A picture containing logo

Description automatically generated

William Zhu

Data Warehousing for Analytics

SQL Assignment 1

Use Google Big Query and the public dataset for this individual assignment.

1. Consider the data set bigquery-public-data.new\_york.citibike\_trips , and the column tripduration which gives the length of the trip in seconds. Use distinct ripduration and ORDER BY to find how long the longest ride was in seconds.

SELECT Distinct tripduration

FROM `bigquery-public-data.new\_york\_citibike.citibike\_trips`

ORDER BY  tripduration desc

Limit 5

| Row | tripduration |  |
| --- | --- | --- |
| 1 | 19510049 |  |
| 2 | 15962256 |  |
| 3 | 15020934 |  |
| 4 | 13931824 |  |
| 5 | 13586276 |  |

1. Use the DISTINCT and ORDER BY on the table bigquery-public-data.new\_york.citibike\_trips , to find the smallest (oldest) birth\_year.

SELECT DISTINCT birth\_year

FROM `bigquery-public-data.new\_york\_citibike.citibike\_trips`

order by  birth\_year

limit 5

|  |
| --- |
|  |
| Row | birth\_year |  |
| 1 | *null* |  |
| 2 | 1874 |  |
| 3 | 1884 |  |
| 4 | 1885 |  |
| 5 | 1886 |  |

1. What are all the possible values from the column usertype in the dataset bigquery-public-data.new\_york.citibike\_trips?

SELECT Distinct usertype

FROM `bigquery-public-data.new\_york\_citibike.citibike\_trips`

Subscriber and customer

1. What were the two dates of the most recent and oldest stoptimes in the dataset bigquery-public-data.new\_york.citibike\_trips ?

SELECT Distinct stoptime

FROM `bigquery-public-data.new\_york\_citibike.citibike\_trips`

order by stoptime

Limit 3

|  |
| --- |
|  |
| Row | stoptime |  |
| 1 | *null* |  |
| 2 | 2013-07-01T00:04:02 |  |
| 3 | 2013-07-01T00:06:21 |  |

SELECT Distinct stoptime

FROM `bigquery-public-data.new\_york\_citibike.citibike\_trips`

order by stoptime Desc

Limit 3

|  |
| --- |
|  |
| Row | stoptime |  |
| 1 | 2018-09-05T14:55:01.167000 |  |
| 2 | 2018-09-05T14:50:33.416000 |  |
|  |  |  |

Write queries to figure out the following questions. All the questions deal with the dataset bigquery-public-data.austin\_bikeshare.bikeshare\_trips .

How many unique rides involved the bike with a bikeid of 446.

SELECT count(bikeid)

FROM `still-totality-327202.12.bikeshare`

where bikeid=446

|  |
| --- |
|  |
| Row | f0\_ |  |
| 1 | 1737 |  |

1. For the bike with bikeid='446', what was the time of its longest ride in minutes?

SELECT bikeid, duration\_minutes

FROM `still-totality-327202.12.bikeshare`

where bikeid= 446

order by duration\_minutes desc

|  |
| --- |
|  |
| Row | bikeid | duration\_minutes |  |
| 1 | 446.0 | 1393 |  |

1. How many rides had a trip duration of 30 minutes or less?

SELECT tripduration

FROM `bigquery-public-data.new\_york\_citibike.citibike\_trips`

where tripduration > 30

53,108,721 rides

1. How many bike rides had a trip duration between 1 and 2 hours (including both 1 and 2 hour trips)?

SELECT tripduration,bikeid

FROM `bigquery-public-data.new\_york\_citibike.citibike\_trips`

where tripduration >=60 and tripduration <= 120

660791 rides where duration of the trip was between 1 and 2 hours

1. Consider the following two types of bike rides:

Started at "ACC - West & 12th Street" and ended at "Zilker Park West"

Started at "Nueces @ 3rd" and ended at "Toomey Rd @ South Lamar"

Of all these types of bike rides, what was the shortest trip duration in minutes?

SELECT start\_station\_name, end\_station\_name, duration\_minutes

FROM `still-totality-327202.12.bikeshare`

where start\_station\_name ="ACC - West & 12th Street"

and end\_station\_name ="Zilker Park West"

order by duration\_minutes

SELECT start\_station\_name, end\_station\_name, duration\_minutes

FROM `still-totality-327202.12.bikeshare`

where start\_station\_name ="Nueces @ 3rd"

and end\_station\_name ="Toomey Rd @ South Lamar"

order by duration\_minutes

Of these two bike rides, the shortest trip duration was Nueces @ 3rd to Toomey Rd @ South Lamar with the duration at 2 minutes.

**Additional Notes:**

 This is an individual assignment.

 Assemble all of the responses to the questions in one MS Word.

 For the SQL Queries, format your query in Courier New font (at least 12 point size). Paste in the results of the query into MS Word.