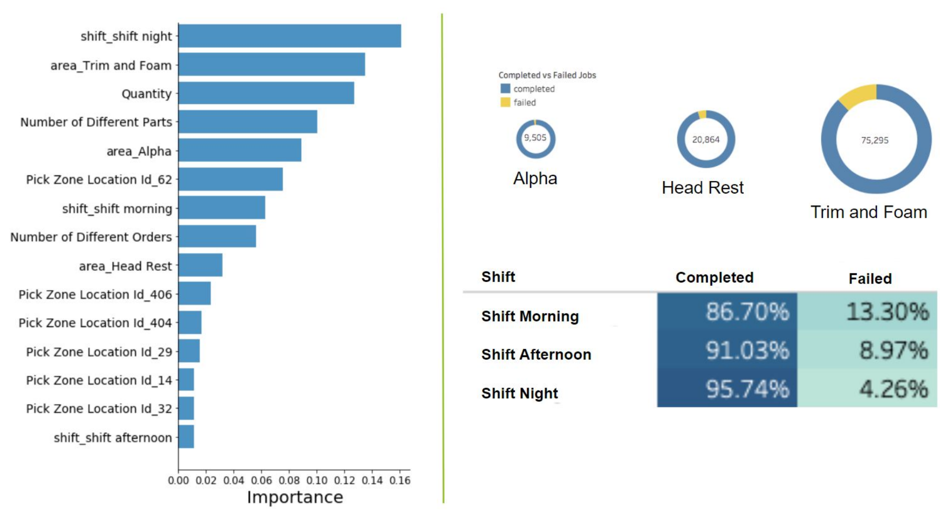


Figure 10: Feature importance analysis of Random Forest model on classifying failed/completed jobs. The result agrees with our KPI analysis in that shift information and area information have critical information in deciding completion and failure of a job.

In order to inspect results deeply, we analyzed feature importance of random forest and coefficients of logistic regression. Feature importance ranking of our Random Forest model shows that shift night, time and foam, quantity, number of different parts, etc are important in descending order. (Fig. 10) To understand this result, we compared the first two important features with previous KPIs' results. Interestingly, shift night had the highest completion rate compare to other shifts, and trim and foam area had the highest failure rate in comparison to other areas. Thus, our modelling results confirms that unbiased modelling agrees with our KPIs, suggesting which feature has to improve its accuracy and efficiency in completion/failure rate.

Description (..	Pick ..	Name	completed	failed
Alpha	14	A Driver Kit ePick	98.57%	1.43%
	15	A Passenger Kit ePick	97.05%	2.95%
	71	A Driver Wires	99.28%	0.72%
	73	A Passenger Wires	99.28%	0.72%
Head Rest	391	1st ROW HR	95.26%	4.74%
	394	2nd ROW HR	95.19%	4.81%
	397	3rd ROW HR	95.69%	4.31%
Trim and Foam	27	L R1 DRV Cush Trim	92.39%	7.61%
	28	L R1 DRV Back Trim	84.90%	15.10%
	29	L R1 PASS Cush Trim	97.05%	2.95%
	30	L R1 PASS Back Trim	92.90%	7.10%
	31	L R2 LH Cush Trim	95.74%	4.26%
	32	L R2 LH Back Trim	96.73%	3.27%
	33	L R2 RH Cush Trim	89.93%	10.07%
	34	L R2 RH Back Trim	95.72%	4.28%
	62	L R3 LH Back Trim	73.10%	26.90%
	64	L R3 LH Cush Trim	81.46%	18.54%
	66	L R3 RH Back Trim	80.60%	19.40%
	68	L R3 RH Cush Trim	92.26%	7.74%
	400	A R1 DRV Back Trim	81.13%	18.87%
	402	A R1 PASS Back Trim	79.86%	20.14%
	404	A R1 DRV Cush Trim	69.50%	30.50%
	406	A R1 PASS Cush Trim	69.36%	30.64%

Figure 11: Feature importance analysis of Logistic Regression model on classifying failed/completed jobs.

Moreover, analysis on coefficients of logistic regression yielded interesting information (Fig. 11). Our logistic regression model considers zone 15, 42, 400, area trim and foam, and zone 406 as top five features to classify a job as failed. Yet, our KPI shows that zone 15 has the highest completion rate, contradicting the logistic regression model and suggesting the reason of numerous false positives. Still, other positive coefficients and negative coefficients agrees to our KPI results that zone 402, 400, 406 and area 406 have high failure rate and zone 31, 34, 32, 71, 73 have high completion rate. Most importantly, having less number of non tracked parts guided logistic regression to label a job to be completed, suggesting that reducing the number of non tracked parts can hugely benefit the kitting completion rate.

Table 4: Result of predictive analysis on binary classification of slow vs fast completion of a pickjob.

	Random Forest	**Logistic Regression**	**XGBoost**
Precision	0.12	0.12	0.12
Recall	***0.90**	0.73	0.89
Accuracy	0.33	**0.49**	0.36
Hyperparameters	N_estimatros: 1000 Criterion: entropy Max_depth: 6	C: 1 Max_iter: 500	Gamma: 5 Subsample: 1 Max_depth: 3

2) Predict slow or fast completion of a pickjob

For further analysis, we performed the same pipeline for classifying slow and fast completion of a job. Employing the same pipeline was reasonable since the data was also imbalanced with binary labels. However, the result was not as promising and interpretable as the first task (Table 4). While recall was higher with Random Forest and XGBoost, precision and accuracy was extremely low that further analysis was not needed. In fact, Accuracy was even lower than random guess (0.5), suggesting that the data is missing critical information for classifying slow and fast completion of a job. We strongly believe that adding locational features such as coordinates and distance between each zone will benefit the algorithms with higher precision and accuracy, since above mentioned information are directly correlated with the amount of time needed to complete a task.

Conclusion and Next Step

Throughout the 10 weeks of capstone project, we successfully finished 11 given KPIs and elucidated four more KPIs. With our established KPIs and Tableau dashboards that we generated, we believe that our customers can easily notice significant days, locations, or even users during certain time period. We strongly believe that our product let customers to access and visualize their ProVIEW data easily and cleanly with rich insights, such that these KPIs will certainly help users to improve their business management. Moreover, we went further than given goals and employed predictive analysis to model failure/completion of a job and fast/slow completion of a job. Our model on failure/completion job confirms our findings from previous KPIs that several zones and areas require better management and necessity to reduce non-tracked parts. Yet, our other model on slow/fast completed job did not produce promising results with low precision and accuracy. We believe that this is due to missing important features that are correlated with completion time. Thus, in the future, we will obtain data from other months with more attributes may yield better insights not only in predictive analysis, but also in making more robust KPIs. Also, we plan to collaborate with replenishment team to combine replenishment data and insights to make holistic analysis, with final goal to build an easy to interpret and interactive visualization tool or program. As a final remark, we would like to thank our sponsors Linh Nguyen, Ravi Nareppa, Cory Albert from Panasonic, and Prof. Ajay Anand, Prof. PJ Fernandez, and Yuexi Wang from University of Rochester for providing great insights and guidance for this project.