**SQL Queries Individual Assignment**

**Query Set 1**

**Query 1**

The first query I found on Quarry is called “Articles without categories on plwiki”, and it is published by The Polish. Here is the link for the query:

https://quarry.wmflabs.org/query/2212 (Date accessed 10/16/2017)

https://quarry.wmflabs.org/The%20Polish. (Date accessed 10/16/2017)

This query represents a non-English Wiki since it starts with “plwiki\_p;” (It is Polish). The query aims to find the articles without categories on Polish Wikipedia. I looked at the person’s Wikipedia profile to understand more about what their intention for this query, and I saw that they have many maintenance related queries on their profile. For example, they had worked on queries such as “Articles without links on plwiki” which I will be using as my second query. Another example for the maintenance related queries include: “articles not connected to Wikidata”, “False articles redirects on plwiki”, and so on. I assume that they were formatting/sterilizing page titles as wherever page title is seen in this format “'# [[',page\_title,']]” change it as to this format “page\_title”. In addition, they also used the clause “LEFT JOIN” to combine columns from other tables.

**Query 2**

The second query (for the query set 1) is “Articles without links on plwiki” and it was published by the same person.

Here is the link for the query:

https://quarry.wmflabs.org/query/2215 (Date accessed 10/16/2017)

https://quarry.wmflabs.org/The%20Polish. (Date accessed 10/16/2017)

This query intends to find articles without links on plwiki on Polish Wikipedia. Since the query was done by the same person, the purpose is the same i.e., maintenance. Again, they were formatting page titles by using the CONCAT clause to format the “page\_title”. Also, they were using the LEFT JOIN clause to relate the columns from other tables.

**Query Set 2**

**Query 1**

My first query for the query set 2 is called “Most active users by edits on plwiki in last 6 months”.

Here is the link for the query:

https://quarry.wmflabs.org/query/1972 (Date accessed 10/16/2017)

https://quarry.wmflabs.org/The%20Polish. (Date accessed 10/16/2017)

They aimed to find the most active users by edits on Polish Wikipedia in 2016. They used the BETWEEN clause to limit the time frame to six months (July 1st, 2016 to December 31st, 2016). They also used INNER JOIN clause to form a new table by using two separate tables. The two tables they used are “revision” and “revision.rev\_user”. By using INNER JOIN, they formed a new table “user.user\_id”. They also used ORDER BY recent\_user\_editcount DESC to sort the results by descending order.

**Query 2**

The second query for the query set 2 is called “Most active users by edits on plwiki in last 12 months”.

Here is the link for the query:

https://quarry.wmflabs.org/query/1318 (Date accessed 10/16/2017)

https://quarry.wmflabs.org/The%20Polish. (Date accessed 10/16/2017)

The goal of this query is to figure out the most active users by edits on Polish Wikipedia in 2016. Since they were looking for a specific timespan, they used the BETWEEN clause to limit it to January 1st, 2016 to December 31st, 2016. They also used CONCAT to bring all types of user names columns ("[[user:",user\_name,"|",user\_name,"]]"), under the user\_name column. They also LIMIT the results by 500 rows. Likewise, the other query they sorted the results by descending order.

**Query Set 3**

**Query 1**

My last query for the query set 3 is “Retrieving Featured Pictures in the Commons”.

https://quarry.wmflabs.org/query/21947 (Date accessed 10/16/2017)

https://quarry.wmflabs.org/Miriam%20(WMF) (Date accessed 10/16/2017)

This is not a query set by a single user. It was a single query from a single user. However, I really enjoyed the query and wanted work and experience with it myself, and turned this into a query set (from different users). I believe the profile (Miriam WMF) belongs to a research scientist. With this query she aimed to retrieve the featured pictures on Wikipedia. The query was made by using SELECT, FROM, LEFT JOIN, ON, WHERE clauses. Also, she did not specify the type or kind of pictures she wanted to retrieve. If she did, she would have used the LIMIT clause to limit the results.

**Query 2**

The second query for the query set 3 is “Retrieving Featured Istanbul Pictures in the Commons”.

https://quarry.wmflabs.org/query/22291 (Date accessed 10/16/2017)

https://quarry.wmflabs.org/Produix (Zeynep Diker) (Date accessed 10/16/2017)

I wanted to practice more on what we did for the class exercise last week, and forked this query. My goal was to retrieve featured pictures about Istanbul on Wikipedia. I did this by adding the LIKE clause “AND page\_title LIKE '%Istanbul%' “ to the query. I was going to add the LIMIT clause as well; however, the results were not that many. I also tried UMD and Maryland instead of Istanbul but I had “zero” results for both UMD and Maryland.

Word Count: 776