**Report of SQL Queries used**

**Google Analytics Case Study**

**Union ALL tables**

**SELECT\***

**FROM `data-analytics-case-study-2022.cyclistic.bike\_tripdata\_202201`**

**UNION ALL**

**SELECT\***

**FROM `data-analytics-case-study-2022.cyclistic.bike\_tripdata\_202202`**

**UNION ALL**

**SELECT\***

**FROM `data-analytics-case-study-2022.cyclistic.bike\_tripdata\_202203`**

**UNION ALL**

**SELECT\***

**FROM `data-analytics-case-study-2022.cyclistic.bike\_tripdata\_202204`**

**UNION ALL**

**SELECT\***

**FROM `data-analytics-case-study-2022.cyclistic.bike\_tripdata\_202205`**

**UNION ALL**

**SELECT\***

**FROM `data-analytics-case-study-2022.cyclistic.bike\_tripdata\_202206`**

**UNION ALL**

**SELECT\***

**FROM `data-analytics-case-study-2022.cyclistic.bike\_tripdata\_202207`**

**UNION ALL**

**SELECT\***

**FROM `data-analytics-case-study-2022.cyclistic.bike\_tripdata\_202208`**

**UNION ALL**

**SELECT\***

**FROM `data-analytics-case-study-2022.cyclistic.bike\_tripdata\_202209`**

**UNION ALL**

**SELECT\***

**FROM `data-analytics-case-study-2022.cyclistic.bike\_tripdata\_202210`**

**UNION ALL**

**SELECT\***

**FROM `data-analytics-case-study-2022.cyclistic.bike\_tripdata\_202211`**

**UNION ALL**

**SELECT\***

**FROM `data-analytics-case-study-2022.cyclistic.bike\_tripdata\_202212`**

* **Cleaning Dataset**

**UPDATE `data-analytics-case-study-2022.cyclistic.bike-tripdata-2022-raw`**

**SET start\_station\_name = "Base - 2132 W Hubbard"**

**WHERE start\_station\_name = "Base - 2132 W Hubbard Warehouse"**

**UPDATE `data-analytics-case-study-2022.cyclistic.bike-tripdata-2022-raw`**

**SET start\_station\_name = "Broadway & Wilson"**

**WHERE start\_station\_name = "Broadway & Wilson - Truman College Vaccination Site"**

**UPDATE `data-analytics-case-study-2022.cyclistic.bike-tripdata-2022-raw`**

**SET start\_station\_name = "Campbell Ave & Montrose Ave"**

**WHERE start\_station\_name = "Campbell Ave & Montrose Ave (Temp)"**

**UPDATE `data-analytics-case-study-2022.cyclistic.bike-tripdata-2022-raw`**

**SET start\_station\_id = "chargingstx6"**

**WHERE start\_station\_id = "chargingstx06"**

**UPDATE `data-analytics-case-study-2022.cyclistic.bike-tripdata-2022-raw`**

**SET start\_station\_id = "chargingstx7"**

**WHERE start\_station\_id = "chargingstx07"**

**UPDATE `data-analytics-case-study-2022.cyclistic.bike-tripdata-2022-raw`**

**SET end\_station\_id = "chargingstx7"**

**WHERE end\_station\_id = "chargingstx07"**

**UPDATE `data-analytics-case-study-2022.cyclistic.bike-tripdata-2022-raw`**

**SET end\_station\_id = "chargingstx6"**

**WHERE end\_station\_id = "chargingstx06"**

**UPDATE `data-analytics-case-study-2022.cyclistic.bike-tripdata-2022-raw`**

**SET start\_station\_name = "Lincoln Ave & Belmont Ave"**

**WHERE start\_station\_name = "Lincoln Ave & Belmont Ave (Temp)"**

**UPDATE `data-analytics-case-study-2022.cyclistic.bike-tripdata-2022-raw`**

**SET start\_station\_name = "WestChi",**

**start\_station\_id = "DIVVY 001"**

**WHERE start\_station\_name = "WEST CHI-WATSON"**

**and start\_station\_id = "DIVVY 001 - Warehouse test station"**

**DELETE**

**FROM `data-analytics-case-study-2022.cyclistic.bike-tripdata-2022-raw`**

**WHERE start\_station\_name IS NULL**

**AND start\_station\_id IS NULL**

**AND end\_station\_id IS NULL**

**AND end\_station\_name IS NULL**

Gathering Information

**Total number of start stations**

**SELECT start\_station\_name**

**FROM `data-analytics-case-study-2022.cyclistic.bike-tripdata-2022-raw`**

**WHERE start\_station\_name <> "null"**

**GROUP BY start\_station\_name**

**Total number of end stations**

**SELECT DISTINCT (end\_station\_name)**

**FROM `data-analytics-case-study-2022.cyclistic.bike-tripdata-2022-raw`**

**WHERE end\_station\_name <> "null"**

**ORDER BY end\_station\_name**

**– Total start charging stations (0-7)**

**SELECT DISTINCT (start\_station\_name)**

**FROM `data-analytics-case-study-2022.cyclistic.bike-tripdata-2022-raw`**

**WHERE start\_station\_id LIKE "chargingstx%"**

**– Total end charging stations (0-7)**

**SELECT DISTINCT**

**end\_station\_id**

**FROM `data-analytics-case-study-2022.cyclistic.bike-tripdata-2022-raw`**

**WHERE end\_station\_id LIKE "chargingstx%"**

**ORDER BY end\_station\_id**

* **Count Distinct Total Rides**

**SELECT COUNT(DISTINCT ride\_id) AS total\_rides**

**FROM `data-analytics-case-study-2022.cyclistic.bike-tripdata-2022-raw`**

* **Annual rides per rider type**

**SELECT member\_type, COUNT (1) AS rides\_per\_member**

**FROM `data-analytics-case-study-2022.cyclistic.bike-tripdata-2022-clean3`**

**GROUP BY member\_type;**

* **Rideable type**

**SELECT DISTINCT (rideable\_type)**

**FROM `data-analytics-case-study-2022.cyclistic.bike-tripdata-2022-raw` LIMIT 1000**

* **Ride\_length\_min and day\_of\_week**

**SELECT**

**ride\_id,**

**rideable\_type,**

**started\_at,**

**ended\_at,**

**start\_station\_name,**

**start\_station\_id,**

**end\_station\_name,**

**end\_station\_id,**

**start\_lat,**

**start\_lng,**

**end\_lat,**

**end\_lng,**

**member\_casual AS membership\_type,**

**EXTRACT(DAYofweek FROM started\_at) AS day\_of\_week,**

**TIMESTAMP\_DIFF(ended\_at, started\_at, minute) AS ride\_length\_mins**

**FROM `data-analytics-case-study-2022.cyclistic.bike-tripdata-2022-raw-copy2`**

**Adding month\_of\_ride column**

**SELECT**

**ride\_id,**

**member\_type,**

**rideable\_type,**

**day\_of\_week,**

**EXTRACT (MONTH FROM started\_at) AS month\_of\_ride,**

**ride\_length\_mins,**

**started\_at,**

**ended\_at,**

**start\_station\_name,**

**end\_station\_name,**

**start\_station\_id,**

**end\_station\_id,**

**start\_lat,**

**start\_lng,**

**end\_lat,**

**end\_lng,**

**FROM `data-analytics-case-study-2022.cyclistic.bike-tripdata-2022-clean2`**

**Total rides per month (no member distinction)**

**SELECT month\_of\_ride, COUNT (\*) AS monthly\_rides**

**FROM `data-analytics-case-study-2022.cyclistic.bike-tripdata-2022-clean3`**

**GROUP BY month\_of\_ride**

**ORDER by month\_of\_ride**

**Total rides per month (with member distinction)**

**SELECT month\_of\_ride, COUNT (\*) AS monthly\_rides,**

**COUNT(CASE WHEN member\_type = 'member' THEN 1 END) AS monthly\_member\_rides,**

**COUNT(CASE WHEN member\_type = 'casual' THEN 1 END) AS monthly\_casual\_rides**

**FROM `data-analytics-case-study-2022.cyclistic.bike-tripdata-2022-clean3`**

**GROUP BY month\_of\_ride**

**ORDER by month\_of\_ride**

**Bike type preferences per user**

**SELECT member\_type, COUNT (\*) AS rider\_type\_total,**

**COUNT(CASE WHEN rideable\_type = 'electric\_bike' THEN 1 END) AS electric,**

**COUNT(CASE WHEN rideable\_type = 'classic\_bike' THEN 1 END) AS classic,**

**COUNT(CASE WHEN rideable\_type = 'docked\_bike' THEN 1 END) AS docked,**

**FROM `data-analytics-case-study-2022.cyclistic.bike-tripdata-2022-clean3`**

**GROUP BY member\_type**

**ORDER by member\_type**

**Rides per weekday per user type**

**SELECT member\_type,**

**COUNT(CASE WHEN day\_of\_week = 1 THEN 1 END) AS Sunday,**

**COUNT(CASE WHEN day\_of\_week = 2 THEN 1 END) AS Monday,**

**COUNT(CASE WHEN day\_of\_week = 3 THEN 1 END) AS Tuesday,**

**COUNT(CASE WHEN day\_of\_week = 4 THEN 1 END) AS Wednesday,**

**COUNT(CASE WHEN day\_of\_week = 5 THEN 1 END) AS Thursday,**

**COUNT(CASE WHEN day\_of\_week = 6 THEN 1 END) AS Friday,**

**COUNT(CASE WHEN day\_of\_week = 7 THEN 1 END) AS Saturday**

**FROM `data-analytics-case-study-2022.cyclistic.bike-tripdata-2022-clean3`**

**GROUP BY member\_type**

**ORDER by member\_type**

**Rider types per start\_stations**

**SELECT DISTINCT start\_station\_name,**

**COUNT (CASE WHEN member\_type = "casual" THEN 1 END) AS casual\_users,**

**COUNT (CASE WHEN member\_type = "member" THEN 1 END) AS member\_users**

**FROM `data-analytics-case-study-2022.cyclistic.bike-tripdata-2022-clean3`**

**GROUP BY start\_station\_name**

**Top 30 stations - casual riders**

**SELECT DISTINCT start\_station\_name,**

**COUNT (CASE WHEN member\_type = "casual" THEN 1 END) AS casual\_users,**

**FROM `data-analytics-case-study-2022.cyclistic.bike-tripdata-2022-clean3`**

**GROUP BY start\_station\_name**

**ORDER BY 2 DESC**

**LIMIT 30**

**Top 30 stations - members**

**SELECT DISTINCT start\_station\_name,**

**COUNT (CASE WHEN member\_type = "member" THEN 1 END) AS member\_users,**

**FROM `data-analytics-case-study-2022.cyclistic.bike-tripdata-2022-clean3`**

**GROUP BY start\_station\_name**

**ORDER BY 2 DESC**

**LIMIT 30**

**Total rides under 1 hr & riders 1hr or longer**

**SELECT member\_type,**

**COUNT (CASE WHEN ride\_length\_mins >= 60 THEN 1 END) AS rides\_1hr\_or\_longer,**

**COUNT (CASE WHEN ride\_length\_mins < 60 THEN 1 END) AS rides\_under\_1hr,**

**FROM `data-analytics-case-study-2022.cyclistic.bike-tripdata-2022-clean3`**

**GROUP BY member\_type**

**Total rides under 1 hr & riders 1hr or longer/AVG ride length per type**

**SELECT member\_type,**

**COUNT (CASE WHEN ride\_length\_mins >= 60 THEN 1 END) AS rides\_1hr\_or\_longer,**

**COUNT (CASE WHEN ride\_length\_mins < 60 THEN 1 END) AS rides\_under\_1hr,**

**MAX (ride\_length\_mins) AS Max\_ride\_length,**

**AVG (ride\_length\_mins) AS Average\_ride\_length,**

**FROM `data-analytics-case-study-2022.cyclistic.bike-tripdata-2022-clean3`**

**WHERE member\_type = "casual"**

**OR member\_type = "member"**

**GROUP BY member\_type**

**Average length rider per day\_of\_week**

**SELECT**

**member\_type,**

**AVG(CASE when day\_of\_week = 2 and member\_type = 'casual' OR member\_type = 'member' AND day\_of\_week = 2 then ride\_length\_mins end) AS Monday\_Avg\_mins,**

**AVG(CASE when day\_of\_week = 3 and member\_type = 'casual' OR member\_type = 'member'AND day\_of\_week = 3 then ride\_length\_mins end) AS Tuesday\_Avg\_mins,**

**AVG(CASE when day\_of\_week = 4 and member\_type = 'casual' OR member\_type = 'member'AND day\_of\_week = 4 then ride\_length\_mins end) AS Wednesday\_Avg\_mins,**

**AVG(CASE when day\_of\_week = 5 and member\_type = 'casual' OR member\_type = 'member'AND day\_of\_week = 5 then ride\_length\_mins end) AS Thursday\_Avg\_mins,**

**AVG(CASE when day\_of\_week = 6 and member\_type = 'casual' OR member\_type = 'member' AND day\_of\_week = 6 then ride\_length\_mins end) AS Friday\_Avg\_mins,**

**AVG(CASE when day\_of\_week = 7 and member\_type = 'casual' OR member\_type = 'member'AND day\_of\_week = 7 then ride\_length\_mins end) AS Saturday\_Avg\_mins,**

**AVG(CASE when day\_of\_week = 1 and member\_type = 'casual' OR member\_type = "member" AND day\_of\_week = 1 then ride\_length\_mins end) AS Sunday\_Avg\_mins,**

**FROM `data-analytics-case-study-2022.cyclistic.bike-tripdata-2022-clean3`**

**GROUP BY member\_type**

**Top 30 Docked bikes annual use and weekday use**

**SELECT DISTINCT start\_station\_name,**

**COUNT (CASE WHEN rideable\_type = "docked\_bike" AND member\_type = "casual" THEN 1 END) AS Total\_docked\_bike\_rides,**

**COUNT (CASE WHEN rideable\_type= "docked\_bike" AND day\_of\_week = 2 THEN 1 END) AS Monday\_Docked\_use,**

**COUNT (CASE WHEN rideable\_type= "docked\_bike" AND day\_of\_week = 3 THEN 1 END) AS Tuesday\_Docked\_use,**

**COUNT (CASE WHEN rideable\_type= "docked\_bike" AND day\_of\_week = 4 THEN 1 END) AS Wednesday\_Docked\_use,**

**COUNT (CASE WHEN rideable\_type= "docked\_bike" AND day\_of\_week = 5 THEN 1 END) AS Thursday\_Docked\_use,**

**COUNT (CASE WHEN rideable\_type= "docked\_bike" AND day\_of\_week = 6 THEN 1 END) AS Friday\_Docked\_use,**

**COUNT (CASE WHEN rideable\_type= "docked\_bike" AND day\_of\_week = 7 THEN 1 END) AS Saturday\_Docked\_use,**

**COUNT (CASE WHEN rideable\_type= "docked\_bike" AND day\_of\_week = 1 THEN 1 END) AS Sunday\_Docked\_bike\_use,**

**FROM `data-analytics-case-study-2022.cyclistic.bike-tripdata-2022-clean3`**

**GROUP BY start\_station\_name**

**ORDER BY 2 DESC**

**LIMIT 30**