

Predicting Customer Habits



instacart

Rob Zhang



kaggle



aws



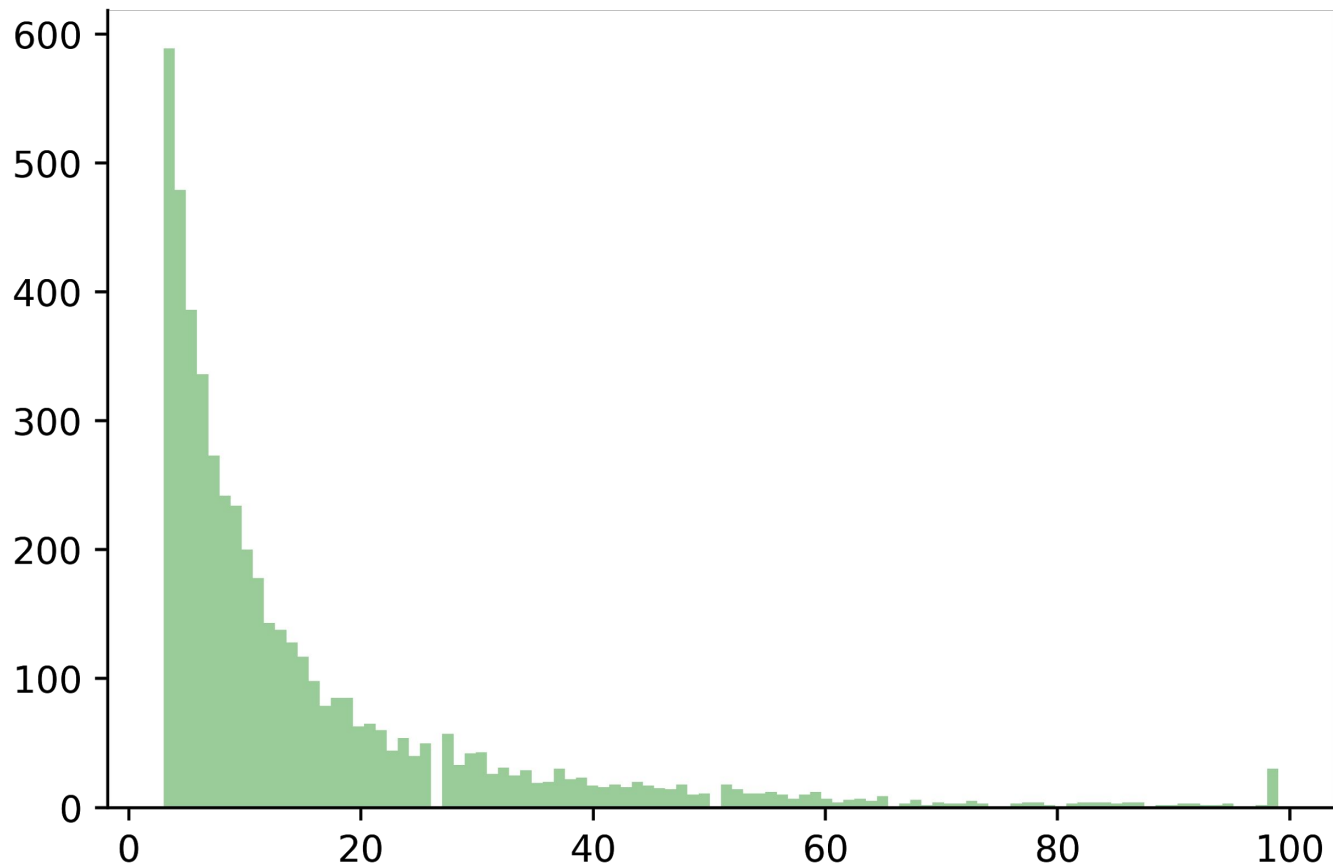
Departments
Aisles
Products
Orders
Order/Products

	user_id	department_id	aisle_id	product_id	order_id	order_number	order_dow	order_hour_of_day	days_since_prior_order	add_to_cart_order	reordered
16	35	3	128	48230	2562704	2	3	16	26.0	4	0
17	35	4	83	2846	2562704	2	3	16	26.0	5	0

- Current and Past orders predict future orders?
- Regular VS Non-regular customers?
- Target a group for promotions?



Total Orders per User



What can help us determine whether or not a customer will re-order in 7 days?

Feature Engineering

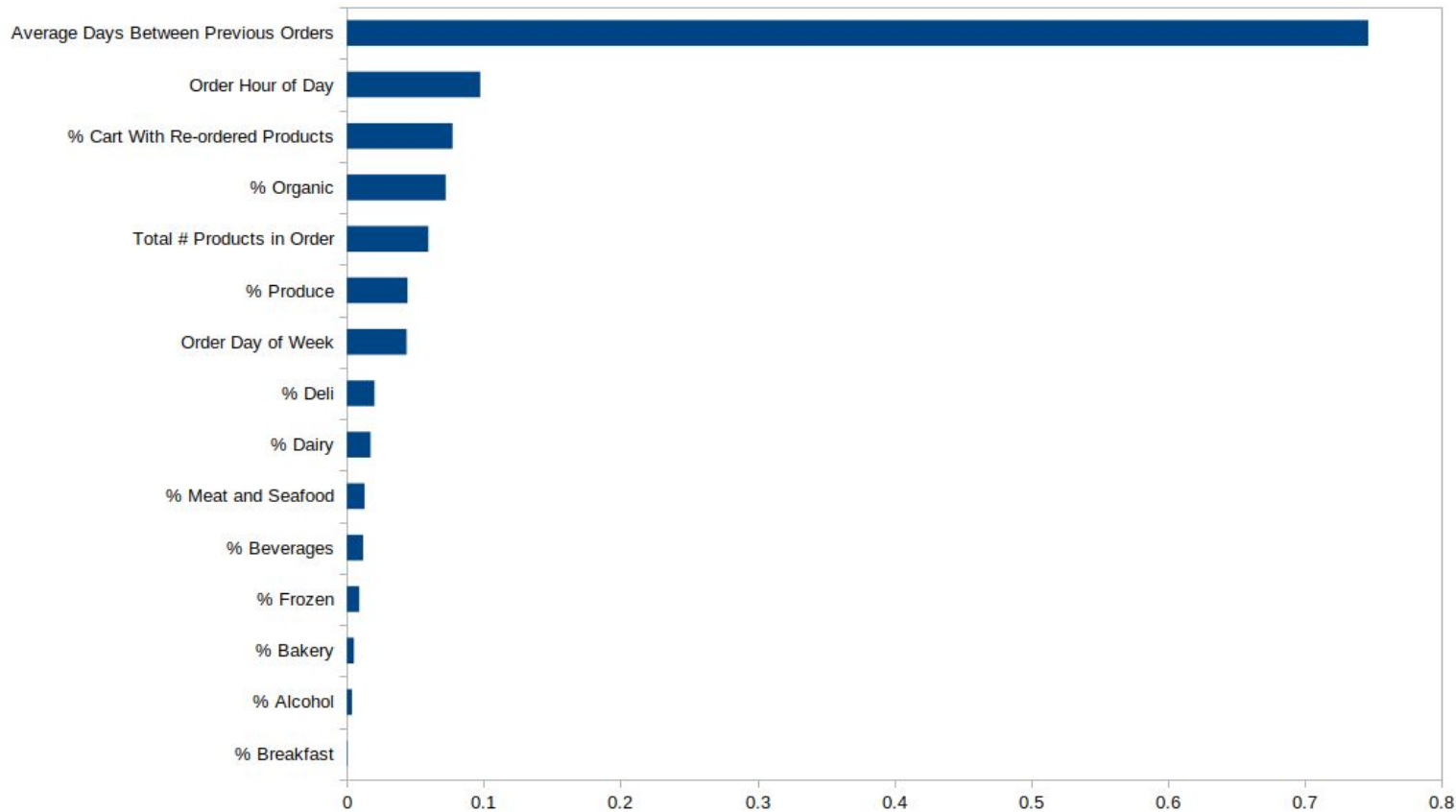
- % of cart that is:

Organic
Produce
Alcohol
Frozen
Bakery

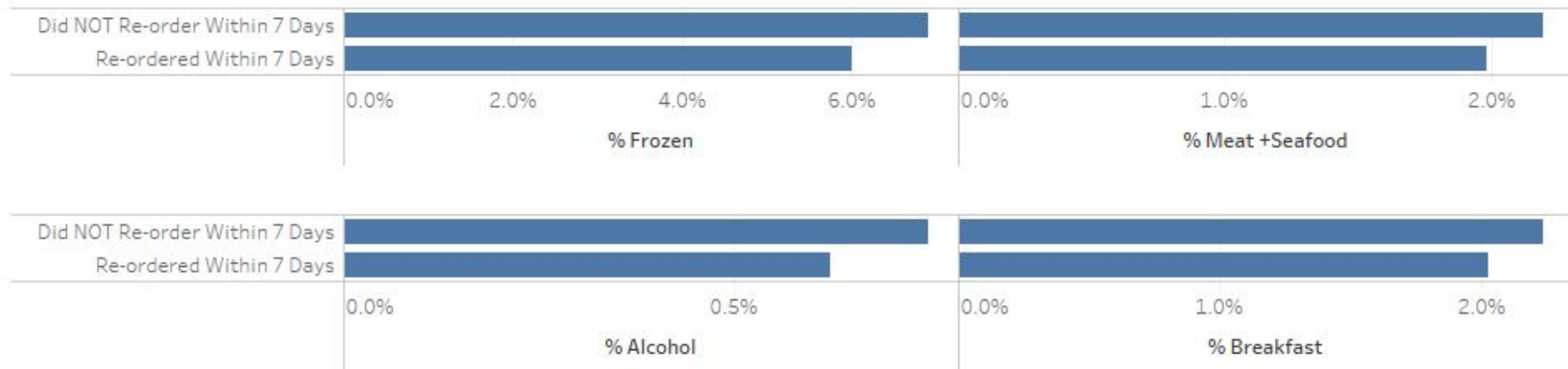
Meat+Seafood
Breakfast
Dairy
Deli
Beverages

- Total # products in each order
- Average # days between previous orders

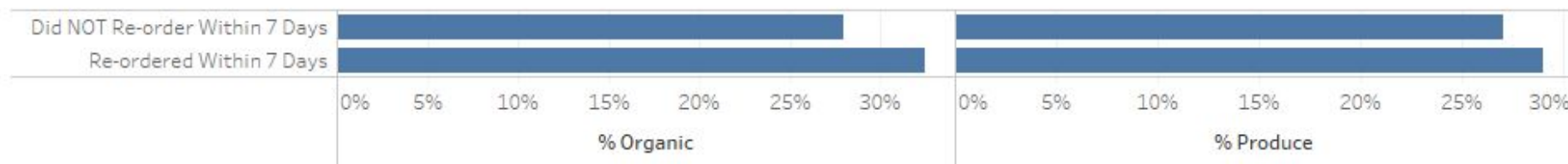
Logistic Regression Feature Coefficients (abs)

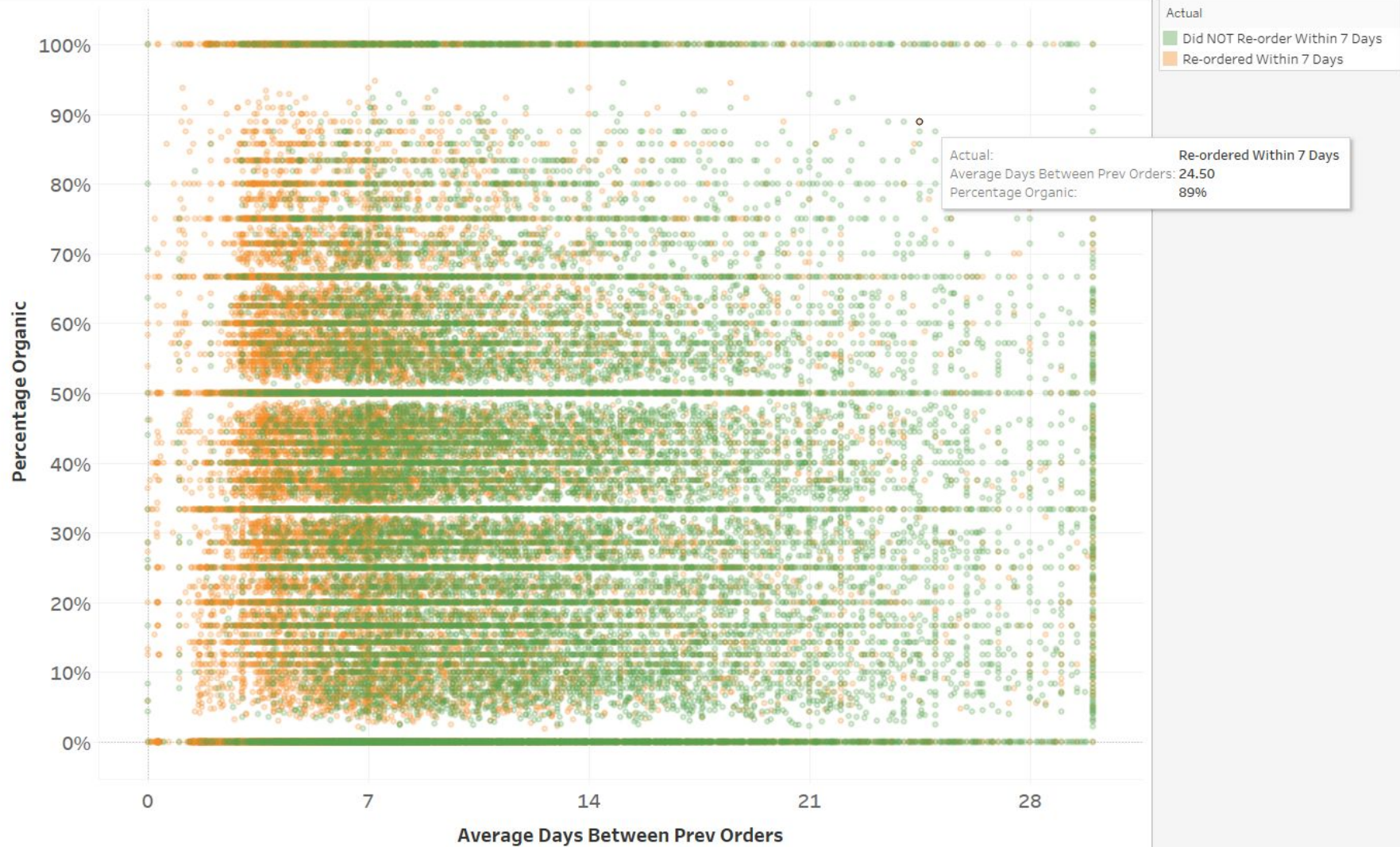


Categories hinting at NO re-ordering within a week



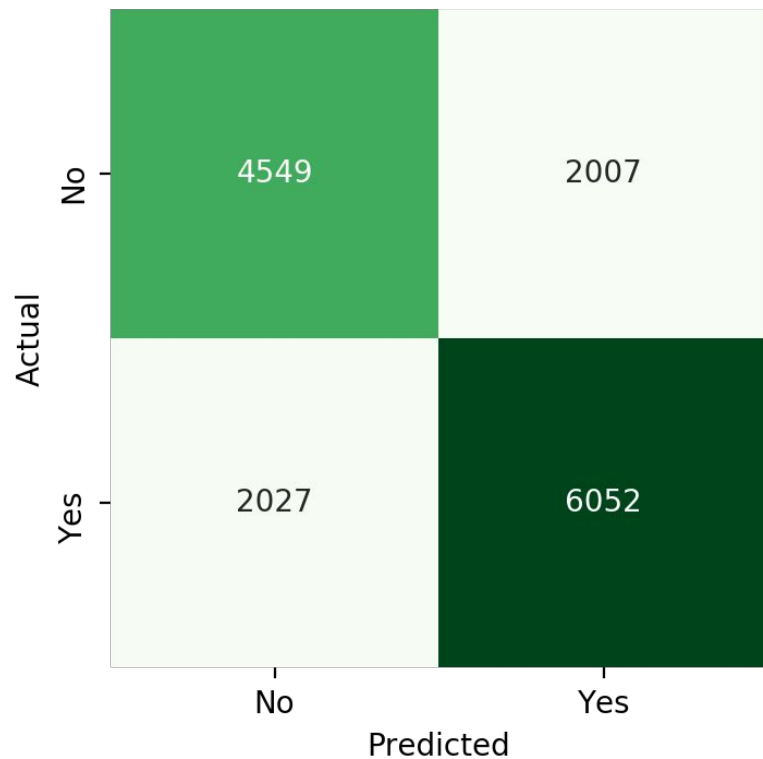
Categories hinting at re-ordering within a week





Model Results

Re-order within 7 days?



$$\frac{2027}{(2027 + 4549)} = 30.8 \%$$

Final Thoughts

- Past order behavior does a great job of predicting future behavior
 - Customers who re-order frequently purchase mostly the same products as those who don't
 - Predictive Model has potential for targeting select group to increase sales
-

—

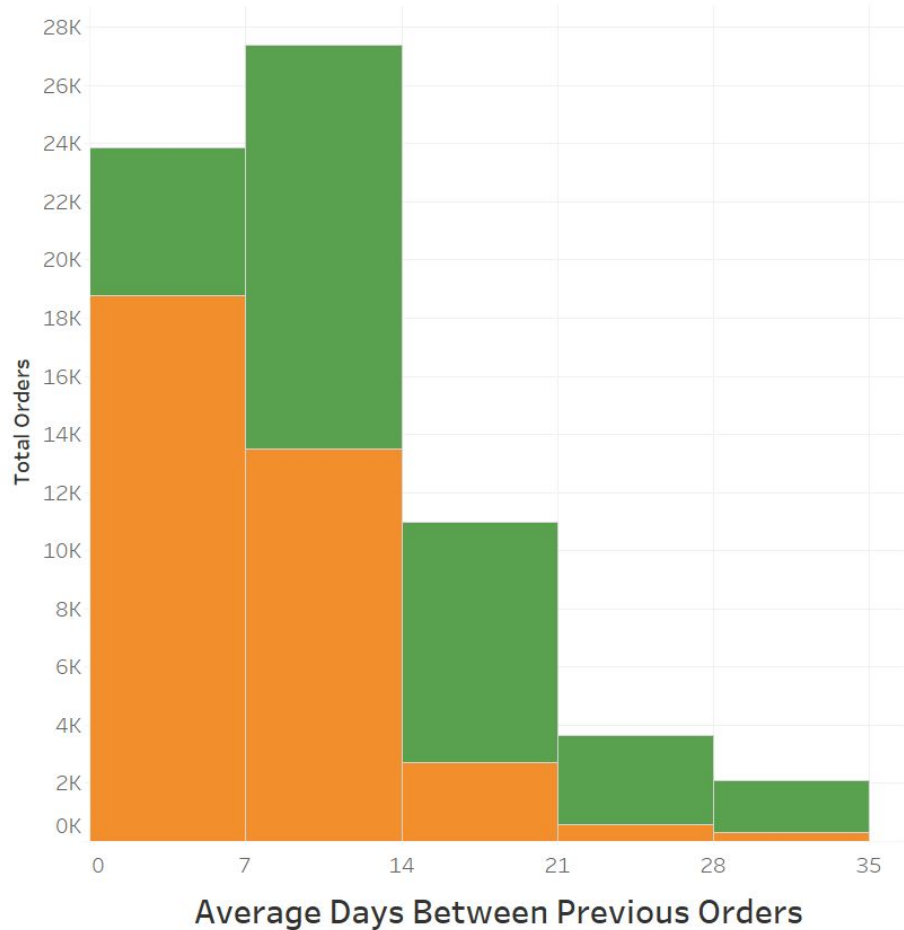
Questions?



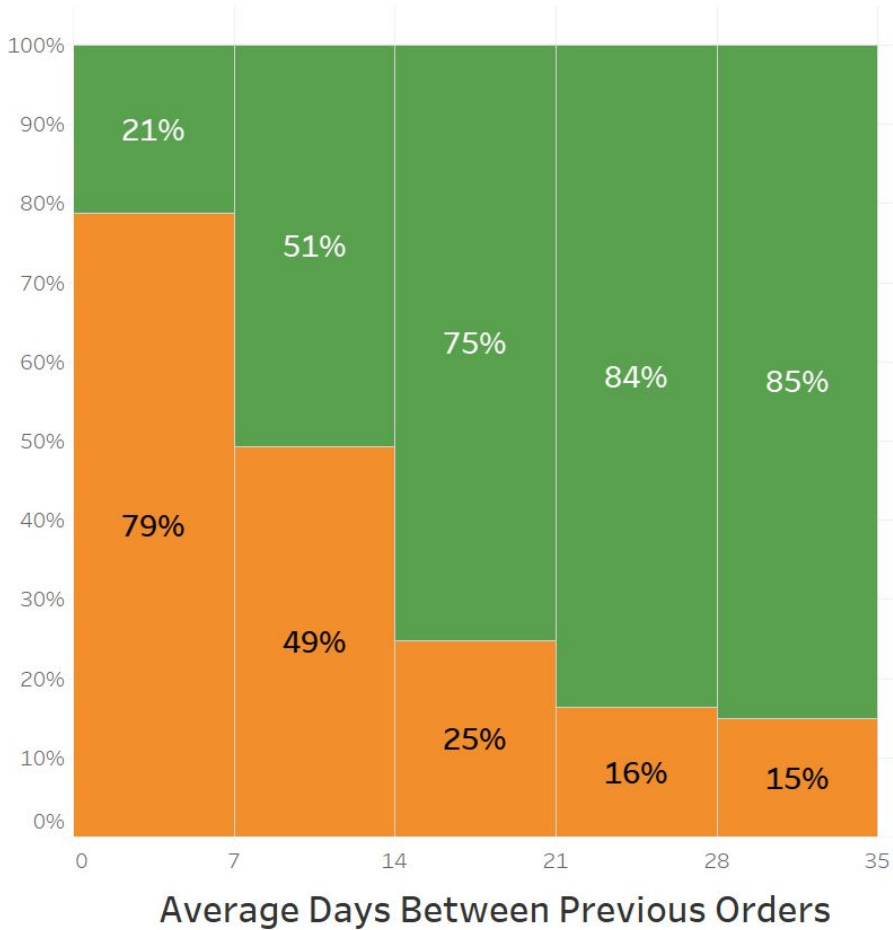
—

Appendix

Total Orders Grouped By
Average Days Between Orders



More Time Between Orders = Less Likelihood of
Re-ordering Sooner



Days Since Previous Order

