J. Crew take-home assessment scenario responses

Exercise 1:

SQL: What are monthly sales by customer type? Customer type defined as Single (1 family size), Couple (2 family size), Family (3-5), Large Family (6+)

Analysis Q Why would J.Crew care about this? If Madewell's sales is high and increasing for Single family types, how should Madewell think about that? What else would you want to know?

My Answer: This is a basic function that would allow J.Crew/Madewell to track which types of customers are making purchases, when those purchases and being made, how many purchases each group is making, and the quantity of items bought and the amount of money spent by each group. This paints a picture of sales behavior which can be leveraged for all sorts of things: marketing campaigns, sales predictions, supply chain forecasting, etc.

If Madewell's sales are high and increasing for Single family types, it shows us that the current marketing, onboarding, product quality & selection and price point are all *working* for this family type right now. Madewell could lean into this by increasing their marketing toward Single people, aiming to cast a wider net and onboard more new customers in this group. It would be important for Madewell to consider which products are most popular with this group and ensure their supply chain is equipped to handle the greater demand that would be the goal of such a marketing campaign. This could be investigated by tweaking the SQL query to include the specifics of which type of items are being purchased by this group over time: their quantity, cost, net profit to the company, etc. We could analyze the sales trend data to uncover insights into product popularity and quantify the bottom line for J.Crew/Madewell. In this way, we could uncover the types and quantity of products the company needs to make to keep up with demand, use data to inform future product development and identify marketing and J.Crew/Madewell crossover opportunities in this customer base.

On the other hand, Madewell could consider the disparities between the Single family type and the other family types; perhaps there is market potential to capture more customers who fall into one of the other family types. I would start by comparing the monthly sales behavior of each group and identify the group(s) with the *most* similarity in purchasing behavior to the Single family type. Then, I'd investigate what the disparities are: do we simply have fewer customers of this group? Are they buying the same products or different products? Would it be reasonable for Madewell to increase production and marketing of the products this group *is* buying?

Exercise 2:

SQL: What is the average cart value by acquisition year cohort (e.g., what is the average cart value for customers acquired in 2019, 2020, 2021, etc.)

Analysis Q: You find that average cart value is going down for later acquisition cohorts v. earlier — What are some hypotheses for why this is happening and what analysis would you run to test those hypotheses?

My Answer: Some hypotheses:

1. Later acquisition cohorts may have less disposable income; e.g. they could be recent college grads vs. more-established working professionals from earlier cohorts. An analysis of the average cart value *over time* for each cohort would reveal any trends in support of this hypothesis. For instance, if we found that the average cart value of earlier cohorts increased over time, it would support that each cohort tends to buy a greater value of product over time. A line plot with average cart value over time, grouped by year cohort and normalized for the amount of time since signing up, would be a really nice way to visualize this. Analyzing demographic data for each cohort would also be helpful for this analysis, if we can get it. I'd ideally want to look at age, education, profession, income bracket, location (and cost of living in location), etc to test out this hypothesis and gain insights on customer behavior.

2. Later cohorts may be biased toward spending *less* at the same time, but how often are they buying? Does this group tend to purchase a lower cart value at a time, but make those purchases more often? How does this behavior (and the net profit to the company) compare to earlier cohorts? I would run an analysis to compare average number of cart purchases per month for each cohort to analyze purchase frequency behavior. I could multiply the avg frequency of purchases by the avg cart value to evaluate how each cohort is contributing to company profit. These figures could be compared between the cohort groups over time to evaluate any real differences in behavior. If we found, for instance, that later groups **do** prefer to purchase less at a time but more frequently, that would give the business valuable insight into how we may be able to increase the purchase volume for this type of customer--would offering free shipping on a lower cart value encourage this type of customer to buy smaller quantities even more frequently, e.g.? We could run an A|B test to find out.

3. If the frequency of purchases is heavily influenced by the cohort year (i.e. earlier cohorts are buying less frequently as time goes on vs. new cohorts), a change to the overall cost of the products over time could explain the observation. An analysis of average product cost along with average cart value over time could indicate whether the observation that later acquisition cohorts are spending less in each purchase is a reflection of changes to the cost-structure over time. I would couple this with an analysis of purchase frequency over time for each cohort, to evaluate the trends in cohort purchasing behavior as a function of how long they've been a customer, similar to in point 1, above.

4. Does our data cut off mid-year, thereby telling an incomplete story about the most recent cohort? If cart values go up significantly around the Holidays, but we don't have that data for the most recent cohort year, the simple explanation could be that we're not comparing apples to apples and need to exclude the current year's cohort from the analysis.

Exercise 3:

SQL: For each customer, what is the difference between the value of their first order and the second order?

Analysis Q: Is this a useful metric? Why or why not? Feel free to suggest alternative metrics that could be interesting here!

My Answer: I'd say this *might* be a useful metric, depending on the question we're trying to answer. If we want to know whether users spend more, less or the same amount on their second order compared to their first order, this is the right metric for that! This could help us understand user behavior to drive business decisions, and could be a concrete metric for A|B Testing, where we might assess whether a new marketing campaign, product launch, UX interface, brick-and-mortar store feature etc. encourages new users to spend more money on their second purchase, and thereby might be more likely to be happy repeat customers.

We might also look at the amount of time between a customer's first and second purchases and the proportion of new customers who make a second purchase as an alternative way to assess our success with making new customers repeat customers.

One of the problems with this metric is that it is most useful for new customers; we can use the data from the older customers as a historic reference. It might be interesting to look at these same metrics but with respect to the customers' *last two* purchases as a means of analyzing shopping behavior for existing customers. In this case our lens would probably be focused on strategies to increase profits through customer retention.