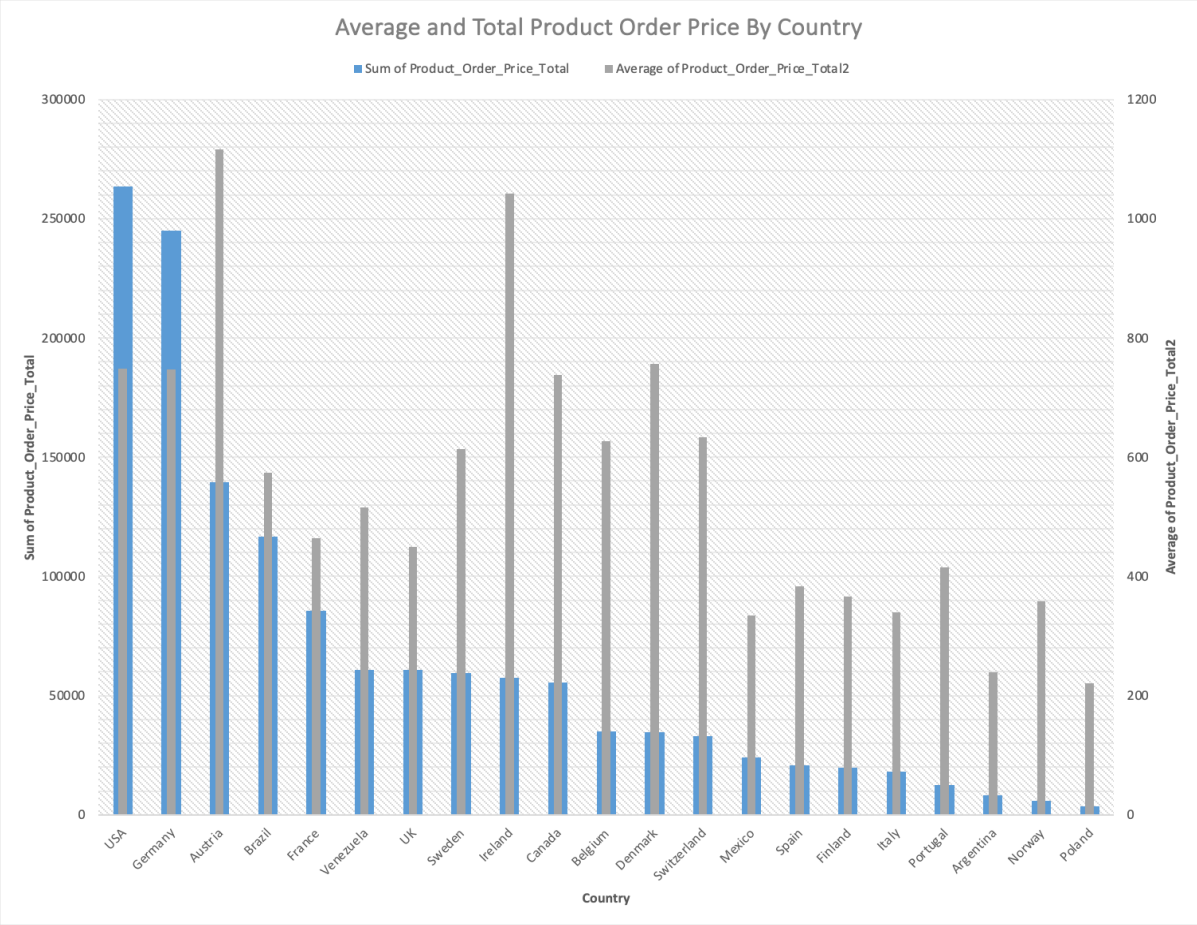
T5 - Dataset Evaluation

Team Information:
Team Name: Team B-4
Instructor Name: M. Guggemos
Section Number: Section (11:15 AM)

Members of Team B-4:
1. Nate Huffman - nmhuffma | nmhuffma@asu.edu
2. Maanas Lalwani - mlalwan1 | mlalwan1@asu.edu
3. Dayton McCann - Djmccann | djmccann@asu.edu
4. Ramiro Morales - rmoral44 | rmoral44@asu.edu
5. Samrat Serikkaliyev - sserikka | sserikka@asu.edu
6. Andrew Vinci - Acvinci | Acvinci@asu.edu

Visualization of Data Subset(1): Product Order Price Total



Data Subset: Product Order Price Total

Why Product Order Price Total?

- Price is an important determinant of increasing sales because it is the baseline from which sales are determined. Total price will allow us to see which country orders the most in total, while average total product order price will allow us to see which country orders the most per order.

What's the story?

- This graph visualizes the average and total product order price for each country. We can see that, indisputably, the United States and Germany have the highest total of product orders by price when summing all orders made by each country. Yet, we can see that they do not represent the highest order total by price on average. This means that even though the United States and Germany buy the most product, they do not buy the most product on average (per order), that title belongs to Austria and Ireland. This tells us that to fully maximize our profits, it would be beneficial to increase order frequency from countries with a high average product order by price, and increase individual order amounts for countries with a high total product order by price.

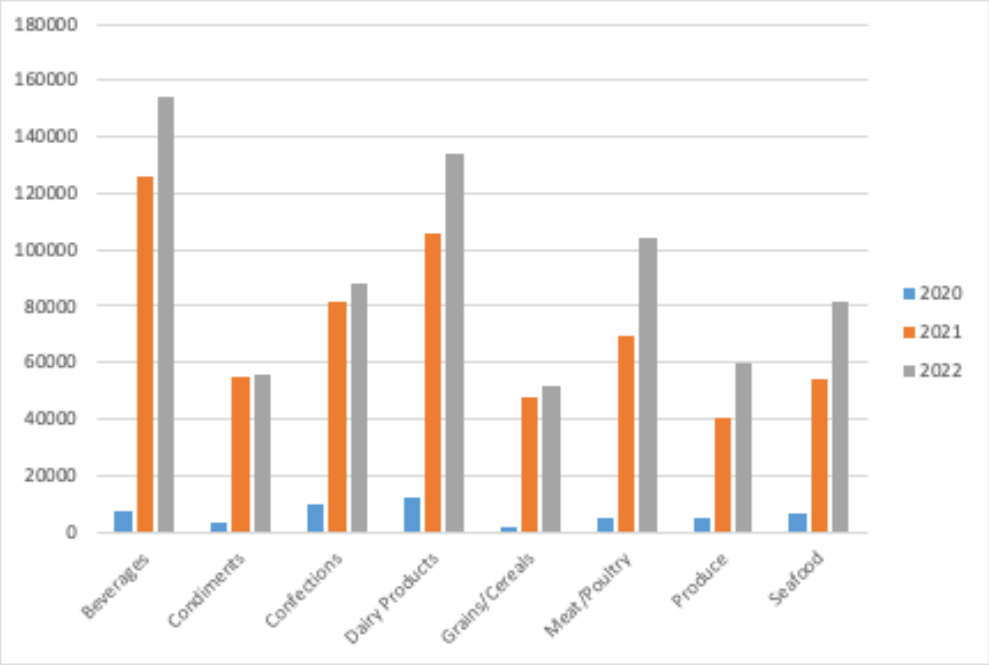
What subsets can we link to this?

- This visualization is best paired with order quantity by country, which would allow us to also visualize how each country utilizes their orders to fit their product needs, how many times the country orders, and how much product each country orders.

What don't we know that we should know?

- To fully understand the best way to increase profit and decrease costs, it is important to know the freight-out costs to fully visualize the true cost and profit that doing business with each country entails.

Visualization of Data Subset(2): Product Category



Data Subset: Product Category

Why Product Category?

- As a team, we chose to analyze the yearly sales of product categories as it will help our business understand what category is being sold the most and which category do we need to improve our sales in. This visualization helps us understand that from 2020 to 2022, the sales by product category have been increasing which is a good sign for our company.

What's the story?

- From this graph, we can conclude that people love spending on beverages and dairy products. On the other hand, we can see that grains/cereals and condiments are sold the least. Another important thing to keep in mind is that the data for 2020 is for 2 months, i.e. November and December. The data for 2022 also consists of the first 9 months only so we cannot give any proper conclusion.

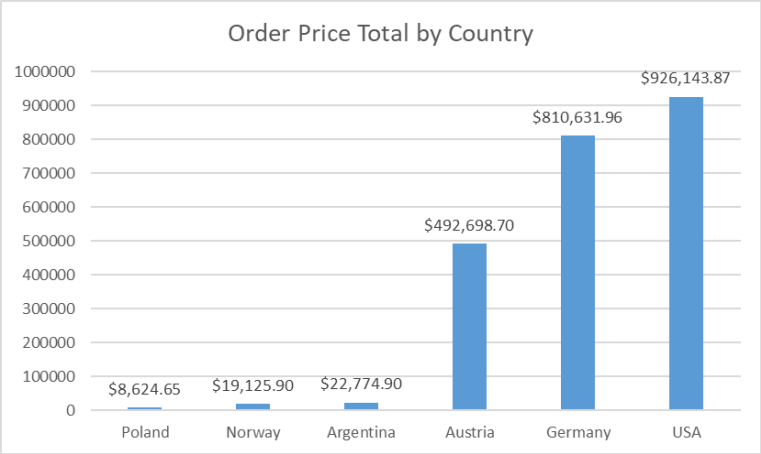
What subsets can we link to this?

- We will need to take the average spendings of each category to see if our insights change or not. We can also include countries in this specific graph but that would be difficult for users to read.

What don't we know that we should know?

- Important information that could be useful for increasing profit while cutting costs is knowledge of the relevancy of different products in different seasons or cultures. The most we can do is assume based on sales data.

Visualization of Data Subset (3): Customer Country



Data Subset: Customer Country

Why Customer Country?

- We chose to analyze the country of customers since understanding which countries are placing the most expensive orders, the greatest amount of orders, and the most frequent orders is essential to increase sales and decrease costs.

What's the story?

- The graph shows the countries with the three highest and lowest order price totals between, between November of 2020 and September of 2022. A customer's country shows us which counties are spending the most and least amount of money on our products.

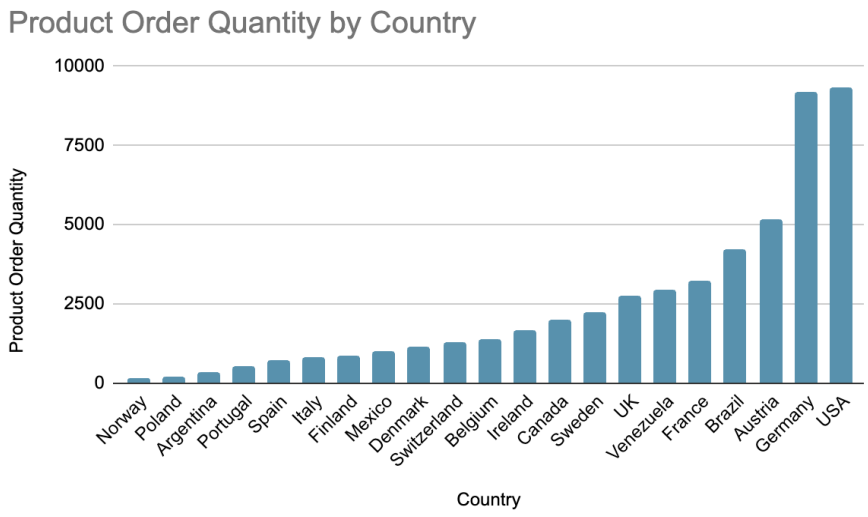
What subsets can we link to this?

- Linking customer's country to product type shows us where certain products are most popular, narrowing this scope by linking country to product name shows us which brand names to prioritize in certain locations, linking country to product unit price shows us the average price that different countries are willing to pay per product, and linking country to product order quantity shows us which countries are purchasing the greatest amount of our products.

What don't we know that we should know?

- Important information that isn't provided in the given data are the different shipping expenses and fees required to export our products to each country.

Visualization of Data Subset(4): Order Quantity



Data Subset: Order Quantity

Why Order Quantity?

- Another aspect we chose to analyze was the order quantity that each country purchased. This is essential because we need to see which countries want the most in quantity to keep promoting and supplying efficiently to those countries and to also focus on the ones that aren't buying as much in quantity and try to promote the products more in said countries.

What's the story?

- There is a major jump from Austria to Germany and the US showing that those two countries are our biggest buyers and both exceed 9000 units of products sold.

What subsets can we link to this?

- To dive in deeper with this aspect we can find the amount of each product that the customers buy and decide which one should be maximized in production based on which is the most popular and to change or revitalize the ones that aren't selling as much.

What don't we know that we should know?

- We don't know individual customer behavior which could provide useful information as to how much a customer orders.

Visualization of Data Subset (5): Month

Data Subset: Month

Why Month?

- We also decided to analyze the impact each month has on order price total. This specific subset of data is better than if we were to use data by day or year for calculating order price total. This narrows it down to a much more simplified way that provides not too much information compared to if the day was being interpreted and not too little information compared to if the year was only being analyzed.

What's the story?

- The story that the visualization shares with us is the time of year most spending occurs across all customers. This data helps provide a good idea for what time of year there are the greatest amounts of orders and what time of year carries the least amount of orders. The more the order price total, the more essential it is to increase sales and decrease cost, but the opposite could be said if the order price total is low. From this subset of data, we know that from months May - August has the highest order price totals and for months October - January, they have the least amount of order price totals.

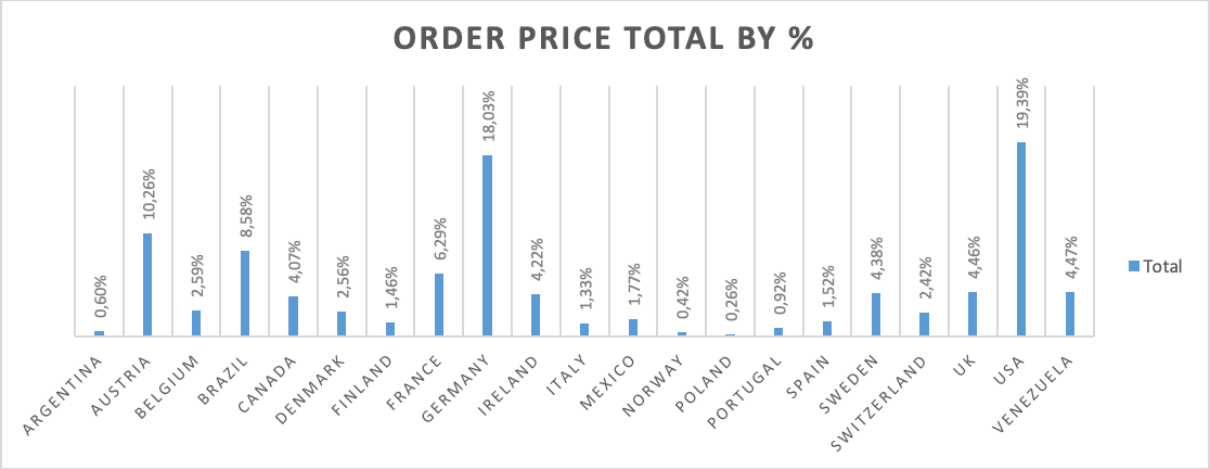
What subsets can we link to this?

- Another data set that may be important to link to this data set is the year, that way we could have each month's order price totals for the years counted in this data set to give an even more descriptive analysis.

What don't we know that we should know?

- What we don't know is that November and December also include order price totals from the year 2020. No other month accounts for order price totals from 2020 except for November and December which is what we would not know just by looking at the graph.

Executive Summary



Why these five subsets?

- The reasoning behind choosing the five analysis areas in the given statements relies on comprehending and optimizing numerous aspects of a business's decision-making and operations. Every department has a distinct function in assisting the company in making decisions that will boost revenue and cut expenses.

Our findings:

- By looking at the countries where clients are from, we can see that the USA, Germany, and Austria are the most profitable countries for the company. Combining this knowledge with our findings that the most expensive purchases are made in the May-August period, and the highest selling products are beverages and dairy products, we know to heavily market these products during the given time period to the most profitable countries. This will lead to a major boost in revenue while allowing us to cut costs in areas where we won't profit as much, such as the winter months in countries like Poland, Norway, and Argentina.

Working Data Set:

https://docs.google.com/spreadsheets/d/1gfzjAfG-loWqi92DxqML0AjRZ1U8Tpyr/edit#gid=6708480

Conclusion

To summarize, each of these specific fields offer a comprehensive viewpoint on consumer behavior, product efficacy, and seasonal patterns. The company may make data-driven decisions to boost sales and cut expenses by connecting various datasets and taking into account variables like order quantity, product categories, and seasonal fluctuations. Nevertheless, the declarations clearly recognize that more details are required in order to completely comprehend the costs and profits related to conducting business with each nation, such as shipping prices and fees.

Keep examining and improving the products offered in the most and least well-liked categories. To encourage growth, allocate resources towards product development and marketing tactics for segments with lower sales.