

University of Colorado Boulder

French C2C E-Commerce Platform: User Engagement Analytics

By: David Hebert, Jamie Liu, Jonathan Dressel, Julia Jungquist, Renee Gagne

MSBX5405 Fall 2020 // Keith Sparling

I: Overview

Background: As the world continues to become increasingly interconnected and dependent on the internet, consumers have access to more products—and information about those products—than ever before. The wake of a global pandemic in 2020 has changed consumer habits and encouraged online and economical shopping. According to Lyst's annual report, search words pertaining to used, vintage, and secondhand products increased by 104%¹. In October of 2020, a new magazine *Display Copy* was first printed, which exclusively features items found secondhand. As buying used items becomes a streamlined habit among consumers, there is a growing window of opportunity for individuals interested in purchasing and reselling items online, as well as for digital platforms that operate on the premise of user resale. One specific business model selling used items online is known as Customer to Customer (C2C), where users sell and trade goods to other customers through using a third party platform². Consumers' interest in buying items online creates a need for better understanding user habits as well as ways to expand online engagement.

Data: Our study investigates a set of 3 data files downloaded from Kaggle, which was collected from a French e-commerce C2C platform in order to better understand present-day consumer and seller habits, such that future platforms and small business owners to make educated decisions regarding their digital sales. The data used in this study originates from an online French C2C platform³ that was launched in 2009. The data features information regarding registered user information for buyers and sellers on the C2C platform. From the data, it can be observed that there is a spectrum of user engagement levels and predominantly female engagement. Based upon these factors, we determined our project objectives.

Project Objectives: The objectives of this project are to determine as follows:

- Determine the attributes of the most active users on the C2C platform.
- Determine data-backed suggestions to increase male user engagement on the platform, as this industry is female-dominated.
- Determine which countries have the most interactive users, and which countries to target for promotion of the platform to increase engagement.
- Determine the valuable insights for businesses in the online resale community.
- Determine most valuable users.

¹Source: Lyst <https://www.lyst.co.uk/year-in-fashion-2020/#movements>

²Source: Investopedia <https://www.investopedia.com/terms/c/ctoc.asp>

³Source: French C2C <https://www.kaggle.com/jmmvutu/ecommerce-users-of-a-french-c2c-fashion-store>

II: Database Organization

Database Overview: The C2C database created on MySQL features 12 tables containing information from the initial 3 data files from Kaggle. There are over 9,000 rows in the database with 65+ columns.

Tables: The 11 tables we derived from categories within the initial 3 data files are as follows: femaleBuyersByCountry, maleBuyersByCountry, femaleSellers, maleSellers, topSellersFemale, topSellersMale, topSellersProduct, seniority, userDetails, userProducts, and userSocialdata. The tables femaleBuyersByCountry, maleBuyersByCountry, femaleSellers, maleSellers were created with data from the initial ‘comparison_of_sellers_by_gender_and_country.csv’ file. These tables all contain information on countries and gender with aggregated data of products sold, listed, etc. The topSellersFemale, topSellersMale, topSellersProduct tables contained data from the initial file, ‘countries_with_top_sellers_fashion_c2c.csv’, broken into Female, Male, and product categories. These tables all contain information about countries with top sellers, broken down into male, female, and products categories. The seniority, userDetails, userProducts, and userSocialdata tables contain data on each user on the platform. The data from these tables came from the initial data file ‘6M-0K-99K.users.dataset.public.csv’ broken up into seniority details, user details, user products, and social data categories. These tables all contain information about users and their social/shopping behavior on the platform. Our final table, countries, was created with lists of countries in both French and English. This was created as a reference table as each country was provided in French.

Functions: We created two stored functions to assist in our analysis. Our first stored function, PopularCountries, was created to return the popularity level of a country based on the total number of products sold. This function takes the total number of products sold as an input. The function labels countries with over 500 products sold as a ‘Popular country with most sought after listings’ and countries with less than 100 products sold as a ‘Unpopular country with least sought after items’. The countries in between those ranges were labeled as a ‘Country with average number of sought after items’. This function provides insights on which countries have the most popular sellers and items, which can be used to inform the e-commerce company about which countries to promote more to encourage sales. Our second function, EnglishTranslate, was created to translate the French country name to English using the ‘countries’ table we created. The ‘countries’ table contains lists of countries in both French and English. This function takes the French country name as an input and retrieves the English version from the ‘countries’ table where the French country input matches the value in the French country column. The function made it easy to retrieve the english country name.

Procedure: We created a stored procedure that calculates two ratios: one that compares the number of products listed to products sold, and another that compares the number of products wished and products purchased. The goal of this procedure is to produce two additional measures that can be used to understand users' buying and selling habits based upon their levels of engagement. The procedure takes in five values-the identifier hash, the number of products liked, sold, purchased, and listed. From these inputs, the procedure divides the number of products listed by the number sold, and the number of liked items by the number of purchases.

Trigger: We created a trigger that is activated when a database user inserts a new country in the 'French Country Name' column in the countries table. Utilizing the 'before insert' syntax allows us to ensure that the first letter of the new country name is capitalized before it is inserted. Using the upper and lower SQL function followed by substring, our trigger identifies the correct character to capitalize and the subsequent characters that should be converted or remain as lowercase. This trigger was created to ensure that our countries are inserted into the database with consistent capitalization.

Indexes: To increase the efficiency of analyses conducted in our database, we created 3 indexes on columns frequently searched against. Our first index, userDemographics, sits on our UserSocialData table that consists of 3,000 rows. This index contains demographic information using the columns 'country', 'gender', and 'language'. We expect these columns to be frequently searched against, as they provide insights on the demographics of their users which can be used to inform segmentation strategies. Our second index, FemaleSellersInfo, sits on the FemaleSellers table and consists of the columns 'totalproductssold', 'totalbought', 'country', and 'sex'. This index can be used to quickly compare information on female users' success on the platform in terms of sales and purchases. Our third index, ProductsListedAndSold, sits on our UserProducts table and consists of the columns 'country', 'productsListed', and 'productsSold'. This index can be used to speed up searches focused on analyzing countries and their sellers' success on selling their listings.

III: Noteworthy Columns & Data Quality Assurance

Important Columns: Out of the 65+ columns in our database, we concluded that the following columns were the most referenced in our analysis: totalProductsSold, totalBought, numberOfSellers, topSellers, Country, Gender, Language, and IdentifierHash. These columns provided the foundation of our queries in that they helped to join tables and group users into relevant bins. Beyond these, we drew upon a variety of unique columns, such as hasProfilePicture and daysSinceLastLogin, that were important to understanding general user engagement and specifically male users. For example, looking at whether users have iOS or Android apps allowed us to have insight as to user access and devices used.

Data Transformations: Our original dataset consisted of three separate files with around 90 columns. The file containing information on users consisted of over 90,000 rows of users, which we decided to take a sample of about 3,000 to use for our analysis. Cutting the user file down to 3,000 ensured that our database would not run into issues regarding inefficiencies due to the size of the file. For better organization of the data within our database, we split the original 3 files into 12 new files for our 12 tables. Before importing the 12 files into their tables, we cleaned our data in Excel to ensure the data was high quality, consistent, and usable for the database. We began by filling in each blank value in the files with zeros in order to be interpreted as null values. As our decimal values for our tables were set to two decimal places, we formatted each column with decimals to have a maximum of two decimal places for consistency. As our initial index column for our users data consisted of large, complex values, we renumbered the identifier from 1-3,000 for ease of analysis. We also removed redundant columns in our user tables, as there were several columns indicating the user's gender. As a few of our columns were in binary format, we re-coded the columns that contained 'TRUE' or 'FALSE' values to also be binary with '1' or '2' values for consistency of our data types. The 'country' column from the original user file was provided in French and contained corrupted values for the countries of Sweden (Suède) and Greece (Grèce). We replaced the corrupted values with the correct French country names once the data was imported to each table. We decided to leave the country names in French since the E-Commerce platform is based in France and would be less useful for the company in English, and created a new table with the French and English translations of the countries to refer to and use with the EnglishTranslate function for our own analysis.

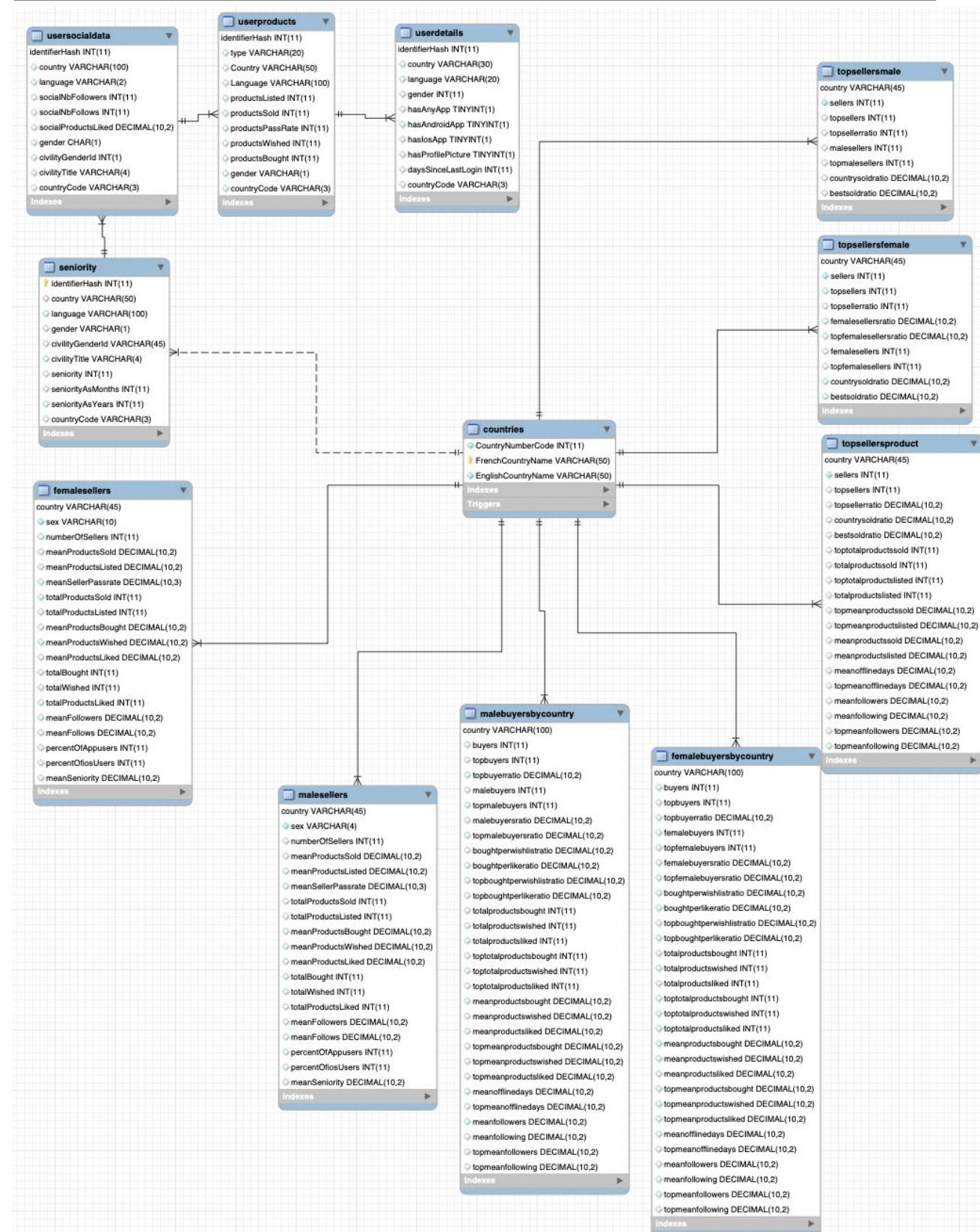
IV: Data Normalization

Normalization: We began normalizing our data by evaluating the best way to break up our initial 3 large dataset files. For the purpose of being more functional and comprehensible, we separated the 3 datasets into 12 tables. If we had imported the dataset without segmenting the initial datasets into 3 tables, it would have been cumbersome and difficult to navigate considering there are over 65 columns. This process led to an improved organization of the data and increased usability. Next, we recognized the need for a column that could serve as an ultimate connecting column. We decided to create a table for the countries specifically, which could act as a reference table. Based on the relationships between all of our tables, we decided the countries table would be the best table to use as the main connecting table. After acknowledging this, we were able to identify the country column as a consistent variable in all of our tables. The following tables were all connected to the countries table: femaleBuyersByCountry, maleBuyersByCountry, femaleSellers, maleSellers, topSellersFemale, topSellersMale, topSellersProduct. The remaining tables: seniority, userDetails, userProducts, and userSocialData, were connected to each other and seniority was used to connect to the countries table as well. After connecting our tables, our database resembled a snowflake schema. Although join complexity and query times increased slightly due to having many tables connected to the countries table with tables containing user information branching off, this design ultimately eliminates redundancy with our country columns as it prevents us from having both French and English versions of the countries in each table. Having a snowflake schema also allowed us to prevent potential data storage and integrity issues.

1NF: Moving on, in order to achieve first normal form, we needed to ensure that each record in our tables were unique and each cell in each column had a single value. Once this was completed, we needed to establish a single column primary key so each record could be referenced uniquely. This is demonstrated by our creation of four separate tables whose data originated from one large dataset containing user information. In order to maintain referential integrity, a unique primary key was generated for each record that would be used to reference other user related tables in our database.

In Addition: To better understand our dataset, we add several columns to the countries table: the country name in English and a number code for each country. We attempted to recode the country name variables to match the number country code we assigned in our new countries table, but found it difficult to ensure that we properly converted these values in the other tables; if we had been able to figure out a solution for this, our database would have been more effectively normalized.

V: ER Model



VI: Business Questions

To accomplish our objectives of providing insights on the E-Commerce platform, we created 25 business questions and corresponding queries to use to inform the company's strategies.

1. Order countries by the number of top sellers female in descending order. What country had the highest number of female top sellers?
 - a. Since we've discovered more females use this app, it is helpful to know what countries the top female sellers are from. This can help the business in marketing because they can target the countries they already have effective users in. It also tells them what users are generating the most profits and what countries are they operating in. *Reference Data Visualization #1*
2. What country are the most users from and what are those countries in english?
 - a. Answering this question will inform the company which countries they are successful in and which countries they need to expand into. It will also tell marketing what country specific preferences to capitalize on when marketing to their most important country segment which would be the country with the most users. In this query we utilized the english translate function which outputs the country name in english. It is useful for businesses whose first language is english to be able to easily retrieve the english translation. *Reference Data Visualization #2*
3. How many android and IOS users are there and what percent of our users have an android or an ios device?
 - a. This question the device types users are interacting with their app on. It can help inform design decisions as well as marketing because it indicates the majority of users have IOS devices; therefore, designing the features to be compatible for these devices as well as placing ads on app stores that sell to IOS devices will be more effective. *Reference Data Visualization #3*
4. How many sales were there? How many sales does each gender have and what percent of total sales was it?
 - a. This question provides information on total sales and segments it based on the gender. Businesses can use this to understand their most valuable users and identify the opportunity to increase sales made by male users.
5. What countries have male and female top sellers and what is their total number of top sellers with their male and female top sellers combined?
 - a. This question allows businesses to see the countries where the highest number of top sellers aka both male and female sellers originate from. This will help them understand the countries where both top male and female sellers reside. This can help them evaluate the reason these countries have successful male and female participating and could help design promotions for targeting both genders in other more segmented countries.

6. Are there more male or female users of the platform? Show total counts in descending order. *Reference Data Visualization #6*
 - a. This question used a group by to provide a quick overview of the distribution of male and female users of the platform. As a goal of this analysis is to determine male engagement rates compared to female, this query will update as additional users are added to the database.
7. Which language is most commonly spoken among users of the platform? Show all languages and the number of speakers in descending order. *Reference Data Visualization #7*
 - a. This question used a group by to provide an overview of the languages spoken by users of the platform. This can be used to inform developers on which languages to support on the platform. This question also used the index, userDemographics, to retrieve the output quickly.
8. Do male or female users have a higher total number of items bought on the platform? *Reference Data Visualization #8*
 - a. This question used a join of the MaleSellers and FemaleSellers tables to provide a quick overview of the total number of items bought by both male and female users. This can be used to analyze the most active gender in terms of sales to potentially inform strategies of which gender to promote the platform to. This question also used the index, userDemographics, to retrieve the output quickly.
9. Which countries are the most popular in terms of having the highest number of items sold (more than 500 sold) by both male and female users in those countries? *Reference Data Visualization #9 & 10*
 - a. This question used a union to display all countries with the highest number of products sold by both male and females. This question used the function, popularCountries, to determine which countries were labelled as popular based on having over 500 sales. This question also used the index, FemaleSellersInfo, in the second select statement for the female countries.
10. What are the top 5 countries based on the highest number of male users?
 - a. This question used a subquery to determine the top 5 number of users and all of the countries that contain those number of users. This query can be used to determine which countries to analyze for factors they share that led to a higher number of male users.
11. What is the max amount bought/sold from male and female users?
 - a. This question allows us to determine the current highest maximum value for products bought and sold for both males and females. *Data Visualization #11*
12. What is the average product pass rate by country ordered by highest to lowest?
 - a. Does the output signal to the business that there needs to be stricter policies/guidelines when users post items online? A low average may indicate the business needs to implement better guidelines/penalize users who are falsely

advertising items or items being sold that are not accurately matching the descriptions. *Data Visualization #12*

13. What are the top 3 countries by total number of application users?
 - a. Allows us to find out what countries account for the highest number of app users. Use this to help determine what is driving higher app usage in these countries. This query utilizes a subquery that returns countries greater than the minimum value of the top 3 countries based on the total number of application users. *Data Visualization #13*
14. Which users have the highest amount of social media followers? What countries are they from and how many products have they sold?
 - a. The output from this question can provide insight into what characteristics popular users have and potentially leverage these popular users with lots of reach in an influencer type campaign. In this query multiple tables were joined to return information like the English country name translation from the countries table. *Data Visualization #14*
15. What is the average seniority by country from the most popular countries (by total users)?
 - a. We want to measure customer loyalty and retention and we can determine which countries have the most loyal customers. Similar to the query above, a join was utilized in order to return values such as English country name translation from other tables such as countries. *Data Visualization #15*
16. What languages are spoken by users in each country (using a group by)? *Reference Data Visualization #16*
 - a. This visualization was used to gain a better understanding of the distribution of languages spoken by users across the world. From this query, it was found that five languages are spread amongst the countries with active users.
17. Which countries do not have male users (using a group by, having)? *Reference Data Visualization #18*
 - a. This query investigates male user engagement based upon geographic location. The results of this query suggest that there is great potential for higher male engagement in the countries that only had female users.
18. How does the number of follows and followers compare between users with and without a profile picture (using a join and union)?
 - a. This query seeks to understand whether a profile picture relates to a user's activity levels measured by number of followers and follows. The strength of the relationship between profile pictures and activity could be used to predict a user's future level of engagement. The query revealed that, based upon the data, users without profile pictures had higher numbers of followers and follows than those with.
19. How do top purchasing users compare based upon whether they have a profile picture (using a join and case when)?

- a. This query investigates whether having a profile picture influences if a user is more likely to purchase items on the platform. This query uses a case when statement to rank users based upon their number of purchases and account having/not having a profile picture.
20. Which users have bought and sold more than 25 products (using a subquery)?
- a. This subquery looks at highly engaged users who both buy and sell items on the platform. This query is useful because it conveys an additional metric of engagement based on buy/sell records.
21. Out of the top female sellers, which countries had the best sold ratio?
- a. Using a limit to see the top 5 countries, we can see which top 5 countries have the highest best sold ratio in order out of female topsellers. This query is useful because we can see the top countries that are able to convert their listings into sales over other countries.
22. What is the total product wished for between males and females?
- a. This select statement gives us an idea about the interactions that each gender has when it comes to products that they wish for. This is helpful because we get an idea as to who is interacting with users websites more often.
23. How many social followers do these top sellers have combined in France, and how many are either Android or IOS app users?
- a. The select statement uses a join with usersocialdata and userdetails, to depict specifically from France, from the top sellers, how many social followers do they have and the cumulative number of Android/IOS app users. Both systems have different interfaces, and users can tailor their software designs based on buyers that are more frequent IOS or Android users within the country of France.
24. What are the top 5 companies that have the highest amount of products sold with 100% product passing rate?
- a. This select statement uses a join from the userdetails table to get data identifying which stores sell the most products with 100% passing rate. Passing rate relates to the product description, and the accuracy it fits when sold. Having more products sold with higher passing rate shows the consistency that these companies provide for their customers
25. How many sellers have a bestsoldratio greater than 1 per country?
- a. Using a subquery, we can compare each country to one another and the amount of sellers these countries have that have a bestsoldratio greater than one. Using this information, we can figure out which countries have more sellers than others, that do more business, potentially relating to these businesses that provide better customer service and high quality products.

VII: Data Visualizations

To accompany our business questions, we created 25 data visualizations in Tableau. These visualizations provide quick insights for the C2C platform to use to guide their strategies. (*See Appendix B for screenshots*)

List of Visualizations:

1. Top Female Sellers by Country
 - a. This visualization illustrates the number of top female sellers from each country ordered by countries with the greatest number of sellers to the least. It was created with a highlight table so we could add color shading to indicate the hierarchy top and bottom countries. It provides an easy to interpret preview of the countries with the top sellers.
2. Users by Country
 - a. This visualization shows the percent of users from each country. A highlight table was used to exhibit the percent of users from each country. The coloring was applied to show rank amongst the countries. We applied the count function and percentage of the total calculator to the identifier hash column which was segmented by countries. This table provides a quick overview of the top countries users in the dataset come from.
3. Android and IOS users
 - a. This visualization displays the number of total users as well as number of Android and IOS users, then the percent of Android and IOS users compared to the total users. A text table was used to present this data. We utilized the sum function on the 'has an android' and 'has an IOS' columns and the count function on the 'identifierHash' to get the total users. We created two calculated fields to create the percentages. This text table is useful for understanding the devices users use.
4. Products Listed, Bought, and Sold by Country
 - a. This visualization depicts the number of products listed, bought, and sold by users in each country. A side by side bar chart was used to exhibit these numbers. For viewing purposes, several countries with none to few products were excluded. The following columns were used to create the chart: country, products listed, products bought, and products sold. The sum function was applied to the product columns then group by country. Color coding was applied to differentiate the products listed, sold, and bought from each other. This chart provides a visual comparison of products listed, bought, and sought by country.
5. Percent of Male vs Female Users
 - a. This visualization is a simple chart comparing the percent of female users to male. A pie chart with added coloring was used to represent these percentages. The quick percent of the total calculator in tableau was used to generate the percent. It is a clear visual breakdown of users' gender.

6. Number of Male vs. Female Users

- a. This visualization provides a quick overview of the number of male and female users on the platform. Created as a pie chart, the distribution of male and female users is displayed with labels of the gender and corresponding count of users per gender. The pie chart is also colored based on gender for quicker understanding of the breakdown of genders.

7. Languages Spoken by Users

- a. This visualization was created to provide a simplified view of the languages spoken by the users of the platform, as well as a count of users per language. This visualization was created as a text table and uses the ‘languages’ column from the UserSocialData table as a row and the count of the ‘languages’ column as the text labels.

8. Total Number of Items Bought per Gender

- a. This visualization provides a simplified view of the total number of items bought by male and female users. It was created as a quick way to view the gender that is most active in terms of buying on the platform. This visualization was created as a text table and uses the aggregated sum of the total number of items bought for both genders using a join of the MaleSellers and FemaleSellers as the measures, and uses the gender values as filters.

9. Countries with the Most Products Sold by Male Users

- a. This visualization was created to display the countries with the most products sold by male users. This can be used to quickly inform the E-Commerce platform on which countries have the most sought-after listings by male users. Created as a map, it provides a quick view of the popular countries using colors to separate the countries. It uses the male gender and aggregated sum of the total products sold over 500 as filters to ensure that it is only displaying countries with a high volume of male products sold.

10. Countries with the Most Products Sold by Female Users

- a. This visualization was created to display the countries with the most products sold by female users. This can be used to quickly inform the E-Commerce platform on which countries have the most sought-after listings by female users. Created as a map, it provides a quick view of the popular countries using colors to separate the countries. It uses the female gender and aggregated sum of the total products sold over 500 as filters to ensure that it is only displaying countries with a high volume of female products sold.

11. Female Sellers by Country

- a. This map visualization displays where most of the female sellers in the database reside. After summing up the number of female sellers, the English translation for countries was pulled in from the countries table for ease of readability. Blue countries indicate countries where there are a higher volume of female sellers.

12. Total Products Bought by Country

- a. This bubble chart displays the total number of products bought per country. For better readability, the English translations for country names were utilized from the countries table. Larger circles indicate a higher volume of total products bought.

13. Total App Users by Country

- a. This visualization shows us the amount of total application users from each country. Application users were first aggregated and were then mapped out using the map type graph. The blue countries indicate ones with higher numbers of total application users.

14. Max Products Bought/Sold by Gender

- a. This visualization provides a breakdown by gender of the highest amount of products sold and bought in the database. This gives insight into the current maximum value of the volume of products both bought and sold by a specific Male and Female.

15. Average Seniority Measured by the Most Popular Countries

- a. This bar chart displays the average seniority (years) of the most popular countries. The most popular countries were determined by counting all users within each country. The seniority can give insight into which countries have the most loyal customers.

16. Breakdown of Languages Spoken By Users in Each Country

- a. This text table lists out every country contained within the dataset, and indicates which among the five languages are spoken by users in that specific country. This format allows for ease of understanding which languages are used among users within a country, and can be used in a business setting to both expand language options and communicate with users. This visualization does not have any filters, but rather aggregates the country data in a column and language data by rows.

17. Countries with the Highest Averages of Number of Days Since Last User Login

- a. This visualization displays the top ten countries with the highest average number of days since last login among its users as a treemap. An average was used as opposed to the instances of the longest last login time period to mitigate the influence of countries with one record. To create this treemap, the countries were filtered by the top 10 average days since last login, using the average of the days since last login for the size, color, and labels for the visualization.

18. Countries with Only Female Users

- a. This visualization displays a world map having countries with only female users shaded in blue. This visualization is useful because it conveys areas where there is great growth potential through engaging more male users. A map was chosen in order to determine if there are any major regional patterns regarding no male engagement. This map is filtered by country where the count is one, to funnel in

on countries with only one gender of users. There is a second filter on gender, which only includes female users so that countries without male users are displayed. Longitude and latitude are used for the row and column values.

19. Product Wishes By Country

- a. This packed bubbles visualization displays countries by size and count of the number of wishes made by its users. This map is useful because it shows a country-based indicator of user engagement based upon a measure other than products bought and sold. Thus, this visualization conveys a representation of user activity levels. This visualization filters by country and has the sum of products liked used as an input for size, color, and labels. English country names are also used as labels.

20. Purchase Density Map

- a. This visualization shows a world map with countries shaded according to the number of purchases made in that country relative to all purchases made. This map is useful because it shows a regional pattern that the majority of purchases are made by users in the United States, Canada, and Europe. This map uses longitude and latitude as row and column inputs, the sum of the products sold as percent of a whole as the color, and country as a detail element.

21. Top 5 Countries with Best Sold Ratio for Females

- a. This bar chart displays the best sold ratio in the top 5 countries in the world, in regards to females. We can visually see which countries are dominating the female market currently. Using the sum of best sold ratio under top sellers female column, as well as filtering on english country names.

22. Products Wished Per Gender

- a. This pie chart displays the total amount of products wished between males and females. The size of the pie chart, for each gender, is the sum of products wished. This gives us a better understanding of which gender is utilizing user apps, and interacting with products they wish to have

23. Average Number of Offline Days from Sellers within the Top 10 Countries

- a. Using a horizontal bar chart, this visual displays the average number of offline days from top sellers, within their respective countries. I applied a filter using countries from Top Sellers Product, setting it to 10, to narrow down the results. Next to each country's name, is the country's number. This visualization can give us insights about countries that have top sellers spending the most time offline from their e-commerce sites (not doing business with consumers).

24. Top 5 Sellers with highest product description in relation to products sold

- a. This visual chart using circle marks gives us insight as to who the top 5 sellers are, in relation to the number of products sold with 100% product pass rate. Product pass rate is the percentage of products sold that match the product description, evaluated by the stores team before the products are shipped. Using a

filter on identifierHash where the condition is the sum of product pass rate = 100, limiting the field to the top 5.

25. Products Liked between Males and Females

- a. Using a barchart, we can visually see the difference between females and male users when it comes to liking products online. This shows who is shopping more often, and engaging with these users websites. Clearly Females are in the lead when it comes to liking products, leaving room for companies to explore marketing efforts towards Males.

VIII: Appendix

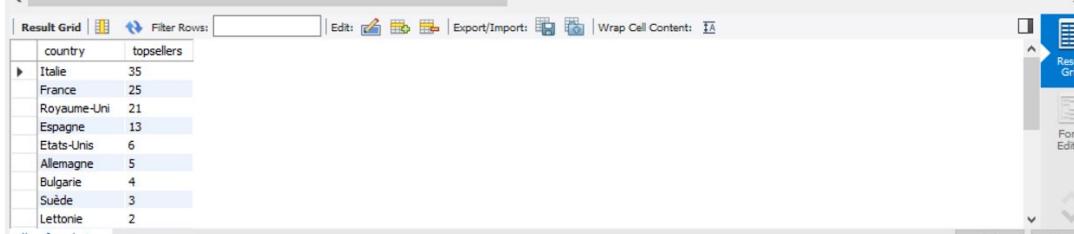
Appendix A: Screenshots from our business question select statements. The numbers in the Business Questions section correspond to the numbers of the images below.

1.

```

3      #Order countries by the number of topsellers female in descending order. What country had the highest number of female topsellers?
4 •   select country, topsellers
5   from topsellersfemale
6   group by country
7   order by topsellers desc;
8
9      #What country are the most users from and what are those countries in english?

```



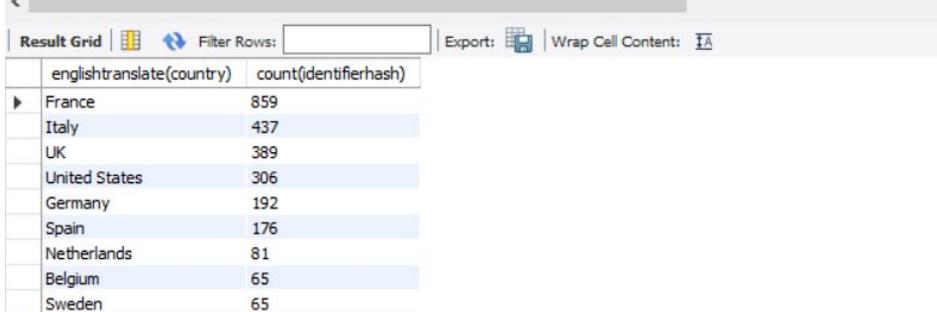
country	topsellers
Italie	35
France	25
Royaume-Uni	21
Espagne	13
Etats-Unis	6
Allemagne	5
Bulgarie	4
Suède	3
Lettonie	2

2.

```

9      #What country are the most users from and what are those countries in english?
10 •  select englishtranslate(country), count(identifierhash)
11   from userdetails
12   group by country
13   order by count(identifierhash) desc;
14

```



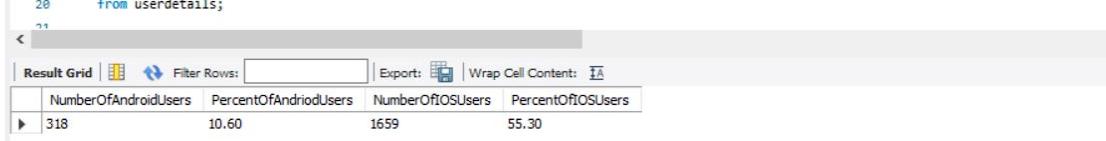
englishtranslate(country)	count(identifierhash)
France	859
Italy	437
UK	389
United States	306
Germany	192
Spain	176
Netherlands	81
Belgium	65
Sweden	65

3.

```

14
15      #How many android and ios users are there and what percent of our users have an android or an ios device?
16 •  select (select count(identifierhash) from userdetails where hasandroidapp=1) as NumberOfAndroidUsers,
17   round(((select count(identifierhash) from userdetails where hasandroidapp=1)/count(identifierhash))*100),2 as PercentOfAndriodUsers,
18   (select count(identifierhash) from userdetails where hasiosapp=1) as NumberOfIOSUsers,
19   round(((select count(identifierhash) from userdetails where hasiosapp=1)/count(identifierhash))*100),2 as PercentOfIOSUsers
20   from userdetails;
21

```



NumberOfAndroidUsers	PercentOfAndriodUsers	NumberOfIOSUsers	PercentOfIOSUsers
318	10.60	1659	55.30

4.

```

33  #what country have male and female top sellers and what is there total number top sellers with their male and female to sellers combined?
34  ● select c.frenchcountryname, f.topsellers, m.topsellers, f.topsellers+m.topsellers as totaltopsellers
35  from topsellersmale m
36  join countries c
37  on m.country=c.frenchcountryname
38  join topsellersfemale f
39  on c.frenchcountryname=f.country
40  group by c.frenchcountryname
41  order by f.topsellers+m.topsellers desc;
42

```

Result Grid | Filter Rows: Export: Wrap Cell Content:

frenchcountryname	topsellers	topsellers	totaltopsellers
Allemagne	5	5	10
Roumanie	2	2	4
Autriche	1	1	2
Cypre	1	1	2
Croatie	1	1	2

5.

```

33  #what country have male and female top sellers and what is there total number top sellers with their male and female to sellers combined?
34  ● select c.frenchcountryname, f.topsellers, m.topsellers, f.topsellers+m.topsellers as totaltopsellers
35  from topsellersmale m
36  join countries c
37  on m.country=c.frenchcountryname
38  join topsellersfemale f
39  on c.frenchcountryname=f.country
40  group by c.frenchcountryname
41  order by f.topsellers+m.topsellers desc;
42

```

Result Grid | Filter Rows: Export: Wrap Cell Content:

frenchcountryname	topsellers	topsellers	totaltopsellers
Allemagne	5	5	10
Roumanie	2	2	4
Autriche	1	1	2
Cypre	1	1	2
Croatie	1	1	2

6.

```

16  #6 Are there more male or female users of the platform?
17  ● SELECT gender, COUNT(gender) as 'Number per Gender' FROM usersocialdata USE INDEX(userDemographics)
18  GROUP BY gender
19  ORDER BY COUNT(gender) DESC;
20

```

100% 56:16

Result Grid | Filter Rows: Search: Export:

gender	Number per Gend...
F	2333
M	667

```

9   #7 Which language is most commonly spoken among users of the platform?
10  #Show all languages and the number of speakers in descending order.
11  • SELECT language, COUNT(language) as 'Number of Speakers' FROM usersocialdata USE INDEX(userDemographics)
12  GROUP BY language
13  ORDER BY COUNT(language) DESC;
14
100% 72:9 |
```

Result Grid Filter Rows: Search Export:

language	Number of Speak...
en	1334
fr	888
it	420
de	191
es	167

7.

```

21  #8 Do male or female users have a higher total number of items bought on the platform?
22  • SELECT SUM(f.totalbought) as 'Female: Total Items Bought', SUM(m.totalbought) as 'Male: Total Items Bought'
23  FROM malesellers m
24  JOIN countries c ON c.frenchcountryname = m.country
25  JOIN femalesellers f ON c.frenchcountryname = f.country;
26
100% 87:21 |
```

8.

```

28  #Which countries are the most popular in terms of having the highest number of items sold
29  #(more than 500 sold) by both male and female users in those countries?
30  • SELECT country, sex, totalproductssold, popularcountries(totalproductssold)
31  FROM malesellers
32  WHERE popularcountries(totalproductssold) ='Popular country with most sought after listing'
33  UNION
34  SELECT country, sex, totalproductssold, popularcountries(totalproductssold)
35  FROM femalesellers USE INDEX(FemaleSellersInfo)
36  WHERE popularcountries(totalproductssold) ='Popular country with most sought after listing'
37  ORDER BY sex, totalproductssold DESC;
38
100% 90:28 |
```

9.

Result Grid Filter Rows: Search Export:

country	sex	totalproductss...	popularcountries(totalproductssold)
France	Female	2365	Popular country with most sought after listings
Italie	Female	1865	Popular country with most sought after listings
Royaume-Uni	Female	1442	Popular country with most sought after listings
Etats-Unis	Female	817	Popular country with most sought after listings
Espagne	Female	786	Popular country with most sought after listings
Italie	Male	955	Popular country with most sought after listings
France	Male	649	Popular country with most sought after listings

```

40  #What are the top 5 countries based on the highest number of male users?
41  #Step 1: Top 5 number of male users
42 • SELECT DISTINCT numberOfSellers FROM malesellers
43      ORDER BY numberOfSellers DESC
44      LIMIT 5;
45  #Step 2: Select min of the top 5
46 • SELECT min(numberOfSellers) FROM
47      (SELECT DISTINCT numberOfSellers FROM malesellers
48      ORDER BY numberOfSellers DESC
49      LIMIT 5) sq;
45  #Step 3: Find countries with number of male users = or > than Step 2
50  • SELECT * FROM malesellers
51      WHERE numberOfSellers >=
52          (SELECT min(numberOfSellers) FROM
53          (SELECT DISTINCT numberOfSellers FROM malesellers
54          ORDER BY numberOfSellers DESC
55          LIMIT 5) sq)
56      ORDER BY numberOfSellers;
57
100%  2:40

```

Result Grid Filter Rows: Search Edit: Export/Import:

country	sex	numberOfSellers	meanProductsSold	meanProductsListed	meanSellerPassrate	totalProductsSold	totalProductsListed
Allemagne	Male	34	2.00	1.00	19.150	68	34
Etats-Unis	Male	41	3.66	3.49	30.310	150	143
Royaume-Uni	Male	64	5.69	1.61	25.310	364	103
Italie	Male	118	8.09	5.13	42.220	955	605
France	Male	233	2.79	2.12	25.890	649	493

10.

```

3  #What is the max amount bought/sold from male and female users?
4 • select max(productsBought) maxProductsBought, max(productsSold) maxProductsSold, gender
5  from userproducts
6  group by gender;
7
100%  26:1

```

Result Grid Filter Rows: Search Export:

maxProductsBoug...	maxProductsSo...	gender
279	174	M
405	170	F

11.

```

10  #What is the average products pass rate by country ordered by highest to lowest?
11 • select identifierHash, avg(productsPassRate) avgpassrate, country
12  from userproducts
13  group by country
14  order by avg(productsPassRate) desc;
15
100%  37:8

```

Result Grid Filter Rows: Search Export:

identifierHash	avgpassrate	country
2476	0.0000	Bahreïn
2631	0.0000	Inde
2562	0.0000	Liban
2681	0.0000	Corée du Sud
2286	0.0000	Lituanie
2506	0.0000	Brésil
1739	0.0000	Islande

12.

```

18      #What are the top 3 countries by total amount of application users?
19 •  select country, count(hasAnyApp) as appusers
20    from userdetails
21   group by country
22   having count(hasAnyApp) >= (select min(app) from
23   (select distinct count(hasAnyApp) as app from userdetails
24   group by country order by app desc
25   limit 3)sq)
26   order by count(hasAnyApp) desc;
27

```

100% 1:17

Result Grid Filter Rows: Search Export:

country	appusers
France	859
Italie	437
Royaume-Uni	389

13.

```

30      #Which users have the highest amount of social media followers?
31      #What countries are they from (returns English translation) and how many products have they sold?
32 •  select identifierHash, socialNbFollowers, sum(productsSold) totprodsold, EnglishCountryName
33    from userproducts p
34   join usersocialdata usd using(identifierHash)
35   join seniority s using(identifierHash)
36   join countries c on s.country = c.FrenchCountryName
37   group by identifierHash
38   order by socialNbFollowers desc
39   limit 5;
40
41

```

100% 1:29

Result Grid Filter Rows: Search Export:

identifierHash	socialNbFollowe...	totprodsold	EnglishCountryName
9	744	104	Italy
159	353	15	France
1120	205	1	France
18	176	81	Spain
58	172	34	Italy

14.

```

43      #What are is the average seniority by country (English translation)
44      #from the most popular countries (by total users)?
45 •  select EnglishCountryName, avg(seniorityAsYears) avgSeniority, count(country) totalUsers
46    from seniority s
47   join countries c on s.country = c.FrenchCountryName
48   group by s.country
49   order by count(country) desc, avg(seniorityAsYears) desc
50   limit 10;

```

100% 9:39

Result Grid Filter Rows: Search Export:

EnglishCountryName	avgSeniority	totalUsers
France	8.6054	859
Italy	8.6293	437
UK	8.5887	389
United States	8.6078	306
Germany	8.7083	192
Spain	8.6534	176
Netherlands	8.6914	81
...

15.

```

21      ## What languages are spoken by users in each country?
22      ## Order by Number of Speakers Desc
23 •   SELECT Country,
24      (CASE WHEN language = 'en' THEN 'Enlgish'
25      WHEN language = 'fr' THEN 'French'
26      WHEN language = 'it' THEN 'Italian'
27      WHEN language = 'de' THEN 'German'
28      WHEN language = 'es' THEN 'Spanish'
29      ELSE 'Other'
30      END) AS Language,
31      COUNT(language) AS 'Number of Speakers'
32  FROM userdetails
33  GROUP BY country, language
34  ORDER BY COUNT(Language) DESC;

```

16.

```

6 •  use _team20_c2c;
7
8  -----
9
10     ##: Which countries do not have male users?
11     -- Assuming 1 on the user details table represents males
12 •  SELECT Country, Gender, count(gender)
13    FROM userSocialData
14    WHERE gender != 'M'
15    GROUP BY country, Gender
16    HAVING COUNT(country) = 1
17    ORDER BY country ASC, Gender Desc;

```

17.

Country	Gender	count(gender)
Bahreïn	F	1
Brésil	F	1
Brunei Darussalam	F	1
Corée du Sud	F	1
Guam	F	1
Honduras	F	1
Îles Canaries	F	1
Islande	F	1
Kazakhstan	F	1
Kiribati	F	1
Liban	F	1

```

54 •   SELECT up.identifierHash, up.productsBought,
55   (
56     CASE WHEN up.productsBought >= 100 AND ud.hasProfilePicture = 1 THEN 'Top Buyer With Profile Pic'
57       WHEN up.productsBought >= 100 AND ud.hasProfilePicture = 0 THEN 'Top Buyer Without Profile Pic'
58       WHEN up.productsBought >= 50 AND up.productsBought < 100 AND ud.hasProfilePicture = 1 THEN 'Mid Level Buyer With Profile Pic'
59       WHEN up.productsBought >= 50 AND up.productsBought < 100 AND ud.hasProfilePicture = 0 THEN 'Mid Level Buyer Without Profile Pic'
60       WHEN up.productsBought >= 25 AND up.productsBought < 50 AND ud.hasProfilePicture = 1 THEN 'Basic Level Buyer With Profile Pic'
61       WHEN up.productsBought >= 25 AND up.productsBought < 50 AND ud.hasProfilePicture = 0 THEN 'Basic Level Buyer Without Profile Pic'
62       WHEN up.productsBought >= 1 AND up.productsBought < 25 AND ud.hasProfilePicture = 1 THEN 'Infrequent Buyer With Profile Pic'
63       WHEN up.productsBought >= 1 AND up.productsBought < 25 AND ud.hasProfilePicture = 0 THEN 'Infrequent Buyer Without Profile Pic'
64     ELSE 'No Purchase Record'
65   END
66   ) a
67   FROM userproducts AS up
68   JOIN userdetails ud ON ud.identifierHash=up.identifierHash
69   WHERE up.productsBought != 0
70   ORDER BY productsBought DESC;
71
72
73
100% 1:49

```

18.

```

38    ## : How does the avg number of follows and followers compare between users with and without profile pictures?
39    -- does 0 or 1 indicate a profile picture?, 0 = No Profile Pic
40 •   SELECT 'Profile Picture', AVG(usd.socialNbFollowers) AS 'Avg Followers', AVG(usd.socialNbFollows) AS 'Avg Follows'
41   FROM usersocialdata AS usd
42   JOIN userdetails ud ON ud.identifierHash=usd.identifierHash
43   WHERE ud.hasprofilepicture = '1'
44   UNION
45   SELECT 'No Profile Picture', AVG(usd.socialNbFollowers) AS 'Avg Followers', AVG(usd.socialNbFollows) AS 'Avg Follows'
46   FROM usersocialdata AS usd
47   JOIN userdetails ud ON ud.identifierHash=usd.identifierHash
48   WHERE ud.hasprofilepicture = '0';

```

19.

```

75    ##: What users have sold and bought more than 25 products? (Subquery)
76
77 •   SELECT IdentifierHash, Country, productsBought, productsSold
78   FROM userProducts
79   WHERE productsBought IN
80     (SELECT productsBought
81      FROM userProducts
82      WHERE productsBought >=25 )
83   AND productsSold IN (SELECT productsSold
84      FROM userProducts
85      WHERE productsSold >= 25)
86   ORDER BY Country ASC;
87

```

20.

Result Grid Filter Rows: Search Export: Fetch rows:

identifierHash	productsBought	a
1145	405	Top Buyer With Profile Pic
211	279	Top Buyer With Profile Pic
325	174	Top Buyer With Profile Pic
23	115	Top Buyer Without Profile Pic
7	105	Top Buyer Without Profile Pic
2964	93	Mid Level Buyer With Profile Pic
2297	87	Mid Level Buyer Without Profile Pic
503	85	Mid Level Buyer Without Profile Pic

Result Grid Filter Rows: Search Export:

Profile Picture	Avg Followers	Avg Follows
Profile Picture	8.7936	11.6373
No Profile Picture	19.4169	41.6430

Result Grid Filter Rows: Search Export:

IdentifierHa...	Country	productsBought	productsSold
17	Allemagne	32	82
38	Espagne	32	50
13	Espagne	36	87
45	Etats-Unis	69	45
23	France	115	69
68	Luxembourg	80	30
7	Suède	105	108

	<pre> 245 ## Which top 5 countries have the best sold ratio from female sellers? 246 • select country, bestsoldratio from topsellersfemale 247 order by bestsoldratio desc 248 limit 5; 249 </pre> <p>100% 71:245</p> <p>Result Grid Filter Rows: Search Edit: </p> <table border="1"> <thead> <tr> <th>country</th><th>bestsoldra...</th></tr> </thead> <tbody> <tr> <td>Monaco</td><td>8.95</td></tr> <tr> <td>Allemagne</td><td>2.85</td></tr> <tr> <td>Lettonie</td><td>2.25</td></tr> <tr> <td>Royaume-Uni</td><td>2.16</td></tr> <tr> <td>Bulgarie</td><td>2.10</td></tr> </tbody> </table>	country	bestsoldra...	Monaco	8.95	Allemagne	2.85	Lettonie	2.25	Royaume-Uni	2.16	Bulgarie	2.10						
country	bestsoldra...																		
Monaco	8.95																		
Allemagne	2.85																		
Lettonie	2.25																		
Royaume-Uni	2.16																		
Bulgarie	2.10																		
21.	<pre> 253 ## What is the total product wished for between males and females? 254 • select gender, sum(productsWished) from userproducts 255 group by gender; 256 </pre> <p>100% 68:253</p> <p>Result Grid Filter Rows: Search Export: </p> <table border="1"> <thead> <tr> <th>gender</th> <th>sum(productsWish...</th> </tr> </thead> <tbody> <tr> <td>M</td> <td>14444</td> </tr> <tr> <td>F</td> <td>69691</td> </tr> </tbody> </table>	gender	sum(productsWish...	M	14444	F	69691												
gender	sum(productsWish...																		
M	14444																		
F	69691																		
22.	<pre> 260 ## How many top sellers in France use Android and IOS apps? How many social followers do these top sellers have combined? 261 • select country, topsellers, sum(socialNbFollowers), sum(ud.hasAndroidApp) as '# of Android App User', sum(ud.hasiosApp) as 262 '# of IOS App User' from topsellersproduct 263 join usersocialdata ud using(country) 264 join userdetails ud using(country) 265 where country = 'France'; 266 </pre> <p>100% 100:260</p> <p>Result Grid Filter Rows: Search Export: </p> <table border="1"> <thead> <tr> <th>country</th> <th>topsellers</th> <th>sum(socialNbFollowe...)</th> <th># of Android App U...</th> <th># of IOS App Us...</th> </tr> </thead> <tbody> <tr> <td>France</td> <td>25</td> <td>6952746</td> <td>90195</td> <td>367652</td> </tr> </tbody> </table>	country	topsellers	sum(socialNbFollowe...)	# of Android App U...	# of IOS App Us...	France	25	6952746	90195	367652								
country	topsellers	sum(socialNbFollowe...)	# of Android App U...	# of IOS App Us...															
France	25	6952746	90195	367652															
23.	<pre> 269 ## What are the top 5 users when it comes to products sold at 100% productPassRate? 270 • select ud.identifierHash, productsSold, productsPassRate AS '% of products meeting product description' from userproducts 271 join.userdetails ud using(identifierHash) 272 where productsPassRate = 100 273 group by identifierHash 274 order by productsSold desc 275 limit 5; 276 </pre> <p>100% 84:269</p> <p>Result Grid Filter Rows: Search Export: </p> <table border="1"> <thead> <tr> <th>identifierHa...</th> <th>productsSo...</th> <th>% of products meeting product descri...</th> </tr> </thead> <tbody> <tr> <td>5</td> <td>125</td> <td>100</td> </tr> <tr> <td>15</td> <td>86</td> <td>100</td> </tr> <tr> <td>35</td> <td>55</td> <td>100</td> </tr> <tr> <td>36</td> <td>54</td> <td>100</td> </tr> <tr> <td>37</td> <td>51</td> <td>100</td> </tr> </tbody> </table>	identifierHa...	productsSo...	% of products meeting product descri...	5	125	100	15	86	100	35	55	100	36	54	100	37	51	100
identifierHa...	productsSo...	% of products meeting product descri...																	
5	125	100																	
15	86	100																	
35	55	100																	
36	54	100																	
37	51	100																	
24.																			

```
280      ## How many sellers with a best sold ratio greater than 1, does each country have?
281 * select country, sellers from topsellersproduct
282   where totalproductssold IN (select totalproductssold from topsellersproduct where bestsoldratio > 1 );
283
100% 83:280 |
```

Result Grid Filter Rows: Search Edit: Export/Import:

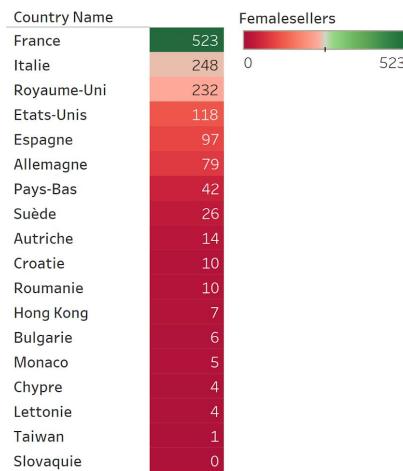
country	sellers
Allemagne	105
Bulgarie	9
Croatie	10
Espagne	119
Etats-Unis	150
France	713
Italie	347

25.

Appendix B: Screenshots from our data visualizations in Tableau. The numbers in the Data Visualization section correspond to the numbers of the images below.

1.

Top Female
Sellers by
Country



Sum of Femalesellers broken down by Country Name. Color shows sum of Femalesellers. The marks are labeled by sum of Femalesellers. The context is filtered on Country Name, which keeps 18 of 18 members.

2.

Users By Country



% of Total Count of
Identifiers (Userdetails)
broken down by Country Name.
Color shows % of Total Count of
Identifiers (Userdetails).
The marks are labeled by % of
Total Count of IdentifierHash
(Userdetails).

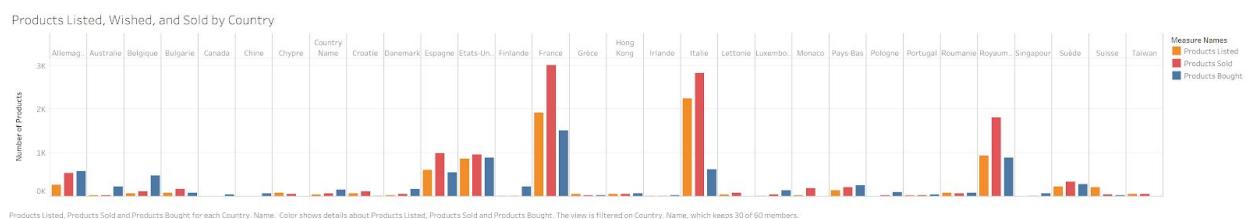
Andriod and IOS Users

Has Android App	318
Has Ios App	1,659
Count of identifierHash (u..	3,000
Percent of Andriod Users	11
Percent of IOS Users	55

Has Android App, Has Ios App, count of identifierHash (userdetails), Percent of Andriod Users and Percent of IOS

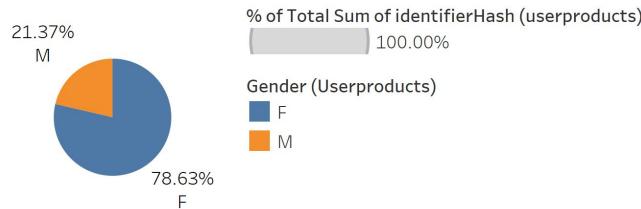
3.

4.



5.

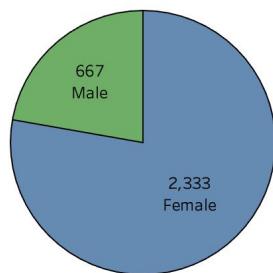
Percent of Male vs Female Users



% of Total Sum of identifierHash (userproducts) and Gender (Userproducts). Color shows details about Gender (Userproducts). Size shows % of Total Sum of identifierHash (userproducts). The marks are labeled by % of Total Sum of identifierHash (userproducts) and Gender (Userproducts).

6.

Total Number of Male vs Female Users



7.

Languages of Users

Language (User..	F
English	1,334
French	888
Italian	420
German	191
Spanish	167

8.

Total Number of Items Bought by Gender

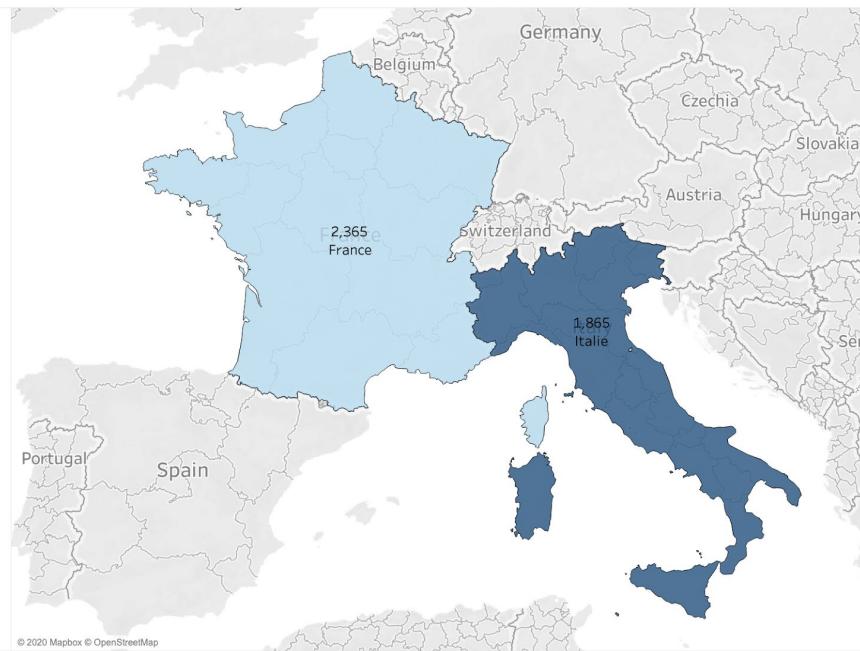
Female Users	4,387
Male Users	1,103

9.

Countries with the Most Products Sold by Male Users

Sex

Male



10.

Countries with the Most Products Sold by Female Users

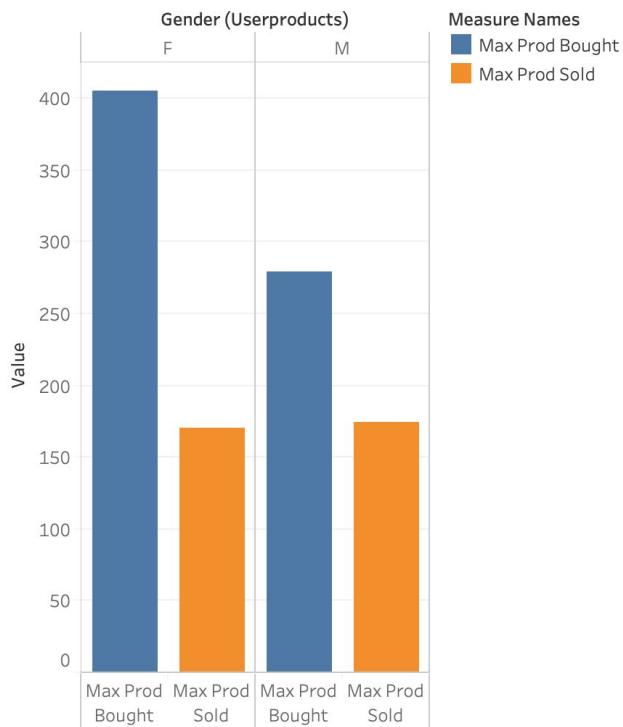
Sex (Females..)

Female



11.

Max Products Bought/Sold By Gender

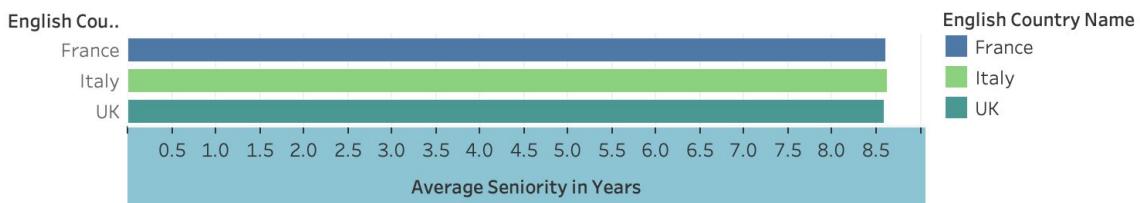


Max Prod Bought and Max Prod Sold for each Gender (Userproducts). Color shows details about Max Prod Bought and Max Prod Sold .

12.

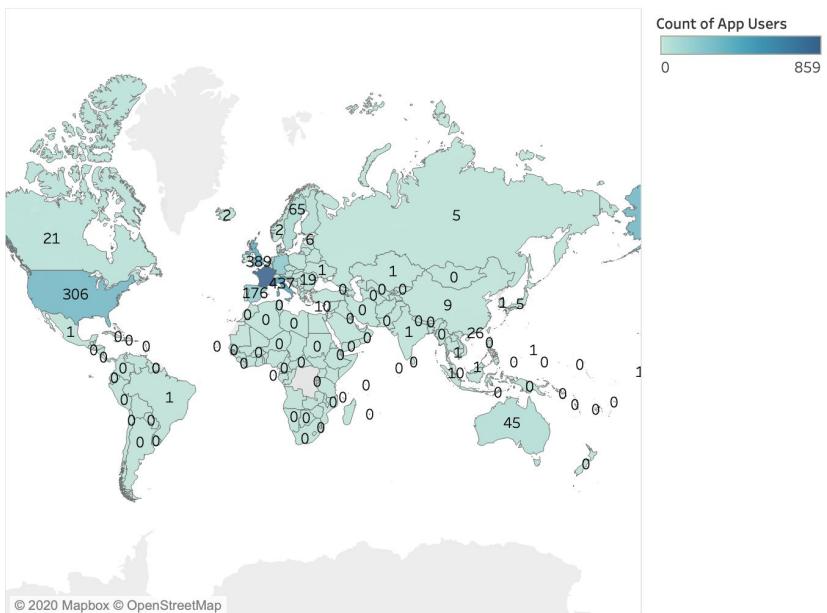
13.

Average Seniority Measured by the Most Popular Countries



Average Seniority in Years for each English Country Name. Color shows details about English Country Name. The view is filtered on Country Count and Average Seniority in Years. The Country Count filter ranges from 350 to 859. The Average Seniority in Years filter ranges from 0.00000 to 9.00000.

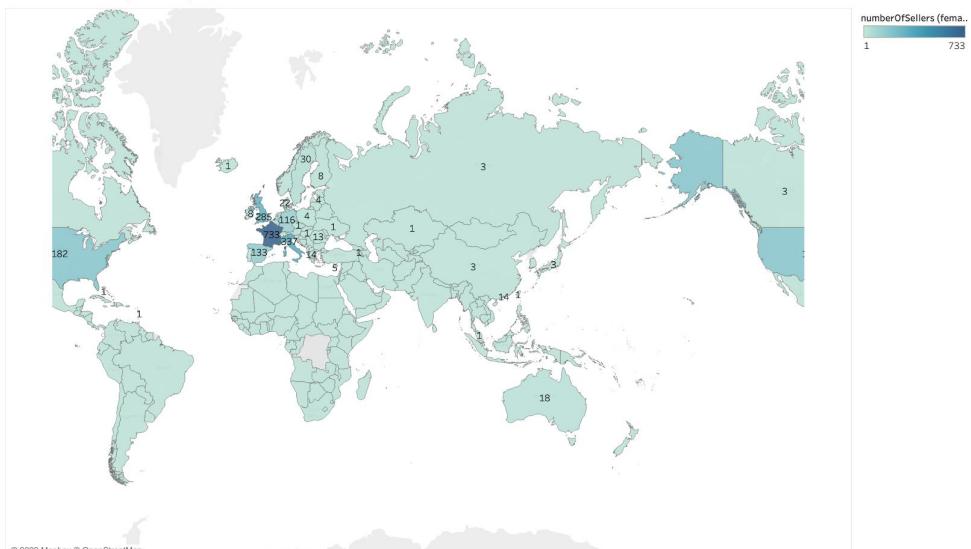
App Users By Country



Map based on Longitude (generated) and Latitude (generated). Color shows Count of App Users. The marks are labeled by Count of App Users. Details are shown for English Country Name.

14.

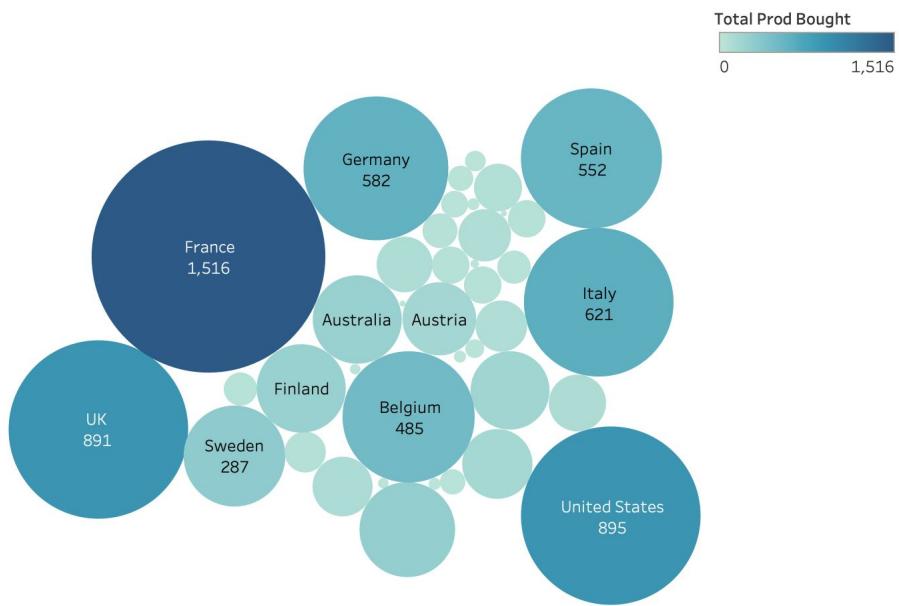
Female Sellers by Country



15.

Map based on Longitude (generated) and Latitude (generated). Color shows sum of number of sellers (femalesellers). The marks are labeled by sum of number of sellers (femalesellers). Details are shown for English Country Name.

Total Products Bought by Country



16. English Country Name and Total Prod Bought . Color shows Total Prod Bought . Size shows Total Prod Bought . The marks are labeled by English Country Name and Total Prod Bought .

Languages Spoken by Users in Each Country

What languages are spoken by users in each country?

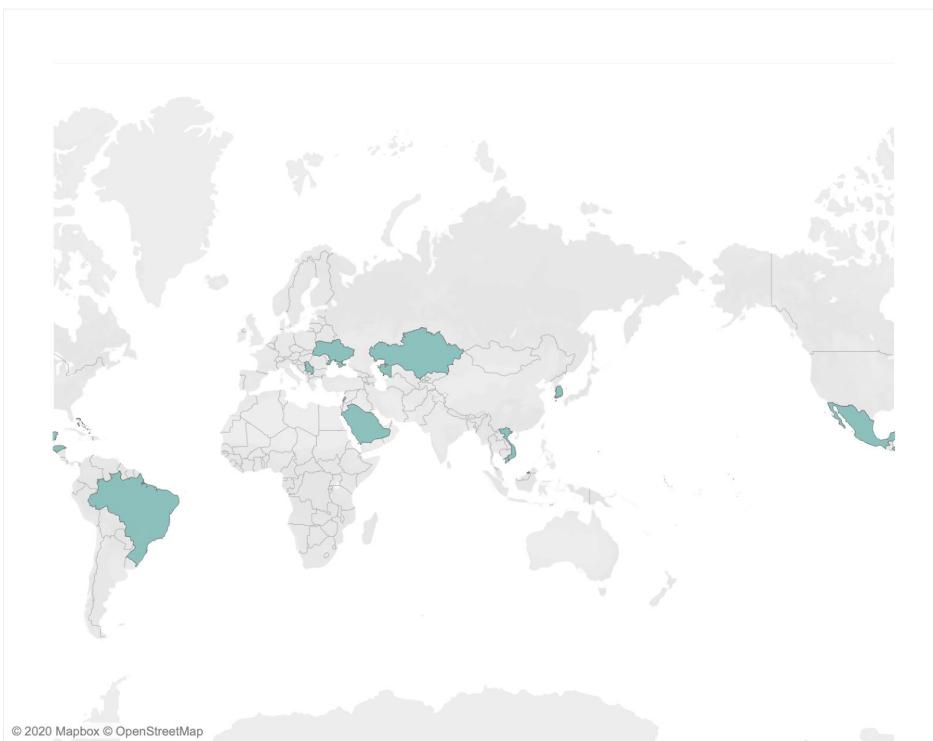
Pertains to MySQL Query by Renee Gagne



The view is broken down by Language (Usersocialdata) vs. Country (Usersocialdata). Details are shown for Country (Usersocialdata).

Countries With Only Female Users

What countries only have female users?
Pertains to MySQL query by Renee Gagne



18.

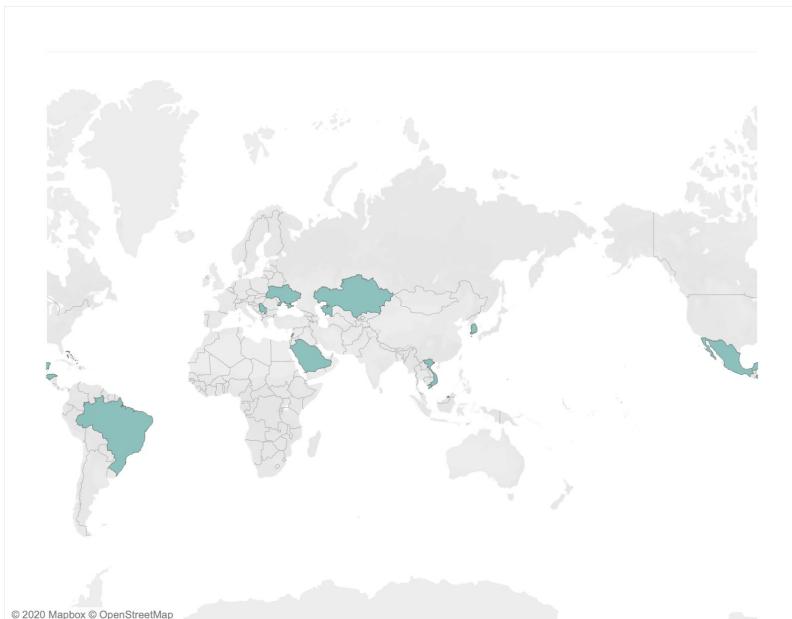
Average Days Since Last Login

Top 10 countries with the highest last log on averages
Renee Gagne



Countries With Only Female Users

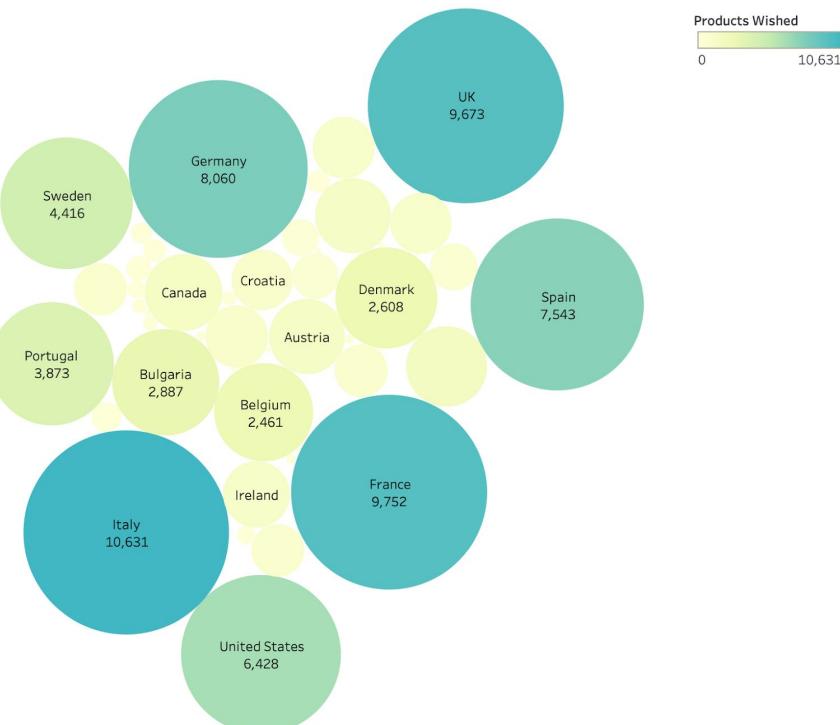
What countries only have female users?
Pertains to MySQL query by Renee Gagne



19.

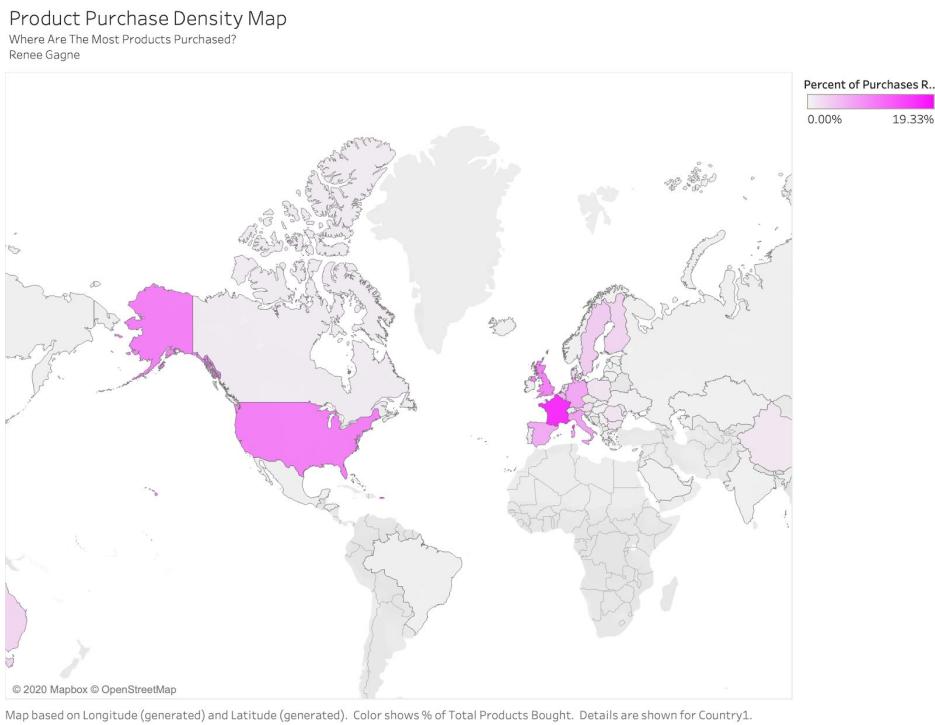
Products Wished By Users

Users in which countries have the most product wishes?
Renee Gagne



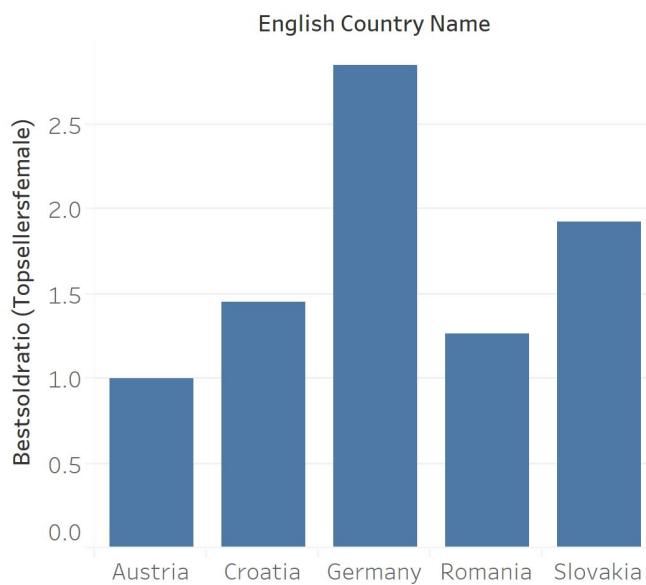
20.

English Country Name and sum of Products Wished. Color shows sum of Products Wished. Size shows sum of Products Wished. The marks are labeled by English Country Name and sum of Products Wished. The view is filtered on English Country Name, which keeps 200 of 200 members.



21.

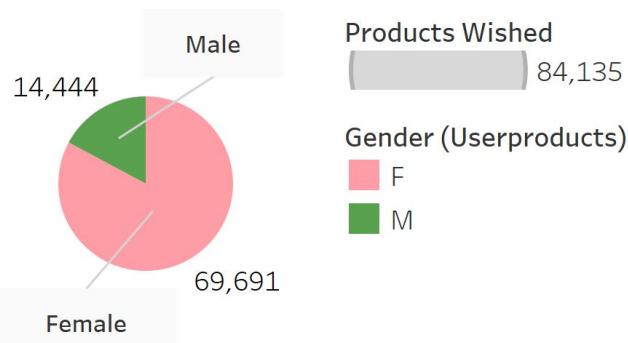
Top 5 countries with best sold ratio for females



Sum of Bestsoldratio (Topsellersfemale) for each English Country Name. The view is filtered on English Country Name, which has multiple members selected.

22.

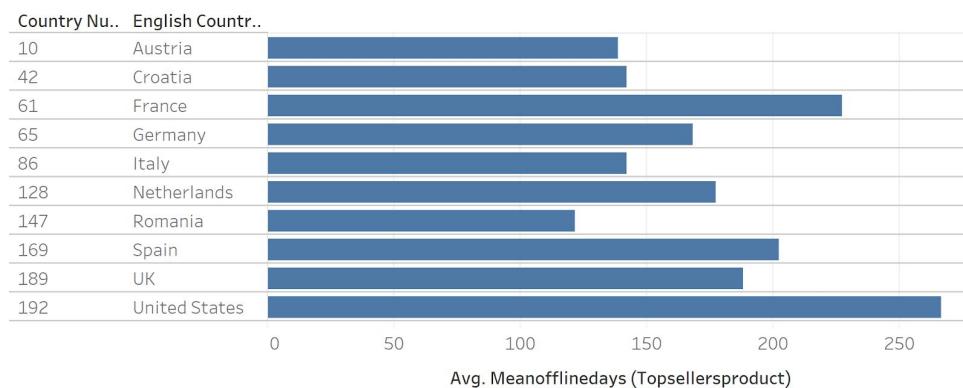
Products "Wished" By Gender



Gender (Userproducts)
(color) and sum of Products
Wished (size).

23.

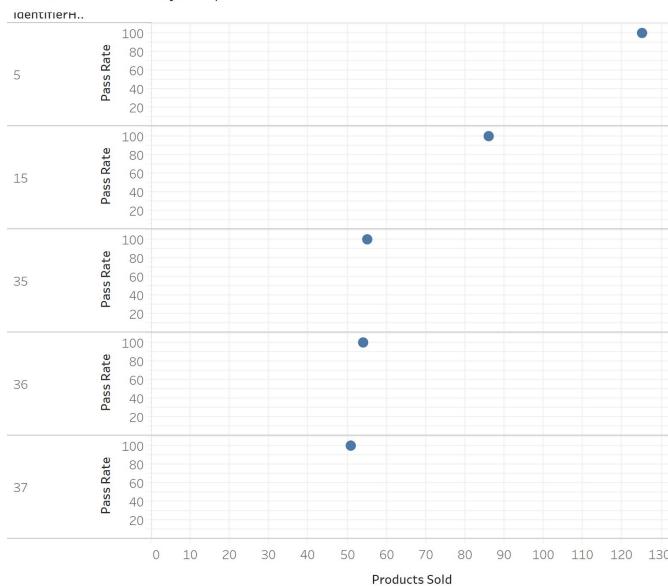
Average Number of Offline Days from Sellers within the Top10



Average of Meanofflinedays (Topsellersproduct) for each English Country Name broken down by Country Number Code.
The data is filtered on Country (Topsellersproduct), which has multiple members selected.

24.

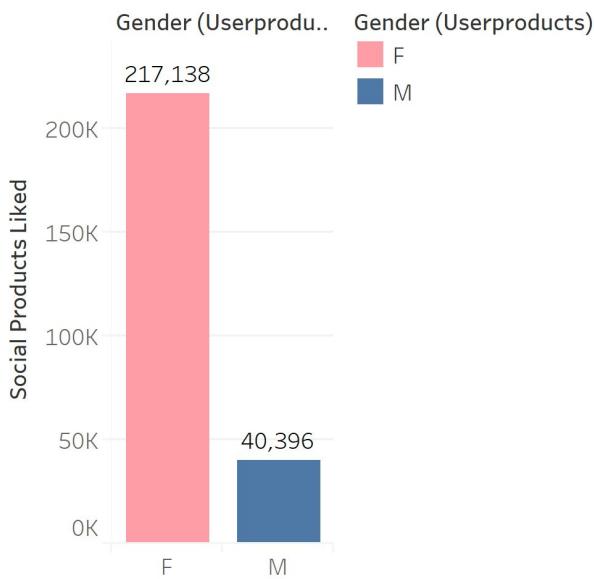
Product Accuracy: Top 5 Sellers



Sum of Products Sold vs. sum of Products Pass Rate broken down by identifierHash (userproducts). The view is filtered on identifierHash (userproducts), which has multiple members selected.

25.

Products Liked between Male & Female



Sum of Social Products Liked for each Gender (Userproducts).
Color shows details about Gender (Userproducts).

Appendix C: Functions, procedure, and trigger

Function 1 Creation:

```

DELIMITER //
• CREATE FUNCTION PopularCountries (totalProductsSold int) RETURNS VARCHAR(50)
DETERMINISTIC
BEGIN
    DECLARE sellerLvl VARCHAR(50);
    IF totalProductsSold > 500 THEN set sellerLvl = 'Popular country with most sought after listings';
    ELSEIF totalProductsSold < 100 THEN set sellerLvl = 'Unpopular country with least sought after items';
    ELSE SET sellerLvl = 'Country with average number of sought after items';
    END IF;
    RETURN (sellerLvl);
END //
DELIMITER ;

```

Calling on function 1: See Business Question #9

Function 2 Creation:

```

1
2      #EnglishTranslate Stored Functions
3
4 •  drop function if exists  EnglishTranslate;
5      delimiter //
6 •  create function EnglishTranslate(input_frenchname varchar(50))
7      returns varchar(50)
8      deterministic
9      begin
10         declare EnglishName varchar(50);
11         set EnglishName=(select englishcountryname
12             from countries where input_frenchname=frenchcountryname);
13         return (englishname);
14     end//
15     delimiter ;
16

```

Calling on function 2: See Business Question #2

Procedure Creation:

```

DROP PROCEDURE IF EXISTS userRatio;
DELIMITER //
CREATE PROCEDURE userRatio (
    IN idNo INT,
    IN prodListed INT,
    IN prodSold INT,
    IN prodWished INT,
    IN prodBought INT,
    OUT sellRatio DECIMAL(10,2),
    OUT buyRatio DECIMAL (10,2)
)
BEGIN
    SET sellRatio = (SELECT productsListed / productsSold
                     FROM userProducts
                     WHERE idNo = identifierHash
                     GROUP BY identifierHash );

    SET buyRatio = (SELECT productsWished/productsBought
                     FROM userProducts
                     WHERE idNo = identifierHash
                     GROUP BY identifierHash );
END //
DELIMITER ;

```

Calling on the procedure:

```

121      ## CALLING on the user ratio funcion for
122          -- user # 1, who has the most products sold (174)
123          -- user # 1145, who has the most products purchased (405)
124
125      CALL userRatio(1, @prodListed, @prodSold, @prodWished, @prodBought, @sellRatio, @buyerRatio);
126 •   SELECT 1, @sellRatio, @buyerRatio;
127
128 •   CALL userRatio(1145, @prodListed, @prodSold, @prodWished, @prodBought, @sellRatio, @buyerRatio);
129 •   SELECT 1145, @sellRatio, @buyerRatio;
130
100%  ◁ 1:118 |
```

Result Grid Filter Rows: Search Export:

	@sellRatio	@buyerRatio
1145	1.00	4.73

Trigger Creation:

```

DELIMITER //
CREATE TRIGGER capCountryName
BEFORE INSERT ON countries
FOR EACH ROW
BEGIN
    SET NEW.FrenchCountryName = CONCAT(UPPER(SUBSTRING(NEW.FrenchCountryName, 1, 1) ), LOWER(SUBSTRING(NEW.FrenchCountryName FROM 2 )));
END //
DELIMITER ;

```