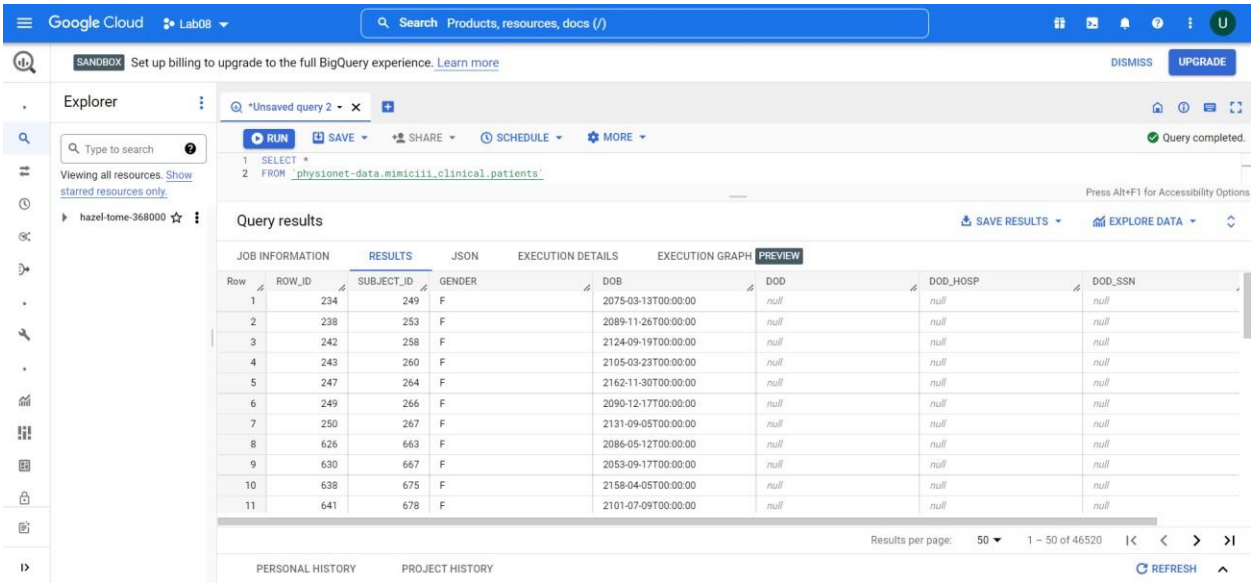
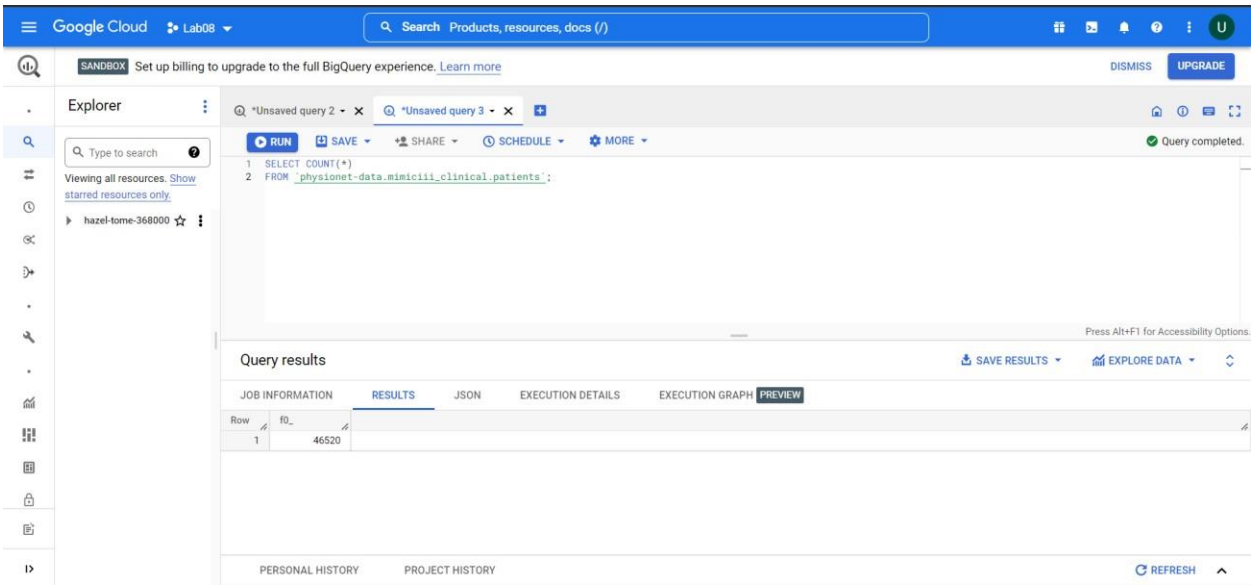
Working with MIMIC Dataset using Google BigQuery

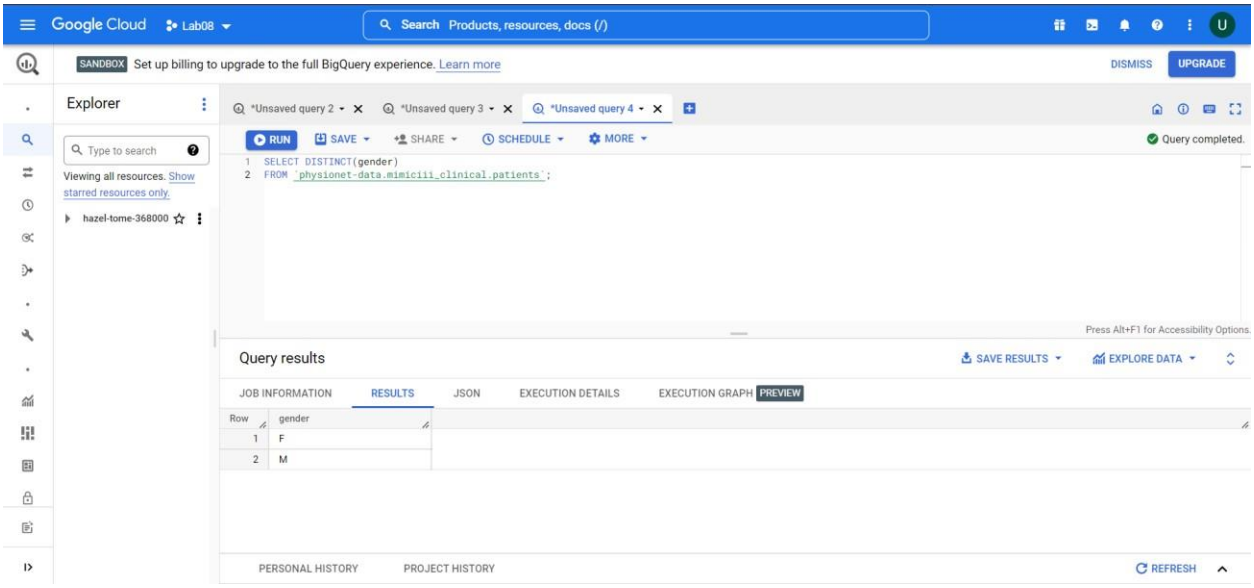
Query 1: Patient Numbers



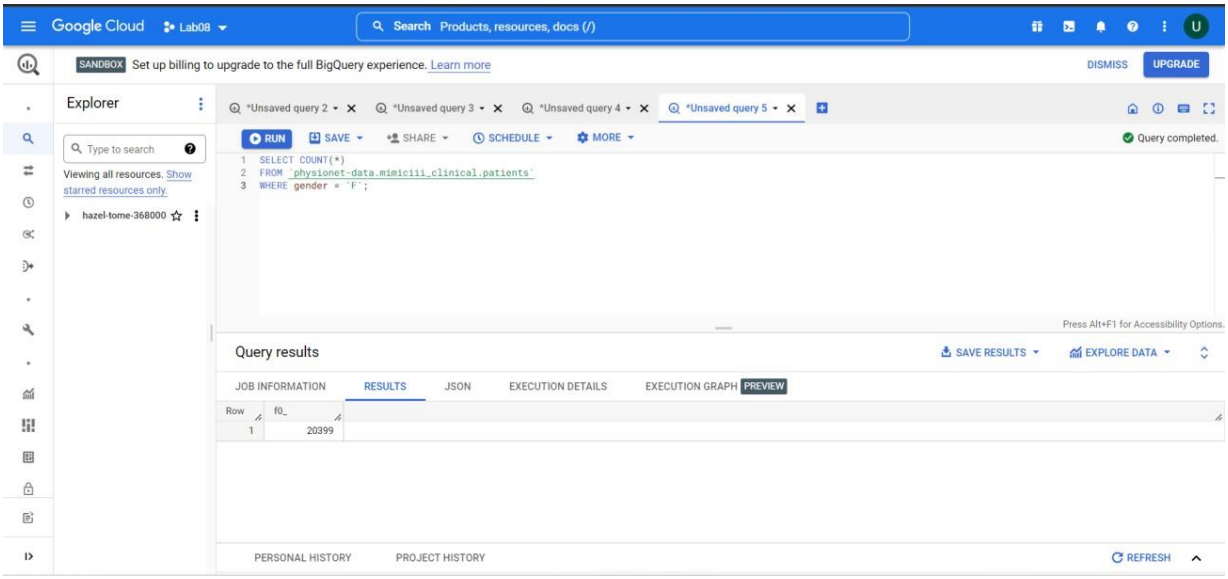
Query 2: Obtain the number of patients with the following query



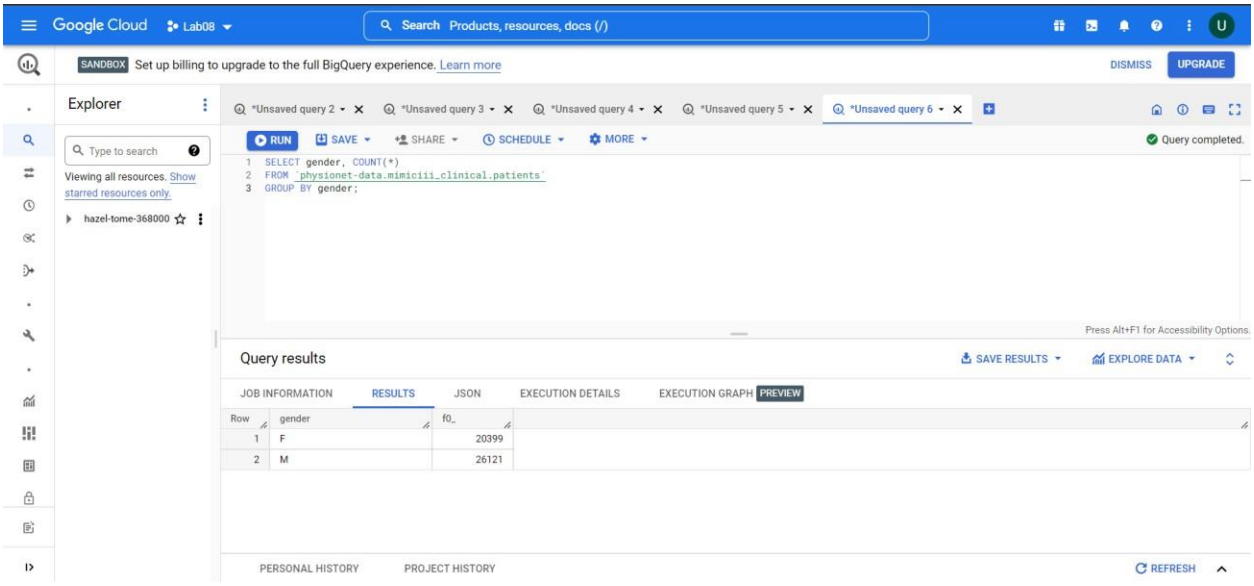
Query 3: The ‘gender’ column identifies the gender of the patient. We can obtain the distinct values used to indicate gender using the following query:



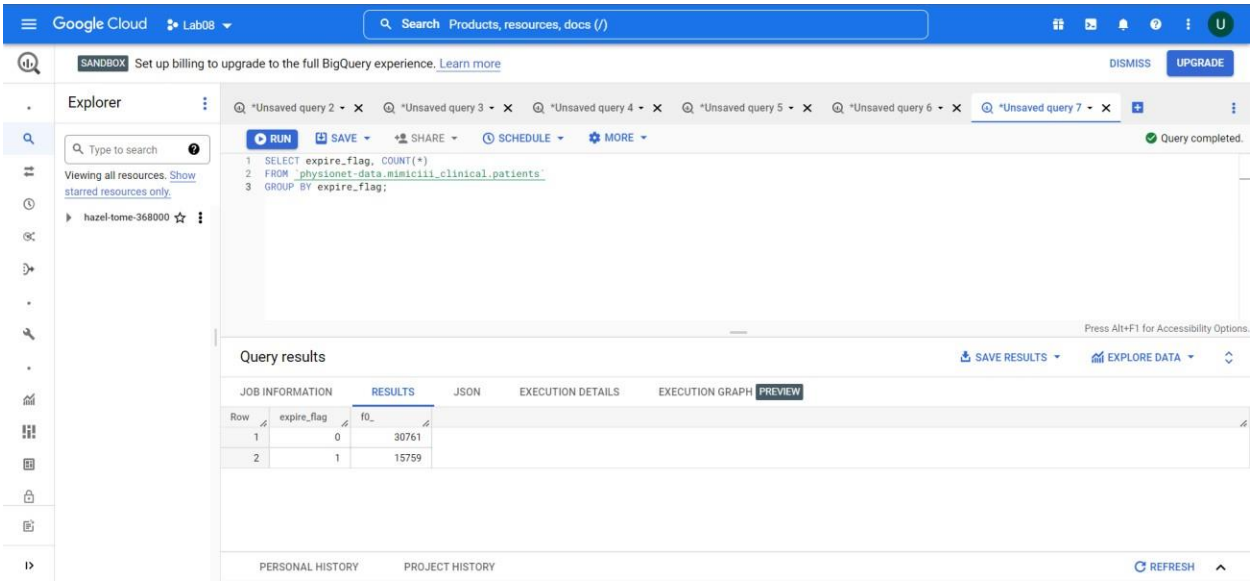
Query 4: We can see that ‘M’ and ‘F’ are the two characters used to indicate patient gender. We can use this information to obtain the number of female patients by adding a condition to select rows where the gender is ‘F’:



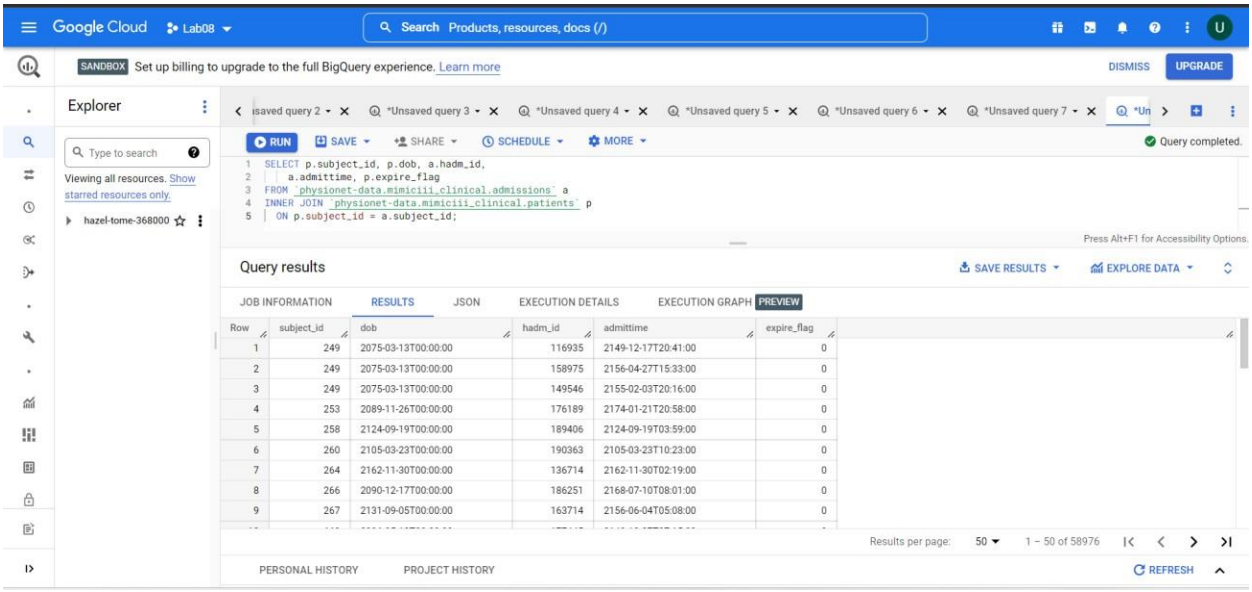
Query 5: the numbers of male and female patients can be obtained using the following query



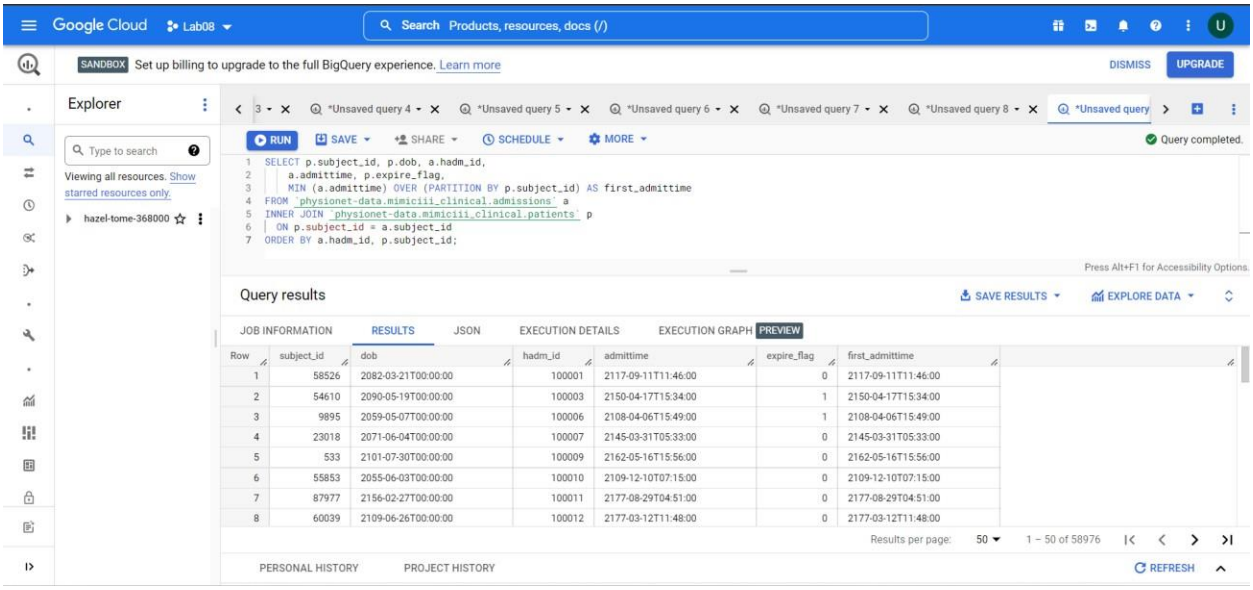
Query 6: **Mortality and admissions-** A flag which records whether or not a patient died in the hospital is stored in the patients table. Count the number of patients who died using the following query:



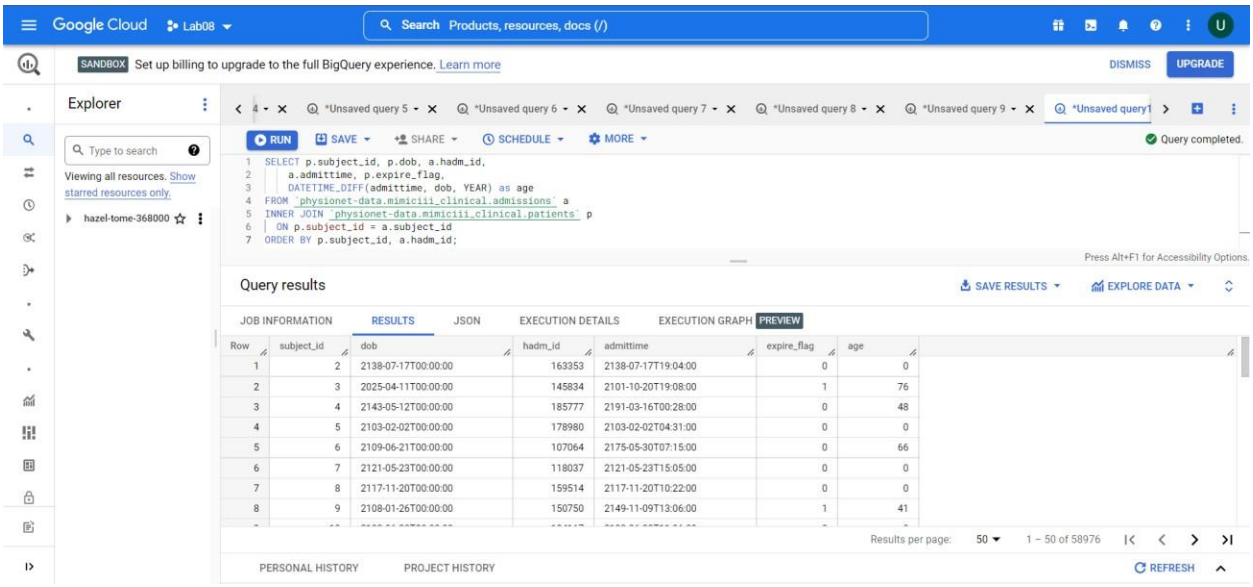
Query 7: **Patient age and mortality**



Query 8: Next we find the earliest admission date for each patient. This requires the use of two functions: the ‘MIN’ function, which obtains the minimum value, and the ‘PARTITION BY’ function, which determines the groups over which the minimum value is obtained. To determine the earliest admission time for each patient:

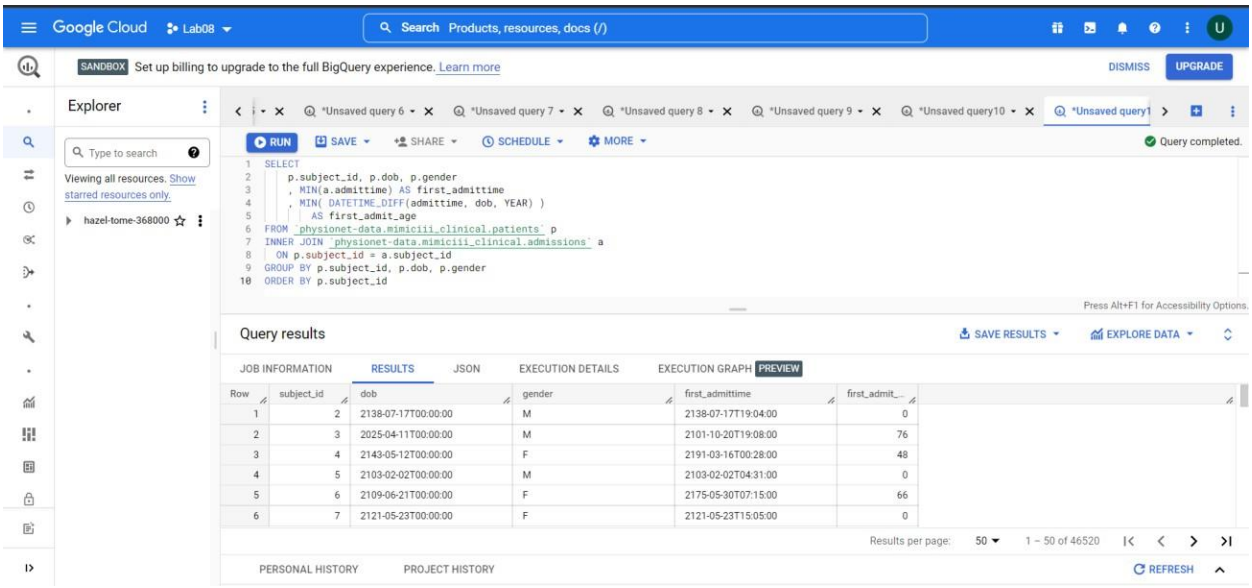


Query 9: Let's calculate the age of patients at their time of admission:

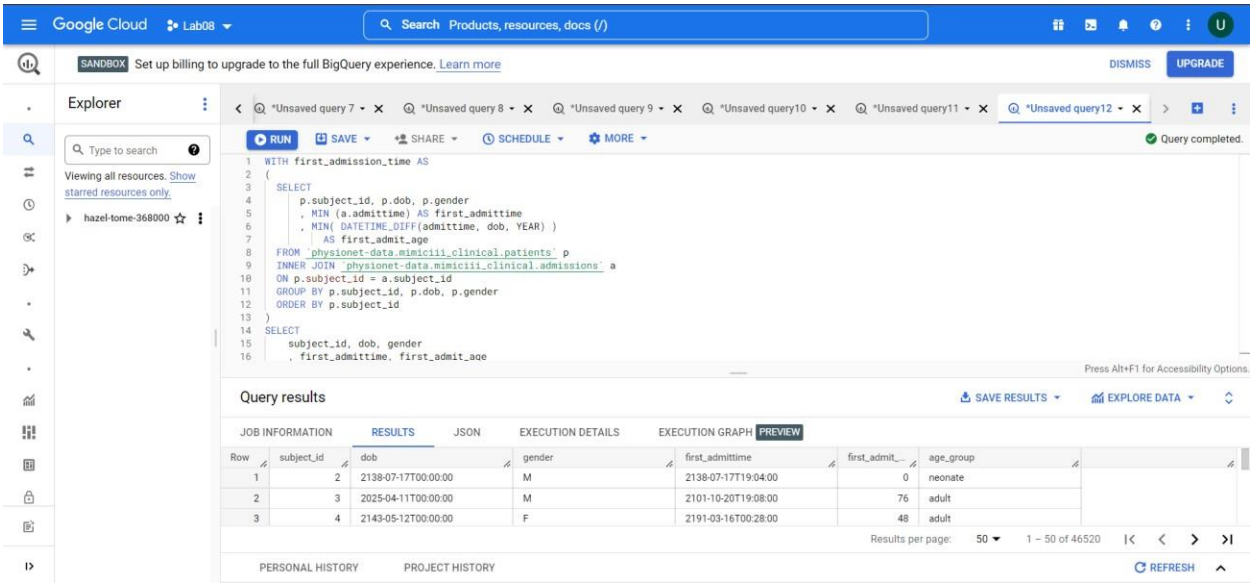


Query 10: If we examine the same patient more than once when calculating a statistic such as mortality, then our estimate will contain "repeated measures". Unless we handle this phenomenon explicitly, our calculation will be biased.

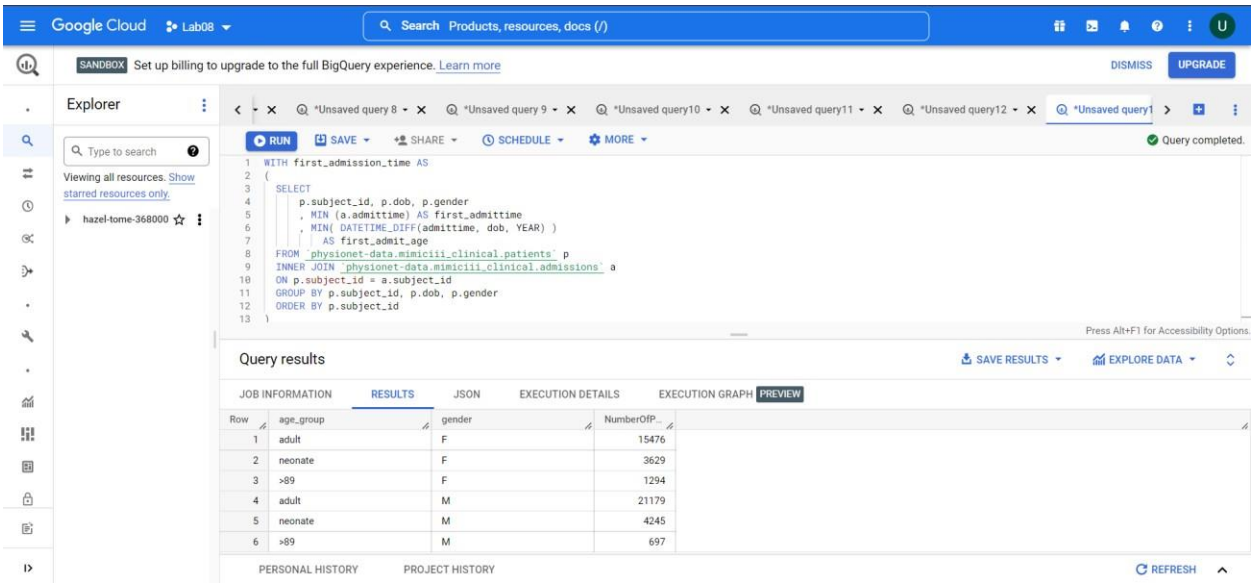
A simple solution is to only examine the first hospitalization for each patient, which we can do with a `GROUP BY` clause.



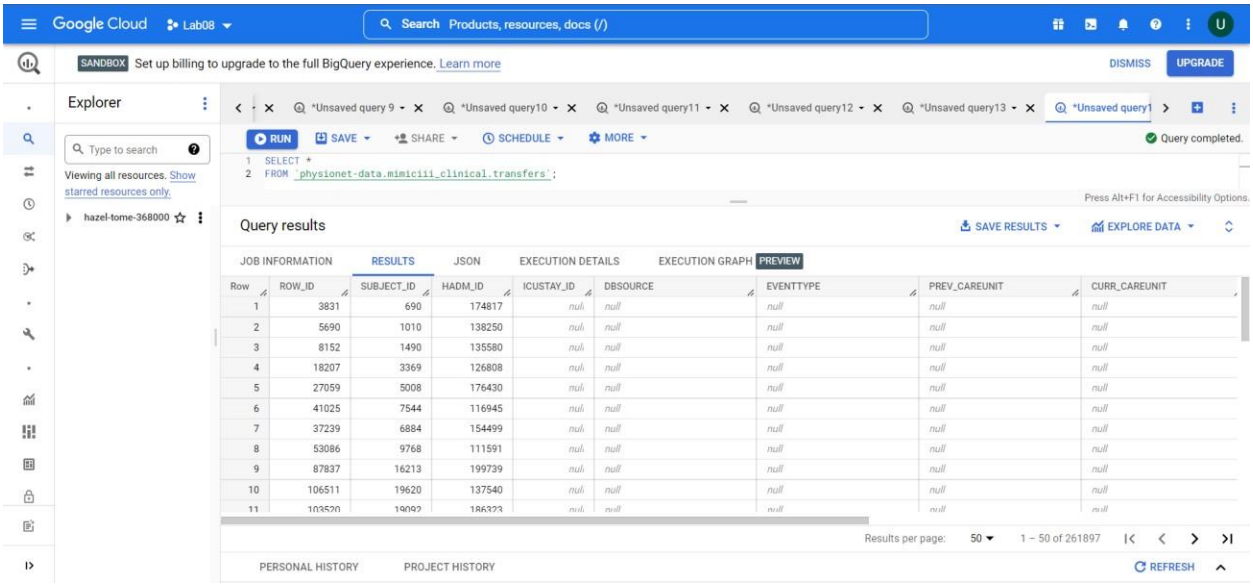
Query 11: Now that we have a set of unique patients with their age, we can group them into sensible categories based upon age and calculate the mortality rate in these categories. Patients with an age >= 15 years old are adults and the rest are assigned to other categories. Note the use of the `WITH` clause, which allows us to make a temporary view which we can query against in subsequent lines.



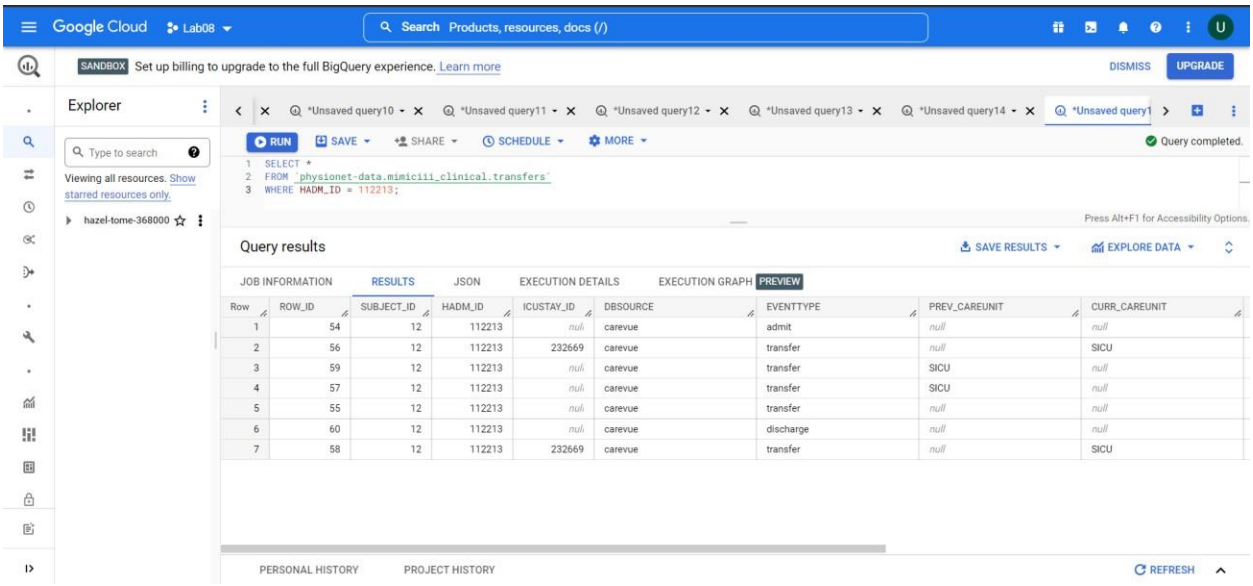
Query 12: The above query can now be combined with the **WHERE** and **COUNT** functions described earlier to determine the number of adult patients, whether or not they died, and therefore, their mortality rate.



Query 13: **ICU stays**

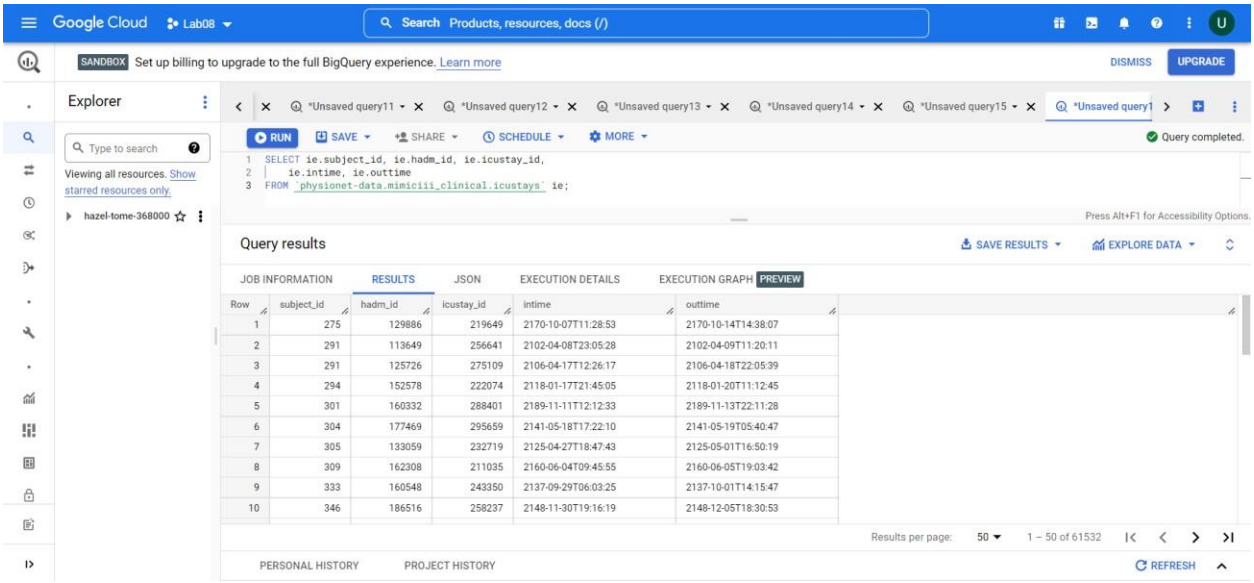


Query 14: The transfers table may have multiple entries per patient to provide detail of all movement between various careunits of the hospital. The first entry in the transfers table for a patient who comes into the ICU will have nothing in the ‘prev_careunit’ column. Similarly, the last entry for a patient will have nothing in the ‘curr_careunit’. Entries that have nothing in both previous and current careunit columns indicate that the patient has been transfered between non intensive care units. An example query for one patient and result from the transfers table is shown below. Note that columns ‘intime’, ‘outtime’, and ‘los’ have been truncated.

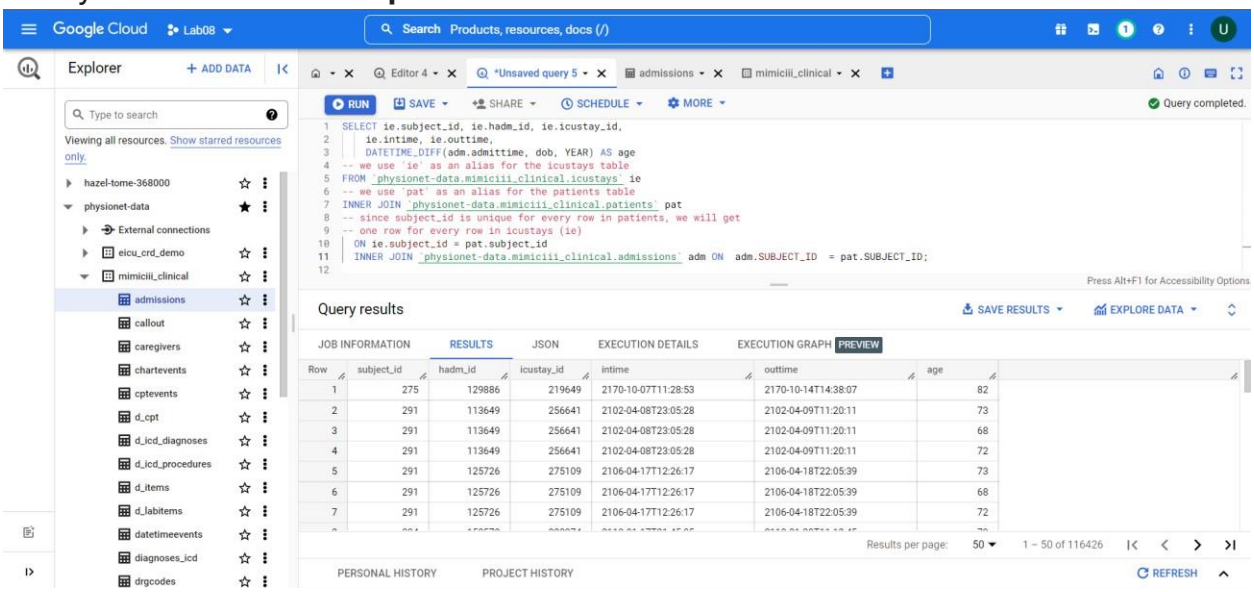


Query 15: **Solutions to the problems in section 8**

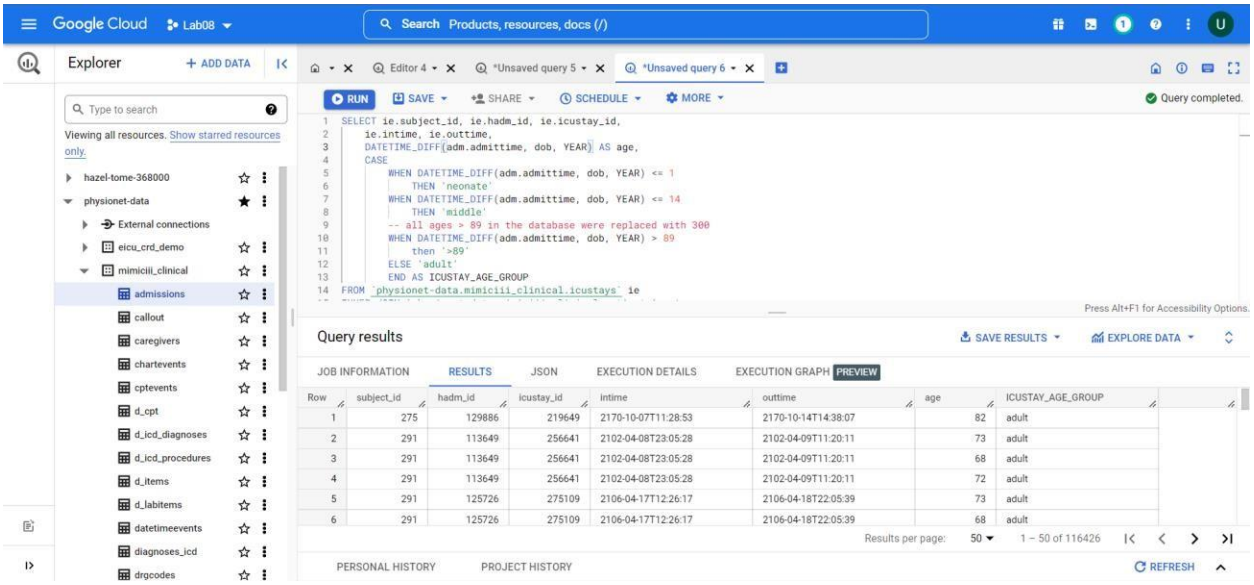
Solution to step 1



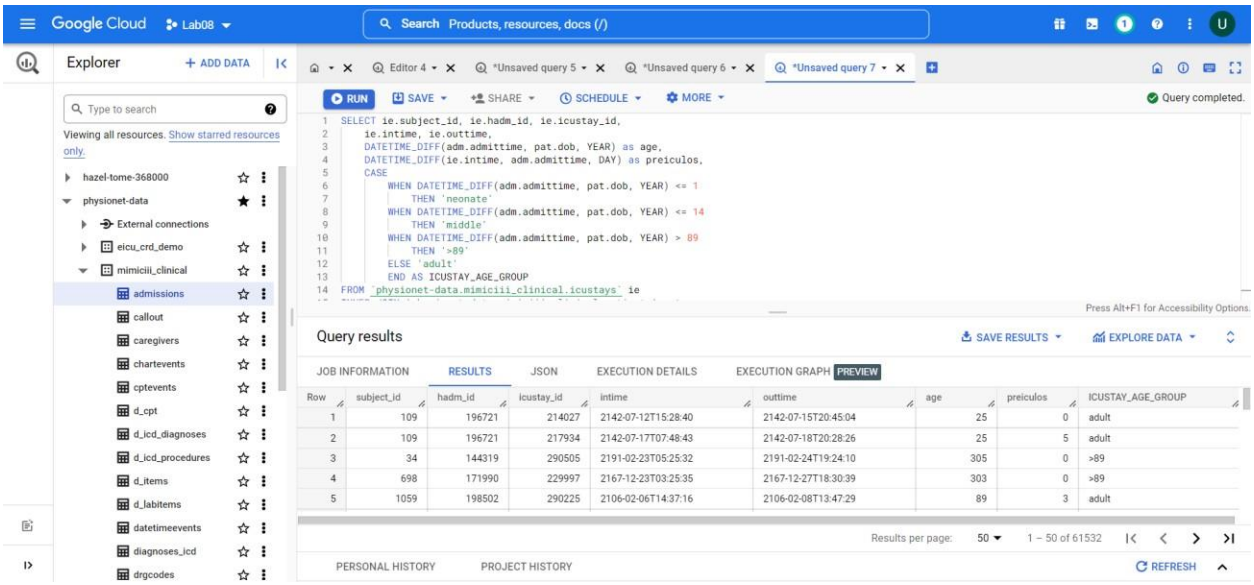
Query 16: **Solution to step 2**



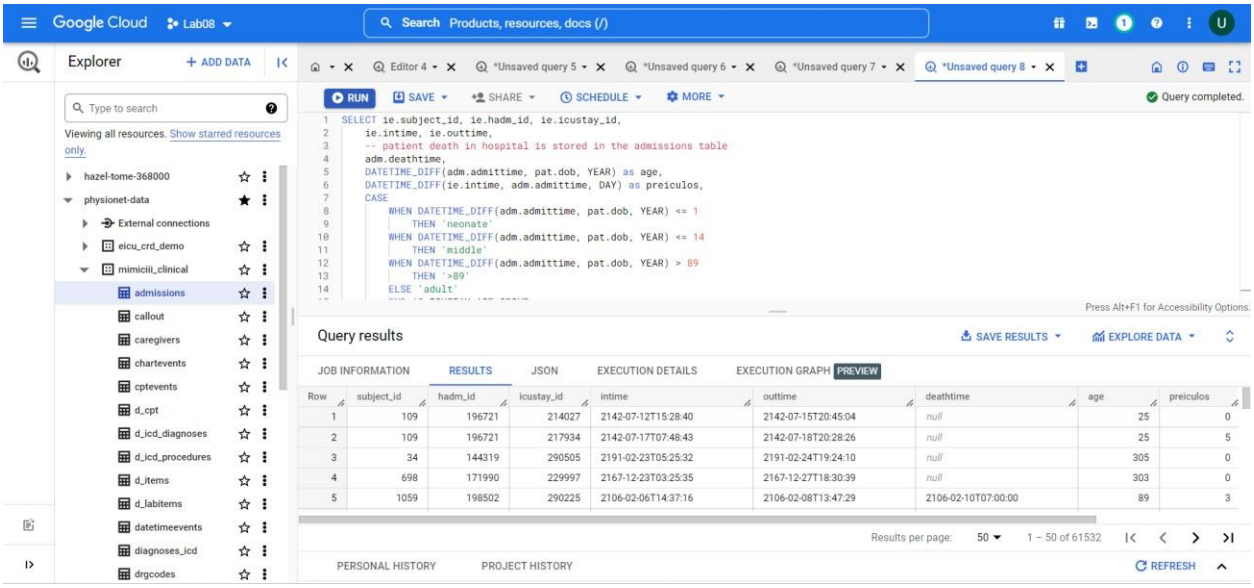
Query 17: **Solution to step 3**



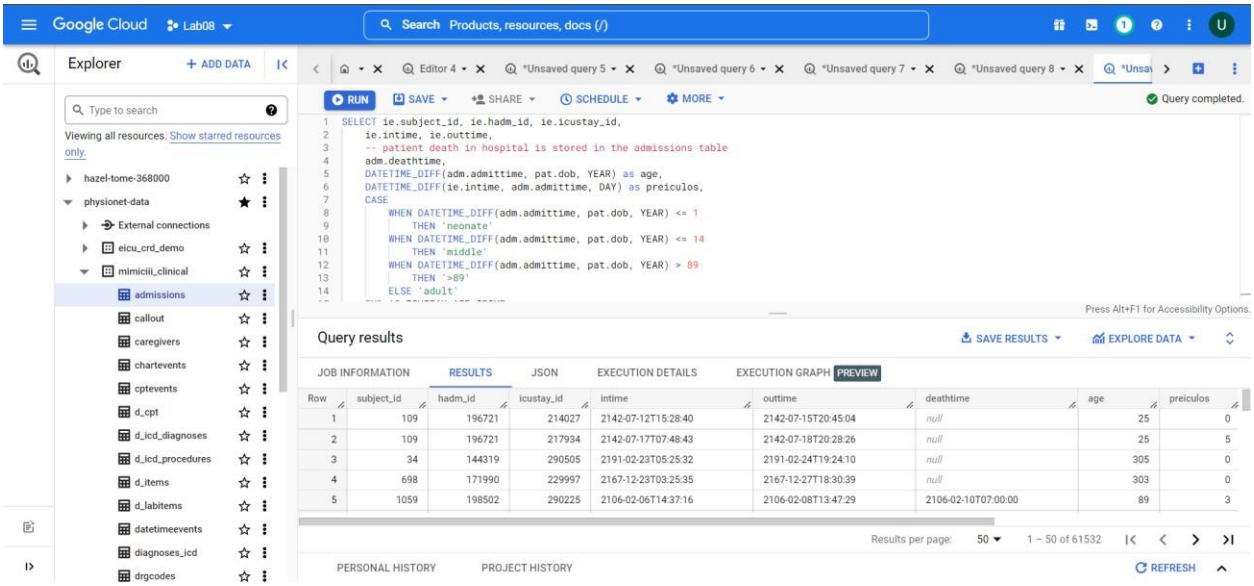
Query 18: **Solution to step 4**



Query 19: **Solution to step 5**



Query 20: **Solution to step 6**



Query 21: **Solution to step 7**