By Elsia Phua

Analysing an Online Restaurant

Quick Intro: Datasets sponsored by an SG online restaurant

- Operates from an online platform
- No brick and mortar outlet
- Serves B2C and B2B customers
- Orders placed via web /mobile application or sales team
- Meals are prepared in an offsite kitchen and delivered to doorstep or pick-up point

Agenda

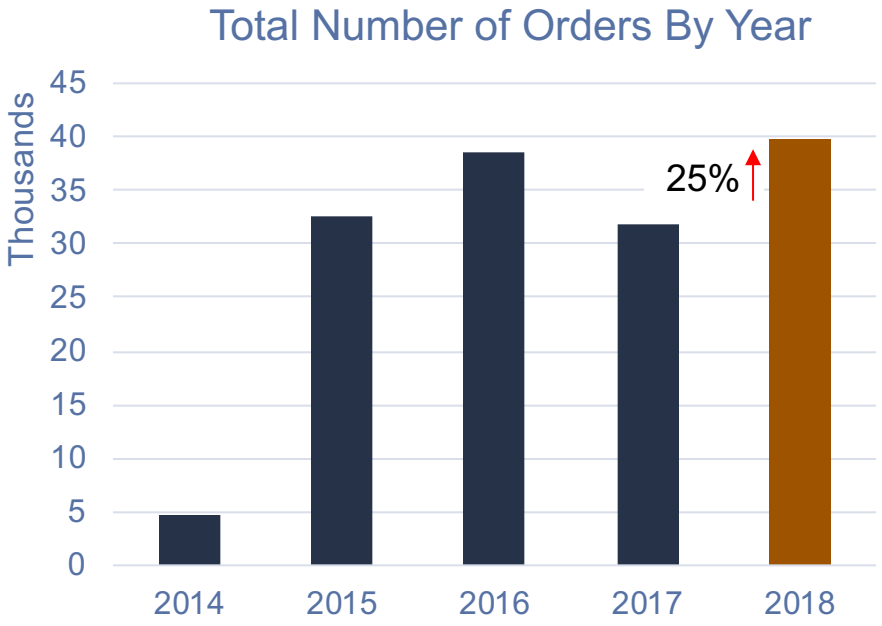
- Business Analysis
- Problem Statement
- Customer Segmentation
- Product Segmentation
- Conclusion & Recommendations

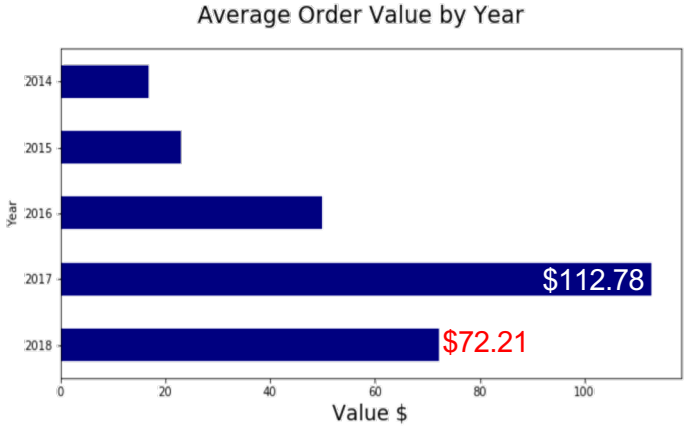
1

Business Analysis

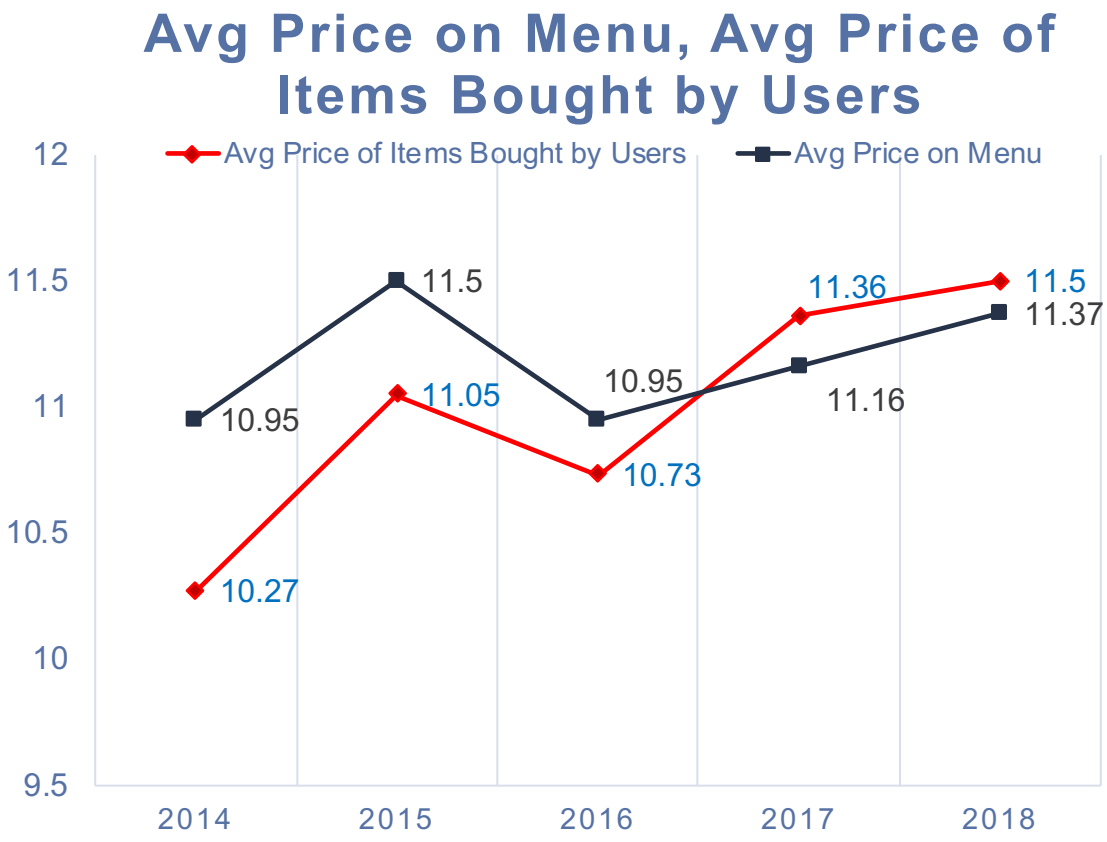
Finding a direction through descriptive statistics

Number of orders peaked in 2018; average value of an order fell 35%

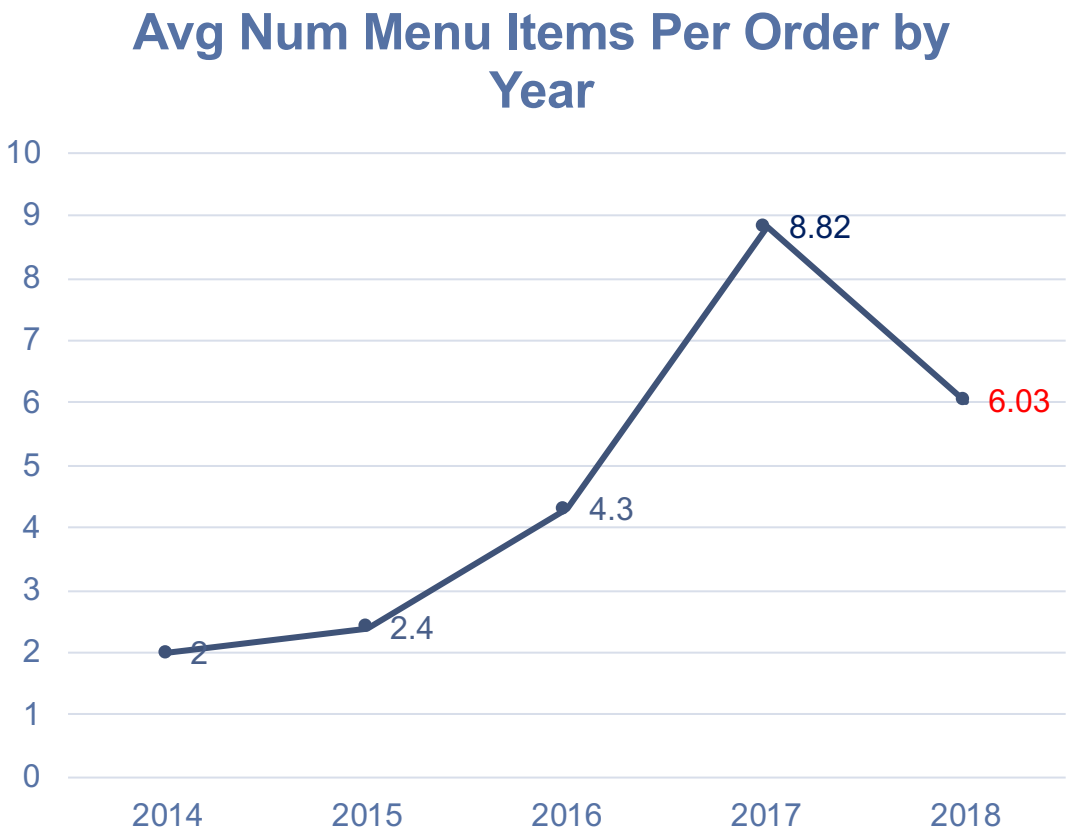




- Sales = Price x Quantity
- Are users purchasing lower priced items?
- Have the meals become cheaper?
- Has quantity per order reduced?



- Products did not get cheaper
  - Menu prices have been inching up from 2016 to 2018.
- Users are paying marginally more.
  - If users are not purchasing lower priced items, has quantity dropped?



- Although number of orders peaked in 2018, number of items per order fell.
- Users were placing more orders but each order had less items.

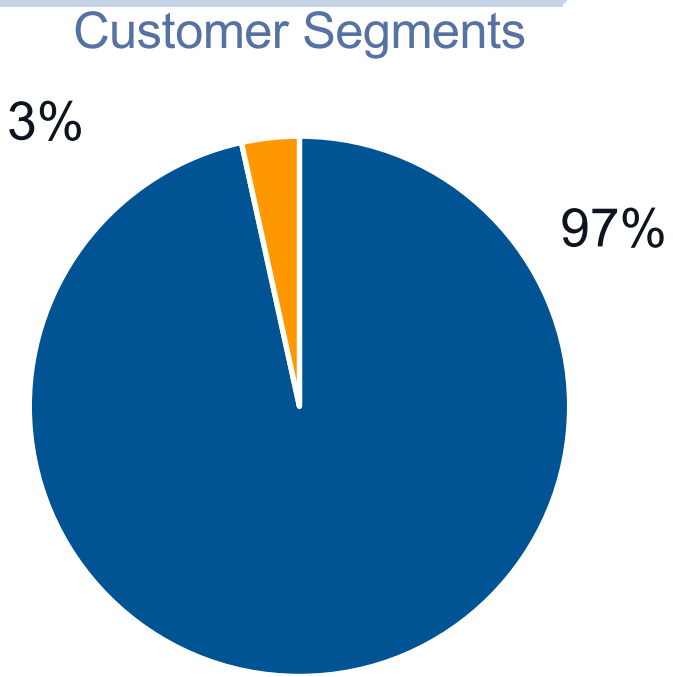
Problem statement: How can we increase revenue by increasing order value?

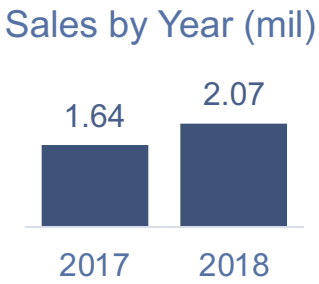
2

Customer Segmentation for Customized Strategies

Segmenting users using
K-Means Unsupervised Clustering

K-Means clustering grouped users into two segments by their purchase behaviour





- K-means clustering is one of the simplest and popular unsupervised machine learning algorithms.
- Group similar data points together and discover underlying patterns.
- A cluster refers to a collection of data points aggregated together because of certain similarities.
- Silhouette score of 0.51 with 2 segments

The Party Organizer (caters for ~ 50 Pax)

- **Avg total order value: $888**
- Avg individual meal price: $17
- Avg delivery fee per order: $44.5
- **100% of the time, users were ordering in quantities for groups**
- Meals time between 11:00 -11:59 am and during off peak
- **95% of orders come through the sales team**
- **Meals from the off-menu category**

The Small Time Influencer (orders for ~ 5 Pax)

- **Avg total order value: $57 (orders above $50 do not incur a surcharge)**
- **Avg individual meal price: $11.70**
- Avg delivery fee per order : $1.85
- 20% sign ups were incentivized by coupon
- 65% of the time, users were ordering in quantities for groups
- Meal time between 11:00am -12:59pm.
- Orders come through web and mobile
- Majority of meals from the regular and highlight category, only ~10% order a drink, dessert or side
- Slight indication of preference for high protein, low fat meals

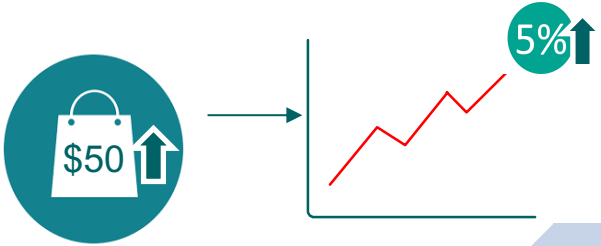
Opportunity: Have the sales team upsell to The Party Organizers

What is in their basket? The Party Organizers currently order the catering sets but are also picking up other à la carte items.

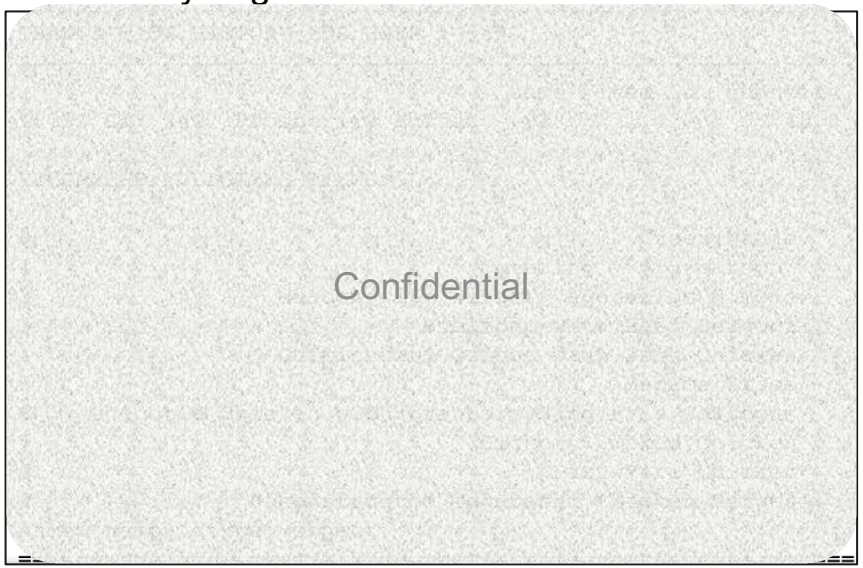
Confidential

Train/ Incentivise sales team to offer and upsell additional items as premium add-ons for the sets

- An increase of $50 per order ($1 per pax) could bring 5% uplift in sales.



- Knowing users in this cluster behave the same way, we can use a recommender system to guide upselling
- The following are examples of predictions of interest for users who belong in The Party Organizer cluster



Referencing the user/item interactions of many other users, a recommendation is made for a user based on the purchases of other users with whom he have bought same items before.

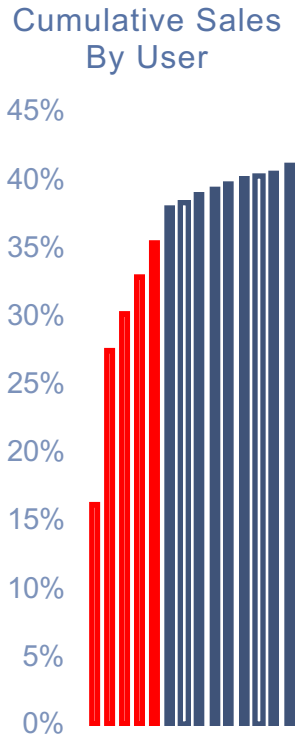
Opportunity: Capitalise group purchase behaviour of The Small Time Influencer

What is in their basket? The Small Time Influencers have their regulars but also take interest in **highlights** including :

Confidential

- **Expand range in Highlights to include dessert, drinks and sides for which take up is currently not high**
- **Offer promotions for group/combo purchase of drinks, sides and dessert e.g. 5 sides at $2 off**
  - 50% conversion in a group of 5 could bring 25% uplift in sales.

One more reason to better understand the customer; 35% of total revenue in 2018 is dominated by the below 5 users



User Id	Delivery address(es)	Monetary Value (2018)	Cluster
28126	Confidential	$460k	The Party Organizer
28133		$325k	The Party Organizer
24270		$78k	The Small Time Influencer
4727		$77k	The Small Time Influencer
40025		$71.5k	The Small Time Influencer

There is also potential to sign term contracts with big spenders.

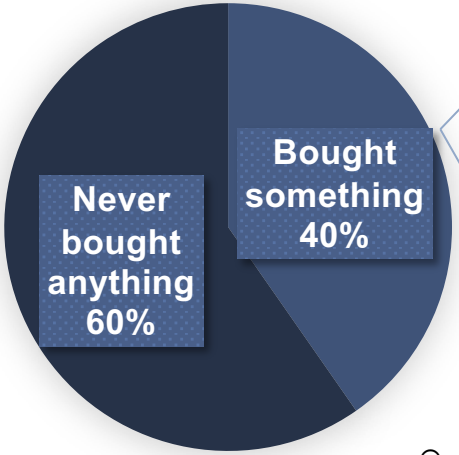
3

Narrow the Target
- Identify low hanging fruits

Using Recency, Frequency,
Monetary Value metrics

Of the 24k users who ever made purchases, 35% have potential to spend more, more frequently

Registered Users



Clusters	Recency (days)	Frequency (delivery orders)	$$ Value (sales)	No. of Users
Loyalist	364	24	1,695.50	3,253
Potential Loyalist	616	9	634	5,402
Hanging by the Thread	774	2	176	6,716
Say Goodbye	1203	1	28	8,814

- **Given limitations of marketing budget targeting The Loyalists and The Potential Loyalists who are more likely to respond positively, could bring a higher ROI in the immediate term.**

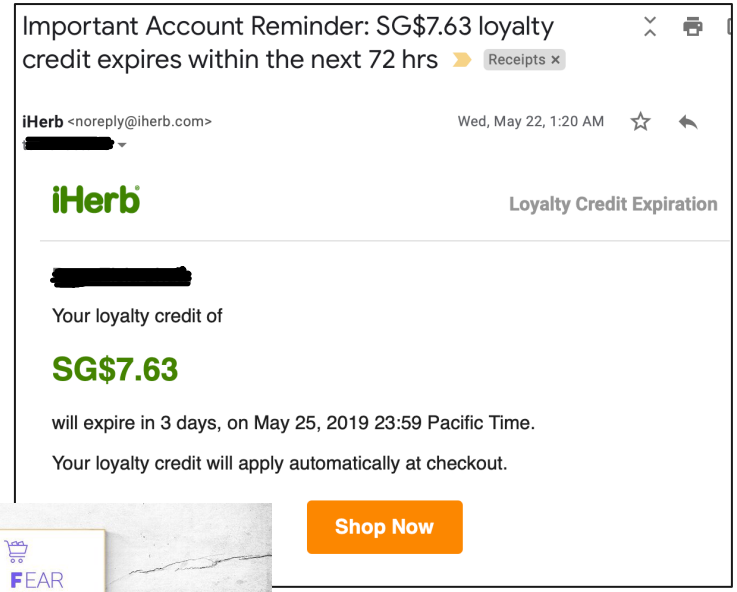
Number of days between each purchase is long

- Half of the users have not made another purchase in more than 2 years
- Transactions were dated from 2014-07-07 to 2019-12-01

Opportunity: Incentivise users to come back sooner

iHerb Loyalty Programme

- Part of the iHerb loyalty programme is a credit value equal to 5% of order value, excluding shipping charges. This credit is automatically applied on the following purchase and **expires in 60 days**.



No one likes to leave money on the table

- To The Small Time Influencer, this could be free delivery ($2.85) that he could shout about.

4

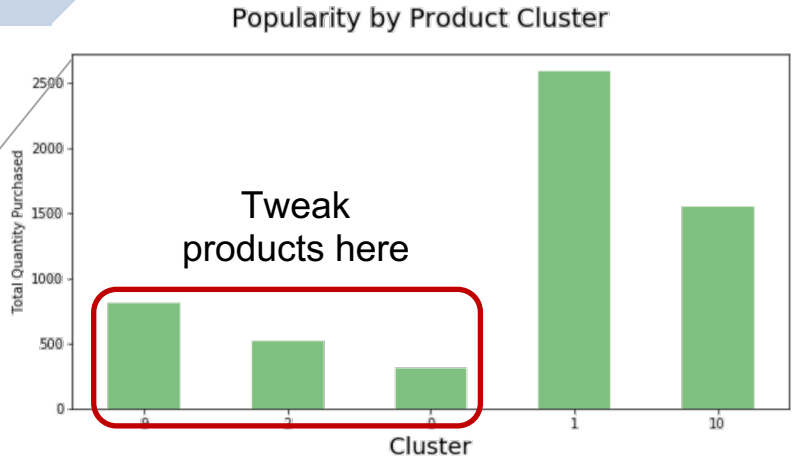
Capture More Value From Products

Segmenting products using K-Means Unsupervised Clustering

Clustering products by popularity, price, nutrition, temperature and category reveals 12 segments

- Products in clusters largely fall into the labelled categories
- 5 segments have prices between $11.60 and $12.70

Opportunity: Tease prices apart by tweaking products to form new, better option



- **Adapt products with popular ingredients used in other popular categories e.g. brown rice, olive oil, mushroom, goji berry to form 'Best of Both Worlds' premium category priced in the $13.50 range.**

Opportunity: Introduce bundle options for products in popular clusters to complete the price ladder

E.g. Bundle Local (popular)

Add Local sides to your Local meal for just $4.
Choose from

Confidential

5

Conclusion

Who, What, When, Where, Why, How?

Conclusion and Recommendations

Why sales dip?

Order value fell 35% y-o-y

What can be done?

Increase order value through understanding of customer purchase behavior and product value

How?

1. Upsell relevant add-ons to The Party Organizer
2. Capitalise on the group purchase behaviour of The Small Time influencer
3. Launch rewards to bring users back more often
4. Strengthen the price ladder

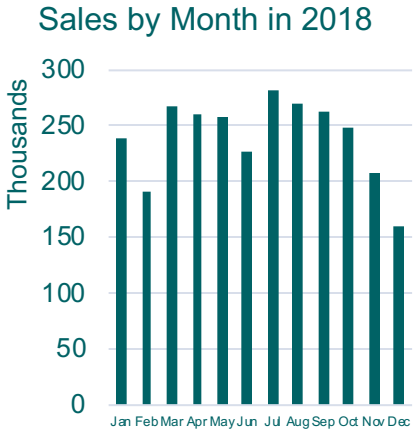
Conclusion and Recommendations

Who to target?

The Loyalists
and
The Potential Loyalists

Where to reach them?

- Sales Team
- Digital Marketing Channels



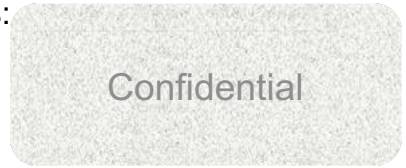
When?

- Launching promotions in Feb, to avoid seasonality
- Train sales team before the year end festivities

Thank you

Annex 1: Challenges in Data Cleaning

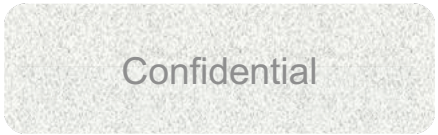
- Many null values for useful information
  E.g. Birthday (82% missing values) – even so, there were birthdays beyond current date
  Price of Menu items – also errors:



- Duplicate columns with inconsistent information
  E.g. delivery_fee and delivery_fees having different values:

	delivery_fees_amt	delivery_fee
4457	0	5500.0
7440	6000	0.0

- Extreme values
  E.g. Unit price of products:



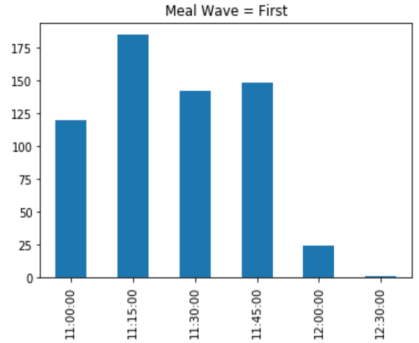
- For User's address, having users indicate home or office can be useful to source demographics information including household income, race(for exploring dietary preference) etc.

Drop down options/ hints/ rules in system can help eliminate inconsistencies

Annex 1: Challenges in Data Cleaning

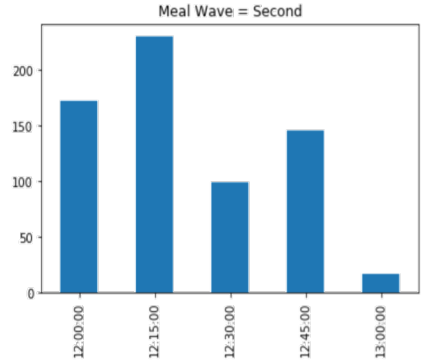
- Inconsistent definitions in data labels
  E.g. Time periods for each meal wave over lapped each other's

```
Meal Wave = First
-------------------------------------------------
11:00:00    120
11:15:00    185
11:30:00    142
11:45:00    149
12:00:00     24
12:30:00      1
Name: serving_date, dtype: int64
```



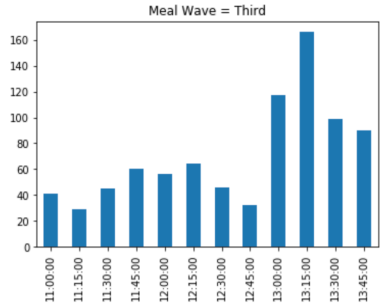
First wave peaks at 11:15am. Busy period is between 11:15-1145am. It will make sense to define first wave as 11:00 to 11:59am.

```
Meal Wave = Second
-------------------------------------------------
12:00:00    173
12:15:00    230
12:30:00     99
12:45:00    146
13:00:00     17
Name: serving_date, dtype: int64
```



Second wave peaks at 12:15am. Busy period is between 12:00-12:15am will make sense to define second wave as 12:00 to 12:59am.

```
Meal Wave = Third
-------------------------------------------------
11:00:00     41
11:15:00     29
11:30:00     45
11:45:00     60
12:00:00     56
12:15:00     64
12:30:00     46
12:45:00     32
13:00:00    117
13:15:00    166
13:30:00     99
13:45:00     90
Name: serving_date, dtype: int64
```



Third wave peaks at 13:15am. Busy period is between 13:00-1315am. It will make sense to define third wave as 13:00 to 13:59am.

Drop down options/ hints/ rules in system can help eliminate inconsistencies

Annex 2: Features used in Customer Segmentation

Definition	Binary
Whether user signed up with a coupon code	X
Average delivery fee user paid across all her deliveries	
Higher boundary of delivery fees user is willing to pay	
Whether user is purchasing for more than self, more than half the time	X
Average value of each delivery order by user	
Higher boundary of value of delivery orders by user	
Average menu item price paid by user	
Whether user ever given feedback	X
% of meals ordered that was low calorie, low fat, high protein, low carb	
% of meals ordered that was a chilled item	
% of delivery orders include item from this category	
% of delivery orders were ordered for this meal wave	
% of delivery orders were ordered on this platform	

Annex 3: Next Steps

1. Have users rate their food!
   - While feedback is almost never received from The Party Organizer, ~20% of The Small Time Influencers provide feedback on the delivery service or the food.
   - **Explicit ratings** of menu items could be collected to build a **recommender system**.
   - An **added service** of recommending menu items based on the user's past interactions with the products can be offered
     - **Simplifies ordering** for The Party Organizer
     - Meets **personalized needs** of The Small Time Influencer
2. Beyond exploratory data analysis, we could set out to deliberately separate users into a set number of clusters according to prior knowledge, to achieve business goals
   - The K-Means unsupervised clustering method allows us to set number of clusters (K)
   - Insights from the clusters can be applicable to develop innovation ideas accordingly