DATABASE DOCS

<https://umich.instructure.com/courses/38100/files/folder/Case_Studies/18_HospitalAdmissions>?

HospitalAdmissionsData

AdmissionLengthDays: Duration of hospital stay (in days) (float)

Death\_1: Indicator of Death (1) survival (0) (int)

Admission\_Type: Type of admission (categorical) (string)

Insurance\_Type: Type of Health insurance (string)

EnglishLanguage\_1: Primary Language English Indicator (1), 0 otherwise (int)

Religion\_Type: Type of Religious Affiliation (string)

Married\_1: Indicator of marital status (Married=1) (int)

Race: Race, categorical (string)

Dx: Diagnosis (string)

Death\_Status

Status\_ID: Unique identifier for the description of the status of the patient. (int)

Status\_Comment: description of the status of the patient. (string)

Insurance\_Type\_Info

Insurance\_Type\_ID: Unique identifier for the description of the status of the patient. (int)

Insurance\_Type\_Comment: description of the type of insurance the patient carries. (string)

Race

Race\_ID: unique identifier for the description of the patients races.

Race\_Type: description of the race of the patient.

Admission\_Type

Admission\_Type\_ID: unique identifier for the type of admission type comment that the patient was

enrolled to the hospital. (int)

Admission\_Type\_Comment: description of the patients admission status when was admited to

the hospital. (string)

Marital\_Status

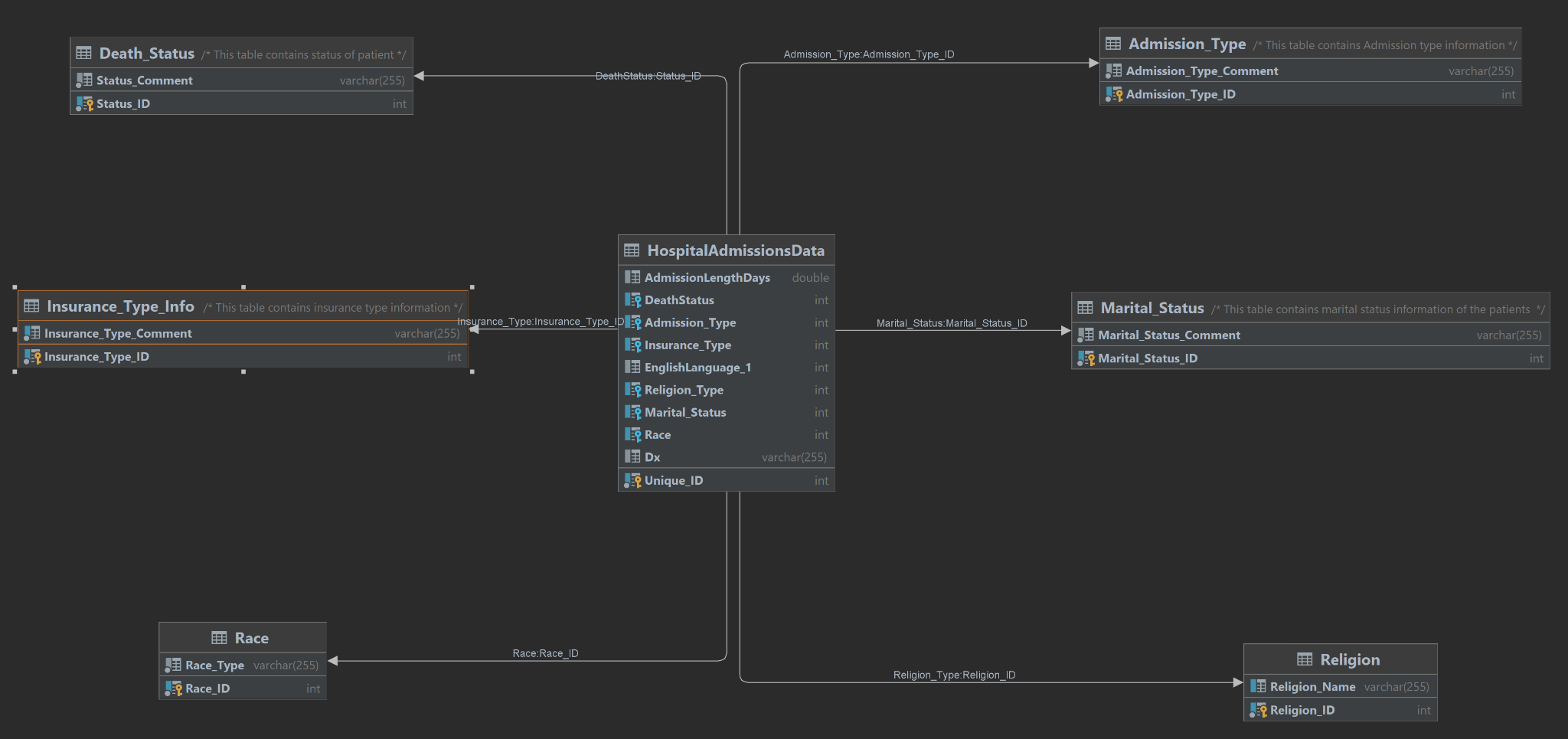
Marital\_Status: unique identifier for the martial status comment of the patient. (int)

Marital\_Status\_Comment: description of patients marital status. (string)

Religion

Religion\_ID: unique identifier for religion name. (int)

Religion\_Name: descrition of religion. (string)



your normalized data should be **at least 2 to 4 tables.** The entire data set should contain **at least 25-100 rows total**. Your data should have properly defined keys.

What to turn in:

1) Your database documentation, including ERD diagrams.

2) A letter from someone at your fictional company outlining important questions they need answered from the dataset. This should have at 4-10 questions in it.

3) The answers to the questions posed in the letter (**including** your queries to get the result).

4) An outline of why you made the database design choices you made.

We made a choice on the database that we did because it had categorical data and we thought it would be easier to normalize the data. At the start we cleaned the data and removed some unnecessary columns that we thought would prove to be redundant, such as ID and made our own by auto incrementing the ID throught the feature in Datagrip.

5) **A .sql file containing your exported database.**

Possible questions

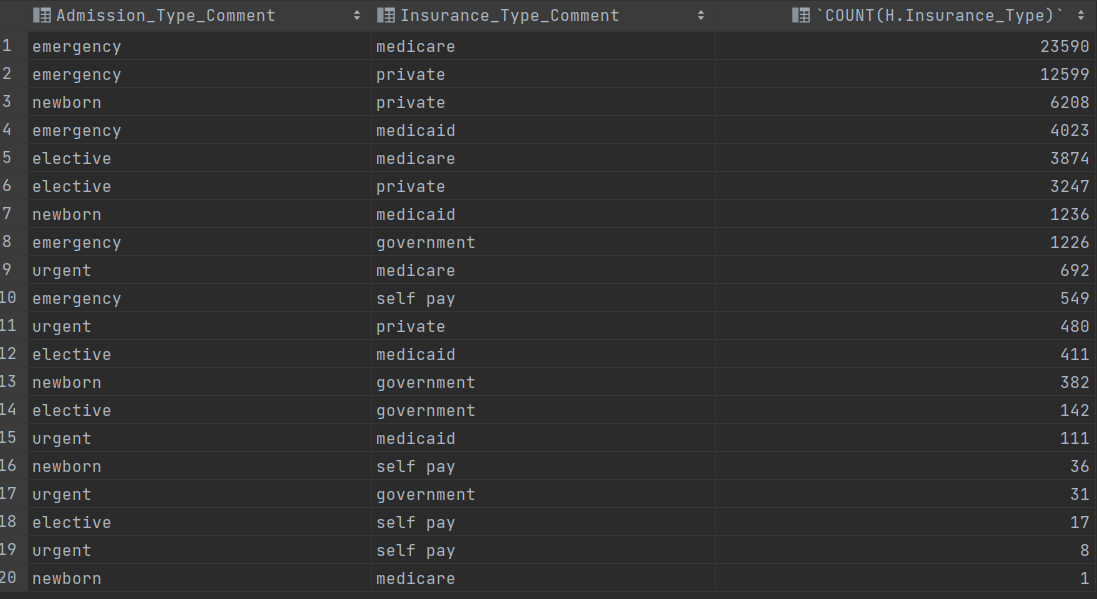
1. Based on the patient's admission, what is the insurance type they carry when grouped by patient admission type?

SELECT A.Admission\_Type\_Comment,ITI.Insurance\_Type\_Comment,COUNT(H.Insurance\_Type) FROM

((HospitalAdmissionsData H LEFT JOIN Insurance\_Type\_Info ITI on H.Insurance\_Type = ITI.Insurance\_Type\_ID) LEFT JOIN Admission\_Type A on A.Admission\_Type\_ID = H.Admission\_Type)

GROUP BY H.Admission\_Type, H.Insurance\_Type

ORDER BY COUNT(H.Insurance\_Type) DESC;

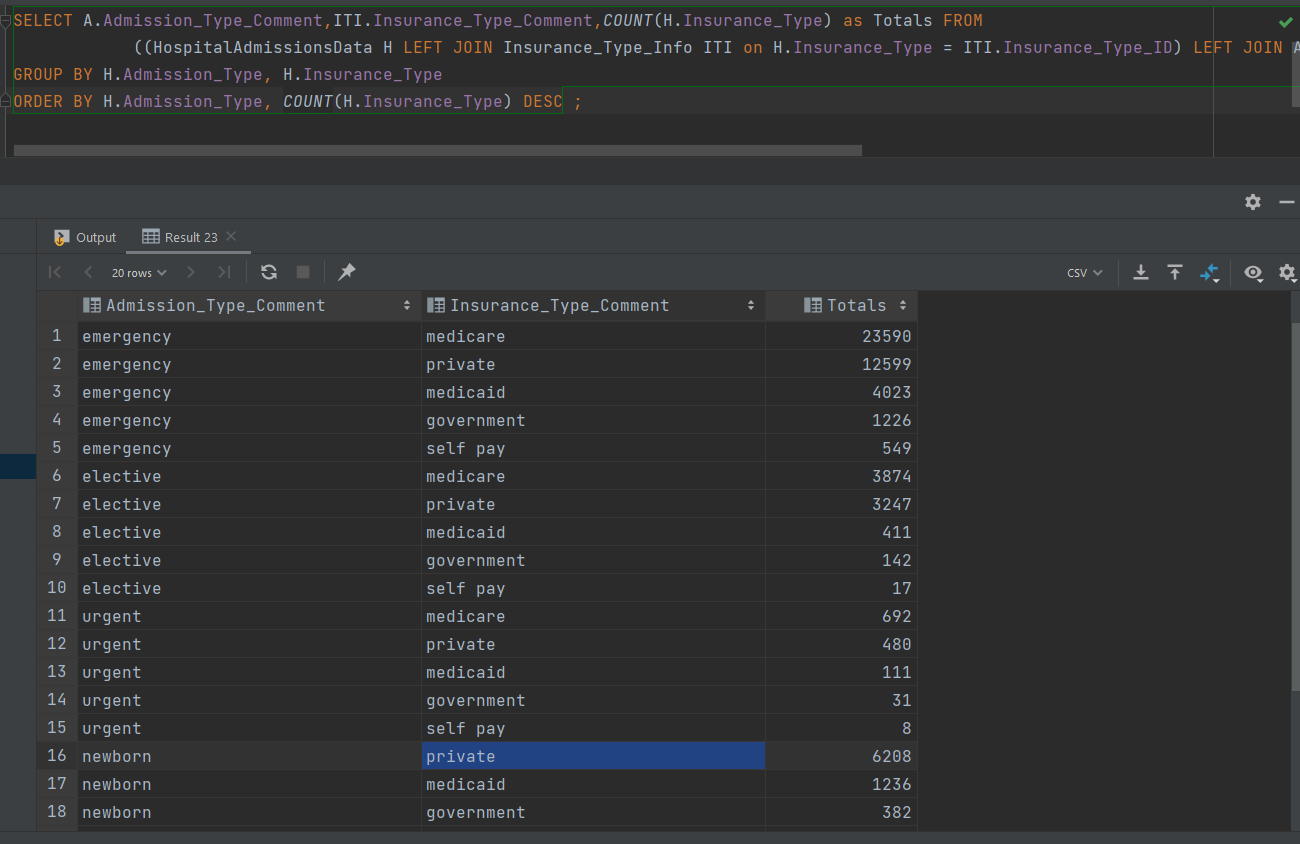


SELECT A.Admission\_Type\_Comment,ITI.Insurance\_Type\_Comment,COUNT(H.Insurance\_Type) as Totals FROM

((HospitalAdmissionsData H LEFT JOIN Insurance\_Type\_Info ITI on H.Insurance\_Type = ITI.Insurance\_Type\_ID) LEFT JOIN Admission\_Type A on A.Admission\_Type\_ID = H.Admission\_Type)

GROUP BY H.Admission\_Type, H.Insurance\_Type

ORDER BY H.Admission\_Type, COUNT(H.Insurance\_Type) DESC ;



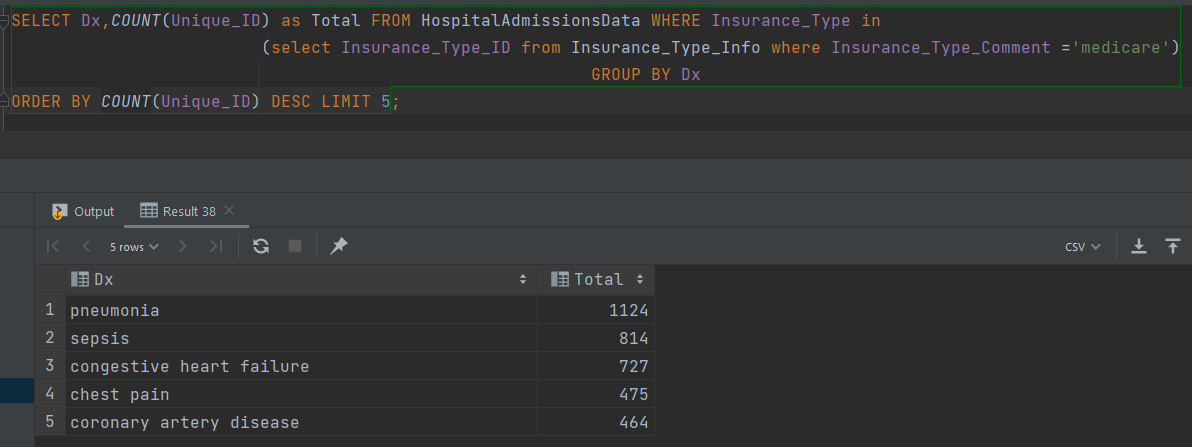
1. What is the most common diagnosis among patients when the insurance type is Medicare?

SELECT Dx,COUNT(Unique\_ID) as Total FROM HospitalAdmissionsData WHERE Insurance\_Type in

(select Insurance\_Type\_ID from Insurance\_Type\_Info where Insurance\_Type\_Comment ='medicare')

GROUP BY Dx

ORDER BY COUNT(Unique\_ID) DESC LIMIT 5;



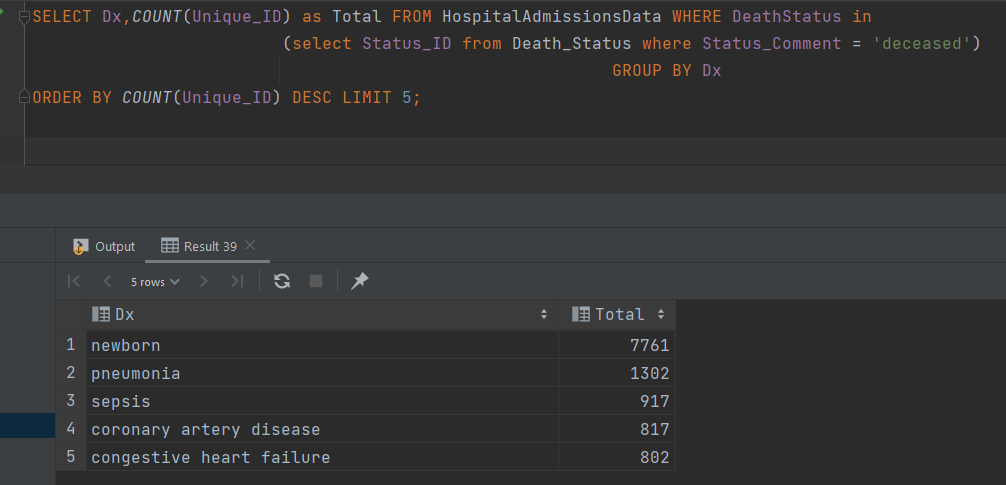
1. On average, what is the diagnosis that causes the most death among patients?

SELECT Dx, COUNT(Unique\_ID) as Total FROM HospitalAdmissionsData WHERE DeathStatus in

(select Status\_ID from Death\_Status where Status\_Comment = 'deceased')

GROUP BY Dx

ORDER BY COUNT(Unique\_ID) DESC LIMIT 5;



1. On average, what is the diagnosis that causes the most number of death among patients if they are admitted under emergency?

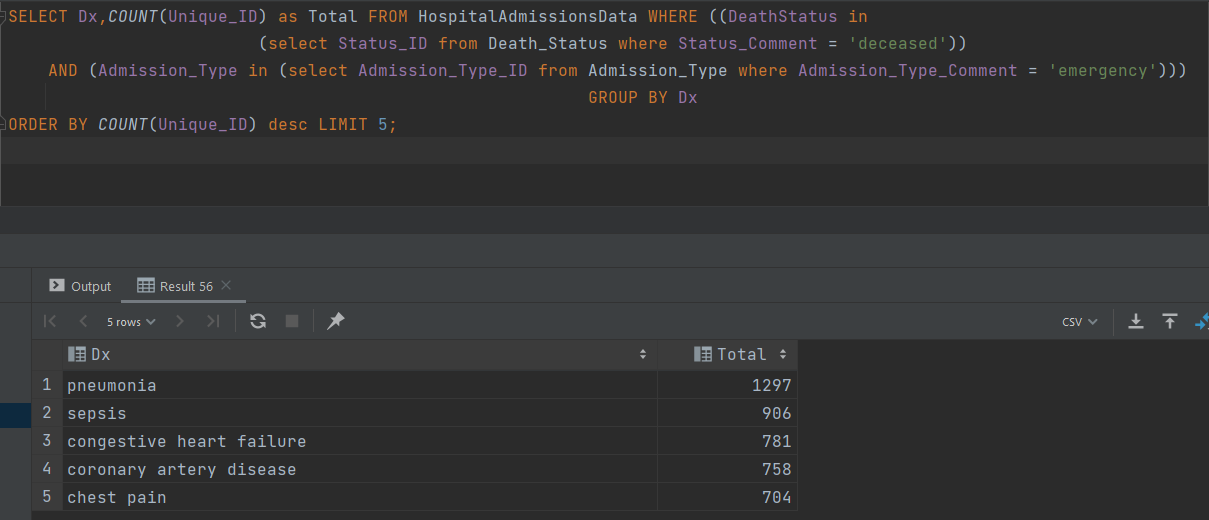
SELECT Dx,COUNT(Unique\_ID) as Total FROM HospitalAdmissionsData WHERE ((DeathStatus in

(select Status\_ID from Death\_Status where Status\_Comment = 'deceased'))

AND (Admission\_Type in (select Admission\_Type\_ID from Admission\_Type where Admission\_Type\_Comment = 'emergency')))

GROUP BY Dx

ORDER BY COUNT(Unique\_ID) desc LIMIT 5;



1. On average, what is the diagnosis which has least length of stay among patients who are admitted to emergency?

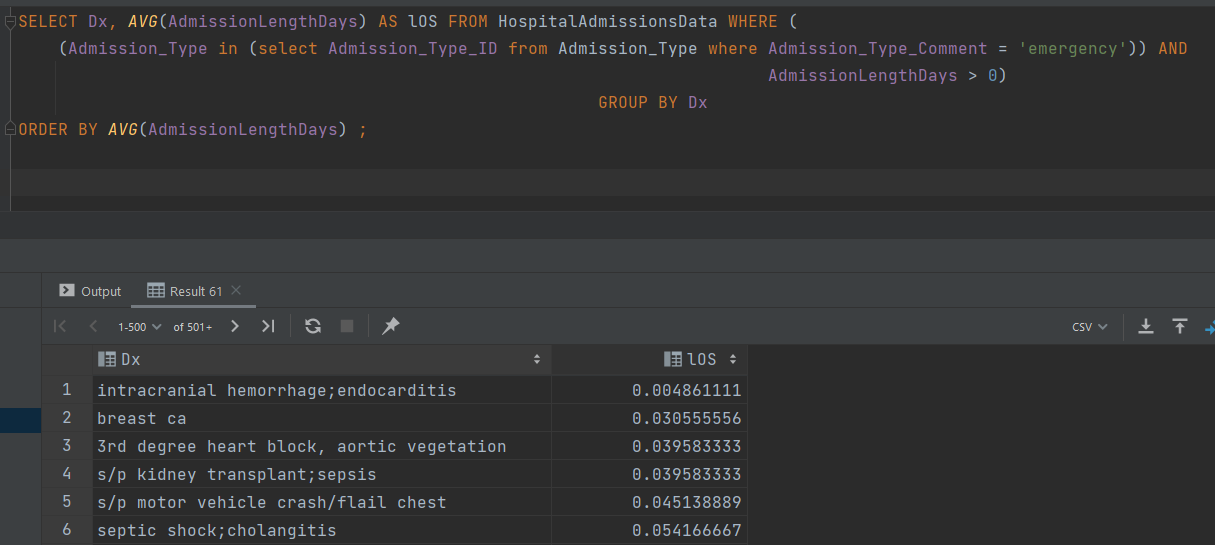
SELECT Dx, AVG(AdmissionLengthDays) AS lOS FROM HospitalAdmissionsData WHERE (

(Admission\_Type in (select Admission\_Type\_ID from Admission\_Type where Admission\_Type\_Comment = 'emergency')) AND

AdmissionLengthDays > 0)

GROUP BY Dx

ORDER BY AVG(AdmissionLengthDays) ;



1. Among the admission type ‘newborn,’ what are the fatality rates?

SELECT (SELECT COUNT(Unique\_ID)

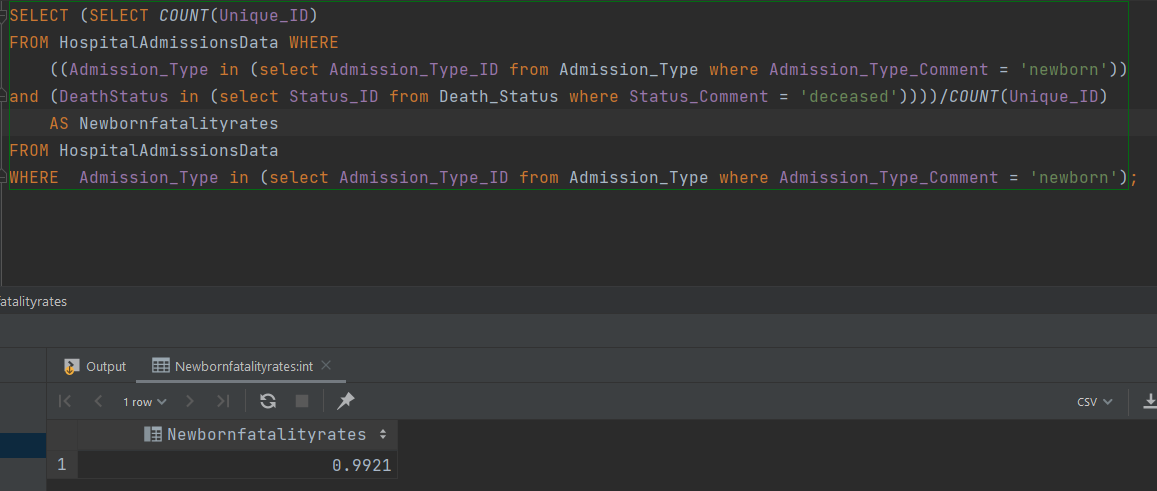
FROM HospitalAdmissionsData WHERE

((Admission\_Type in (select Admission\_Type\_ID from Admission\_Type where Admission\_Type\_Comment = 'newborn'))

and (DeathStatus in (select Status\_ID from Death\_Status where Status\_Comment = 'deceased'))))/COUNT(Unique\_ID) AS Newbornfatalityrates

FROM HospitalAdmissionsData

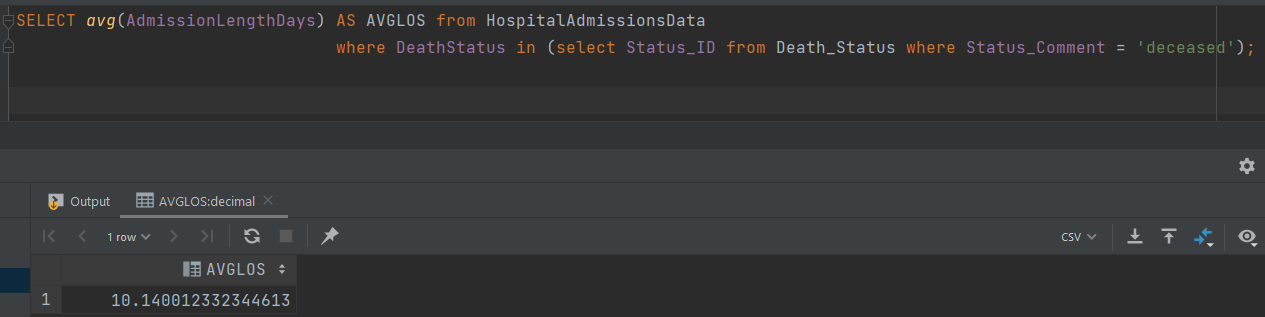
WHERE Admission\_Type in (select Admission\_Type\_ID from Admission\_Type where Admission\_Type\_Comment = 'newborn');



1. When patient mortality occurs, what is the average length of admission?

SELECT avg(AdmissionLengthDays) AS AVGLOS from HospitalAdmissionsData

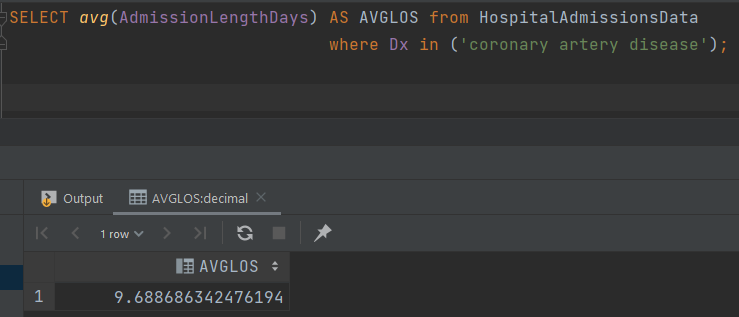
where DeathStatus in (select Status\_ID from Death\_Status where Status\_Comment = 'deceased');

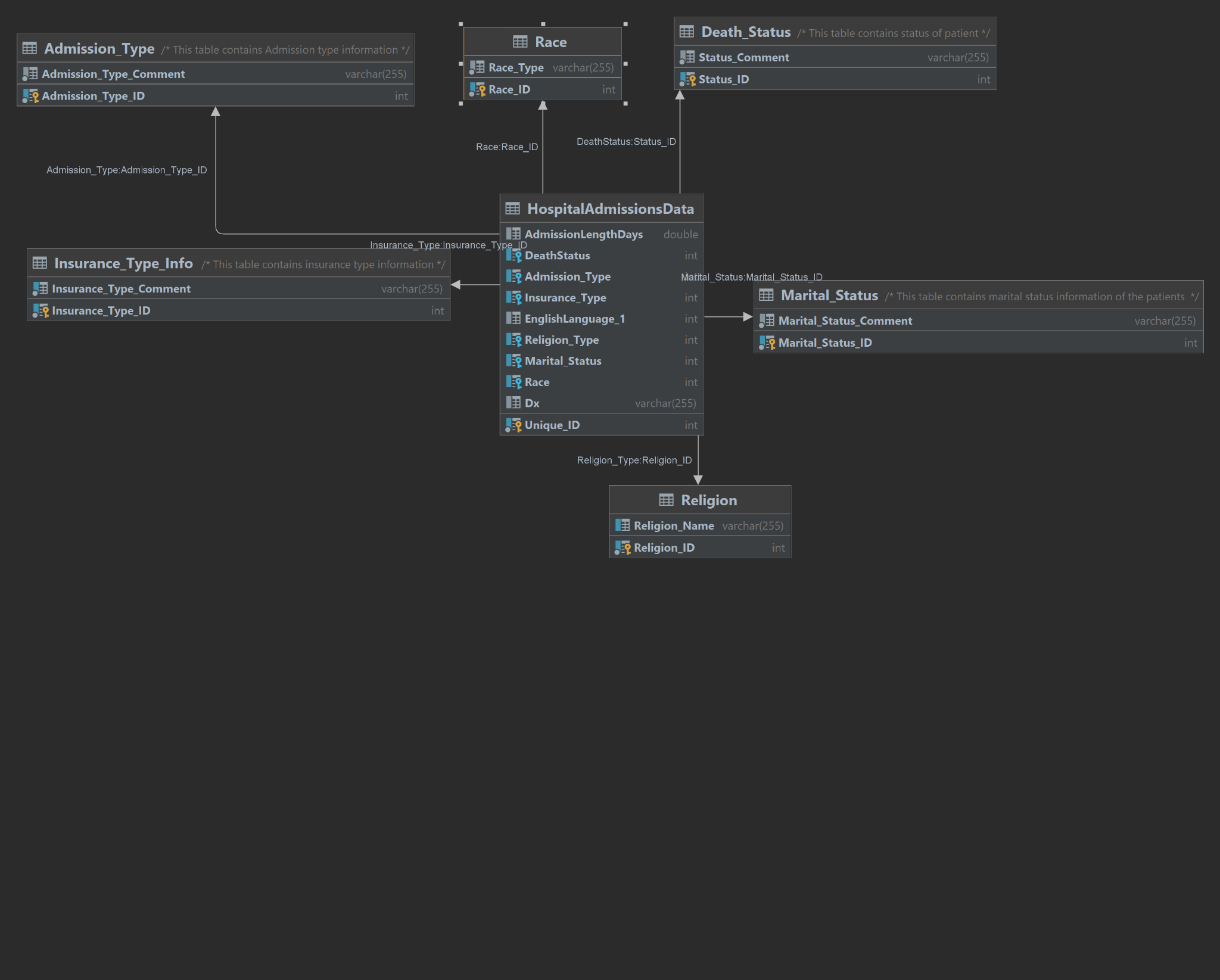


1. What is the average length of stay for a patient who has a Dx of coronary artery disease?

SELECT avg(AdmissionLengthDays) AS AVGLOS from HospitalAdmissionsData

where Dx in ('coronary artery disease');





Dear Borromean team,

This is Bob Blahman, I am excited to be working side by side with your company. We at Blah-Blah Data INC are very proud to bring our work with local hospitals and require the following to be done by analysts from your company:

1. Based on the patient's admission, what is the insurance type they carry when grouped by patient admission type?
2. What is the most common diagnosis among patients when they are grouped by insurance type?
3. On average, what is the diagnosis that causes the most death among patients?
4. On average, what is the diagnosis that causes the least death among patients?
5. On average, what is the diagnosis which has least length of stay among patients who are admitted to emergency?
6. Among the admission type ‘newborn,’ what are the fatality rates?
7. When a patient mortality occurs, what is the average length of admission?
8. What is the average length of stay for a patient who has a Dx of coronary artery disease?

Once these tasks are complete please message me back so we can arrange a company to company mandatory fun day. We plan on throwing a large party in honor of our future collaborations together for the betterment data science.

Best regards,

Bob Blahman

CEO, Blah-Blah Data INC