**Summary of data understanding for Instacart data**

All products : 49,688

All departments : 21

All aisles : 134

Total orders given : 3.2 million

Total products in orders : 32.4 million

* 30k products are distributed in 6 departments
* 34.7 % products are in top 5 aisles. 46% of products are in top 10 aisles.
* As the add to cart order is decreasing, reorder rate is also decreasing
* Top 2500 (5%) items appear for 69.96 % of total, Top 7500 (15%) items appear for 87.45 % of total . So top 2500 are considered as one model and next 5000 as next model. Top 15000 (30%) items appear for 95.1 % of total. So next 10,000 as one model and discard the rest
* Number of products in each order is peaking at 5 products per order. 62.2% orders have less than 10 products each. If we remove users who have orders > 30 days, still the percentage is same at 62.3%
* Days since prior order is peaking at 7,14, 21 and 30 days. >30 days can be considered as noise. Users who have more than 2-3 such orders can be considered inconsistent
* Number of users who have particular number of orders graph is decreasing one. Highest number of users have exactly 4 orders. 62% of users have less than 10 orders
* 4 products are ordered 1.6 million times
* 12.08 % of the orders in the Prior set has no-reordered products and the rest of them has at least one reordered products.
* 75% of products purchased are in six departments
* Reordered ratio for pets and alcohol is very high even though their % of total is very low
* Weekly users are the highest by number among total users followed by users who are between 24-30 days
* 0 and 1 have the highest day of the week in all orders
* Approximately 61% of products are reordered. Rest all are new
* Alcohol and babies are highly reordered in the range of 0-3 days. Rest all are reordered mostly in 7 days.