|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Explaining | Talk detail to programmer | Talk to team | Talk to client |
| names = ['Jessica','John','Bob','Jessica','Mary','John','Mel','Mel']  grades = [95,78,76,95,77,78,99,100]  GradeList = list(zip(names,grades))  df = pd.DataFrame(data = GradeList,  columns=['Names', 'Grades'])  df |  | Create a list of students name, and a list of their grades.  Then zip those lists into one list, and use DataFrame function in Pandas to make it into a dataframe name df, naming column 1 ‘Names’, and column 2 ‘Grades’. | Create a name list, and a grade list.  Then zip them together,  Then create dataframe with columns name = Names and Grades | Create a table that contain students name and grades |
| dupe = df.duplicated() #creates list of True/False values |  | Use duplicate function to create a list of duplicate row in df dataframe and store it in variable ‘dupe’ | Create list of duplicate data and name it ‘dupe’ | Create another database/table for duplicate data |
| df[dupe] |  | Show rows of duplicate data | Show rows of duplicate data | Show rows of duplicate data |
| nodupedf = df.drop\_duplicates()  nodupedf.count() | Pass the drop\_duplicates function to the dataframe and call it nodupedf  Then verify the work by counting the value in the new dataframe | Use drop\_duplicates() functions to drop duplicates value from | Drop duplicate value and count the data after that to see if the number of data decrease |  |
| Passing argument into csv function |  |  |  |  |