**Assignment 4 - Analytical Functions**

Using the us\_states table, select the total confirmed cases, by state, as of 3/1/2022. Be  
sure to show the date. Order by state name. You might be able to avoid an analytical  
function in this part, and this is more to get you familiar with the COVID dataset.

Your Answer:

**Query :**

SELECT  state\_name,date,SUM(confirmed\_cases) total\_cases

from `bigquery-public-data.covid19\_nyt.us\_states`

where date = '2022-03-01'

group by state\_name,date

order by state\_name

![Graphical user interface, text, application, email

Description automatically generated]()

**Question 2**

Let’s rewind back to last year when COVID was just spreading in the US. Using the  
us\_states table, show the new confirmed cases by day, by state. In this case, new cases  
are the differences of two days: day(n) - day(n-1). Sort by state and date in ascending  
order.

Please paste your code here.

Your Answer:

Query:

SELECT state\_name, date,confirmed\_cases AS cummulative\_cases,

confirmed\_cases - lag(confirmed\_cases)

  OVER (

    PARTITION BY state\_name

    ORDER BY date ASC

  ) AS new\_cases

FROM `bigquery-public-data.covid19\_nyt.us\_states` where state\_name = 'New York'

![Graphical user interface

Description automatically generated]()

**Question 3**

Using the results in question 2, show the 7 day moving average of new cases and  
cumulative cases, for Utah. Keep in mind, a 7 day moving average is inclusive of the  
current date and the 6 days prior. Be sure to round your answer to the second decimal  
place.

Your Answer:

Query:

with  ques3 as

(

SELECT state\_name, date,

confirmed\_cases AS cummulative\_cases,

confirmed\_cases - lag(confirmed\_cases)

  OVER (

    PARTITION BY state\_name

    ORDER BY date ASC

  ) AS new\_cases,

  from `bigquery-public-data.covid19\_nyt.us\_states` where state\_name = 'Utah'

)

select

state\_name, date,

cummulative\_cases,new\_cases,

ROUND(moving\_7day\_newcases,2) AS moving\_7day\_newcases,ROUND(moving\_7day\_cumu,2) AS moving\_7day\_cumu

FROM (

  SELECT

    \*,

    AVG(new\_cases) OVER(ORDER BY UNIX\_DATE(date) RANGE BETWEEN 6 PRECEDING AND CURRENT ROW) AS moving\_7day\_newcases,

    AVG(cummulative\_cases) OVER(ORDER BY UNIX\_DATE(date) RANGE BETWEEN 6 PRECEDING AND CURRENT ROW) AS moving\_7day\_cumu,

  FROM

    ques3 )

![Graphical user interface, table

Description automatically generated]()

**Question 4**

Using the us\_counties table, rank the top 3 counties, by state, for confirmed cases.  
Remember struct and array\_agg? You may find them useful for this question.  
The data should be on or before 3/1/2022.

Your Answer:

Query :

with CTE as

(select state\_name,

array\_agg(struct(county,confirmed\_cases)limit 3) as county\_cases,

RANK() OVER(PARTITION BY state\_name ORDER BY confirmed\_cases desc) as rank

from `bigquery-public-data.covid19\_nyt.us\_counties`

where date = '2022-03-01' group by state\_name,confirmed\_cases)

select state\_name,county\_cases  from CTE  where state\_name = 'California' order by rank limit 3

![Text

Description automatically generated with medium confidence]()

Almost right, you would use your CTE solely to build your LAST\_VALUE() window function vs RANK(). And in your main query you would use that same window function 'confrimed\_cases' in your array\_agg() window function. You would have your LIMIT at the end of your window function. Also we would want to see all states, not only CA