**CUSTOMER SEGMENTATION & RFM TASK**

**Your tasks are:**

* Use only one year of data, 2010-12-01 to 2011-12-01.
* Use SQL for calculation and data selection.
* Calculate recency, frequency and money value and convert those values into R, F and M scores by using Quartiles, 1 to 4 values. In BigQuery, a function APPROX\_QUANTILES is used to set the quartiles. You can check your results with rfm\_value table and rfm\_quantiles. Those tables contain intermediate calculations and are used in next steps(1 step: calculate RFM values, possible answer - rfm\_value table, 2 step: calculate RFM quantiles from RFM values, possible answer - rfm\_quantiles). **Important note: the answers in the tables are one of the possible answers. The results might vary due to data filtering options.**
* Calculate recency from date 2011-12-01.
* Calculate common RFM score. An example of a possible answer is given in the table rfm\_score.
* Segment customers into Best Customers, Loyal Customers, Big Spenders, Lost Customers and other categories.
* Present your analyses with a dashboard by using one of these tools: Tableau/Power Bi/Looker Studio.
* Present some insights which customer group/customer groups should the marketing team get focus on.

**CLV TASK**

**The task**

You got a follow up task from your manager as he read some article that calculating CLV using Shopify’s formula is too simplistic. He has heard somewhere that using cohorts produces more reliable and actionable results. Once again please use [turing\_data\_analytics.raw\_events](https://console.cloud.google.com/bigquery?authuser=3&project=tc-da-1&ws=!1m14!1m3!8m2!1s756497109418!2s1b8fa7cb3d73401f81fea1591fa615a7!1m4!4m3!1stc-da-1!2sturing_data_analytics!3sstackoverflow_posts!1m4!4m3!1stc-da-1!2sturing_data_analytics!3sraw_events&d=turing_data_analytics&p=tc-da-1&t=raw_events&page=table" \t "_blank) table to answer follow up questions from your manager.

TIP: imagine that current week is 2021-01-24 (the last weekly cohort you have in your dataset).

Additionally, he identified 2 problems with your previous analysis:

1. You included only customers who purchased something, while marketing is counting all user registrations that they manage to bring to your ecommerce site. Thus, you need to adjust your calculations to include all users who have been on your website, not only the one who purchased something.
2. Your average customer does not tend to stay with your ecommerce site for too long. He wants to see weekly picture using cohorts. He expects customers to stay no longer than 12 weeks with your ecommerce site.

As the first step you should write 1 or 2 queries to pull data of weekly revenue divided by registrations. Since in this particular site there is no concept of registration, we will simply use the first visit to our website as registration date (registration cohort). Do not forget to use user\_pseudo\_id to distinguish between users. Then divide revenue in consequent weeks by the number of weekly registration numbers.

Next you will produce the same chart, but the revenue / registrations for a particular week cohort will be expressed as a cumulative sum. For this you simply need to add previous week revenue to current week’s revenue. Down below you will calculate averages for all week numbers (weeks since registration). Down below that you will calculate percentage growth, which will be based on those average numbers

Next, we will focus on the future and try to predict the missing data. In this case missing data is the revenue we should expect from later acquired user cohorts. For example, for users whose first session happened on 2021-01-24 week we have only their first week revenue which is 0.19$ per user who started their first session in this week. Though we are not sure what will happen in the next 12 weeks.

For this we will simply use previously calculated Cumulative growth % (red marked area in chart above) and predict all 12 future weeks values (ex. for this cohort we can calculate expected revenue for week 1 as 0.19 USD x (1 + 23.29%) = 0.24 USD, for week 2 as 0.24 USD x (1 + 12.26%) = 0.27 USD). Using Avg. cumulative growth for each week we can calculate that based on 0.19$ initial value we can expect 0.35 USD as revenue on week 12. Provide a chart which calculates these numbers for all future weeks (up till week 12).

The chart above is our final result. You should calculate the average of cumulative revenue for 12th week for all users who have been on your website. This not only provide better estimate of CLV for all your users who have been on your website (including the ones who did not purchase anything) but also allows you to see trends for weekly cohorts. Have a look at the conditional formatting you added to all 3 charts and be prepared to answer follow up questions