Dear Dr. Toboggan,

Let me tell you about how an outstanding company can become even more outstanding and profitable. Maybe you've heard of the company. It's called Beautify, and it's yours. You sell products that can be loosely labeled as either oral hygiene products or hair products. Currently you are flourishing. Over the six month period from the end of July 2014 to the end of January 2015, Beautify sold over 10,000 oral hygiene products and over 11,000 hair products. Surprisingly, more unique consumers bought oral hygiene products than hair products over that time. To explain this discrepancy, we look at how many products each individual consumer buys at a time. As it turns out, the average consumer buying Beautify hair products purchases more per week than their oral hygiene buying counterparts.

To make Beautify thrive, I have three major recommendations for you stemming from three key insights. 1. Most of your consumers buy both types of products. Of your 1200 consumers, just under 1000 of them have bought both products at some point. This tells me consumers tend to love both products as opposed to just the hair or just the oral hygiene products. Given this, I'd recommend adopting a marketing strategy of product bundling to further incentivize customer loyalty. Give a discount to consumers buying both types of products. This will increase consumer satisfaction while increasing sales at the same time. Eventually, the increased demand will allow you to decrease the average production cost per product by scaling your operation.

2. The above analysis has only looked at Beautify in a vacuum. However, to conquer the market, we'll need to take a broader view. Amongst the ten categories falling under the Health and Beauty banner, the two Beautify product categories accounted for (roughly) an astounding 30% of the total products sold. Beautify alone sells more hair products than all other brands sell personal health care products combined. Beautify can take advantage of this in your advertising. Therefore, I recommend letting people know how much people love your products compared to your competitors in your advertisements and commercials. Find celebrities who already love your products to endorse them. Their natural excitement will show through sincerely.

3. Previously, we have talked about *how* to sell and market your products, but an equally important issue remains, *where* do you sell your products? Of the total visits consumers made to non-restaurant establishments, the most frequented store, by far, was Walmart. Intuitively, contacting Walmart to sell your products there makes sense to reach the broadest audience possible. Similarly, Target and Kroger are good establishments to target as the second and third most visited establishments with Kroger being to most frequented grocery store. However, outside of those three, most of the remaining stores had a roughly similar number of total visits. So, who would I recommend contacting to sell your products? Ralphs! The people who shopped at Ralphs visited the store over two and a half times per week. Every other store didn't even reach two visits per week. This shows loyalty to Ralphs and to the brands Ralphs carries. Furthermore, Ralphs is owned by Kroger. Once Beautify's products' excellence shows through and sells well at Ralphs, it won't be difficult to approach Kroger.

Thank you for reading our insights and we look forward to having the opportunity to work together.

Best,

Philip Klein-Rodick

InMarket Methodology

First, I'd like to clarify that due to the ambiguous nature of the column "time_week", I'm assuming that the data given in both tables is for a week long period as opposed to the specific date listed. This is supported by the fact that the dates vary by week long intervals. I also gave approximations such as "over 11,000" instead of the exact number (11,247 in this case) most of the time because the prose flowed better as a letter to the CMO and is less technical as requested. The actual numbers are listed at the bottom in order of their use in the letter above.

To analyze the data, I initially attempted to cross reference the two tables to gain insight about what stores the consumers were buying specific products at by joining the tables using the "time_week" and "device_id" attributes. However, this approach quickly appeared to be unfruitful. For example, on time week 10/26/14, we know the consumer with device id "0VxSq/Hs3/m80M8VjkhCVkD7Esk=" visited a pizza hut from the store_visits table. However, we also know the same consumer bought quite a variety of goods notably not sold at a pizza hut from the store_purchases table. So, we have no way to know where these items were purchased. Similar problems occur for consumers who visited multiple stores in the same week. We have no way of knowing which purchase was made at which store.

Thus, at this point, I decided to approach the two tables independently from one another. The store_purchases table easily gave insights into the two products sold, such as how much of each was sold over the given 6 month period, how many items were sold per week on average, how did they sell compared to other items of the same item_category, how many of each were sold at a time, etc. I used this information to make two recommendations for Beautify's CMO.

The other table proved more challenging to attain insights from. This table gave no direct insights into the products themselves. However, we could still find out about the different chains that would be good targets to sell our products at. With this in mind, I focused on the stores that would sell Beautify's products, that is big box and grocery stores. I looked at information such as the total visitors and average number of visitors per week. Additionally, I independently researched which stores were associated with each other.

Exact statistics used for the executive summary:

-11,247 oral hygiene products were sold

-10,192 hair products were sold

-1,109 unique consumers bought oral hygiene products

-1,067 unique consumers bought hair products

-1200 consumers total bought either oral hygiene or hair products from Beautify.

-976 shared consumers

-On average, an individual buying hair products bought 2.37 per week

-On average, an individual buying oral hygiene products bought 1.96 per week

-Beautify sold 21,439 products out of 75,640 health and beauty products; or, Beautify sold a total of 28.35% of all health and beauty products.

-Walmart had 18,489 total visitors

-Target had 9373 total visitors

-Kroger had 5289 total visitors

-The rest, outside of Fred Meyer and Vons, had 1000-1500 total visitors

-Ralphs had 2.55 visitors per customer per week

Data Insights Challenge – Part 2

1)
```sql
select device_uuid from table_2
where home_state = 'FL'
```

2)
```sql
select device_uuid from table_2
where home_city = 'Miami' and home_state = 'FL'

/*
Note, specifying home_city isn't enough as there are other Miami's, ie Miami, OH
*/
```

3)
```sql
select device_uuid from table_2
where birthdate between '1983-01-01' and '1998-01-01'
```

4)
```sql
select device_uuid from table_1
where visited_store_1 = 'TRUE' and visited_store_2 = 'FALSE'
```

5)
```sql
select a.device_uuid from table_1 a join table_2 b on
a.device_uuid = b.device_uuid
where parent = 'TRUE' and visited_store_2 = 'TRUE'
and home_city like 'G%'
```

6)
```sql
select home_city, count(home_city) as population from table_2
group by home_city, home_state
order by home_city desc

/*
Note , some cities may seemingly show up twice if there are two cities with the same name
in different states.  Ie, Miami, OH and Miami, Fl will both have a Miami entry.  The
question specifies "each city" implying the populations should not be combined into a
single Miami entry.
*/
```

7)
```sql
select top 10 visited_store_1, visited_store_2 from table_1
order by income desc

/*
Note, my query assumes no more than 10 people are tied for having the top ten highest
incomes.  In the case of 2 or more people tying for 10th, my query would effectively
randomly pick the 10th person listed.  The "with ties" clause or additional tie breaking
parameters can be added in this eventuality */
```