How many 1-way trips are there and from where to where?

```
#standardSQL
select start_station_name, end_station_name, count(*) as num_trips
from `bigquery-public-data.san_francisco_bikeshare.bikeshare_trips`
group by start_station_name, end_station_name
order by num_trips desc
```

#How many such trips have a greater than 30 min duration?

```
#standardSQL
select start_station_name, end_station_name, round(avg(duration_sec)/60, 2) as
trip_duration_min, count(*) as num_trips
from `bigquery-public-data.san_francisco_bikeshare.bikeshare_trips`
where duration_sec/60 >= 30
group by start_station_name, end_station_name
order by num_trips desc
```

#How many of such trips take place during rush hour?

```
#Assumptions: Rush hour defined as 7:00 a.m.-9:00 a.m., and 4:30 p.m.-6:30 p.m
select start_station_name, end_station_name, count(*) as num_commuter_trips
from `bigquery-public-data.san_francisco_bikeshare.bikeshare_trips`
where (
  (time(start_date) > '07:00:00' and time(start_date) < '09:00:00') or
  (time(start_date) > '16:00:00' and time(start_date) < '18:30:00')) and
  duration_sec/60 >= 30
group by start_station_name, end_station_name
order by num_commuter_trips desc
```

#How many such trips take place during the week (Mon-Fri)?

```
select start_station_name, end_station_name, count(*) as num_commuter_trips
from `bigquery-public-data.san_francisco_bikeshare.bikeshare_trips`
where (
  (time(start_date) > '07:00:00' and time(start_date) < '09:00:00') or
  (time(start_date) > '16:00:00' and time(start_date) < '18:30:00'))
  and duration_sec/60 >= 30
  and extract(DAY FROM start_date) in (2,6)
group by start_station_name, end_station_name
order by num_commuter_trips desc
```

#1-How many trips take place on weekdays between all the stations?

```
#SELECT
#start_station_name,
#end_station_name,
#count(*) as total_trips,
```

```
#SUM(IF(extract(DAY FROM start_date) in (2,6) and (time(start_date) >= "05:00:00" and
time(start_date) <= "23:00:00"), 1, 0)) as weekday_trips
#FROM `bigquery-public-data.san_francisco.bikeshare_trips`
#GROUP BY start_station_name, end_station_name
#ORDER BY weekday_trips desc
```

#2-What is the distribution of week day trips across the day? Determine AM & PM "commuter" windows

```
#SELECT
#count(*) as total_trips,
#SUM(IF(extract(DAY FROM start_date) in (2,6) and (time(start_date) >= "05:00:00" and
time(start_date) <= "23:00:00"), 1, 0)) as weekday_trips,
#SUM(IF(extract(DAY FROM start_date) in (2,6) and (time(start_date) >= "05:00:00" and
time(start_date) <= "06:00:00"), 1, 0)) as trips56,
#SUM(IF(extract(DAY FROM start_date) in (2,6) and (time(start_date) >= "06:00:00" and
time(start_date) <= "07:00:00"), 1, 0)) as trips67,
#SUM(IF(extract(DAY FROM start_date) in (2,6) and (time(start_date) >= "07:00:00" and
time(start_date) <= "08:00:00"), 1, 0)) as trips78,
#SUM(IF(extract(DAY FROM start_date) in (2,6) and (time(start_date) >= "08:00:00" and
time(start_date) <= "09:00:00"), 1, 0)) as trips89,
#SUM(IF(extract(DAY FROM start_date) in (2,6) and (time(start_date) >= "09:00:00" and
time(start_date) <= "10:00:00"), 1, 0)) as trips910,
#SUM(IF(extract(DAY FROM start_date) in (2,6) and (time(start_date) >= "10:00:00" and
time(start_date) <= "11:00:00"), 1, 0)) as trips1011,
#SUM(IF(extract(DAY FROM start_date) in (2,6) and (time(start_date) >= "11:00:00" and
time(start_date) <= "12:00:00"), 1, 0)) as trips1112,
#SUM(IF(extract(DAY FROM start_date) in (2,6) and (time(start_date) >= "12:00:00" and
time(start_date) <= "13:00:00"), 1, 0)) as trips1213,
#SUM(IF(extract(DAY FROM start_date) in (2,6) and (time(start_date) >= "13:00:00" and
time(start_date) <= "14:00:00"), 1, 0)) as trips1314,
#SUM(IF(extract(DAY FROM start_date) in (2,6) and (time(start_date) >= "14:00:00" and
time(start_date) <= "15:00:00"), 1, 0)) as trips1415,
#SUM(IF(extract(DAY FROM start_date) in (2,6) and (time(start_date) >= "15:00:00" and
time(start_date) <= "16:00:00"), 1, 0)) as trips1516,
#SUM(IF(extract(DAY FROM start_date) in (2,6) and (time(start_date) >= "16:00:00" and
time(start_date) <= "17:00:00"), 1, 0)) as trips1617,
#SUM(IF(extract(DAY FROM start_date) in (2,6) and (time(start_date) >= "17:00:00" and
time(start_date) <= "18:00:00"), 1, 0)) as trips1718,
#SUM(IF(extract(DAY FROM start_date) in (2,6) and (time(start_date) >= "18:00:00" and
time(start_date) <= "19:00:00"), 1, 0)) as trips1819,
#SUM(IF(extract(DAY FROM start_date) in (2,6) and (time(start_date) >= "19:00:00" and
time(start_date) <= "20:00:00"), 1, 0)) as trips1920,
#SUM(IF(extract(DAY FROM start_date) in (2,6) and (time(start_date) >= "20:00:00" and
time(start_date) <= "21:00:00"), 1, 0)) as trips2021,
```

```
#SUM(IF(extract(DAY FROM start_date) in (2,6) and (time(start_date) >= "21:00:00" and
time(start_date) <= "22:00:00"), 1, 0)) as trips2122,
#SUM(IF(extract(DAY FROM start_date) in (2,6) and (time(start_date) >= "22:00:00" and
time(start_date) <= "23:00:00"), 1, 0)) as trips2223
#FROM `bigquery-public-data.san_francisco.bikeshare_trips`
```

#3-What are the 5 most popular commuter routes (AM & PM combined)?

```
#SELECT
#start_station_name,
#end_station_name,
#count(*) as total_trips,
#SUM(IF(extract(DAY FROM start_date) in (2,6) and (time(start_date) >= "05:00:00" and
time(start_date) <= "21:00:00"), 1, 0)) as weekday_trips,
#SUM(IF(extract(DAY FROM start_date) in (2,6) and ((time(start_date) >= "07:00:00" and
time(start_date) <= "10:00:00") or (time(start_date) >= "16:00:00" and time(start_date) <=
"19:00:00")),1,0)) as tot_commuter_trips,
#SUM(IF(extract(DAY FROM start_date) in (2,6) and (time(start_date) >= "07:00:00" and
time(start_date) <= "10:00:00"), 1, 0)) as am_commuter_trips,
#SUM(IF(extract(DAY FROM start_date) in (2,6) and (time(start_date) >= "16:00:00" and
time(start_date) <= "19:00:00"), 1, 0)) as pm_commuter_trips
#FROM `bigquery-public-data.san_francisco.bikeshare_trips`
#GROUP BY start_station_name, end_station_name
#ORDER BY tot_commuter_trips desc
```

#4-What are the 5 most popular AM commuter routes?

```
#SELECT
#start_station_name,
#end_station_name,
#count(*) as total_trips,
#SUM(IF(extract(DAY FROM start_date) in (2,6) and (time(start_date) >= "05:00:00" and
time(start_date) <= "21:00:00"), 1, 0)) as weekday_trips,
#SUM(IF(extract(DAY FROM start_date) in (2,6) and ((time(start_date) >= "07:00:00" and
time(start_date) <= "10:00:00") or (time(start_date) >= "16:00:00" and time(start_date) <=
"19:00:00")),1,0)) as tot_commuter_trips,
#SUM(IF(extract(DAY FROM start_date) in (2,6) and (time(start_date) >= "07:00:00" and
time(start_date) <= "10:00:00"), 1, 0)) as am_commuter_trips
#FROM `bigquery-public-data.san_francisco.bikeshare_trips`
#GROUP BY start_station_name, end_station_name
#ORDER BY am_commuter_trips desc
```

#5-What are the 5 most popular PM commuter routes?

```
#SELECT
#start_station_name,
#end_station_name,
#count(*) as total_trips,
```

```
#SUM(IF(extract(DAY FROM start_date) in (2,6) and (time(start_date) >= "05:00:00" and
time(start_date) <= "21:00:00"), 1, 0)) as weekday_trips,
#SUM(IF(extract(DAY FROM start_date) in (2,6) and ((time(start_date) >= "07:00:00" and
time(start_date) <= "10:00:00") or (time(start_date) >= "16:00:00" and time(start_date) <=
"19:00:00")),1,0)) as tot_commuter_trips,
#SUM(IF(extract(DAY FROM start_date) in (2,6) and (time(start_date) >= "16:00:00" and
time(start_date) <= "19:00:00"), 1, 0)) as pm_commuter_trips
#FROM `bigquery-public-data.san_francisco.bikeshare_trips`
#GROUP BY start_station_name, end_station_name
#ORDER BY pm_commuter_trips desc
```

#6-What is the distribution of customers and subscribers on the top 5 commuter routes based on total commuter trips?

```
#SELECT
#start_station_name,
#end_station_name,
#count(*) as total_trips,
#SUM(IF(extract(DAY FROM start_date) in (2,6) and (time(start_date) >= "05:00:00" and
time(start_date) <= "21:00:00"), 1, 0)) as weekday_trips,
#SUM(IF(extract(DAY FROM start_date) in (2,6) and ((time(start_date) >= "07:00:00" and
time(start_date) <= "10:00:00") or (time(start_date) >= "16:00:00" and time(start_date) <=
"19:00:00")),1,0)) as tot_commuter_trips,
#SUM(IF
#(subscriber_type = "Customer" and
#((time(start_date) >= "07:00:00" and time(start_date) <= "10:00:00") or (time(start_date) >=
"16:00:00" and time(start_date) <= "19:00:00")) and
#extract(DAY FROM start_date) in (2,6),1,0)) as customers,
#SUM(IF
#(subscriber_type = "Subscriber" and
#((time(start_date) >= "07:00:00" and time(start_date) <= "10:00:00") or (time(start_date) >=
"16:00:00" and time(start_date) <= "19:00:00")) and
#extract(DAY FROM start_date) in (2,6), 1, 0)) as subscribers
#FROM `bigquery-public-data.san_francisco.bikeshare_trips`
#GROUP BY start_station_name, end_station_name
#ORDER BY tot_commuter_trips desc, customers desc
```

#7-What is the distribution of customers and subscribers on the top 5 commuter routes based on AM commuter trips?

```
#SELECT
#start_station_name,
#end_station_name,
#count(*) as total_trips,
#SUM(IF(extract(DAY FROM start_date) in (2,6) and (time(start_date) >= "05:00:00" and
time(start_date) <= "21:00:00"), 1, 0)) as weekday_trips,
```

```
#SUM(IF(extract(DAY FROM start_date) in (2,6) and ((time(start_date) >= "07:00:00" and
time(start_date) <= "10:00:00") or (time(start_date) >= "16:00:00" and time(start_date) <=
"19:00:00")),1,0)) as tot_commuter_trips,
#SUM(IF(extract(DAY FROM start_date) in (2,6) and (time(start_date) >= "07:00:00" and
time(start_date) <= "10:00:00"), 1, 0)) as am_commuter_trips,
#SUM(IF
#(subscriber_type = "Customer" and
#(time(start_date) >= "07:00:00" and time(start_date) <= "10:00:00")  and
#extract(DAY FROM start_date) in (2,6),1,0)) as customers,
#SUM(IF
#(subscriber_type = "Subscriber" and
#(time(start_date) >= "07:00:00" and time(start_date) <= "10:00:00") and
#extract(DAY FROM start_date) in (2,6), 1, 0)) as subscribers
#FROM `bigquery-public-data.san_francisco.bikeshare_trips`
#GROUP BY start_station_name, end_station_name
#ORDER BY am_commuter_trips desc, customers desc
```

#8-What is the distribution of customers and subscribers on the top 5 commuter routes based on PM commuter trips?

```
SELECT
start_station_name,
end_station_name,
count(*) as total_trips,
SUM(IF(extract(DAY FROM start_date) in (2,6) and (time(start_date) >= "05:00:00" and
time(start_date) <= "21:00:00"), 1, 0)) as weekday_trips,
SUM(IF(extract(DAY FROM start_date) in (2,6) and ((time(start_date) >= "07:00:00" and
time(start_date) <= "10:00:00") or (time(start_date) >= "16:00:00" and time(start_date) <=
"19:00:00")),1,0)) as tot_commuter_trips,
SUM(IF(extract(DAY FROM start_date) in (2,6) and (time(start_date) >= "16:00:00" and
time(start_date) <= "19:00:00"), 1, 0)) as pm_commuter_trips,
SUM(IF
(subscriber_type = "Customer" and
(time(start_date) >= "16:00:00" and time(start_date) <= "19:00:00")  and
extract(DAY FROM start_date) in (2,6),1,0)) as customers,
SUM(IF
(subscriber_type = "Subscriber" and
(time(start_date) >= "16:00:00" and time(start_date) <= "19:00:00") and
extract(DAY FROM start_date) in (2,6), 1, 0)) as subscribers
FROM `bigquery-public-data.san_francisco.bikeshare_trips`
GROUP BY start_station_name, end_station_name
ORDER BY pm_commuter_trips desc, customers desc
```

#Q3: Which stations have greater than 10 unused bikes in the docking station on average? Which stations top the list?

```
SELECT *, (q1.total_bikes-q1.avg_bikes_avail) as unused
FROM (SELECT station_id, max(total_bikes) as total_bikes, round(avg(bikes_available)) as
avg_bikes_avail
    FROM `loyal-flames-240223.bike_trip_data.total_bikes`
    group by station_id
    order by avg_bikes_avail desc) as q1

WHERE q1.total_bikes-q1.avg_bikes_avail > 10
ORDER BY unused desc

#Answer: Stations 91, 25 and 26 have the most unused bikes on average.
```

#Question 1: Which are the station pairs with the highest number of trips?

```
#Answer: Harry Bridges Plaza (Ferry Building) to Embarcadero at Sansome (9150 trips)
#       San Francisco Caltrain 2 (330 Townsend) to Townsend at 7th (8508)

#SQL Query:

#standardSQL
select start_station_name, end_station_name, round(avg(duration_sec)/60, 2) as
trip_duration_min, count(*) as num_trips
from `bigquery-public-data.san_francisco_bikeshare.bikeshare_trips`
group by start_station_name, end_station_name
order by num_trips desc
```

#Question 2: Which are the station(s) with the highest number of trip starts?

```
#Answer: San Francisco Caltrain (Townsend at 4th) followed by San Francisco Caltrain 2 (330
Townsend) and Harry Bridges Plaza (Ferry Building)

#SQL Query:
#standardSQL
select start_station_name, count(*) as num_starts
    from `bigquery-public-data.san_francisco_bikeshare.bikeshare_trips`
    group by start_station_name
    order by num_starts desc
```

#Question 3: Where do the trips that begin at the stations with the highest number of starts end? (Refer to Question 2 to obtain top 3 stations. What is their distribution in terms of number of trips?

#Answer: San Francisco Caltrain (Townsend at 4th) followed by San Francisco Caltrain 2 (330 Townsend) and Harry Bridges Plaza (Ferry Building) were the top 3 stations from Question 2. The output of question 3 shows the distribution of trips originating at these points to the various end stations.

#SQL Query:

```
#standardSQL
select start_station_name, end_station_name, count(*) as num_trips
from `bigquery-public-data.san_francisco_bikeshare.bikeshare_trips`
where trim(start_station_name) in ('San Francisco Caltrain (Townsend at 4th)', 'San Francisco Caltrain 2 (330 Townsend)', 'Harry Bridges Plaza (Ferry Building)')
group by start_station_name, end_station_name
order by start_station_name, num_trips desc
```

How many 1-way trips are there and from where to where?

```
#standardSQL
select start_station_name, end_station_name, count(*) as num_trips
from `bigquery-public-data.san_francisco_bikeshare.bikeshare_trips`
group by start_station_name, end_station_name
order by num_trips desc
```

#How many such trips have a greater than 30 min duration?

```
#standardSQL
select start_station_name, end_station_name, round(avg(duration_sec)/60, 2) as trip_duration_min, count(*) as num_trips
from `bigquery-public-data.san_francisco_bikeshare.bikeshare_trips`
where duration_sec/60 >= 30
group by start_station_name, end_station_name
order by num_trips desc
```

#How many of such trips take place during rush hour?

```
#Assumptions: Rush hour defined as 7:00 a.m.-9:00 a.m., and 4:30 p.m.-6:30 p.m
select start_station_name, end_station_name, count(*) as num_commuter_trips
from `bigquery-public-data.san_francisco_bikeshare.bikeshare_trips`
where (
  (time(start_date) > '07:00:00' and time(start_date) < '09:00:00') or
  (time(start_date) > '16:00:00' and time(start_date) < '18:30:00')) and
  duration_sec/60 >= 30
group by start_station_name, end_station_name
order by num_commuter_trips desc
```

#How many such trips take place during the week (Mon-Fri)?

```
select start_station_name, end_station_name, count(*) as num_commuter_trips
from `bigquery-public-data.san_francisco_bikeshare.bikeshare_trips`
where (
  (time(start_date) > '07:00:00' and time(start_date) < '09:00:00') or
  (time(start_date) > '16:00:00' and time(start_date) < '18:30:00'))
  and duration_sec/60 >= 30
  and extract(DAY FROM start_date) in (2,6)
group by start_station_name, end_station_name
order by num_commuter_trips desc
```

#0-1- What is the distribution of trip durations?

#0-2- What is the distribution of trips over the year?

#3 - What is the distribution of total "commuter" trips between the various bike stations?

```
SELECT
 start_station_name,
 end_station_name,
 count(*) as total_trips,
 SUM(IF(extract(DAY FROM start_date) not in (2,6), 1, 0)) as weekend_trips,
 SUM(IF(subscriber_type = "Customer" and (time(start_date) > "05:00:00" and time(start_date)
< "21:00:00") and extract(DAY FROM start_date) not in (2,6),1,0)) as weekend_Customers,
  SUM(IF(subscriber_type = "Subscriber" and (time(start_date) > "05:00:00" and
time(start_date) < "21:00:00") and extract(DAY FROM start_date) not in (2,6), 1, 0)) as
weekend_Subscribers,
  SUM(IF(extract(DAY FROM start_date) in (2,6), 1, 0)) as weekday_trips,
  SUM(IF(subscriber_type = "Customer" and (time(start_date) > "05:00:00" and time(start_date)
< "21:00:00") and extract(DAY FROM start_date) in (2,6),1,0)) as weekday_Customers,
  SUM(IF(subscriber_type = "Subscriber" and (time(start_date) > "05:00:00" and
time(start_date) < "21:00:00") and extract(DAY FROM start_date) in (2,6), 1, 0)) as
weekday_Subscribers

FROM `bigquery-public-data.san_francisco.bikeshare_trips`
GROUP BY start_station_name, end_station_name
ORDER BY weekday_trips DESC
```

#4 - What are the 5 MOST POPULAR commuter trips based on total number of weekday trips?

```
SELECT
 start_station_name,
 end_station_name,
```

```
 count(*) as total_trips,
  SUM(IF(extract(DAY FROM start_date) in (2,6), 1, 0)) as weekday_trips,
  SUM(IF(subscriber_type = "Customer" and (time(start_date) > "05:00:00" and time(start_date)
< "21:00:00") and extract(DAY FROM start_date) in (2,6),1,0)) as weekday_Customers,
  SUM(IF(subscriber_type = "Subscriber" and (time(start_date) > "05:00:00" and
time(start_date) < "21:00:00") and extract(DAY FROM start_date) in (2,6), 1, 0)) as
weekday_Subscribers

FROM `bigquery-public-data.san_francisco.bikeshare_trips`
GROUP BY start_station_name, end_station_name
ORDER BY weekday_trips DESC
LIMIT 5
```

#5 - What are the 5 MOST POPULAR commuter trips based on total number of customers?

```
SELECT
 start_station_name,
 end_station_name,
 count(*) as total_trips,
  SUM(IF(extract(DAY FROM start_date) in (2,6), 1, 0)) as weekday_trips,
  SUM(IF(subscriber_type = "Customer" and (time(start_date) > "05:00:00" and time(start_date)
< "21:00:00") and extract(DAY FROM start_date) in (2,6),1,0)) as weekday_Customers,
  SUM(IF(subscriber_type = "Subscriber" and (time(start_date) > "05:00:00" and
time(start_date) < "21:00:00") and extract(DAY FROM start_date) in (2,6), 1, 0)) as
weekday_Subscribers

FROM `bigquery-public-data.san_francisco.bikeshare_trips`

GROUP BY start_station_name, end_station_name
ORDER BY weekday_customers DESC
LIMIT 5
```

#6 - What are the 5 MOST POPULAR commuter trips based on total number of daily subscribers?

```
SELECT
 start_station_name,
 end_station_name,
 count(*) as total_trips,
 SUM(IF(extract(DAY FROM start_date) not in (2,6), 1, 0)) as weekend_trips,
 SUM(IF(subscriber_type = "Customer" and (time(start_date) > "05:00:00" and time(start_date)
< "21:00:00") and extract(DAY FROM start_date) not in (2,6),1,0)) as weekend_Customers,
  SUM(IF(subscriber_type = "Subscriber" and (time(start_date) > "05:00:00" and
time(start_date) < "21:00:00") and extract(DAY FROM start_date) not in (2,6), 1, 0)) as
weekend_Subscribers,
  SUM(IF(extract(DAY FROM start_date) in (2,6), 1, 0)) as weekday_trips,
```

```
  SUM(IF(subscriber_type = "Customer" and (time(start_date) > "05:00:00" and time(start_date)
< "21:00:00") and extract(DAY FROM start_date) in (2,6),1,0)) as weekday_Customers,
  SUM(IF(subscriber_type = "Subscriber" and (time(start_date) > "05:00:00" and
time(start_date) < "21:00:00") and extract(DAY FROM start_date) in (2,6), 1, 0)) as
weekday_Subscribers

FROM `bigquery-public-data.san_francisco.bikeshare_trips`

GROUP BY start_station_name, end_station_name
ORDER BY weekday_Subscribers DESC
LIMIT 5
```