
Data Challenge by Sandra Guo

Executive Summary

Exploratory Analysis

Observation and Recommendation

Staff the Customer Support Team

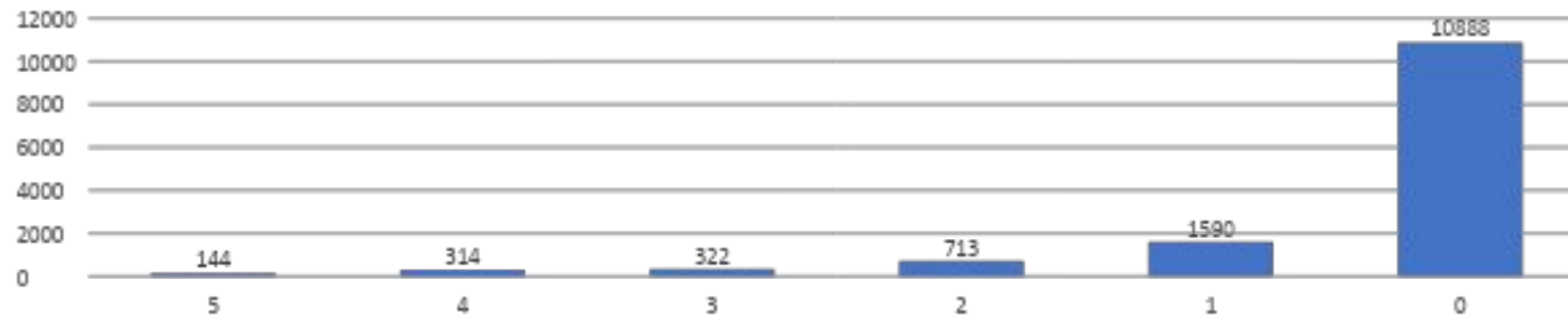
Next Step

Data Exploratory Analysis

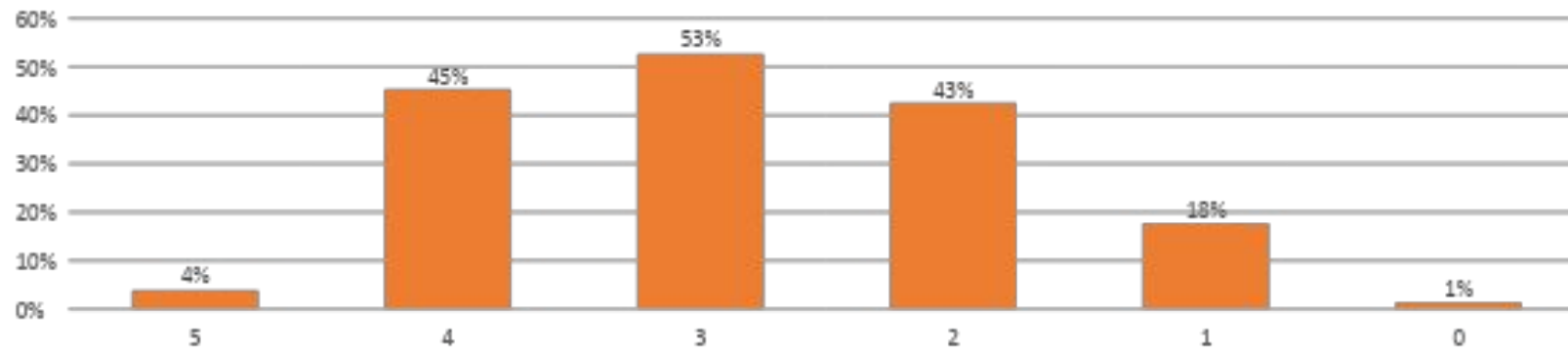
- Number of Region:3
- Region Names: Chi, SF, NYC
- Time Frame: 2014-05-01 08:54:00 to 2014-06-02 06:28:37
- Total number of orders: 13845
 - Chi orders: 5973
 - SF orders: 6670
 - NYC orders: 1202
- Rating Scale: 0 to 5, the higher the rating, the better the review
- Issue reported: Wrong item, Item charged incorrectly, Damaged or spoiled, Item missing, Other Order Issue, Poor replacement, Poor service

Observation 1:

number of orders across ratings



number of orders with issues reported across rating



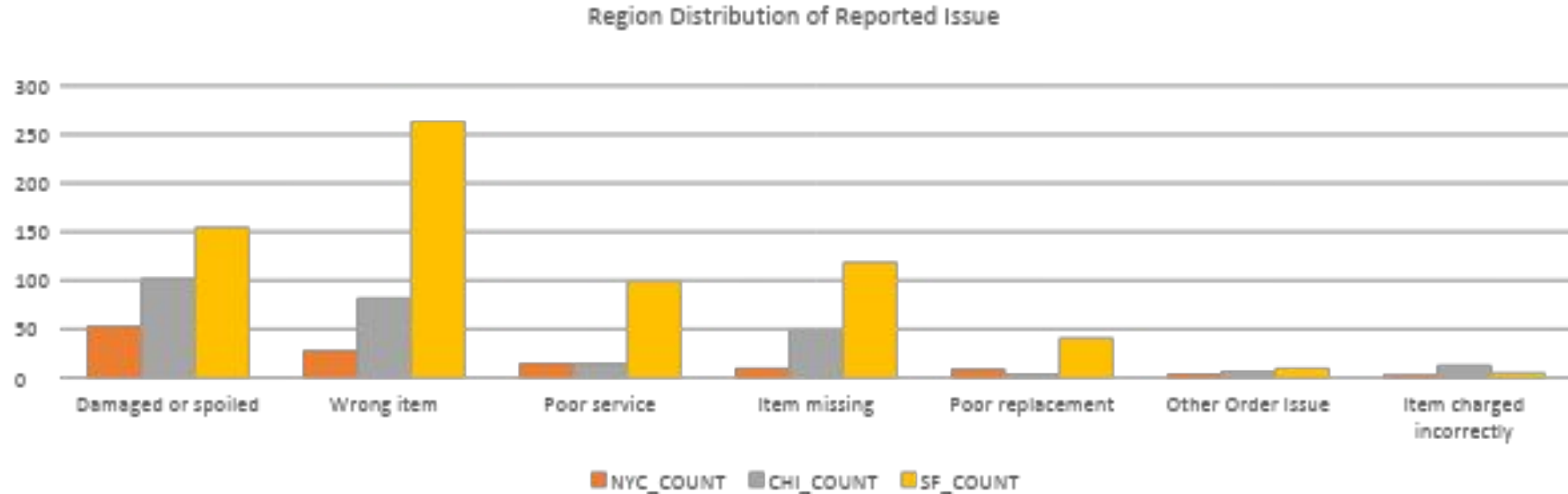
- Insight:
 - 79% of orders get a rating of 5/5
 - As the rating goes down, percentage of issues reported goes up
 - The rating of 0 has only 4% of issues reported
- Follow up analysis/recommendation:
 - The find out retention rate of each rating
 - The rating of 0 has only 4% of issues reported could be customers are too unhappy about the experiences to leave a reported issue, more follow up survey to people with rating of 0 see why no issue reported
 - Do a logistic regression modeling to see which reported issues impact the most on retention(the response variable is retention(dummy variable of whether shopper return within 30 days of order, the features can be reported issues)

Observation 2:



- Insight:
 - 6% of Chicago orders with rating < 4, 13 % of NYC orders with rating < 4, 15% of SF orders with rating < 4
 - Chicago's customer satisfaction is much higher than NYC and SF
- Follow up analysis/recommendation:
 - Look up the customers retention(eg, next order within 30 days of first order) and the of these three regions, to see if the customer of Chicago it's happier of the shopping experiences and more likely to return

Observation 3:



Insight:

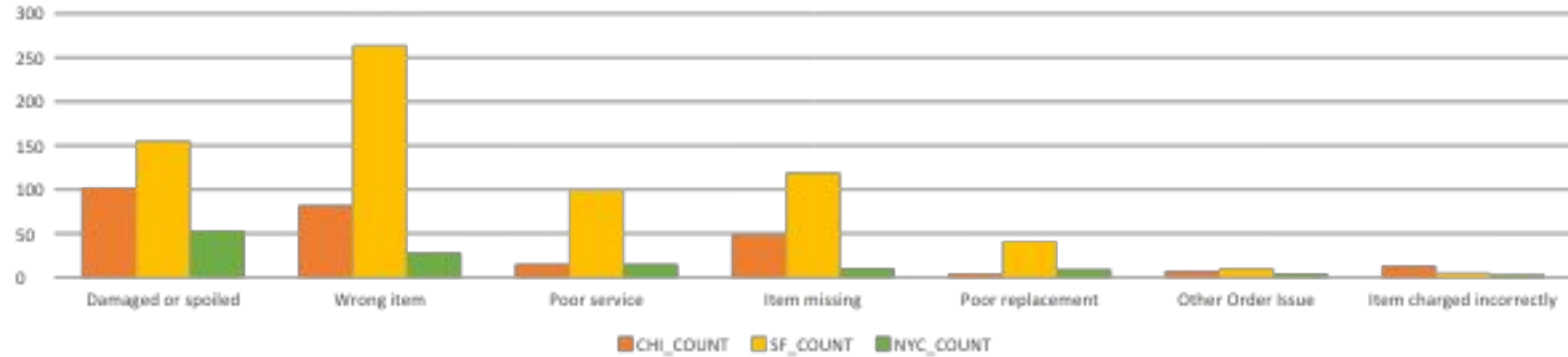
- Of all the reported issues, the damaged and spoiled, and wrong items are the top reported issues

Follow up analysis/recommendation:

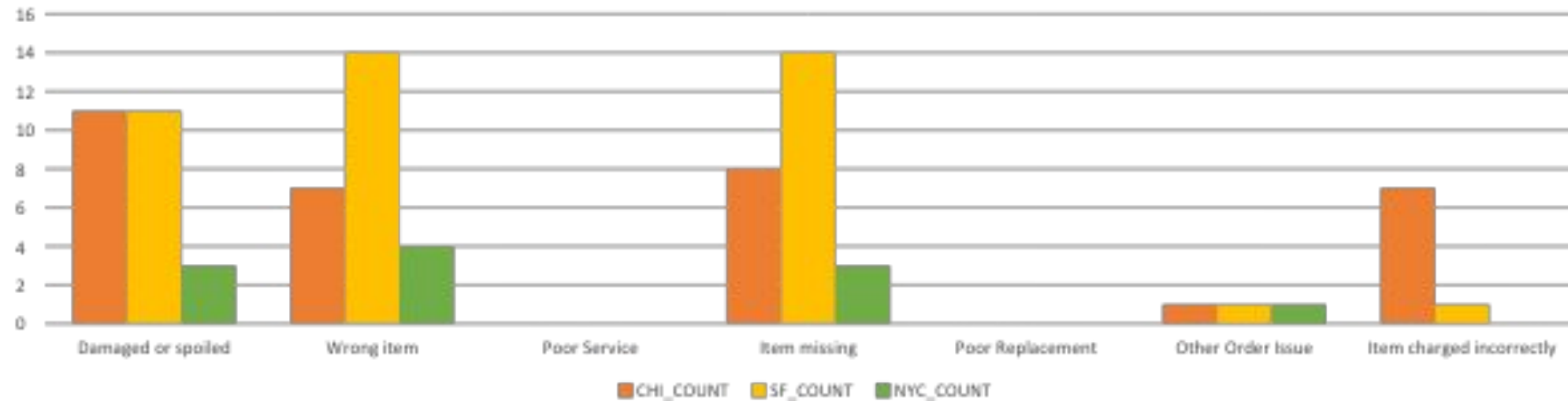
- can increase the QA of item placement for shoppers
- Decrease the tip for shoppers who reported with damaged items and wrong items issues frequently in short period of time
- Develop fast track/squad teams to investigate and refund

Observation 4

Distribution of issues reported over 3 regions

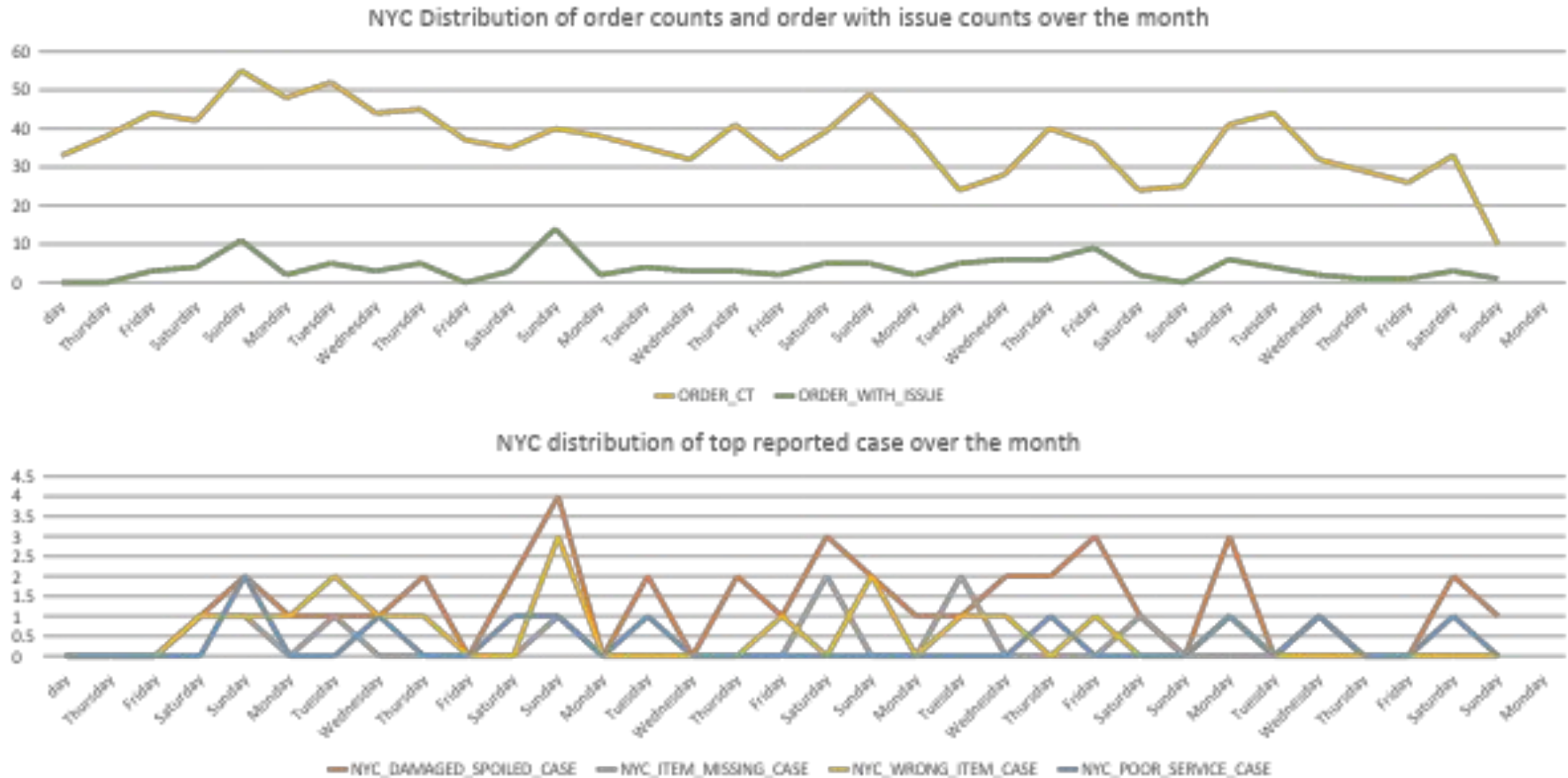


For the order with rating of 5, distribution of issue reported over 3 region



- Insight:
 - Rating of 5 will still have issue reported
 - The distribution of issue reported across 3 region for all orders with issue reported and orders with rating = 5 are following same trend, except poor service.
 - Some customers give rating of 5 with issue reported, might consider the rating as service rating
- Follow up analysis/recommendation:
 - Analyze the retention rate(next order within 30 days of order with issue reported and rating = 5)
 - Combination of rating and issue reported to score customer experiences
 - Rating of 5 can't be guaranteed as great experiences, need to give some weight to issue reported with scoring shoppers

Observation 4a---NYC

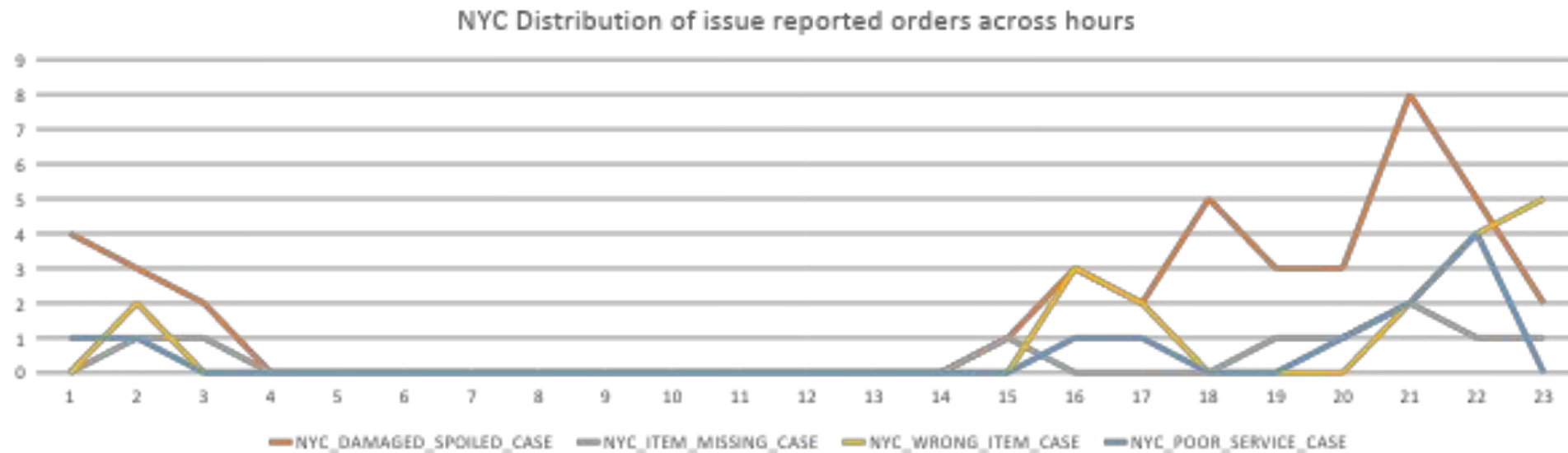
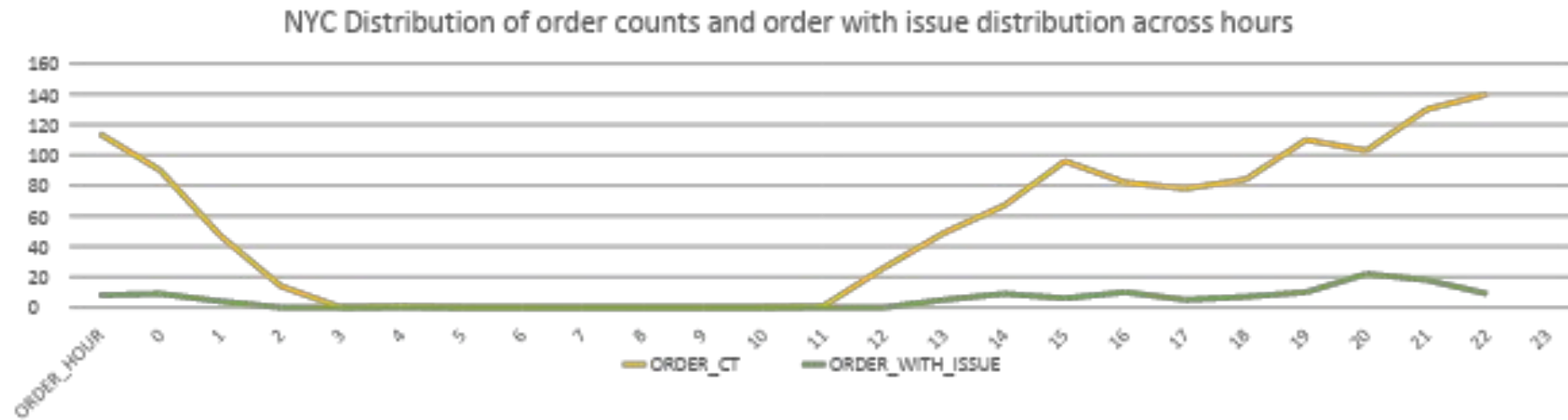


*The date of both chart is from 05-01 to 06-02

* Only include top reported cases for the distribution

- Insight:
 - In NYC, the trends of total order count and the reported issue order are consistent, which means more orders, more likely to get reported issues
 - Peak hours of order placed is hours of 15 to 0
 - Generally speaking, in NYC weekends have more orders than weekdays
- Follow up analysis/recommendation:
 - Incentives to encourage shoppers to work to delivery more orders with weekend

Observation 4b---NYC



Insight in NYC:

- No order placed between hours of 4-10
- No issue reported between hours of 5-13
- Around hour of 0-1 and hour of 15-22 is the peak of issue reported
- The trend of damaged items is very consistent with trend of order placed
- During downtime (hours of 4-16) of orders, less reported issues. The reason of issue reported could be cause by peak hours too many orders, such as squeeze items to cause damage, and too busy to cause missing items

Follow up analysis/recommendation:

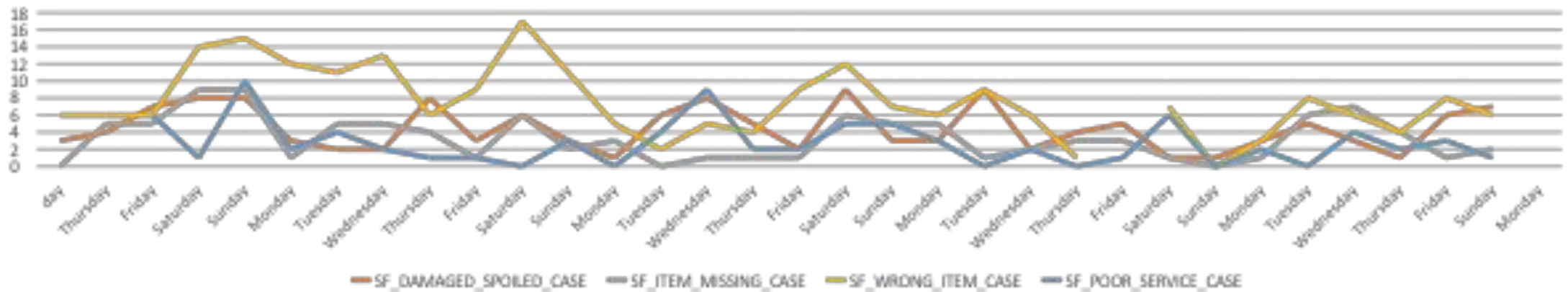
- Incentivize shoppers at peak hours, to place items during transition more carefully. Also, can come up with some assortment facilities in cars to improve item storage quality
- Put more QA/customer support at peak hours

Observation 5a--SF

SF Distribution of order counts and order with issue counts over the month



NYC distribution of top reported case over the month



*The date of both chart is from 05-01 to 06-02

* Only include top reported cases for the distribution

Insight:

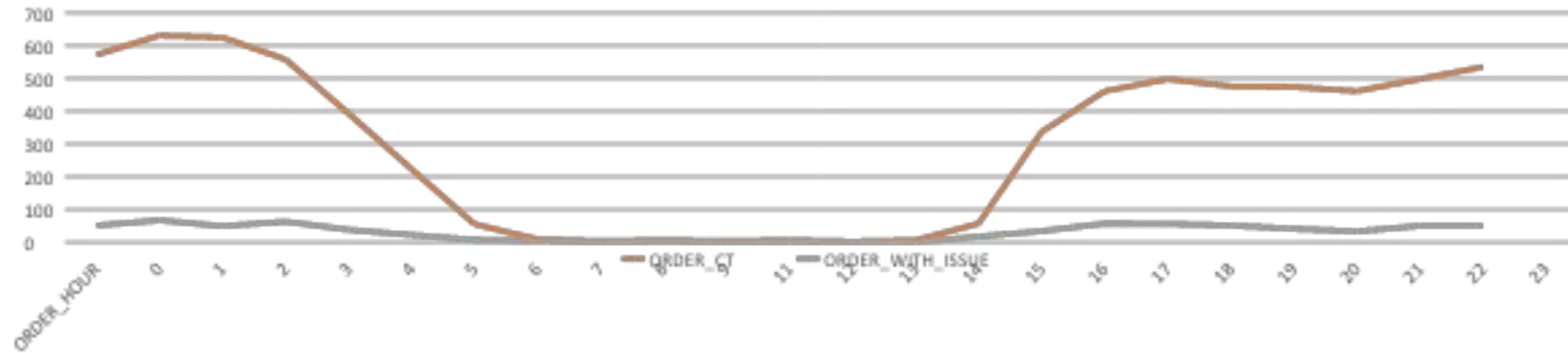
- The distribution of orders in SF with issue reported is flat
- However, the wrong item placement and damage and spoiled case trend are more consistent with distribution of orders
- Generally speaking, in SF, weekends have more orders than weekdays

Follow up analysis/recommendation:

- Incentives to encourage shoppers to work to delivery more orders with weekend
- it is important to QA the placement of items during transition in peak hours, since issue reported could be cause by peak hours with too many orders, such as too busy to cause missing items
- Follow up analysis on category of items spoiled/damaged, ask shoppers to pay extra attention on delicate items

Observation 5b--SF

SF Distribution of order counts and order with issue distribution across hours



SF Distribution of issue reported orders across hours



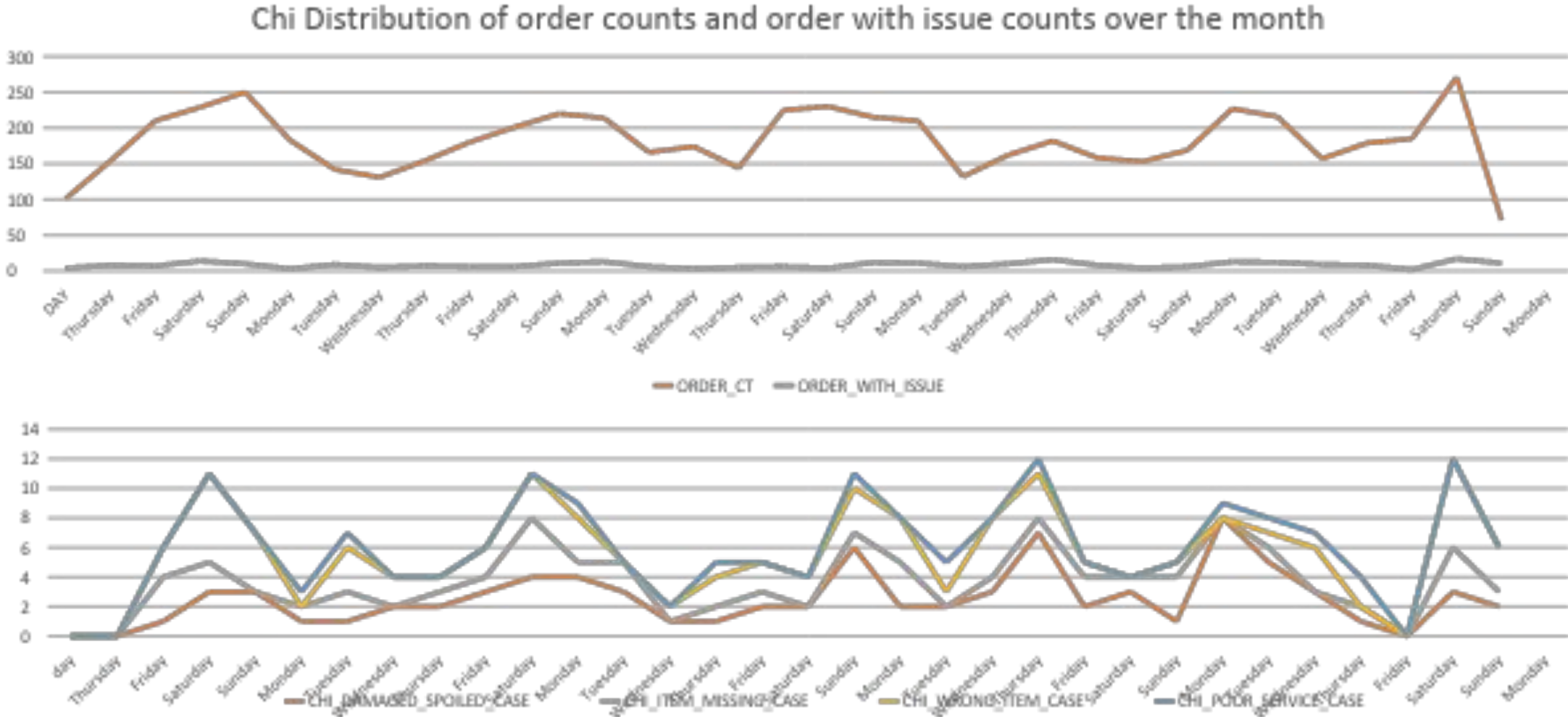
Insight in SF:

- No order placed between hours of 6-13
- No issue reported between hours of 7 – 13
- Around hour of 0-2 and 14-22, are the peak of issue reported
- During downtime (hours of 3-13) of orders, less reported issues.
- The trend of wrong item case is very consistent with order placed

Follow up analysis/recommendation:

- Incentivize shoppers at peak hours, to pick up right items carefully
- Put more QA/customer support at peak hours

Observation 6a--Chi



* The date of both chart is from 05-01 to 06-02
* Only include top reported cases for the distribution

Insight:

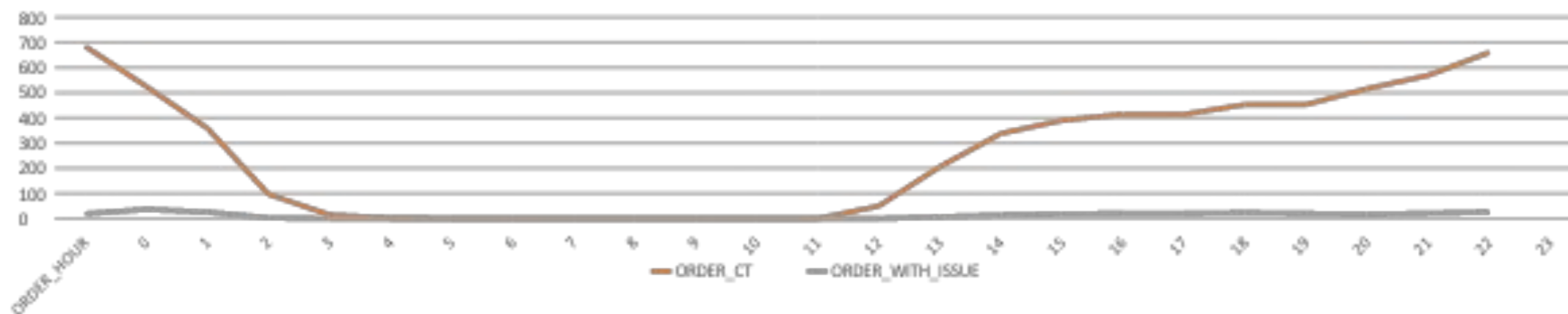
- The distribution of orders in Chi with issue reported is flat
- However, damage and spoiled case trend are more consistent with distribution of orders
- Generally speaking, in Chi, weekends have more orders than weekdays especially during Wed, Tue, and Thursday, there are less orders

Follow up analysis/recommendation:

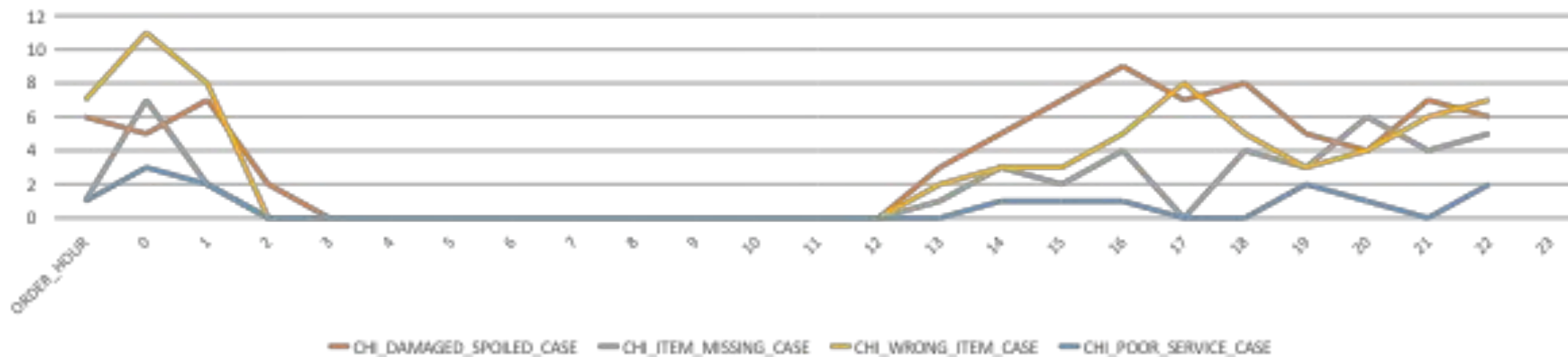
- Incentives to encourage shoppers to work to delivery more orders with weekend
- the damage and spoiled case trend are more consistent with distribution of orders, it is important to QA item placement quality during peak hours, since issue reported could be cause by peak hours too many orders, such as squeeze items to cause damage
- Follow up analysis on category of items spoiled/damaged, ask shoppers to pay extra attention on delicate items

Observation 6b--Chi

Chi Distribution of order counts and order with issue distribution across hours



Chi Distribution of issue reported orders across hours



Insight in Chi:

- No order placed between hours of 3-11
- No issue reported between hours of 3 – 12
- Around hour of 0 – 1 and 14 to 22, are the peak of issue reported
- During downtime (hours of 2 to 13) of orders, less reported issues. The reason of issue reported could be cause by peak hours too many orders, such as squeeze items to cause damage, and too busy to cause missing items
- The trend of wrong item case is consistent with distribution of order placed

Follow up analysis/recommendation:

- Incentivize shoppers at peak hours, to pick up orders more carefully
- Put more QA/customer support at peak hours

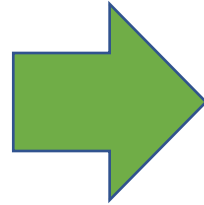


Part III

Staffing the Customer
Support Team

Chi total order of 6670, order with issue reported 6%

- Weekend has more order than weekday
- 3-11 no orders
- 0-1, 14 - 23 are order peak hours
- 0, 16-23 are issue reported peak hours



Chi:

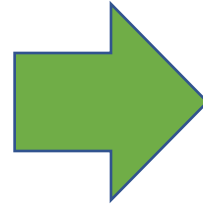
Assume reply to support issues within 24 hours

For weekdays:

- 3 people to support item missing and item damage report during 0-1, 14-23 (consider 8 hours a day, take shift)
- 2 people to support item missing report/item damage report /price incorrect/other issue reported during 0-3, 11-23 (consider 8 hours a day, take shift)
- 1 data scientist/analyst to analyze poor item replacement and poor service
 - Send survey or analyze comment for poor service
- For weekends, 2x customer support staffs for all the above team

SF total order of 5973, order with issue reported 15%

- Weekend has more orders than weekday
- 6-13 no orders
- 0-2, 16-23 are order peak hours
- 0-2, 16-23 are issue reported peak hours



SF:

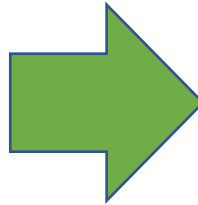
Assume reply to support issues within 24 hours

For weekdays:

- 6 people to support item missing and item damage report during 0-2, 16-23 (consider 8 hours a day, take shift)
- 4 people to support item missing report/item damage report /price incorrect/other issue reported during 0-6, 13-23 (consider 8 hours a day, take shift)
- 2 data scientist/analytic analyst to analyze poor item replacement and poor service
 - Send survey or analyze comment for poor service
- For weekends, 2x customer support staffs for all the above team

NYC total order of 1202, order with issue reported 13%

- Weekend has more orders than weekday
- 4-10 no orders
- 0, 15-22 are order peak hours
- 0-1 and hour of 15-22 is the peak of issue reported



NYC:

Assume reply to support issues within 24 hours

For weekdays:

- 2 people to support item missing and item damage report during 0-2, 16-23(consider 8 hours a day, take shift)
- 2 people to support item missing report/item damage report /price incorrect/other issue reported during 0-6, 13-23 (consider 8 hours a day, take shift)
- 1 data scientist/analytic analyst to analyze poor item replacement and poor service
 - Send survey or analyze comment for poor service
- For weekends, 2x customer support staffs for all the above team

Next steps:

- Due to limited amount of time, unable to further analyze for staffing questions. If I have more time, I will analyze the average number of order during weekend/weekdays/peak hours to proportionally staff the customer support team
- Conduct follow up analysis
- Combine the data with customer behavior/characteristics data, to further analyze retention behavior and uncover growth opportunities
- Combine the data with order data, too see the value loss brought by issue reported/low rating orders, to staff the customer support team accordingly

