



# Cognizance Task - 6

**Name:** Kumar Laxmikant

**Roll No:** CH.EN.U4CSE20039

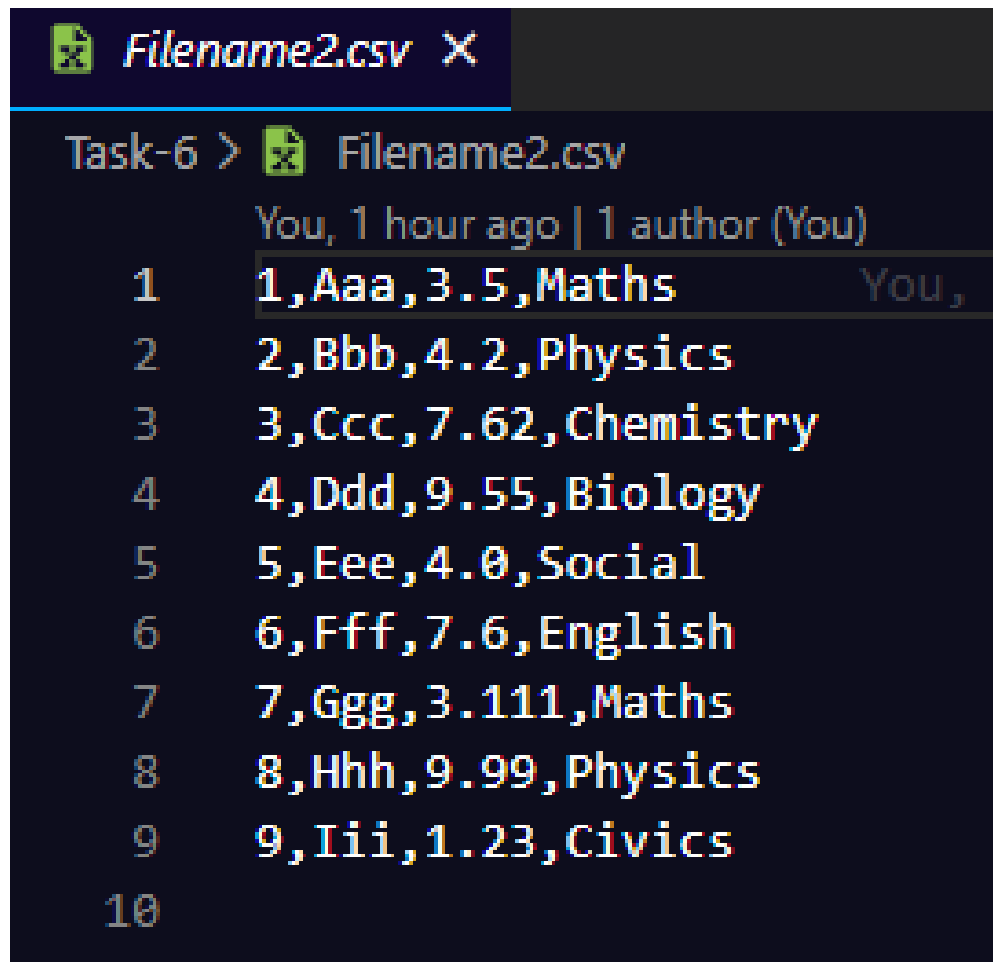
**Batch:** 2020-24

**Department:** CSE-A

## Question-1

(Screenshot of Result):

(OUTPUT File is uploaded to GitHub Repository)



The screenshot shows a file viewer interface. At the top, there is a tab labeled 'Filename2.csv' with a close button. Below the tab, the file content is displayed in a dark-themed editor. The content is a CSV file with 10 rows of data. The first row is highlighted. The data is as follows:

1	1,Aaa,3.5,Maths	You,		
2	2,Bbb,4.2,Physics			
3	3,Ccc,7.62,Chemistry			
4	4,Ddd,9.55,Biology			
5	5,Eee,4.0,Social			
6	6,Fff,7.6,English			
7	7,Ggg,3.111,Maths			
8	8,Hhh,9.99,Physics			
9	9,Iii,1.23,Civics			
10				

## Question-2

### (Screenshot of Result):

The percentage of NaN values in columns with NaN values:

```
LotFrontage    14.141414
Alley          93.939394
BsmtQual       3.030303
BsmtCond       3.030303
BsmtExposure   3.030303
BsmtFinType1   3.030303
BsmtFinType2   3.030303
dtype: float64
```

The 'Alley' column has too many NaN values so such columns would be inefficient for the model to train on. After dropping 'Alley' column, the percentage of NaN values in columns with NaN values:

```
LotFrontage    14.141414
BsmtQual       3.030303
BsmtCond       3.030303
BsmtExposure   3.030303
BsmtFinType1   3.030303
BsmtFinType2   3.030303
dtype: float64
```

Checking if any NaN entry is left in any column:

```
MSSubClass      False
MSZoning         False
LotFrontage     False
LotArea         False
Street          False
LotShape        False
LandContour     False
Utilities       False
LotConfig       False
LandSlope       False
Neighborhood    False
Condition1      False
Condition2      False
BldgType        False
HouseStyle      False
OverallQual     False
OverallCond     False
YearBuilt       False
YearRemodAdd    False
RoofStyle       False
RoofMatl        False
Exterior1st     False
Exterior2nd     False
MasVnrType      False
MasVnrArea      False
ExterQual       False
ExterCond       False
Foundation      False
BsmtQual        False
BsmtCond        False
BsmtExposure    False
BsmtFinType1    False
BsmtFinSF1      False
BsmtFinType2    False
dtype: bool
```

### Question-3

(Screenshot of Result):

```
The words with atleast 6 letters:
```

```
Python  
almost  
aspect  
scientific  
computing  
America  
Python  
crunch  
financial  
Facebook  
Python  
library  
Pandas  
analysis  
libraries  
available  
perform  
analysis  
Python  
Pandas  
Matplotlib
```

```
The most frequent word is: Python
```