Team 6 Final Report

**Caroline Baer, Steven Luong, Anjaneya Malpani, Dipali Mistri, Hewitt Trinh**

## Executive Summary

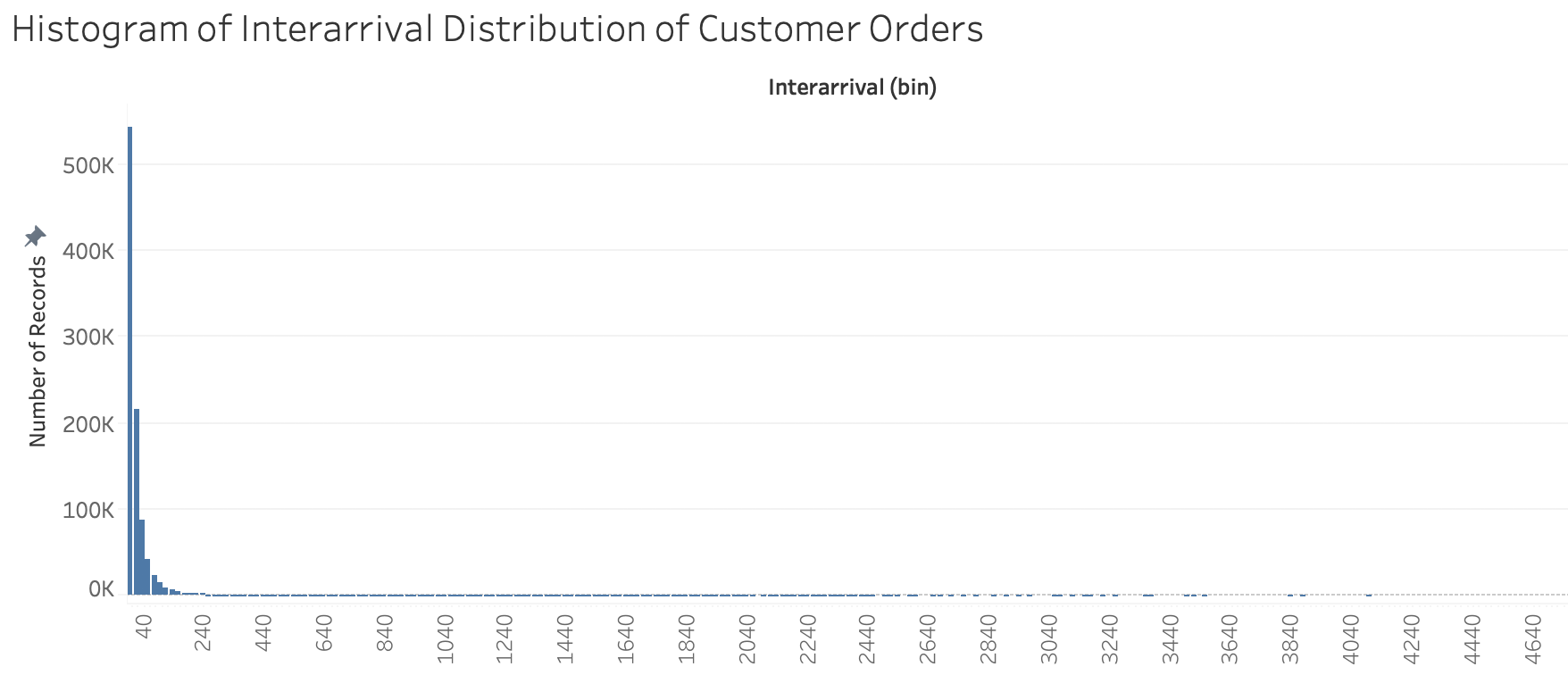
Using simulation modeling to measure and improve the performance of a system has been the central idea of applying prescriptive analytics in the stochastic world. What are the best policies to operate and manage the supply chain for an e-Commerce fulfillment center? The main objective of this report is to explain our approach in analyzing the operational problems and attempting to simulate them to arrive at the optimal policies. We compartmentalized the whole operation into two individual subsystems: inbound processing and outbound processing. Given that the two subsystems communicate with each other through the state of inventory level, separating the two processes allowed us to experiment different policies and details of the outbound process without needing to change the details or sequences occurring in the inbound process. We realized that our optimal policies are only as good as our simulation and the simulation would not be valid without accurate modeling of customer order data. Thus we concluded that while simulation results are critical for us to understand the performance of different components in the system, we should look at the final key performance metrics in the context of the as-is data generating process.

## Modeling the Order Data

In order to properly model the incoming customer orders, two main arrival rates needed to be recreated: arrival rate for orders and arrival rate for product quantities (per order).

### ***Order Arrival Rate***

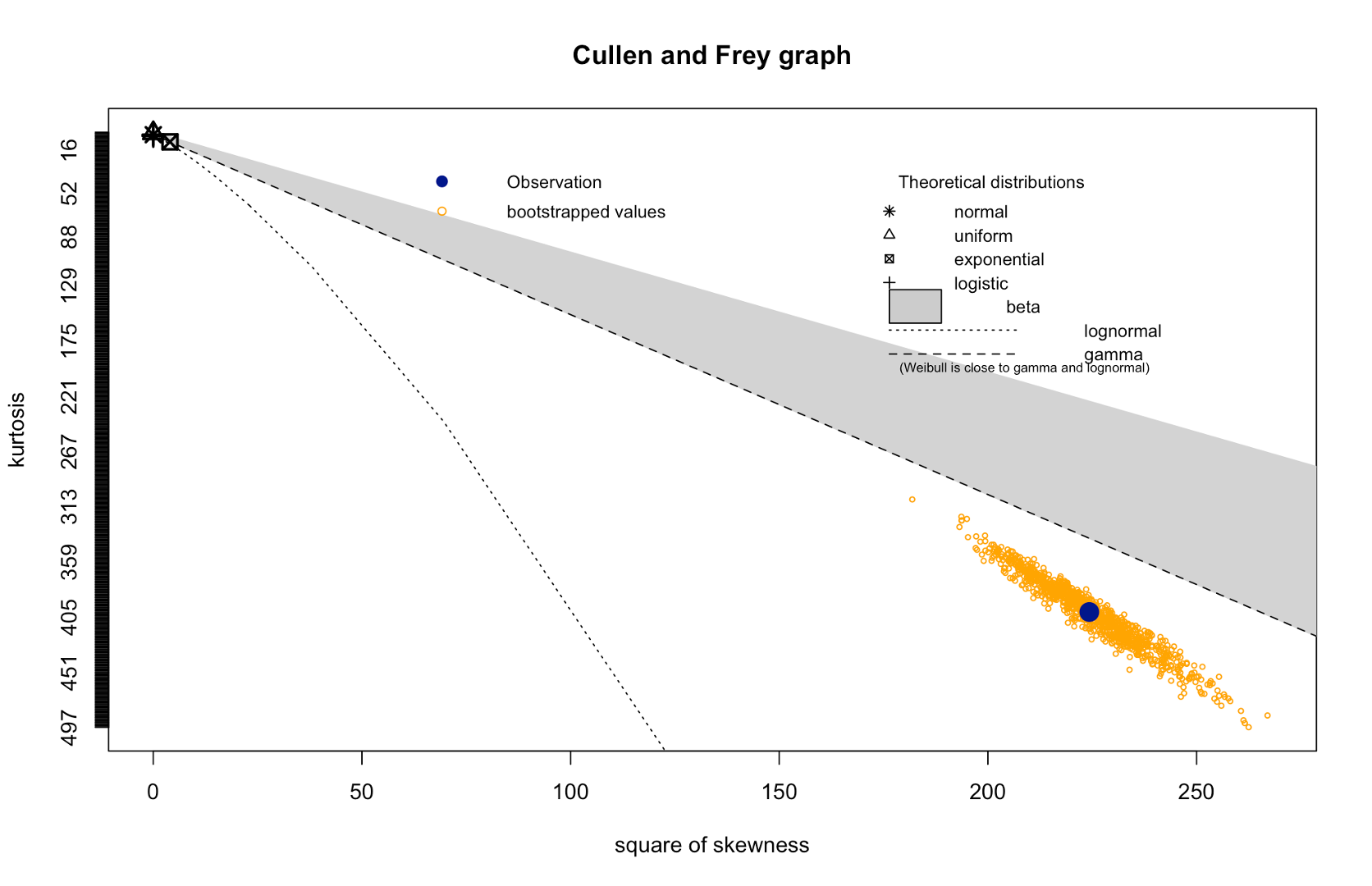
For order arrival rate, the historical dataset was first analyzed as a whole to see if there was an obvious trend in interarrival rates. Shown in the histogram below, the data represents a sharply downward sloping distribution with a long right tail. This suggested that an exponential distribution may best represent the data.

**Histogram of Interarrival Distribution of Customer Order Arrivals**

To further justify or refute this idea, we calculated the interarrival rates and then visualized the data in a Cullen and Frey graph using R, shown below. The blue and orange points represent the order data, while the lines, black points, and shaded areas represent various distributions, including normal, exponential, and gamma distributions. If the data of interest had a similar skew and tail behavior as one of these distributions, the orange points would lay along a distribution line or shaded area. Because the orange points are not aligned with the labeled distribution lines, the order data is not well reflected by a single distribution. This is due to its skewness and kurtosis (behavior of the tail).

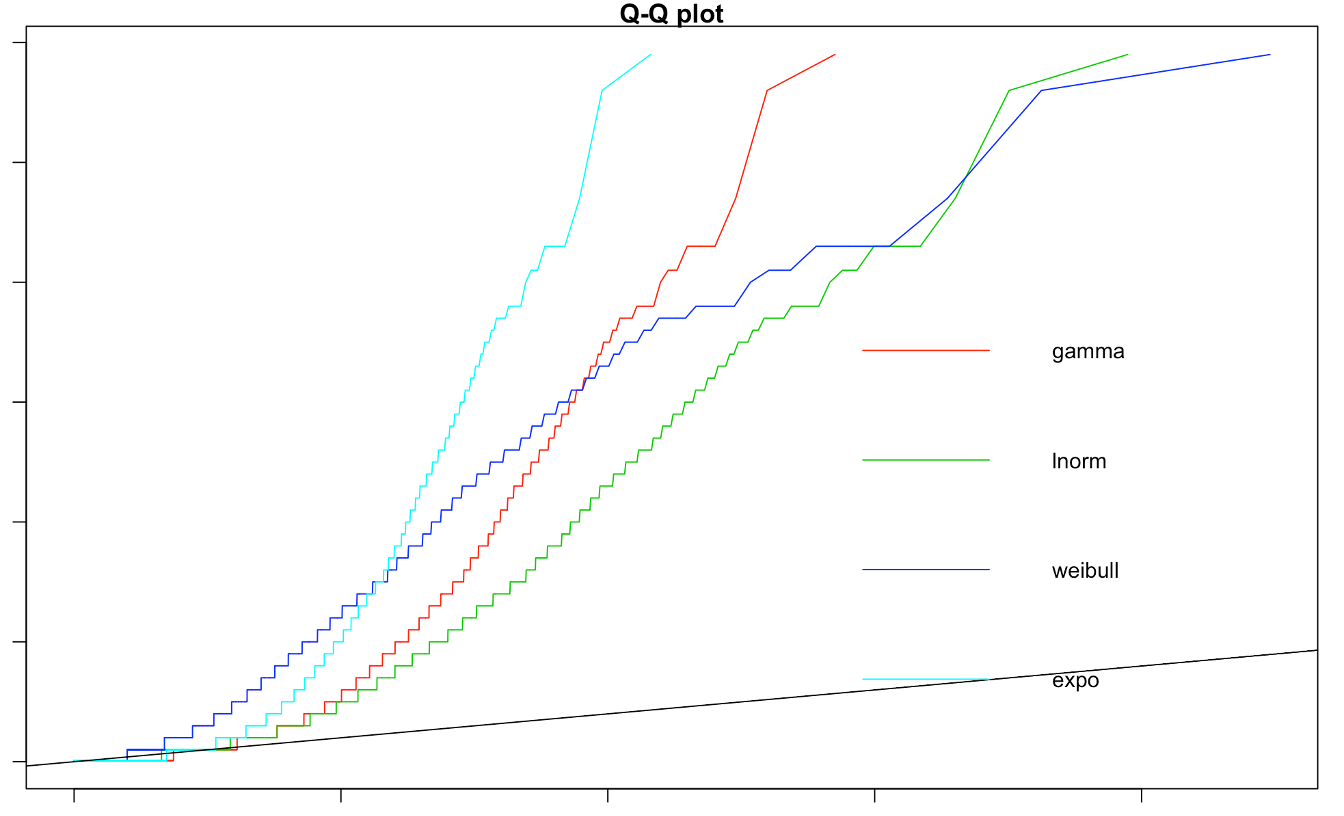
**Cullen and Frey Graph:**

**All Interrarrival Data Points**

In addition, the full data set was attempted to be fit by a variety of these distributions and graphed on a QQ-plot, shown below. As seen, none of the selected distributions are viable options to model the order arrival as a whole, and a deeper analysis was required to build the best arrival distribution to reflect the historic data.

**QQ-Plot:**

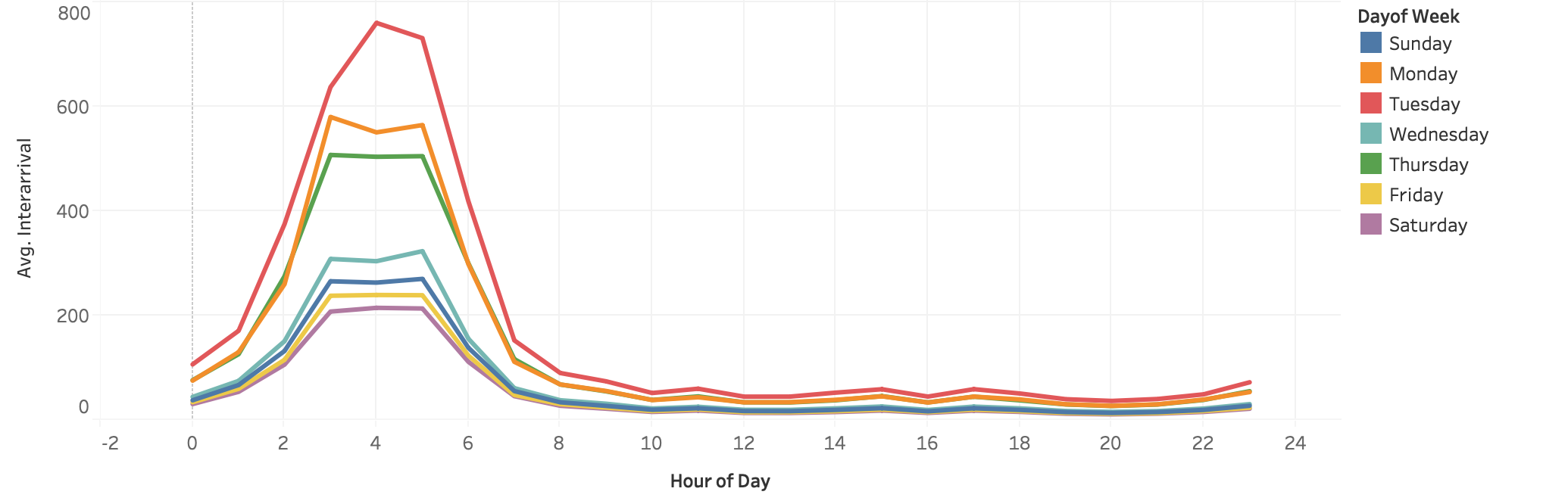
**All Interarrival Data Points**



After looking at multiple “slices” of the data, it was found that there were significant differences in the interarrival rates when segmenting the data by time of the day as well as day of week.

Tableau allowed us to clearly plot and visualize these distinctions. Therefore, it was decided to model each hour and each day using different distribution parameters.

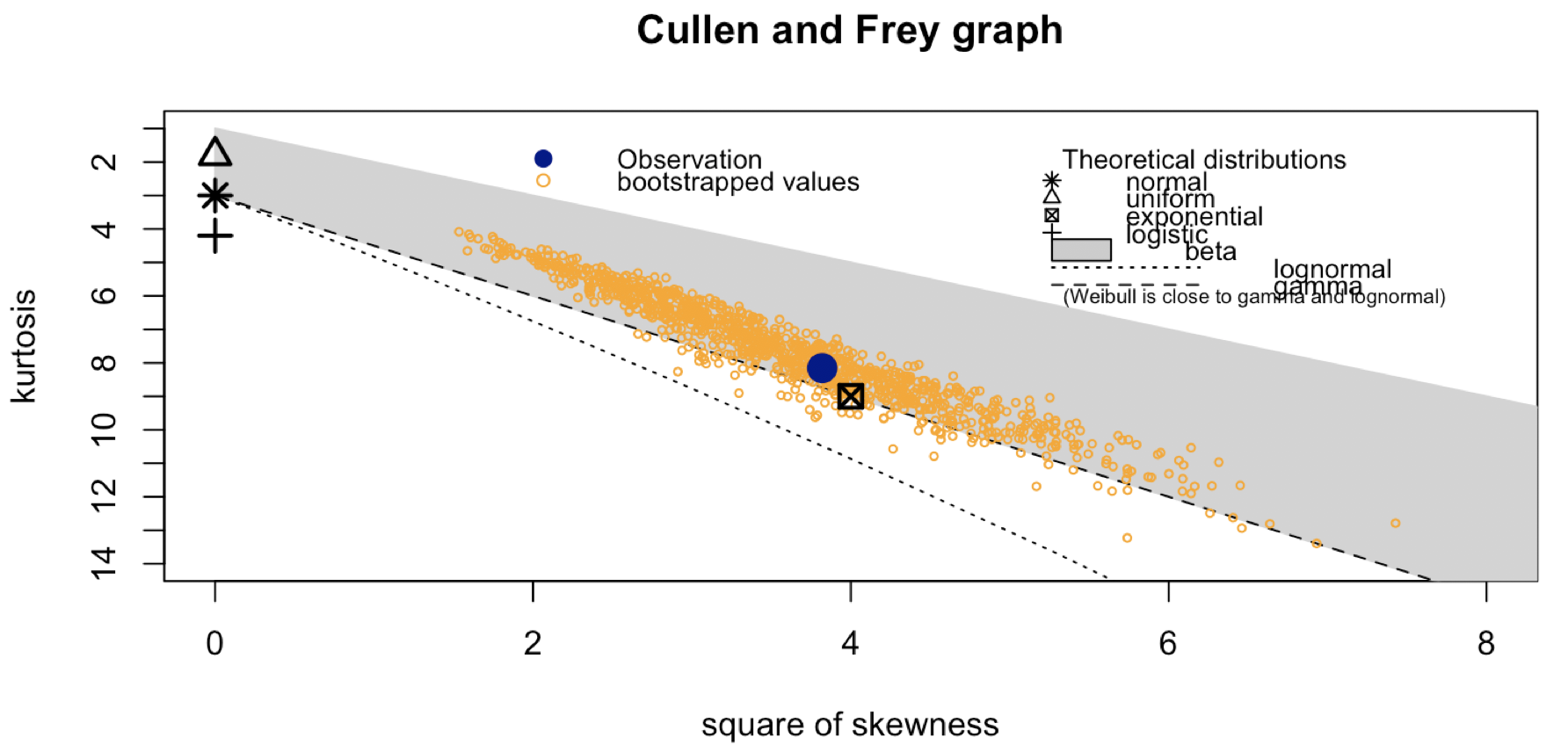
**Interarrival Differences by Day and Hour**



The sliced data was analyzed in R using the same Cullen and Frey graphing technique to gain insight into distribution types. An example is shown below for the hour of day as 5:00am and day of week as Thursday. Following this, QQ-plots were analyzed for the individual slices to confirm the best distribution among exponential, weibull, and gamma distributions. The sliced data was confirmed to follow a solid exponential distribution and, therefore, was modeled using the lambda parameters. As an additional validation step, Tableau was used to validate the lambda parameters found in R with 1/average interarrival. The lambda parameters that were assigned to each hour of the day and day of the week were documented in a table and fed into the simulation in order to best recreate the arrival of new customer orders.

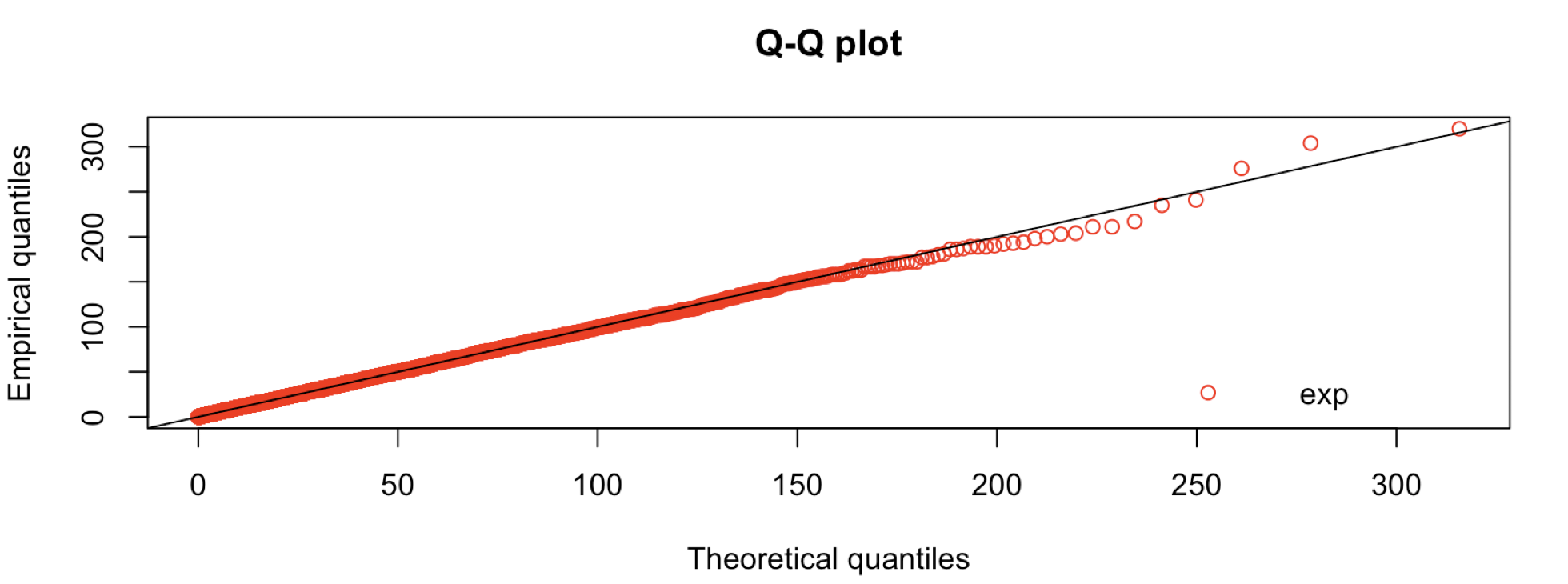
**Cullen and Frey Graph**

**Thursdays 5:00am**

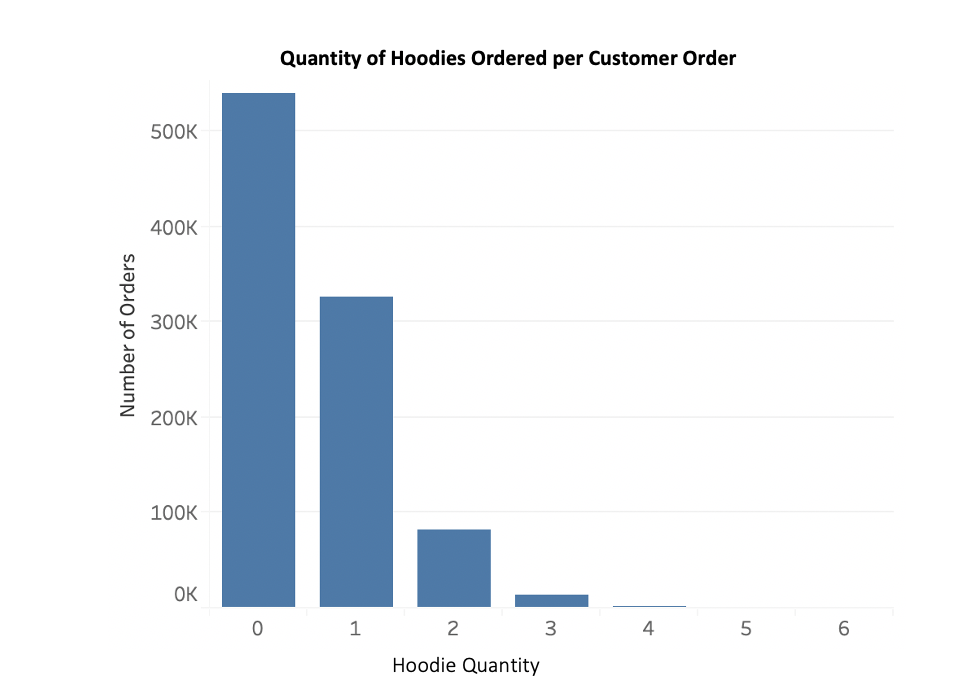


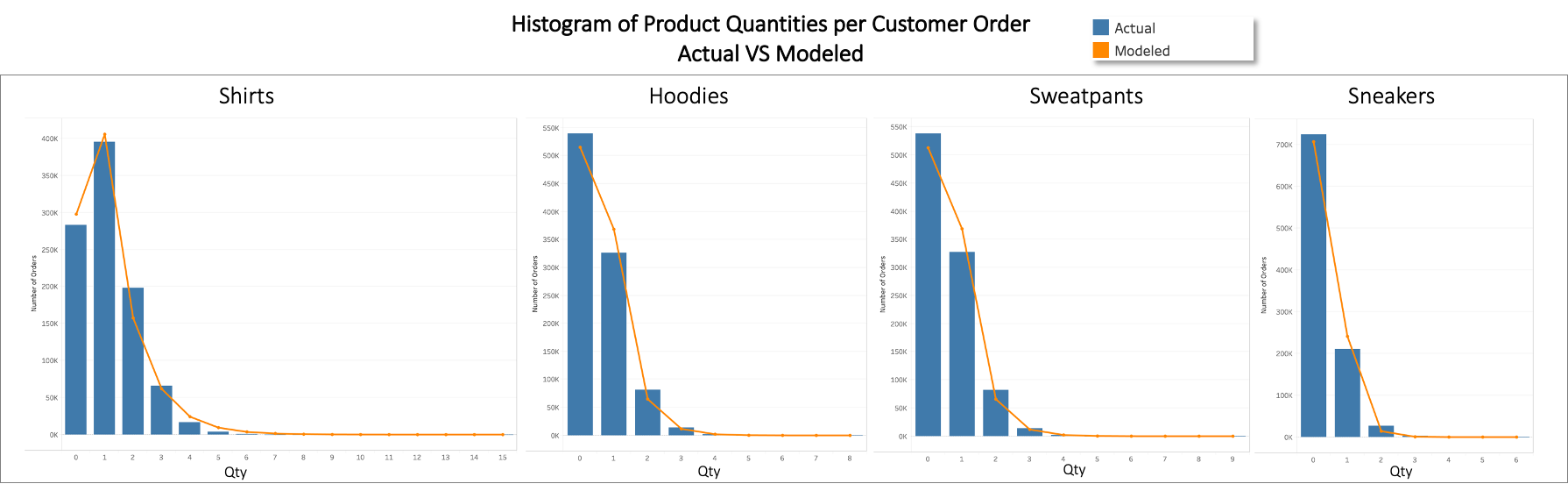
**QQ Plot**

**Thursdays 5:00am**



### ***Product Quantity Arrival Rate (Per Order)***

For the product quantity arrival rate (per order), it was hypothesized that the interarrival rates could be modeled using an exponential distribution based on visual inspection. The histogram on the right is an example of the quantity of sweatpants orders in each customer order. The lambda parameters were found using R, as before. Although this was found to be sufficient, it is important to note that random selection from this exponential distribution must be rounded in a purposeful manner in order to collect integer quantities. The selected number is rounded to the nearest integer, whether that be up or down. Each product has its own exponential distribution, and therefore it assigns an individual lambda value. Once these lambda values were modeled in the simulation, the output could then be compared to the historic data to see how well they represented the data set. The histograms below compare the actual vs modeled product quantity arrival rates, which prove that our modeling is sufficient to be used in the simulation.

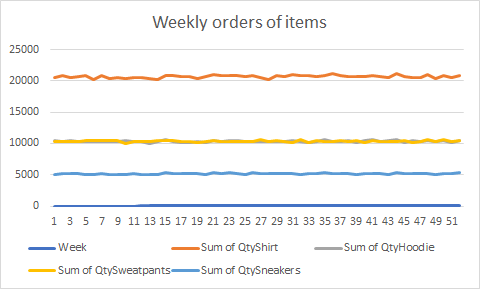


## Deciding the Delivery Schedule

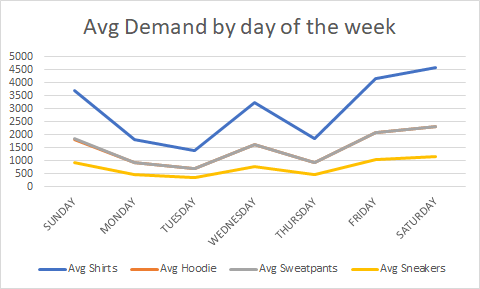
### ***Analyzing demand to decide the delivery schedule***

We explored the delivery options (daily vs weekly schedule) by investigating the orders coming in. We found two things that impacted our decision:

1. The orders per week remained relatively stable throughout the year without too much fluctuation. This was an important finding since it allowed us to use averages without worrying too much that we might not end up fulfilling some of the customer demands.

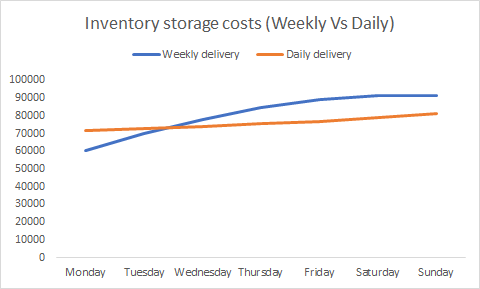


1. There is a difference in the number of orders on different days of the week. As shown in the graph below, we found that Saturdays had the highest number of orders coming in (~10,300 units of shirts, hoodies, sweatshirts, and sneakers), while the slowest day of the week was Tuesday (~3,000 units of items). This daily fluctuation is an important finding, especially for a weekly order schedule. If deciding on a weekly order schedule, we would need to hold enough inventory for high demand days starting on Monday. Alternatively, we would need to plan the schedule such that the peak demand is fulfilled later at the beginning of the week after the order has been placed.



### ***Using storage costs to decide the delivery schedule***

On completing this analysis, we could now analyze the costs of using different delivery schedules. We worked under the assumption that we would try to fulfill the demand as soon as possible from the order coming in, especially to avoid lost sales penalties. In order to do this, we would need to plan for the deliveries estimated for any day of the week before the orders were placed. For example, Tuesday’s predicted demand would be delivered on Monday. This led us to the following costs shown below:



The fixed costs of the weekly deliveries ($50,000) are lower than the weekly daily delivery cost ($70,000). But when we factor in the daily storage costs, there is an increased variable cost for weekly deliveries since more items need to be held through the week to meet the orders coming in on each day. This variable storage cost for the weekly deliveries is further increased because the peak orders come in on the weekends (Friday, Saturday, and Sunday) but the items are stored through Monday, Tuesday, Wednesday, and Thursday.

The daily delivery schedule provides us reduced inventory storage costs because the lower lead time of 1 day allows us to create a policy where each day we will receive the delivery for the expected orders for day N+1. This delivery will then be moved throughout the day N such that before the orders start arriving on day N+1, we have all the items ready for the pickers. This allows us to keep the inventory levels to a minimum by having only the units stored equivalent to day N+1. On performing the calculations as shown above and in the excel workbooks, we concluded that for our system, the daily delivery schedule works the best.

The daily delivery schedule also provides a few more benefits. First, having smaller deliveries on each day lowers the probability that any items would be returned to the shipper. Second, adding flexibility in the system can help adapt to changing customer demand better if we had a dynamic order scheduling contract with the supplier, this would then lower the chance of running into a penalty of not fulfilling the customer orders during unexpected busy periods.

### ***Estimated Delivery and Storage Costs***

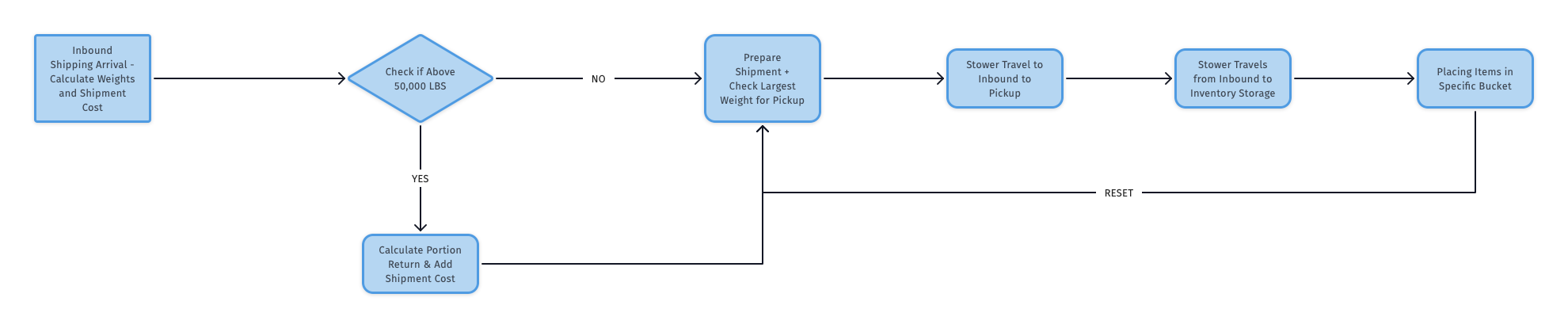
Based on the above decisions, we created a delivery schedule that changed over a week but remained constant throughout each day of the year. This can be found in the schedule.csv file. Total cost per year for delivery and storage can be found below:



## Inbound Processing

### ***Simulating Stower Operations***

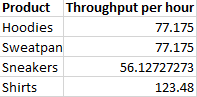
The inbound simulation starts with the delivery of goods, which enters the inbound parking area. Afterwards, we calculate the total weight, combining both the new shipment and the existing inventory in the parking area. If the total weight exceeds 50K, then we return a proportional amount of shipment back, paying a 10K cost. During the time period in between the arrival and delivery of goods, the stowers will transport the goods from inbound parking to the inventory storage. Overall, each stowing operation time takes 4 minutes to travel in between the inventory storage and the inbound parking area plus 10 seconds to drop off per item. For instance, a stower carrying 24 t-shirts will take 4 minutes to travel and additional 4 minutes to drop off all the items. Once the item is dropped off, the whole process repeats itself until the inventory parking area is depleted (meaning all product quantities in the inbound parking are 0). See diagram below.



The inbound processing was modeled in two steps. First, we tried to estimate the number of stowers required based on the delivery schedule and the amount of time taken to stow each item. Second, used these calculations in a simulation to verify whether the number of stowers made sense. We provide some of the metrics for our simulation below.

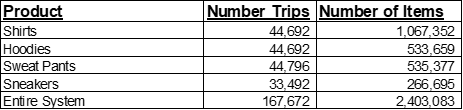
***Metrics and Estimations for the Inbound Processing***

Below, we show the total time spent along with assuming a 24 hour shift. Since the schedules are fixed, our throughput varies inversely to the weight of each item. Below is the estimated throughput for the system.



Since our schedule is fixed with certainty, the service level is 100%.

In total, our stowing operation made 167,672 trips transporting all four goods. Most of the trips were transporting sweatpants, followed closely by shirts and hoodies. Sneakers had the fewest number of trips.



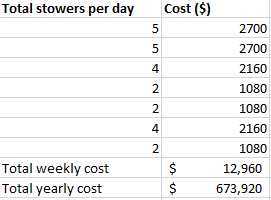
One of the major decisions was to determine the number of stowers based on the goods coming into the inbound terminal. We discuss our rationale for scheduling the number of stowers on a daily basis.

### ***Estimating Stower Schedule***

The estimation process worked in the following steps:

1. Calculate the amount of time taken to stow a unit of item
2. Based on that, find the number of items that can be stowed by a stower during a 8-hour shift
3. Find the number of stowers required for each product based on the deliveries coming in for each day of the week
4. Add a 20% extra buffer to ensure they run at a lower than 100% utilization
5. Convert the number of stower shifts into the 3 shifts equally to find number of stowers required per day
6. Calculate the estimated cost based on the total number of stowers estimated

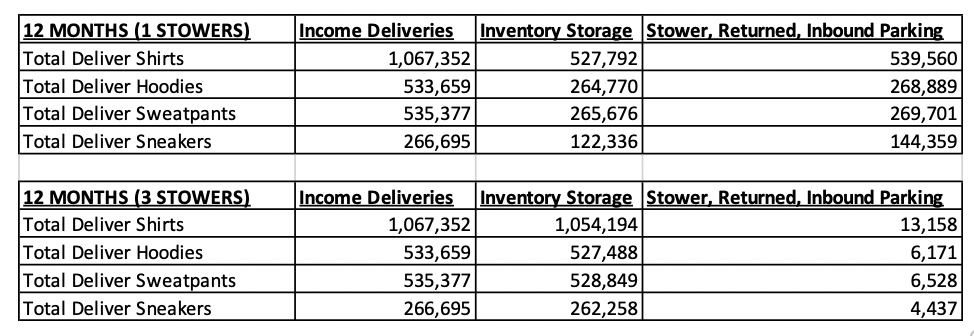
The estimated cost for each day is shown in the table below. More details on the estimation are provided in the schedule.csv.



One major point to note is that the estimation does not take into account that the stower assigned to an item is based on the amount of work left as this additional cost was factored into our simulation. One of the decisions we had to make was choosing between a fixed stower versus a variable stower schedule. We will discuss in further detail why we made the decision to use a variable stowing schedule.

***Rationale for Choosing a Variable Stower Schedule***

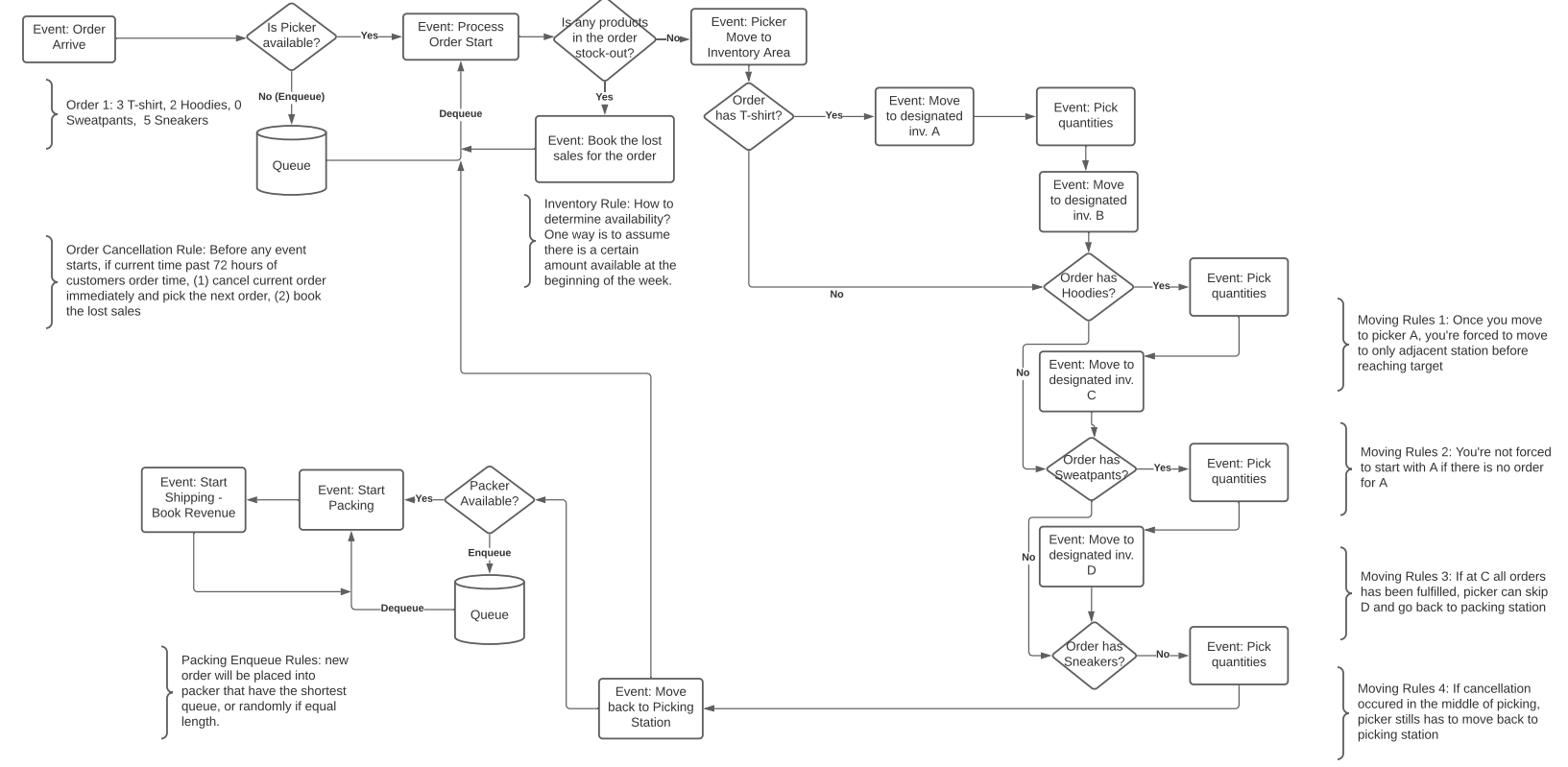
Based on the variability of our schedule, the average number of stowers is 3.42 stowers per day, assuming a 24 hour schedule. The rationale for choosing this value is because our goal is to make sure we move all the inventory from the inbound parking to the inventory storage while trying to minimize the total employment cost per stower. Here below with one or three stowers per a 24 hour shift, at the end of the year, we will still have inventory leftover either still moved by a stower or in the inbound parking area. See the charts below. Note that the “Stower, Returned, Inventory Parking” refers to inventory that resides on a stower, in the inbound terminal or has been returned because of the inbound parking constraints. In other words, this is the combined inventory that did not make it to the inventory storage.



If we were to choose 4 stowers, we will move the entire inbound shipment to the inventory by year end. However, we will incur higher costs as we will pay for additional stowers. In other words, varying the number of stowers per day allows us to minimize the number of stowers while making sure that all inventory is moved from inbound parking to inventory storage.

## Outbound Processing

Outbound processing begins with arrival of customer orders which are processed in a FIFO (first in, first out) fashion. The processing is completed by two sets of resources - pickers and packers. The pickers kick off order processing by gathering the necessary products from the inventory storage area(s). The collected products are then sent off to the packers to pack the order for shipment. Details regarding the operations in the outbound process, as well as major rules, are captured in the following diagram:



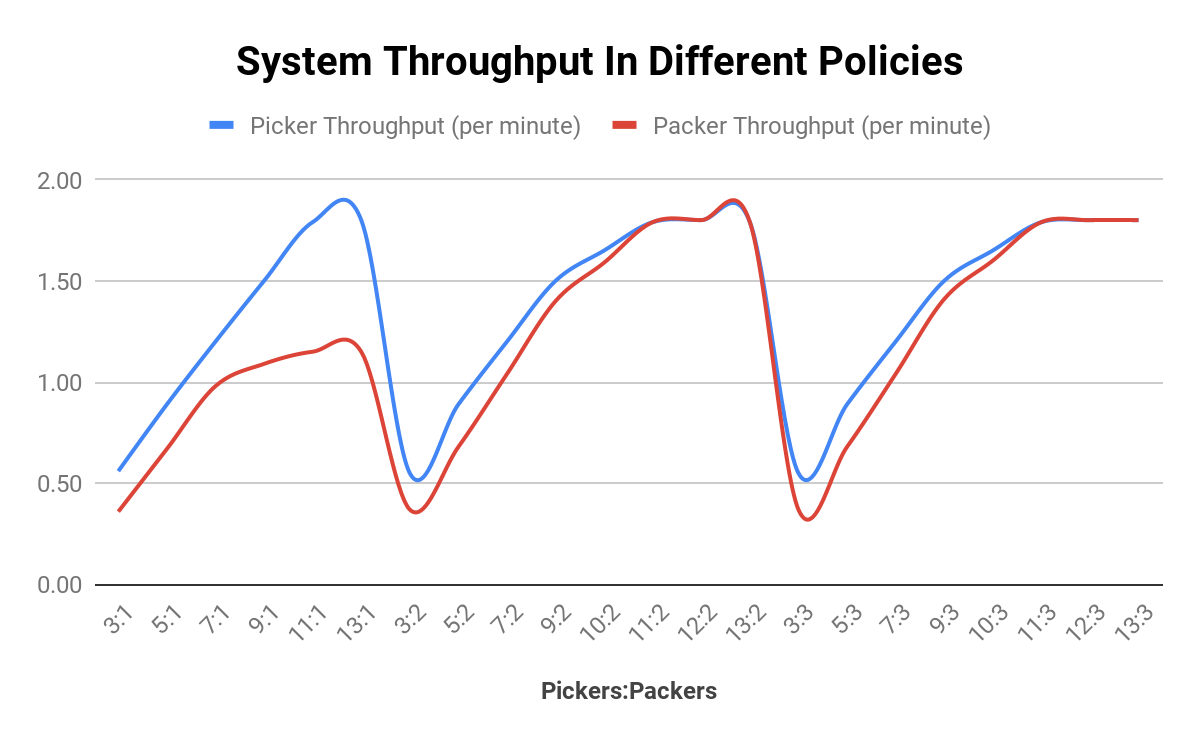
***Major Calculation Rules and Assumptions in Simulation Codes***

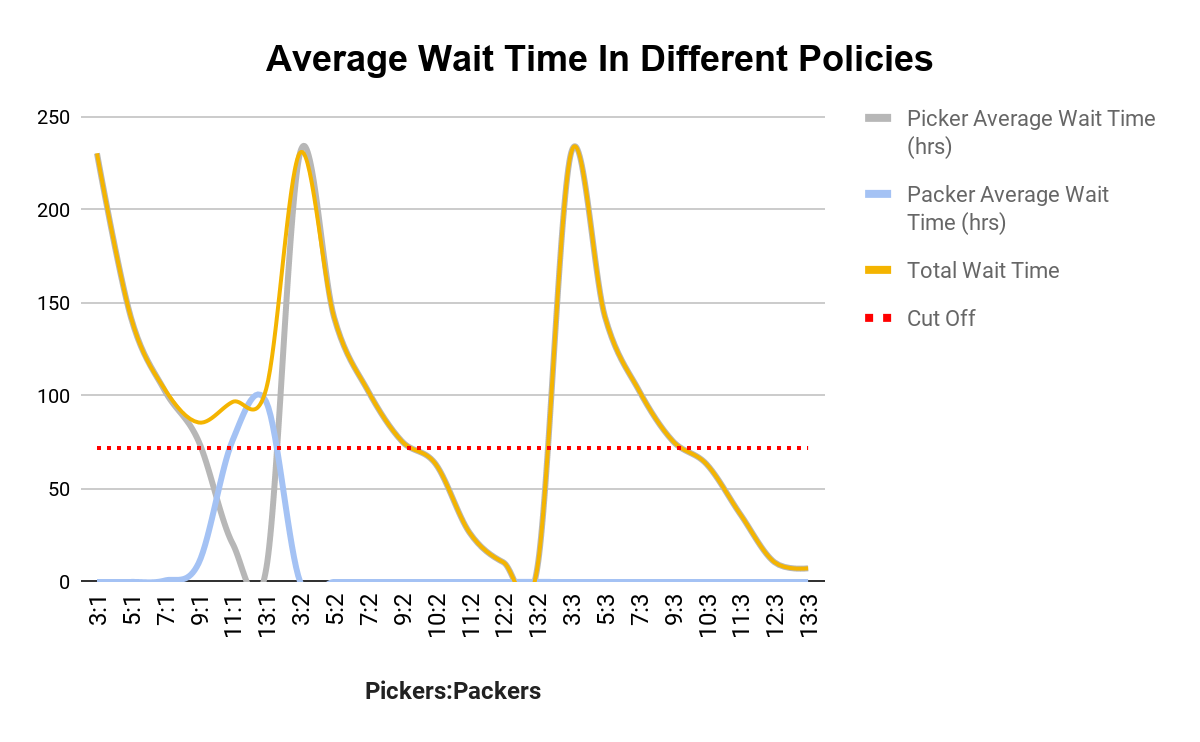
* We factor in picking and packing time into an order’s total waiting time to determine to process that order or not. This simplifies the cancellation process, especially in the middle of picking or packing operations. Furthermore, if an order is canceled after the picking operation begins, the inventory is lost. Determining beforehand whether to process that orders would help us avoid discarding the inventory for orders that are close to the 72 hour fulfillment deadline.
* Items are stored in their designated storage areas. 1: T-Shirts, 2: Hoodies, 3: Sweatpants, 4: Sneakers.
* We assume that there are constant inflows of inventory for each product. In the simulation, we update the inventory with that constant rate once every minute for each product. The rate is determined by the delivery schedule for the day divided by the number of minutes in 24 hours.
* Inventory Holding Cost was calculated by taking the snapshots of current inventory at the end of each day and then multiplying with daily inventory holding cost for each product
* We assume the same number of pickers and packers are working throughout all days and weeks of the year.

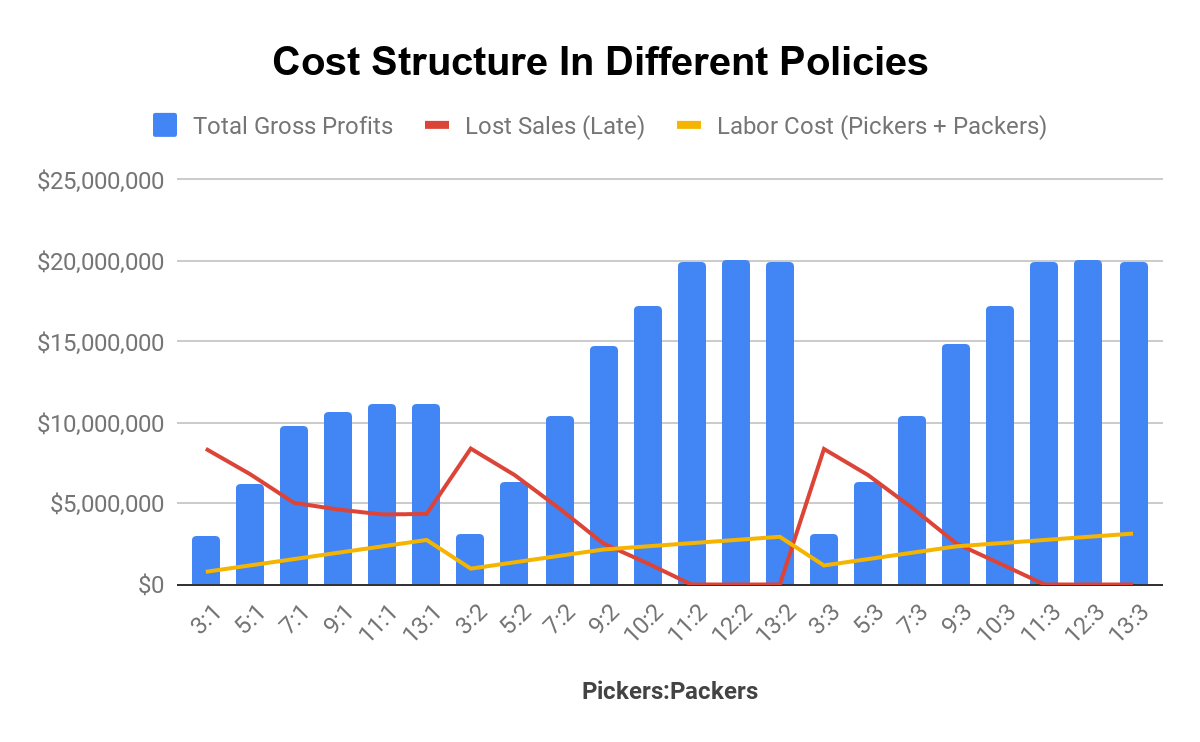
### ***Deciding Picker and Packer Schedule***

To find an optimal number of pickers and packers for our outbound processing, we simulated different runs at different policies to examine pickers and packers’ throughput. As we are looking at the throughput of the whole outbound system only, we use the same inventory inflow rate and at a rate that ensures no stockouts.

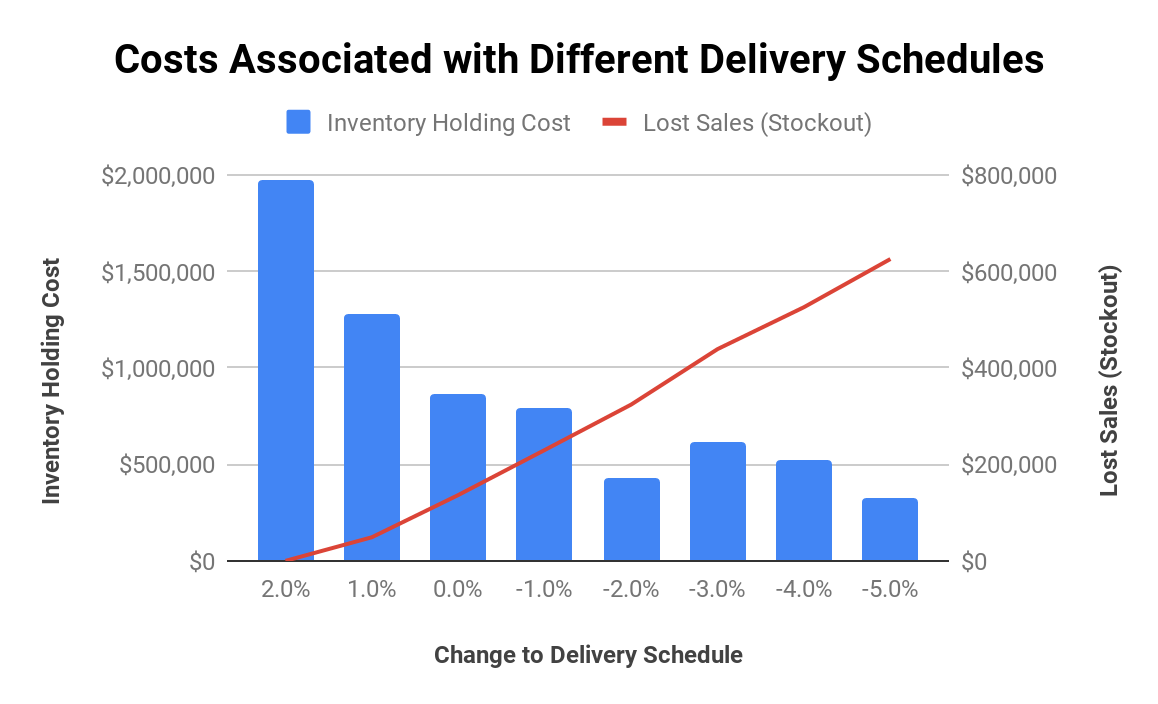
We can see that pickers’ throughput peaked at 1.8 orders per minute even if we increased the number of pickers. That indicates that the order arrived at the same rate. Since, the packing process followed what pickers left as the output, packers’ throughput peaked at the same rate even if we increased the number of packers. To maximize gross profits, we are looking at policies that would perform at or close to peak.

In regards to service level, only policies that gave us total waiting time significantly below 72 hours are worth looking at. At the peak throughput, only 4 policies are left (11:2,12:2,11:3,12:3).

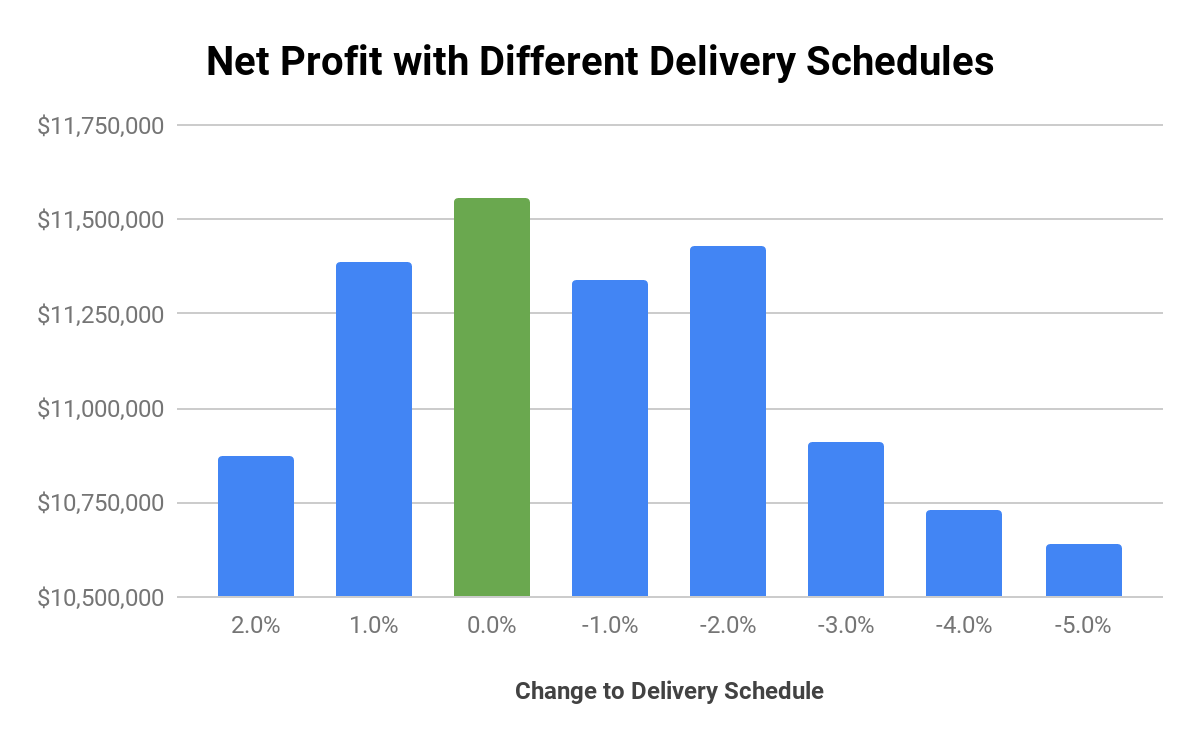


Finally, we looked at two main trade-offs that are directly influenced by the number of pickers and number of packers: Lost Sales due to late processing and Labor Costs. At the 12:2 policy, we see the most favorable combination of gross profits increase, labor cost increase and lost sales decrease. It is then our conclusion on the optimal number of pickers and packers in our system.

### ***Validating optimal inventory holding cost***

Once we determined the optimal number of pickers and packers, we used our simulation to see if we could further refine our delivery schedule. Our goal was to see whether we could optimize the various costs by changing the quantities of items within the delivery schedule. For example, we wanted to see whether reducing the quantities of items by 1% would positively impact costs. As expected, the inventory holding cost reduced as we reduced the number of items in our delivery schedule, whereas the lost sales due to stockout increased.

However, changing the delivery schedule also impacts gross profit. In order to see the net effect of changing the delivery schedule, we calculated the net profit at different changes to the delivery schedule. Ultimately, we found that our delivery schedule did not need to be further refined, as a 0% change yielded the maximum net profit.

Net profit excludes costs associated with stower labor expenses and delivery expenses (as well as possible penalties)

## 

## Putting it all together

|  |  |
| --- | --- |
| **Gross Profit (Yearly)** | **$ 20,419,036** |
| Less)  Delivery Expense and Possible Penalties | $ 3,630,000 |
| Less)  Lost Sales Penalty (Late Processing) | $ 0 |
| Less)  Lost Sales Penalty (Stock-outs) | $ 142,078 |
| Less)  Labor Expense (Pickers + Packers) | $ 2,751,840 |
| Less)  Labor Expense (Stowers) | $ 675,771 |
| Less)  Facilities Fixed Cost | $ 5,000,000 |
| Less)  Packing Station Expense | $ 100,000 |
| Less)  Inventory Holding Cost | $ 592,176 |
| ===  **Total Profit** | $ **7,527,171** |