Visual Analytics - Assignment 1 - Report

Petridou Elena (s2029197), Spiro Andrew (s4015355)

March 2024

1 Explaining the Dashboard

Prompt: Explain the dashboard functionalities (each of the components) in the report. No need to list actual code in the report.

The dashboard consists of 4 interactive widgets. Worldwide Data summarizes global ratings and sales, Ratings and Stability explores the relationship between customer satisfaction and application crashes, Sales describes sales volume across markets and over time, and Sales by Attributes segments sales volume by product and country. All widgets are tab objects, showing different metrics and data in each tab panel. All widgets include sidebars with icons which can be toggled to enable or disable interactive features. At a minimum, all widgets support zooming, dragging, and resetting. Others showing multiple data on the same plot have icons to enable or disable hiding data by clicking on their corresponding labels on the legend.

1.1 Worldwide Data

The Worldwide Data widget consists of four tabs showing geographic data on a color-coded world map. The **Cumulative Sales Revenue per Day** tab shows the total number of transactions made per country from June 1, 2021, up to the date indicated on the slider. Data collection stops at January 1, 2022. In the seven month period, total transactions range from 0 to just over 5,000. The **Historical Average Ratings** tab, shows the total rating of the application per country, over all time, up to January 1, 2022 and on a scale from 1 to 5. The **Average Rating per Month** tab, similar to the **Historical Average Ratings** tab, shows the rating of the application per country averaged from the earliest recording, but only up until the month selected by the slider. Finally, the **Recent Average Ratings** tab shows the average of all ratings between June 1, 2021 and January 1, 2022.

For all maps, hovering over a country provides country-specific information relevant to the current tab and, if applicable, the slider position.

1.2 Ratings and Stability

The Rating and Stability widget consists of three panels: **Daily Ratings**, **Total Ratings**, and **Ratings vs Crashes**. The **Daily Ratings** tab shows both daily rating and number of crashes as a function of time with days on the horizontal axis. Two vertical axes indicate the number of daily crashes and the daily average rating. The colors and labels indicate which vertical axis corresponds to which plot. The **Total Ratings** tab similarly shows ratings and crashes as a function of time, but the rating is plotted as the total average calculated over all time, up to the corresponding date on the horizontal axis. (Accordingly, the second vertical axis shows the total average rating.) For both the **Daily Ratings** tab and the **Total Ratings** tab hovering over points or points on a line shows the exact value associate with the point. Also, data can be hidden and un-hidden by clicking on the corresponding label in the legends. **Ratings vs Crashes** is a scatter plot showing the daily average rating (vertical axis) as a function of the daily number of crashes (horizontal axis). Hovering over a plotted point also shows the date of the crash.

1.3 Sales

The Sales widget describes sales volume by markets in two tabs and over time in the Sales per Day tab.

The **Emerging Markets** tab shows the total number of downloads of the app between June 1, 2021 up to the date on the Day axis for the 6 countries with the highest number of downloads (except the US). Data is segmented by country, yielding 6 lines on the same plot. A legend indicates which line corresponds to which country and a hover tool enables viewing the exact values associated with a point on any of the country lines.

The **High Vol Countries** tab is essentially the same as the **Emerging Markets** tab, but it includes data from the US. While the **High Vol Countries** tab is generally more comprehensive by including the US, the **Emerging Markets** tab provides management with insights on other markets possibly worth targeting.

The **Sales per Day** tab shows the sales volume as a function of time with days on the horizontal axis, and the total daily sales volume in Euros on one vertical axis and the total number of daily transactions on a second vertical axis. Clicking on the interactive legend toggles between hiding and showing a plot, enabling a user to focus on one dataset at a time. Twin axes also allow for simultaneous and accurate rendering of the two different metrics. A hover tool allows users to check the exact value of a point on either plot. Finally the lines, legend, and axes are all color-coded.

1.4 Sales by Attributes

The Sales Volume Segmentation widget has three panels, showing total sales volume distributed over the type of in-app purchase (identified by Sku Id), the product Id, and the country of the buyer. Similar to the **Sales per Day** tab, two vertical axes and an interactive legend are used to plot sales volume in number of transactions and in Euros at the same. Hovering and hiding data are also enabled, allowing a user to easily check the exact quantity for any bar.

2 Visualisation

Prompt: For each component of your dashboard, explain how you mapped data attributes to visual attributes, and why you made various visualization or interaction design choices. How do these choices help management understand the data without having to understand your code?

2.1 Worldwide Data

2.1.1 Cumulative Sales Revenue by Date

The motivation behind this map was to show in an intuitive manner, the way global downloads of the app are distributed across countries during the period from June to December 2021. For this reason, the default state of the map shows all countries in white, and we chose a monochromatic, sequential palette, with a higher number of cumulative sales being coloured with a deeper blue. No other colours were used to ensure that no qualitative difference between countries is implied. An important caveat has to be noted here: the sales volume in the United States far outweighs the sales volume for any of the other countries in the datasets. For this reason, the values for the maximum value for revenue from the US had to be clamped in order for the colour differential of other countries to be visible. For this reason, we envisioned this plot more as an intuitive visualisation of sales globally, and not as an accurate tool for assessing turnover.

Moving on to the selection of the measure: the reason for plotting a cumulative sum of revenue instead of per-day sales revenue is that through this, the stakeholders can more easily understand where and by how much sales grew over the period. This is more user-friendly than a sales-per-day plot, where the colour of the previous day has no bearing on the colour of the next. To produce this measure, we summed the revenue earned across all products, using the *Amount (Merchant Currency)* for **months up to November**, and then converted the local currencies using the local currency to Euro exchange rate posted for the day of the transaction (if available, and the closest available date otherwise), for the other months. Each day's cumulative value is thus the cumulative sum of revenue up to and including that date. This then determines the colour of the country shape on the map.

As for the slider, the update frequency was set to the day to ensure a smooth gradient over time.

2.1.2 Country Ratings Maps

We have plotted three different maps pertaining to average rating, and the reasons behind this decision were informed both by the measures available in the data, but also by the questions we wished to answer in this report. In the provided datasets, two measures of average rating were available: the *Total Average Rating*, and the *Daily Average Rating*. The first measure is the mean score of all reviews for the app from all time up to that date, the second measure is the mean of all review scores left for the app on a particular day. The **Average Country Ratings per Month choropleth** plots the first measure, up to the month selected with the slider. The **Historical Average Rating by Country map** plots the first measure also, but up to the end of the period, and is static. The **Recent Average Rating by Country** plots the mean of only the latter measure, in order to explain user opinions of the app during the period from June to December 2021, without taking into account older reviews of the app. Under each heading, we explain the design choices and motivation for the homonymous map in more detail.

2.1.3 Average Country Ratings per Month

The **Average Country Rating Per Month choropleth** shows how the *Total Average Rating* for the app changed during the period across countries. In fact, what the graph visualises is that average ratings have been rather constant from June to December 2021, apart from a few noteworthy differences which we discuss more at length in the section of the report titled *Client Satisfaction in a Geographical Context*.

The colour palette selected here is also sequential, but in this palette, slight tonal differences between the colours are also visible. This is permissible – and even desirable – as review scores have an ordinal interpretation, as opposed to the continuous range of sales volume.

The slider was set to have months as steps so that management can quickly toggle between clearly defined timeframes. For instance: it is not very meaningful to talk about the difference in average rating between 1st June and 20th June, whereas a more common question might be to wonder if there are differences between the average rating in July and the average rating in November.

2.1.4 Historical Average Rating by Country

The **Historical Average Rating by Country choropleth** visualises the mean rating per country including both the reviews left during the period from June to December 2021, and the reviews from before this time. In fact, this static choropleth is the same as the last slide of the per-month choropleth. However, a separate static one was plotted chiefly to make discussion in the **Client Satisfaction in a Geographical Context** section of this report easier to follow. The reason for this is that it provides a clear distinction between the average reviews of the app over all time versus the average reviews of the app only in the period we are examining.

2.1.5 Recent Average Rating by Country

The **Recent Average Rating by Country** choropleth plots the mean rating per country, calculated only using the user ratings left over the period from September to December 2021. This was plotted in order to help management see what users think of the app only during the examined period, and not taking into consideration ratings for the app left at earlier dates. The colour palette is the same as the preceding maps to make comparisons between the three easier.

2.2 Ratings and Stability

Prompt: For each component of your dashboard, explain how you mapped data attributes to visual attributes, and why you made various visualization or interaction design choices. How do these choices help management understand the data without having to understand your code?

2.2.1 Daily Ratings and Total Ratings

The **Daily Ratings** plot and the **Total Ratings** plot were created from a ColumnDataSource with the date used as an index, and with Total Average Rating, Daily Crashes, Daily ANRs, Daily Average Rating, and Cumulative Crashes (the total number of crashes up to a date) as columns. Plots were created using the date, Total Average Rating, Daily Average Rating, and Daily Crashes.

Both the **Daily Ratings** tab and the **Total Ratings** tab consist of two overlaid plots, both functions of time: a measure of the app's ratings and the daily number of crashes. In the **Daily Ratings** tab, a scatter plot showing the daily ratings was overlaid on top of a line plot showing the daily number of crashes. A scatter plot was chosen to represent the daily rating as data for this metric is not available for every date. Also, plotting the points as a line results in apparently random line segments, not revealing any clear relationship between the points. For the total ratings plot, however, each point shows the total average rating up to the date corresponding to its x-coordinate. As the total average rating for any given day is simply the average of all ratings before that day, data can be generated for all days and plotting a line reveals a clear (nearly constant) relationship between the data points.

The objective of the **Daily Ratings** plot is to show, for a given period, how the number of crashes affected the rating in that period. As supported by the scatter plot (described below), a clear relationship between the number of crashes and the daily rating of the app is not present except that both the total number of ratings and the number of 5/5 ratings is higher when the number of crashes is below a certain threshold (approximately 30). (Further, no trend is discernible in the daily rating data alone, independent from the crashes.)

For the **Total Ratings** plot, the objective is to show how the number of crashes have affected the long-term rating of the app. The near-constant total rating compared with the jump in daily crashes might suggest the two metrics are mostly independent.

Of the two plots, the **Daily Ratings** plot may be of more use to management as it suggests that less than about 30 daily crashes increases 5/5 ratings. It also shows that more ratings in total happen with fewer crashes, possibly suggesting higher user involvement when the app is more functional. This relationship is clearer still in the **Ratings vs Crashes** tab, discussed below. The **Total ratings** plot, on the other hand implies that the two are almost, if not completely, independent.

2.2.2 Ratings vs Crashes

For the **Ratings v Crashes** tab, a ColumnDataSource object was made containing three columns: the date, daily crashes, and the daily average rating. Each element in the daily average rating column was plotted against the corresponding (same date) element in the daily crashes column. Rows for which there was no rating were dropped from the ColumnDataSource.

The plot was created as a scatter plot with circles as markers. A scatter plot was chosen as it clearly shows the relationship between the number of crashes and the daily rating for any given data-point without imposing any interpretation on the dataset as whole (as a line plot may do by generating information wherever data is absent). The initial intent had been to add a line of best fit to infer a trend from the scatter plot, however the line of best fit was approximately constant. Being constant, the trend line thus showed no correlation between the two metrics. The large distance of the majority of points from the trend line further suggested the line did not accurately represent the behavior of the data, so it was omitted from the plot. Despite the constant trend line implying low/no correlation, the management can see at a glance from this plot that the majority of 5/5 ratings occur with fewer than 25 daily crashes. The plot also shows that more ratings in total occur with fewer daily crashes. Again, this may suggest that users are more involved when the app experiences fewer daily crashes.

2.3 Sales

Moving on to the Sales widget, we decided to group three related but different graphs, the first two of which have to do with the global origin of clients, and the last a line graph showing sales per day for all products (of the ddfive type), broken down by transaction count and revenue.

2.3.1 Emerging Markets

Beginning with the emerging markets graph, our aim here was to show how sales volume progressed through the period in the largest markets for the app outside the US. The reason was to answer managements' questions regarding the targeted advertising campaign which they wish to run in three emerging markets.

To calculate the measure used for this graph, we counted the unique transactions recorded in the sales per month datasets, and grouped them by country. We then populated each date's entry with the cumulative transaction count up to and including that day. Finally, we sorted the transactions by highest volume, and took out the US. The resulting plot is a line chart for the cumulative downloads for the app per country up to and including the date displayed when hovering over the line. The motivation for plotting this particular measure is to explain how fast and by how much download volume increased in each of the emerging markets over the period.

A different colour was used for each country, so stakeholders can easily differentiate between each market of interest. The hover feature displays both the amount of downloads up to and including the date, as well as the date itself, so management can see during which periods of time downloads grew the fastest, or slowed down. The ticker at the bottom of the graph shows the months, so that hovering over the last day of the month essentially provides the cumulative number of downloads up to and including that month.

2.3.2 Cumulative Downloads to Date (High Volume Countries)

This graph – found under the 'High Vol' tab is essentially the same as the above, only with the US sales volume included as well. The aim of this was to show management that sales volume in the US far surpasses the following emerging countries, and that sales in this market grow at a fast and steady rate. This graph is mostly for visualisation/intuition purposes and not meant to give much information about the emerging markets themselves.

The measure used for the cumulative download volume is the same as the one described above, only with the US also included.

2.3.3 Sales per Day

The Sales per Day graph is different from the above two, as it does not segment sales per country, and instead shows the download volume (number of transactions), and revenue (amount in EUR) per day. This plot was created to illustrate the changes in sales volume between June 1, 2021 and January 1,2021. The management may benefit from such a visualization in various ways including gaining insights into long-term trends, cyclic patterns, or anomalies in sales. It may also be used in conjunction with other temporal data, for example the release date of a feature, to understand the impact of a particular event on sales volume.

The two metrics measuring sales volume in this plot are the number of daily transactions and the daily sales revenue. The plot combines sales data from two products: the character manager and DM tools. If the same number of these two items were sold every day, the revenue would be directly proportional to the number of transactions. In fact, revenue would equal half the number of transactions multiplied by the sum of each product's individual cost:

$$revenue = (c_1 + c_2) \cdot \frac{1}{2}(n_1 + n_2),$$

where c_1 and c_2 are the costs of the two products, and n is the number of items sold. The vertical re-scaling of one of the datasets would result in a complete overlap between the two plots. After scaling both plots, the Daily Amount line tends to be above the Daily Transactions line. This suggests that the more expensive product is slightly outselling the less expensive product. Management may interpret this as higher cost not discouraging users from buying a more valuable product.

To create the Daily Transactions plot, we filtered for all unique transactions and calculated the number that occurred on the same day for every given day. The Daily Amount plot was created similarly, though the sum of the amount in the merchant's currency was calculated for every day. Both of these were plotted against the transaction date.

2.4 Sales by Attribute

The bar plots in this section were plotted with the aim to help management understand how sales are split according to certain attributes

2.4.1 Per SKU ID

This plot depicts the total count of each in-app feature sold over the studied period, and the total revenue in Euros made from each feature over the period. To produce the metrics seen on the graph, we grouped the data by *SKU Id*, then counted the unique transactions to come up with the count metric. To calculate the Revenue, we summed over the *Amount (Merchant Currency)* for the months preceding November, whereas for the months of November and December, we used a currency conversion library to convert the Item Price from the Currency of Sale to Euros, using the recorded exchange rate between the two countries on the date of the transaction – or if that was not available, the temporally closest exchange rate to the given day.

The resulting plot, we hope, helps management determine which in-app features are of most interest to users, and also which in-app features produce the most revenue for the company.

By plotting both measures on the same graph, we wished to keep the dashboard clutter-free, and thematically group measures which answer similar questions together. As such, using only one graph, the stakeholders can answer both the question of: which in-app purchases are the most bought, and also which in-app purchases are the most profitable?

2.4.2 Per Product Title

This plot, grouped in the same widget due to the topic similarity, splits app purchases based on the name of the app.

The measures we used were calculated in the same way as the counts and revenue for SKU Ids, as discussed above, though grouped by *Product ID* instead.

We used the same colours are for the previous graph since SKU-ID and Product ID refer to the same things, but plotted it for ease of use of management.

2.4.3 Per Country

While quite similar to the choropleth plots, we decided to also plot revenue and sales count per country in a bar plot, since bar plots are easier to glean exact data from than choropleths (though perhaps not as beautiful or intuitive).

As with the previous two plots, we grouped by *Country of Sale* and followed the same procedure to produce the Counts and Revenue measures.

3 Satisfied and Critical Customers

For the following two sections, we have provided several queries which can be used in the submitted jupyter notebook to produce tables with descriptive statistics. They are also listed here inline.

3.1 Client Satisfaction in a Geographical Context

Prompt: Which country or region has the most users that rate the app negatively, and which country or region has the highest average rating?

Client satisfaction with the app varies to a noticeable but not dramatic extent on the geographical axis. **Query 1** produces a table which shows the average rating for the app per continent, demonstrating that the most satisfied region is **Oceania** with a mean rating of **4.541** (to 3 decimal points). In contrast, the continent with the least satisfied users appears to be **Africa** with a mean rating of **3.842** (to 3 decimal points).

```
# Query 1: Average Rating per Continent
continent_ratings = average_country_rating.merge(country_codes,
 how = 'outer', left_on='Country', right_on='alpha-2')
average_continent_ratings = continent_ratings.groupby('region')
.mean()
average_continent_ratings.to_latex(escape = False)
```

However, an important caveat to note is that no new reviews have been left from users in Africa during the studied period. As such, when considering the user satisfaction with the app exclusively over the studied period, it is perhaps more accurate to conclude – using the table produced by **Query 2** – that the most satisfied continent is **Asia**, with a **mean daily rating of 4.6**, calculated from five reviews left in this period. In contrast, the least satisfied continent during this period is **America**, with a **mean daily rating of 3.615** (to 3 decimal points), resulting from a total of **52 reviews** submitted during this period.

```
# Query 2: Daily ratings per continent
continent_daily_ratings = per_country[per_country["Daily Average Rating"].notna()].merge(country_codes, how = 'outer', left_on = 'Country', right_on = 'alpha-2')
continent_daily_ratings_summary = continent_daily_ratings.groupby('region').mean().sort_values(by = "Daily Average Rating")
continent_daily_ratings_summary["Total Daily Reviews per Region"] = continent_daily_ratings.loc[:,["Daily Average Rating", "region"]].groupby('region').count()
continent_daily_ratings_summary
```

As can be gleaned from the above numbers, not many new ratings per country were recorded from June to December 2021. The reviews left during this period per country are viewable via **Query 3** which breaks down the ratings grouped by score and by country. We see that the greatest number of positive reviews, as well as negative reviews were left by customers in the **United States (US)**.

```
# Query 3: Counts of daily ratings by score, per country over the period
lowest_ratings = per_country.sort_values(by = 'Daily Average Rating')
lowest_ratings_by_country = per_country.loc[:,["Country","Daily Average Rating","Total Average Rating"]].groupby(["Country", "Daily Average Rating"]).count()
lowest_ratings_by_country = lowest_ratings_by_country.rename(columns = {"Total Average Rating":"Count of Score"})
lowest_ratings_by_country_sorted = lowest_ratings_by_country.sort_values(by = ['Daily Average Rating', 'Count of Score'])
lowest_ratings_by_country_sorted
```

Perhaps a clearer image of the average rating per country recorded only during the period of June to December 2021 is given by the **Recent Average Rating per Country map**, which plots the average review scores per country for this period. Hovering over each country reveals the average review score given by users. From this, it is clear that the most displeased country with the app during this period was **France (FR)**, with an average review score of **2.33**. This is based on **three ratings** by users, which is a very low number, but not unusually low for this dataset. The frequency table for the counts of ratings collected during the period per country is accessible via **Query 4**, and demonstrates that the number of ratings from users in only one country reached double digits: the **United States (US)** with **45 new ratings** from June to December 2021. As for the most satisfied countries during this period, there are many countries with only new **5-star** ratings. Namely, **Iran (IR), Turkey (TR), Germany (DE), Spain (ES), Colombia (CO), and Indonesia (ID)**, countries from which only one review was submitted during the examined period.

```
# Query 4: Frequency table of ratings per country registered
over the period
count_of_scores_by_country = per_country.loc[:,["Country","
Daily Average Rating"]].groupby(["Country"]).count()
count_of_scores_by_country = count_of_scores_by_country.rename(
columns = {"Daily Average Rating":"Review Counts"})
print(count_of_scores_by_country.loc["FR"])
sorted_score_counts = count_of_scores_by_country.sort_values(by
 = "Review Counts", ascending = False)
sorted_score_counts.head(n = 35)
```

For the rest of this section, we discuss both the **Total Average Rating per Country** and the **Recent Average Rating per Country**. The first measure (Total Average Rating) is tallied over a wider period of time than the months of June to December 2021 and reflects the historical average opinions of users per country. This is plotted on the **Historic Average Rating by Country choropleth**, and the scores for individual countries discussed in the report are available by hovering over the shape of the country on the map. In contrast, the **Recent Average Rating per Country** has been calculated using only the **Daily Average Rating** measures, found in the stats_ratings_2021mm_country datasets. This is plotted on the **Recent Average Rating by Country choropleth** and only reflects user opinions in the span of June to December 2021.

3.1.1 North America

Beginning with the Americas, users in the overwhelming majority of countries in the two subcontinents rated the app decidedly positively. There are a few notable exceptions to discuss, however: Hovering over the lighter values in North America, we observe the lowest ratings in **Guatemala (GT)** with a **2.5 average rating**, and the **Dominican Republic (DO)** with an **average rating of 2**. However, using the **Recent Average Rating per Country map** we can see that these two values have been carried over from an older period through the **Total Average Rating** measure, as there has been no new rating posted from a user in either of the two countries in the examined time period. Moreover, **Query 5** shows that no new purchase has been made by a customer in either of the two countries within the studied period, meaning that we could potentially theorise that the reason for the abnormally low rating observed in these two countries as opposed to the generally high ratings among clients from other countries in the region could be a low number of total reviews. This would mean that the low average rating could be spuriously caused by only a handful of bad reviews. Unfortunately, it is not possible to examine this theory, as it falls outside the scope of the available data (i.e.: it is based on reviews posted before the beginning of the examined time period).

```python
# Query 5: Sales volume for low-rating countries in North America (Dominican Republic and Guatemala)
sales_by_country_per_day_temp = sales_by_country_per_day.loc[:,["Buyer Country", "Amount (Merchant Currency)"]].groupby(['Buyer Country'], dropna = False).count()
sales_by_country_per_day_temp = sales_by_country_per_day_temp.rename(columns = {"Amount (Merchant Currency)": "Units Sold"})
sales_all_countries = country_codes.merge(sales_by_country_per_day_temp, left_on='alpha-2', right_on = 'Buyer Country', how = 'outer')
sales_all_countries = sales_all_countries.fillna(0)
sales_all_countries = sales_all_countries.set_index("alpha-2")
sales_DO = sales_all_countries.loc["DO"]
sales_GT = sales_all_countries.loc["GT"]
sales_all_countries_sorted = sales_all_countries.sort_values(by = "Units Sold", ascending = False)
sales_all_countries_sorted.head(n=50)
```

On the other hand, countries in North America with an overall high average rating tend to also have a high(er) volume of downloads, meaning that this is more likely to be an accurate reflection of the opinions of current or long-time users of the app. For example, **Mexico (MX)** registers a **total mean rating of 4.49** as seen by hovering over it on the **Historical Average Country Rating plot**. While **Query 6** shows that no new ratings were left from users in **Mexico** during this time period, the purchasing volume in units was higher than for the two lowest-rating countries, at **23 units** over the studied period, as can be seen by hovering over **MX** on the **Sales per Country bar plot** and confirmed via **Query 6** also.

```python
# Query 6: Units sold over the period and daily ratings for Mexico
mexico_sales = sales_all_countries.loc["MX"]
mexico_daily_ratings = sorted_score_counts.loc["MX"]
print(mexico_daily_ratings)
```

Another country with overall positive reviews in North America is **Canada (CA)** with a mean rating of **4.12**. As can be observed via the **Average Country Rating per Month plot**, the mean rating for this country changes throughout the studied period, from **4.12** in **June** to **4.13** from **September to October**, then slightly decreases back to **4.12 in December**. This is caused by the reviews left in July, November and December, which are viewable using **Query 7**. Two of these ratings have a score of **3.0**, and were both left in November, whereas a **5.0** was given in July and a **4.0** was given in December. As shown in the **Average Rating VS Number of Crashes line plot** the number of crashes the app experienced began to rise suddenly at the beginning of November, pointing to a possible reason for the less-than-outstanding reviews from these users left in November.

```
# Query 7: Units sold over the period and daily ratings for Canada
canada_sales = sales_all_countries.loc["CA"]
canada_daily_ratings = lowest_ratings_by_country_sorted.loc["CA"]
print(canada_daily_ratings)
```

Similar patterns of availability of daily reviews persist for other continents as well, so we limit the following discussion only to the scores per country broken down by continents. Querying the data using **Query YOUROWN** and entering the Alpha-2 code of the country of interest within the square brackets, yields similar insights to the above for any country.

```
# Query YOUROWN: Units sold over the period and daily ratings for any country of your choosing
custom_country_sales = sales_all_countries.loc["enter alpha-2 here"]
custom_country_daily_ratings = lowest_ratings_by_country_sorted.loc["enter alpha-2 here"]
print(custom_country_sales)
print(custom_country_daily_ratings)
```

3.1.2 South America

The most satisfied country with the app in South America *based on new reviews* was **Colombia (CO)**, which we have already discussed above. Using the *historical average*, however, users in **Peru (PE)- 4.86**, **Argentina (AR) - 4.20**, and **Brazil (BR) - 4.34**, also rated the app highly. For the former, no new reviews were submitted, whereas for the latter two, new reviews showed scores lower than their historical average. These are viewable using **Query 8**. Despite this, only 1 each new review was submitted, so no conclusive evidence can be drawn. The least satisfied country with the app in South America based on new reviews was, in fact, **Brazil (BR)**, with a **review score of 3.0** based on the one review submitted during the period. Taking historical as well as recent data, the least satisfied country in South America appears to be **Paraguay (PY)** with a total rating score of 1. Since this is a score reflected only in the Total Average Ratings measure, it is impossible to say what sample size it is based on, however.

```
# Query 8: New Reviews for clients in South America
argentina_sales = sales_all_countries.loc["AR"]
argentina_daily_ratings = lowest_ratings_by_country_sorted.loc["AR"]
print(argentina_sales)
print(argentina_daily_ratings)
brazil_sales = sales_all_countries.loc["BR"]
brazil_daily_ratings = lowest_ratings_by_country_sorted.loc["BR"]
print(brazil_sales)
print(brazil_daily_ratings)
```

3.1.3 Africa

As stated in the beginning of this section, no new reviews were submitted by users in Africa during the period of available data. However, using the historical averages, it appears the most satisfied users with the app were located in **Egypt (EG)**, with an average review score of **4.86**, whereas the least satisfied users seem to be from **Botswana (BW)**. Using **Query 9**, we see that during the studied period, there were no new purchases of the app from either of these countries.

```
 # Query 9: Download volume in most and least satisfied countries in Africa
egypt_sales = sales_all_countries.loc["EG"]
print(egypt_sales)
botswana_sales = sales_all_countries.loc["BW"]
print(botswana_sales)
```

3.1.4 Asia

As noted above, **Turkey (TR), Iran (IR)**, and **Indonesia (ID)**, were among the global most highly-rating users from June to December 2021, with only 5-star reviews. This trend is also reflected in the historical averages, with the average scores being **5.0, 5.0** and **4.57** respectively. The sales volume in **Query 10**, however, shows that no units were bought by people in Iran during the months of June to December 2021, and **4 units** each were bought by people in **Turkey** and **Indonesia**. Moreover, another country with a high total average rating when considering the historical data, is **China (CN)**, with a **total average score of 5.0**, which, however, might be based on a low count of ratings seeing as no new sales were made in **China** within the five months of available data. In contrast, the least satisfied users in Asia based on total reviews are found in **Iraq (IQ)** with a total average rating of **2.0**. As for sales in this country, no new units were bought during the months studied. It is difficult to say which was the least satisfied country in Asia based on recent scores, as all of them were very positive, and not many new reviews from Asia were submitted during the five months of interest.

```
# Query 10: Download volume in most and least satisfied countries in Asia
turkey_sales = sales_all_countries.loc["TR"]
iran_sales = sales_all_countries.loc["IR"]
indonesia_sales = sales_all_countries.loc["ID"]
china_sales = sales_all_countries.loc["CN"]
iraq_sales = sales_all_countries.loc["IQ"]
print(turkey_sales)
print(iran_sales)
print(indonesia_sales)
print(china_sales)
print(iraq_sales)
```

3.1.5 Europe

According to reviews collected during the five months of data availability, the most satisfied countries in Europe were **Germany (DE)** and **Spain (ES)** with an **average rating based on daily reviews of 5.0**, based on **one review each**. According to **Query 11**, **Germany** also had a relatively high number of sales as compared to other countries outside the Anglosphere, with **47 units sold**. The least satisfied country in Europe during the review period was **France (FR)** with an average review score between June and December 2021 of **2.33**. The individual review scores from France during this period were **two counts of a 1-star, and one count of a 5-star**, and the download volume was **3 units** also.

```
# Query 11: Download volume and ratings in most and least satisfied countries in Europe
germany_sales = sales_all_countries.loc["DE"]
spain_sales = sales_all_countries.loc["ES"]
france_daily_ratings = lowest_ratings_by_country_sorted.loc["FR"]
france_sales = sales_all_countries.loc["FR"]
```

3.1.6 Oceania

Finally, in Oceania we see that the most satisfied country in only the period studied is **Australia (AU)** with an average rating of **4.00**, based on **two recent reviews**, and a **download volume of 42 units** during the latter 5 months of 2021. The only other country with a review score in Oceania is **New Zealand (NZ)** with a **historical average rating of 4.43**, and no new reviews or sales recorded in the available data. You can see this using **Query 12**.

```
# Query 12: Download volume in most and least satisfied countries in Oceania
australia_daily_ratings = lowest_ratings_by_country_sorted.loc["AU"]
australia_sales = sales_all_countries.loc["AU"]
new_zealand_daily_ratings = lowest_ratings_by_country_sorted.loc["NZ"]
new_zealand_sales = sales_all_countries.loc["NZ"]
```

3.2 Crashes and Daily Average Rating

Prompt: In the 5-month duration, several app versions have been released. In one of the releases was a serious bug that affected all android 7 users. Around which date do you think this happened? Look at the crashes and the average daily rating.

To try and understand around which date the version including the bug was released, we use **Query 13** to look for reviews mentioning a crash. The only two which come up are on **13th July** and **17th August**. An additional review on **12th June** makes reference to the app being "buggy". When scanning the **Average Rating VS Number of Crashes line plots**, and hovering over the line for exact dates, we do not notice a peak on either of those day but what is evident are three local maxima in the number of crashes per day: the first and greatest peak is on **27th June, with 57 crashes**; the second is on **7th July with 37 crashes**, which could have motivated the first review talking about crashes, and the third local maximum is on **16th August with 26 crashes**, which may have been the reason for the second review mentioning crashes.

```python
# Query 13: Reviews mentioning a crash
pd.set_option('display.max_colwidth', None)
new_reviews = reviews.loc[:, ["Review Text", "Developer Reply Text", "Review Submit Date and Time", "Star Rating", "Device"]]
new_reviews = new_reviews.drop_duplicates(["Review Submit Date and Time"])
review_samples = new_reviews[new_reviews["Review Text"].notna()].sort_values(by = ["Star Rating"]).set_index("Review Submit Date and Time")
reviews_mentioning_crash = review_samples[review_samples["Review Text"].str.contains("crash")]
reviews_mentioning_crash["Review Text"]
```

Keeping our eyes on the **Average Rating VS Number of Crashes line plots**, we notice that the number of crashes displays an upwards trajectory **from the beginning of November**, which reaches an absolute maximum on **9th December** with a total of **82 crashes** recorded on that date. The number of crashes appears to be taking a downward direction at the end of the studied period, however, not enough data points are available to fully ascertain this. Continuing our search for the approximate release date of the version including the bug, we use **Query 14** to look at all 1-star reviews. Indeed, most of the reviews mention other things about the app that they are displeased with, but not a crash. However, a review left on 25th November reads:

> *"This app has been down for two weeks now. Such a bummer was my go to app. Happy thanksgiving"*

Indeed, the mentioned time period lines up with a huge spike in the number of crashes on **8th November** (we assume that the reviewer was not exactly accurate to the day when mentioning a time-period of two weeks), as well as subsequent spikes on **18th and 23rd November**. This indicates that the timeframe around which the bug was released is in November. We use **Query 15** to pull up reviews from around that period, regardless of score, as users may be experiencing the bugs but still enjoying the overall product, thus not leaving a negative review score. Indeed, here we see that from **20th to 25th November** reviews reference a crash and a subsequent update to the app that fixes it. For example, a review from **20th November** states:

> *"The newest update fixed it, re downloaded the additional content and so far it works great once again!"*

While another reviewer on 21st November writes:

> *"I did what you told me to uninstall n re instal [sic] it but it still did not work. :("*

Finally, **Query 16** pulls up developer replies to reviews around this period, and four replies in a row from **20th to 25th November** contain the phrase: "resolved the issue", then advise users to update the app in a few hours. Finally, the reviews alluding to reliability issues with the app were posted using devices with the following device codenames (also viewable via the aforementioned Query): mido, s5neoltecan , nobleltetmo, gtesltetmo, which are all codenames for android phones.

All this evidence combined leads us to the conclusion that the major bug affecting Android 7 users must have been released with an update sometime during November, and perhaps on November 8th, a day when a local maximum number of crashes was recorded, at 60 crashes.

```
# Query 14: Bad Reviews and Developer replies
bad_reviews = review_samples[review_samples["Star Rating"] == 1]
bad_reviews["Review Text"]
bad_reviews["Developer Reply Text"]
```

```
# Query 15: Reviews from November and December
reviews_end_of_year = review_samples.reset_index(names = "Review Date and Time")
reviews_end_of_year["Review Date and Time"] = pd.to_datetime(reviews_end_of_year["Review Date and Time"])
reviews_end_of_year = reviews_end_of_year[reviews_end_of_year["Review Date and Time"].dt.month >= 11]
reviews_end_of_year = reviews_end_of_year.sort_values(by="Review Date and Time")
reviews_end_of_year
```

```
# Query 16: Developer Replies and phone codenames
developer_replies = reviews.loc[:, ["Review Text", "Developer Reply Text", "Review Submit Date and Time", "Star Rating", "Developer Reply Date and Time"]]
developer_replies = review_samples.reset_index(names = "Developer Reply Date and Time")
developer_replies["Developer Reply Date and Time"] = pd.to_datetime(developer_replies["Developer Reply Date and Time"])
developer_replies = developer_replies[developer_replies["Developer Reply Date and Time"].dt.month >= 11]
developer_replies = developer_replies.sort_values(by = "Developer Reply Date and Time")
developer_replies
```

4 Decision Making

4.1 Considerations for advertising in emerging markets

Prompt: For next year, company management has budget to start a marketing campaign in three countries. Based on the data, what are the emerging countries in which marketing could be worthwhile? State how you define 'emerging'. Try to do more than counting transactions. For example, look for a trend, or also incorporate customer satisfaction.

As can be seen from the **Cumulative Sales Revenue per Day choropleth**, the sales volume for countries where the majority native language is English appears to be the highest during the examined period. However, there is less of a clear skew towards English-speaking countries when the US is not considered. As can be ascertained from the aforementioned choropleth and **Query 17**, the download volume from the United States greatly surpasses the download volume from any other country in the dataset from the beginning of the period. The **Cumulative Downloads to Date (High Volume Countries) line plot**, moreover, shows that the download volume from the United States is high and continues to rise steadily over the entire period. As such, we conclude that it would not be advisable to pursue advertising in the United States, if the goal is to expand the reach of the product to a new global audience. However, if the aim of a future advertising campaign was to increase sales in general, we recommend targeting the ads to the United States, as consumers in this market already seem to resonate with the app.
Turning our attention exclusively to emerging markets now, the **Cumulative Downloads to Date (Emerging Markets) line plot**, delivers insights into the countries where download volume increased from June to December 2021.

```
# Query 17: Cumulative downloads to date for countries with the
 highest volume of downloads (8 top markets)
cumulative_transactions_per_country_p_sorted.head(n = 8)
```

4.1.1 Emerging Market 1: United Kingdom (UK)

In particular, the **United Kingdom (UK)**experienced steady growth over the period, with a total 220 units downloaded until the end of December 2021 – which can be seen by hovering over the UK line on the right-hand side of the **Cumulative Downloads to Date (Emerging Markets) line plot. Query 17** also shows that sales in the UK tended to grow at a more-or-less steady rate over time. Users in the UK also appear to be fairly satisfied with the app over the long-term, with the **Total Average Rating**, viewable by hovering over the UK on the Historical Average Rating per Country choropleth, coming in at **4.19**. More recent ratings from users in this country, however, show that satisfaction in the months with available data was lower than this total average. **Query 18** shows that this score is the result of **five total reviews** left by UK users in this period, **3 of which award the app a 5.0**, whereas the **2 remaining rate it**

a 1.0. The high polarisation in these scores raises the question of which aspects of the app users found unappealing, or whether the negative ratings were the result of the app's performance or reliability issues. To explore this question in the future before embarking on the advertising campaign, we recommend that the client tries to acquire data regarding the reviewer's country when collecting data of the 'reviews_2021mm.csv' format.

```
    # Query 18: Daily Ratings UK
    uk_daily_ratings = lowest_ratings_by_country_sorted.loc["GB"]
    uk_daily_ratings
```

4.1.2 Emerging Market 2: Canada (CA)

Returning our attention to the **Cumulative Downloads to Date (Emerging Markets) line plot**, we see that the second emerging market in terms of download volume is **Canada (CA)**. Taken together, the aforementioned line plot, and **Query 17** (from above) show that unlike the UK, Canada seems to have experienced a more sudden increase in the number of downloads starting in July 2021, and tapering off in the Autumn months. As discussed in the preceding section, Canadian users had an overall positive experience with the app, especially when considering the long-term **Total Average Rating**, viewable on the **Historical Average Rating by Country choropleth**. As discussed in *Client Satisfaction in a Geographical Context*, the recent ratings from Canada were overall skewing positive, with **two ratings of a 3.0** bringing down the average score. As such, judging by the relatively quick growth in downloads from clients in this country, as well as their overall positive perceptions of the product historically and recently, we believe Canada would be an ideal emerging market to target with advertising.

4.1.3 Emerging Market 3: Germany (DE)

Lastly – and once again informed by the **Cumulative Downloads to Date (Emerging Markets) line plot** in conjunction with the **Cumulative Sales Revenue per Day choropleth** – we suggest Germany as a good candidate for a targeted advertising campaign. **Query 17** once again demonstrates that the **total number of downloads** from customers based in Germany by 31st December 2021 reached **94**. Compared to the app's primary market (the **US**, with a total of **2616**) on the same day, this is a small number which could stand to grow from a targeted advertising campaign. Moreover, we see via the **Cumulative Downloads to Date (Emerging Markets) line plot**, that there are periods of time with big jumps in the numbers of downloads, as well as several plateaus, which could show that the final plateau observed from 23rd December 2021 to 31st December 2021 is likely not an indication that the trend is plateauing indefinitely. As discussed in the *Client Satisfaction in a Geographical Context* section of this report, users from **Germany** are also some of the most satisfied with the app, with a **4.157 Total Average Rating**, viewable by hovering over Germany on the **Historical Average Rating by Country**

choropleth, and a **5.0 Recent Average Rating**, viewable on the **Recent Average Rating by Country choropleth**. Due to this as well as the upward trend in the number of downloads from Germany, this would be a great target market for an advertising campaign, as recently users from Germany have left only positive reviews, and are perhaps likely to keep rating the app positively.