**Final Assignment**

**Question 1**

**Date Parsing**

The table `is6850.final\_assignment\_Spring\_2021.date\_strings` is publicly available. Each row of the table contains a single date formatted as a string, but the dates are not formatted consistently. Write a query to parse these strings into dates. Your results should contain a single  
column of type DATE with 41,047 rows and no nulls.

Hint: This is straight from the homework. If you get stuck, follow the guidance in the homework.

Please paste your query for question 1 below and submit a screenshot of the results

Your Answer:

Query :

SELECT  
date\_string,  
  COALESCE(SAFE.PARSE\_DATE('%Y-%m-%d', date\_string),   
  SAFE.PARSE\_DATE('%b-%d-%Y', date\_string),  
SAFE.PARSE\_DATE('%Y-%m-%d', date\_string),  
SAFE.PARSE\_DATE('%Y/%m/%d', date\_string)) as date\_strings  
FROM  `is6850.final\_assignment\_Spring\_2021.date\_strings`

![Graphical user interface, table

Description automatically generated with medium confidence]()

**Question 2**

**Query JSON**

The table `is6850.final\_assignment\_Spring\_2021.census\_json` contains census data in JSON format. Use the JSON\_QUERY function to write the queries below.

1. Write a query that selects the zip code from each row.

2. Note that the “zip\_details” element of each row contains an array. Write a query that selects the geo\_id from the first array element of “zip\_details” in each row.

Paste both of your queries below and submit screenshots of the results.

Your Answer:

select json\_extract(json\_string\_data,'$.zipcode') as zipcodes from `is6850.final\_assignment\_Spring\_2021.census\_json`

![Graphical user interface, text

Description automatically generated]()

Query 2:

WITH cte as   
(  
  SELECT  \*  FROM `is6850.final\_assignment\_Spring\_2021.census\_json`,  
    UNNEST(ARRAY['$.zip\_details']) AS zip\_details

)  
SELECT JSON\_EXTRACT(json\_string\_data, '$.zip\_details[0]') from cte

![Graphical user interface, text, application, email

Description automatically generated]()

missing the geo field after your array look up. this would provide the correct results. Same as assignment 3 with the cards questions.

**Question 3**

**Nest Data**

Write a query that selects from `bigquery-public-data.census\_bureau\_usa.population\_by\_zip\_2010` and nests this data into the form found in `is6850.final\_assignment\_Spring\_2021.census2010`. Your query should produce results that match exactly. (The order of your results won’t match in general.)

Paste your query below and submit a screenshot of the results.

Your Answer:

Query :

select

zipcode,array\_agg(struct(geo\_id,population,minimum\_age,maximum\_age,gender) ) as zip\_details

from `bigquery-public-data.census\_bureau\_usa.population\_by\_zip\_2010`  group by zipcode order by zipcode

SELECT \* from is6850.final\_assignment\_Spring\_2021.census2010  order by zipcode

![Table

Description automatically generated]()

![Table

Description automatically generated]()

**Question 4**

**Unnest Data**

Write a query that selects from `is6850.final\_assignment\_Spring\_2021.census2010` and unnests this data into the form found in `bigquery-public-data.census\_bureau\_usa.population\_by\_zip\_2010`. Your query should produce results that match exactly. (The order of your results won’t match in general.)

Paste your query below and submit a screenshot of the results.

Your Answer:

with cte as

(

    select zipcode,zip\_details

    from `is6850.final\_assignment\_Spring\_2021.census2010`

)

select geo\_id,zipcode, population, minimum\_age,maximum\_age,gender

from cte ,unnest(zip\_details)

order by geo\_id

select \* from `bigquery-public-data.census\_bureau\_usa.population\_by\_zip\_2010`

order by geo\_id

![Graphical user interface, text, application, table

Description automatically generated]()

![Graphical user interface, table

Description automatically generated]()

No need for a CTE.

**Question 5**

**Analytical Functions**

For this problem, you’ll use `is6850.final\_assignment\_Spring\_2021.employee\_table`.

1. Find the earliest and last employees in the employee table, by hire date. There may be more than one employee hired for a date, meaning there may be duplicate dates. Use an analytical function to provide the following - id, first name, last name, and hire date.

Hint - remember the different ranking types we reviewed?

2. What quarter and year were each of the employees in 5.1 hired?

3. Using the results in 5.1 and 5.2, how many days separate the first hire date from the last hire date? Be sure to order by the number of days difference between hire dates.

Hint - the first value of the difference might be null.

Paste your queries below and submit screenshots of the results.

Your Answer:

Query 1 :

select \*,extract(year from hire\_date) as Year,

extract(quarter from hire\_date) as Quarter

 from `is6850.final\_assignment\_Spring\_2021.employee\_table`

where (hire\_date = (select max(hire\_date)

 from `is6850.final\_assignment\_Spring\_2021.employee\_table`)) OR

(hire\_date = (select min(hire\_date)

from `is6850.final\_assignment\_Spring\_2021.employee\_table`))

![Graphical user interface, text

Description automatically generated]()

Query 2 :

Quater 2 and year 2021 and 2018

Query 3:

Though you pulled back the results, you are missing many various elements. Make sure to read to question. Missing DENSE\_RANK(), correct WHERE clause, missing date difference.