for value in ords\_prods\_merge["orders\_day\_of\_week"]:

if value == 0:result.append("Busiest day") elif value == 1:result.append("Second busiest day") elif value == 4:result.append("Slowest day") elif value == 3:result.append("Second slowest day") else:result.append("Regularly busy day")

if value == 0:result.append("Busiest day") elif value == 4: result.append("Least busy") else: result.append("Regularly busy")

most\_orders\_threshold = order\_counts.quantile(0.67)

fewest\_orders\_threshold = order\_counts.quantile(0.33)

if order\_counts[hour] > most\_orders\_threshold: return 'Most orders'

elif order\_counts[hour] < fewest\_orders\_threshold: return 'Fewest orders'

else: return 'Average orders'

[ords\_prods\_merge['max\_order'] > 40, 'loyalty\_flag'] = 'Loyal customer'

[(ords\_prods\_merge['max\_order'] <= 40) & (ords\_prods\_merge['max\_order'] > 10), 'loyalty\_flag'] = 'Regular customer'

[ords\_prods\_merge['max\_order'] <= 10, 'loyalty\_flag'] = 'New customer'

user\_avg\_prices = ords\_prods\_merge.groupby('user\_id')['prices'].mean().reset\_index()

user\_avg\_prices['spending\_flag'] = user\_avg\_prices['prices'].apply(lambda x: 'Low spender' if x < 10 else 'High spender')

ords\_prods\_merge = pd.merge(ords\_prods\_merge, user\_avg\_prices[['user\_id', 'spending\_flag']], on='user\_id', how='left')

user\_order\_frequency = ords\_prods\_merge.groupby('user\_id')['days\_since\_prior\_order'].median().reset\_index()

def order\_frequency\_flag(median\_days): if median\_days > 20: return 'Non-frequent customer'

elif median\_days > 10: return 'Regular customer'

else: return 'Frequent customer'

user\_order\_frequency['order\_frequency\_flag'] = user\_order\_frequency['days\_since\_prior\_order'].apply(order\_frequency\_flag)

ords\_prods\_merge = pd.merge(ords\_prods\_merge, user\_order\_frequency[['user\_id', 'order\_frequency\_flag']], on='user\_id', how='left')

AGE INCOME FAM STATUS

# Create profiling variables

def profile\_customer(row):

if row['Age'] < 18: return ‘Under 18’

elif row['Age'] >= 18 and row['Age'] <= 35: return ’18-35’

elif row['Age'] > 35 and row['Age'] <= 65:

return ‘35-65’

elif row['Age'] > 65:

return ‘65+’

elif row['Age'] > 50 and row['Income'] < 50000 and row['Number of Dependents'] == 0:

return 'Senior Citizen'

elif row['Income'] > 100000:

return 'High Income'

elif row['orders\_day\_of\_week'] in [0, 1]:

return 'Weekend Shopper'

elif row['order\_hour\_of\_day'] >= 5 and row['order\_hour\_of\_day'] <= 9:

return 'Early Morning Shopper'

else:

return 'Other'

assigned states to regions with this list:

<https://simple.wikipedia.org/wiki/List_of_regions_of_the_United_States>

Key Questions

● Thesalesteamneedstoknowwhatthe busiest days of the week and hours of the day are (i.e., the days and times with the most orders) in order to schedule ads at times when there are fewer orders.

● Theyalsowanttoknowwhether there are particular times of the day when people spend the most money, as this might inform the type of products they advertise at these times.

● Instacart has a lot of products with different price tags. Marketing and sales want to use simpler price range groupings to help direct their efforts.

● Aretherecertain types of products that are more popular than others? The marketing and sales teams want to know which departments have the highest frequency of product orders. Note: Instacart is a real company that’s made their data available online. However, the contents of this project brief have been fabricated for the purpose of this Achievement. Page 1

● Themarketing andsales teams are particularly interested in the different types of customers in their system and how their ordering behaviors differ.

For example:

○ What’sthedistribution among users in regards to their brand loyalty (i.e., how often do they return to Instacart)?

○ Aretheredifferences in ordering habits based on a customer’s loyalty status? ○ Aretheredifferences in ordering habits based on a customer’s region?

○ Isthere aconnection between age and family status in terms of ordering habits?

○ Whatdifferent classifications does the demographic information suggest? Age? Income? Certain types of goods? Family status?

○ Whatdifferences can you find in ordering habits of different customer profiles? Consider the price of orders, the frequency of orders, the products customers are ordering, and anything else you can think of