**ASSESSMENT**

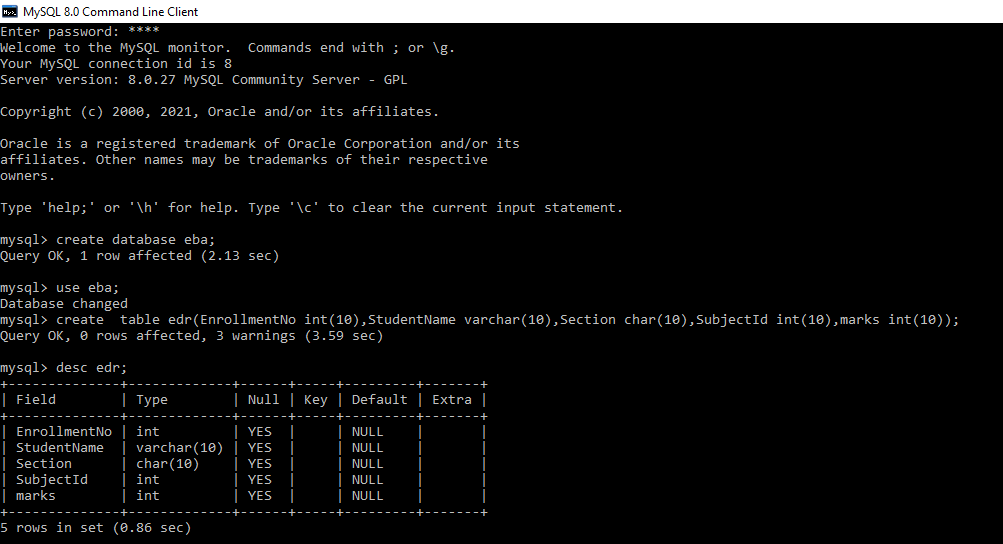
1. **SQL**

To print section wise number of students who got more than or equal to 75 in exam.

**STEP1:** Create database.

**STEP2:** Create a table.

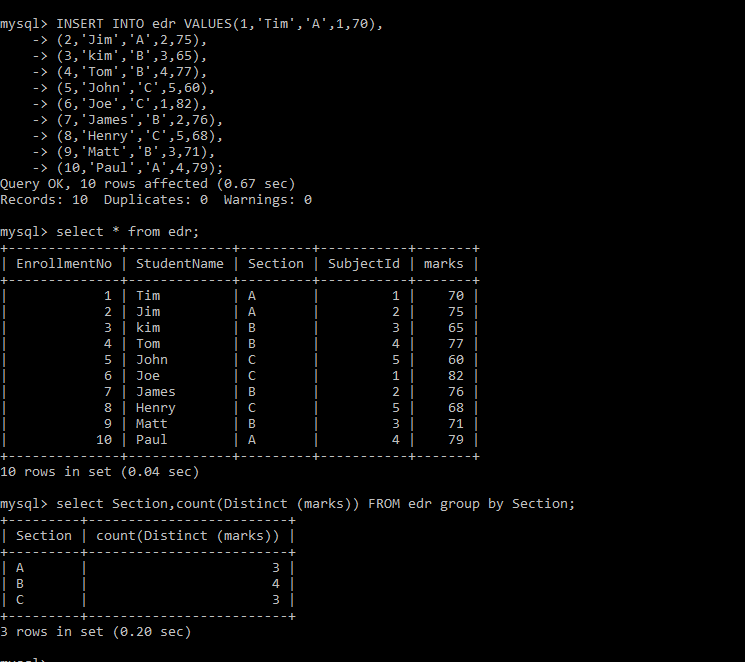
**STEP3:** Insert values into that table.



**STEP4:** To get the given output, we have to write this query-

***Query:***

select Section, count(Distinct (marks)) FROM table\_name GROUP BY Section;



1. **TABLEAU**

**STEP1:**

I took online tableau public to do the given problem.

Firstly, we need the data in CSV format so I entered the data in excel sheet and I kept the file type as CSV format and save it.

**STEP2:**

Import or Load the CSV file in tableau.

**STEP3:**

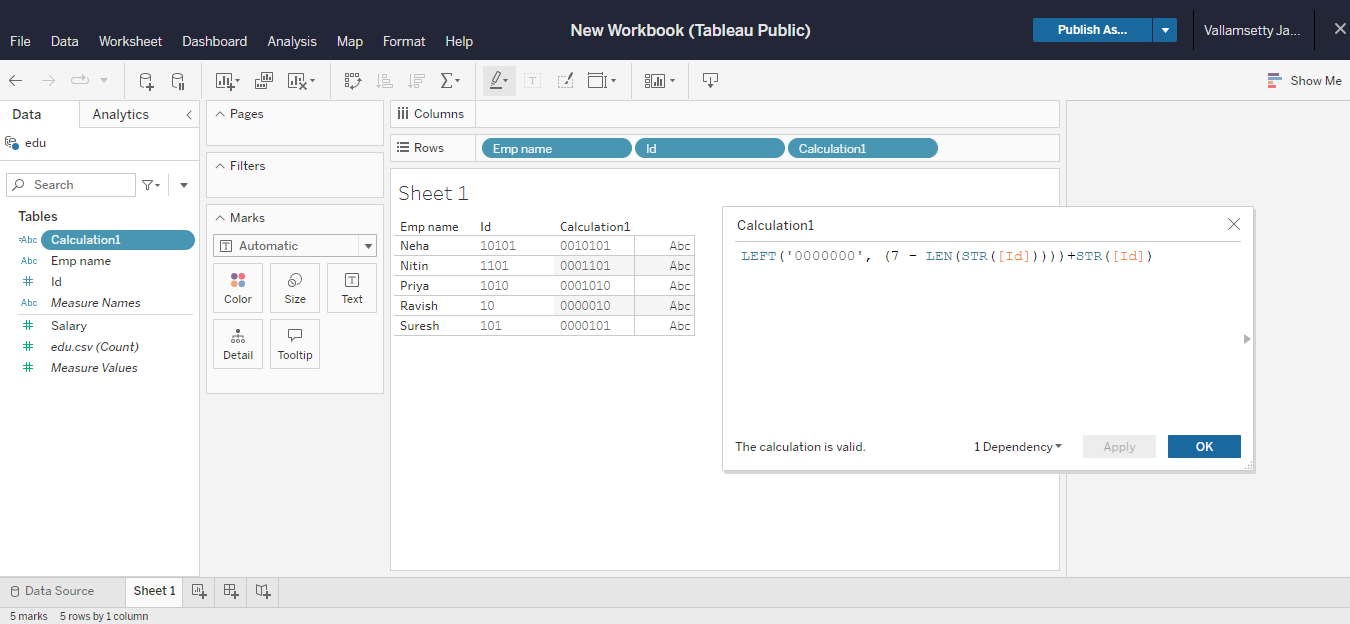
Add all rows in rows shown column.

**STEP4:**

Go to analysis then click on create Calculation field to write the code.

**LEFT('0000000', (7 - LEN(STR([Id]))))+STR([Id])**

After writing the required code apply it and click ok. So the required solution will be there.



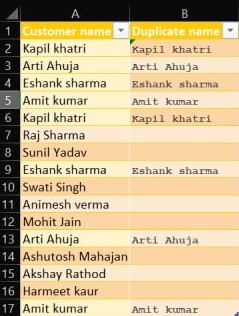
1. **EXCEL**

The given data must be entered in excel sheet. There must duplicate field that occurs more than once. To get given output for the following input.

Enter the formula:

**=IF(COUNT($A2:$A$17,$A2)>1,A2,””)**

Here is the output –



1. **Machine Learning**

**STEP1:**

Download the given datasets. Then import them in the google colab. we have to read the CSV files with the following code-

import pandas as pd

train = pd.read\_csv("/content/test.csv")

test = pd.read\_csv("/content/train.csv")

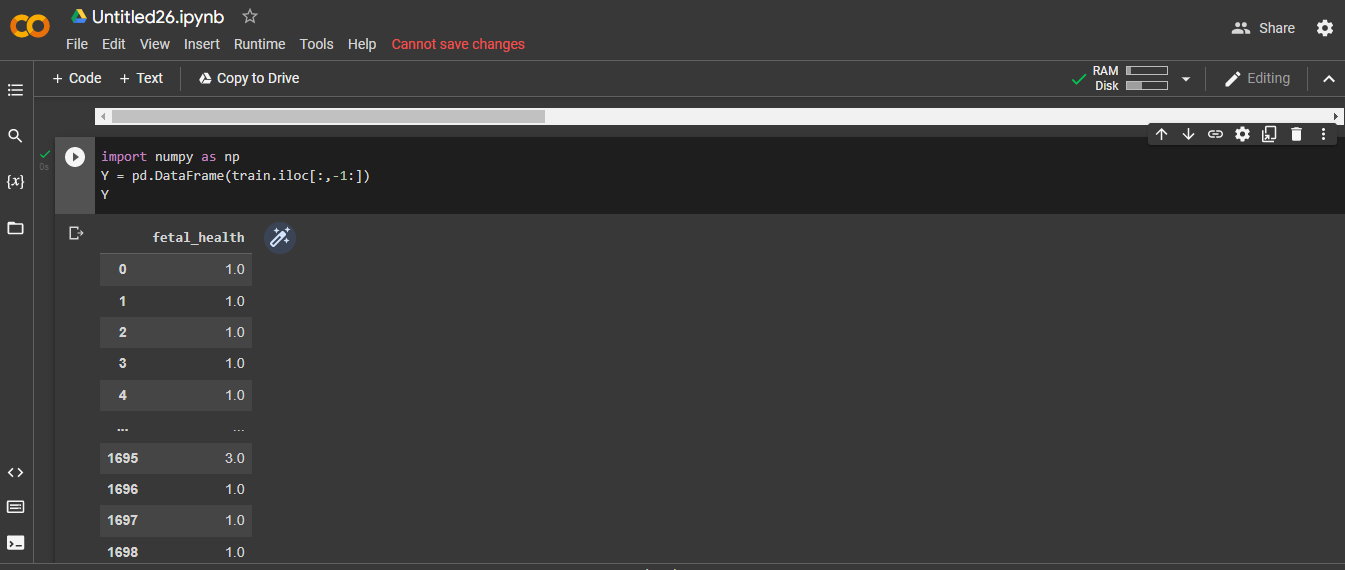
test

train

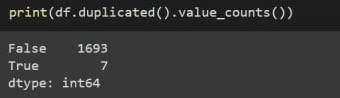
DATASETS WILL BE PRINTED.

**STEP2:**

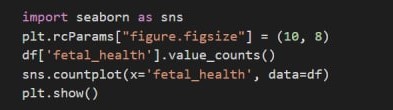
To print the last column of train dataset the following code must be written-



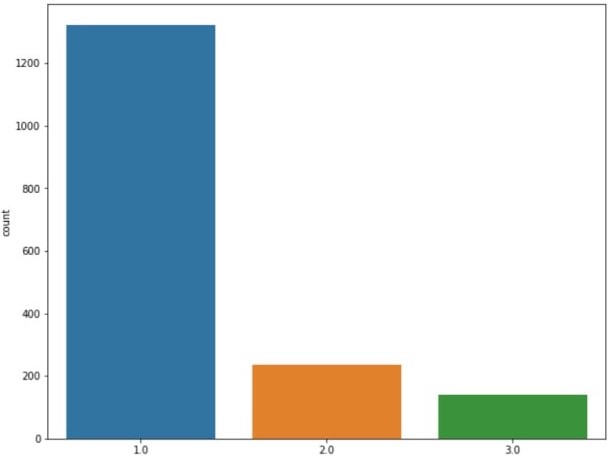
**STEP3:** We have to remove duplicates from the dataset to find approximate accuracy of a given dataset.

****

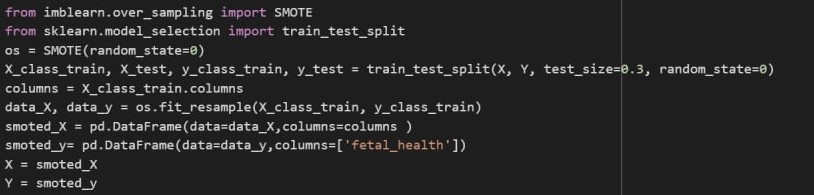
**STEP4:** There is an imbalance in fetal\_health**.**

****

**OUPUT FOR THE ABOVE CODE:**

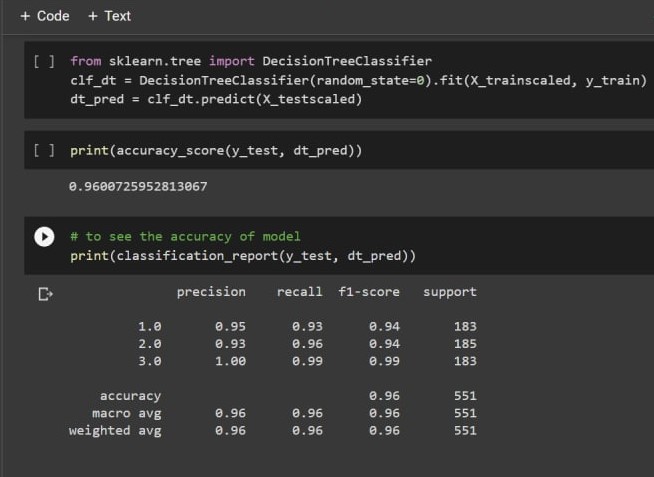
****

**STEP5:** We have to smote the fetal\_health column which has 1.0, 2.0, 3.0 imbalance by doing this, it will get all the values count will be equal.



**STEP5:** To find the accuracy of the dataset there are some ways by using machine learning algorithms. I used Decision tree algorithm to find it.

We can see from the below picture that the accuracy of the given dataset is above 75% which is 0.96.

****