In [50]:
```
#Task-1 Making the three different dataframes in python and saving them as .cs
```

In [129…
```
import pandas as pd    # Importing the pandas as 'pd' to read the tables
employee_data = {
    "ID": ["A001", "A002", "A003", "A004", "A005"],           # updating the a
    "Name": ["John Alter", "Alice Luxumberg", "Tom Sabestine", "Nina Adgra", "
    "Gender": ["M", "F", "M", "F", "F"],
    "City": ["Paris", "London", "Berlin", "Newyork", "Madrid"],
    "Age": [25, 27, 29, 31, 30]
}
employee = pd.DataFrame(employee_data) # saving the table as employee data


employee.to_csv("employee.csv", index=False) # saving the table as csv
```

In [130…
```
seniority_data = {
    "ID": ["A001", "A002", "A003", "A004", "A005"],  # updating the all the ir
    "Designation_Level": [2, 2, 3, 2, 3]
}
seniority = pd.DataFrame(seniority_data) # saving the table as seniority
seniority.to_csv("seniority.csv", index=False) # saving the table as csv
```

In [131…
```
project_data = {
    "ID": ["A001","A002","A003","A004","A005","A002","A005","A003","A001","A00
    "Project": [
        "Project 1","Project 2","Project 3","Project 4","Project 5",
        "Project 6","Project 7","Project 8","Project 9","Project 10",
        "Project 11","Project 12","Project 13","Project 14"  # updating the "P
    ],
    "Cost": [
        1002000,2000000,4500000,5500000,None,
        680000,400000,350000,None,300000,
        2000000,1000000,3000000,200000   # updating the "Cost" information
    ],
    "Status": [
        "Finished","Ongoing","Finished","Ongoing","Finished",
        "Failed","Finished","Failed","Ongoing","Finished",
        "Failed","Ongoing","Finished","Finished"  # updating the "Status" info
    ]
}

project = pd.DataFrame(project_data) # saving the table as Project_data

project.to_csv("project.csv", index=False) # saving the table as csv
```

In [132…
```
pd.set_option("display.max_rows", None)  # displaying the all the "Columns" an
pd.set_option("display.max_columns", None)


print("Employee DataFrame:\n", employee.to_string(index=False))
print("\nSeniority DataFrame:\n", seniority.to_string(index=False))
print("\nProject DataFrame (with missing values):\n", project.to_string(index=
```

```
Employee DataFrame:
   ID            Name Gender    City  Age
A001      John Alter      M   Paris   25
A002 Alice Luxumberg      F  London   27
A003   Tom Sabestine      M  Berlin   29
A004      Nina Adgra      F Newyork   31
A005       Amy Johny      F  Madrid   30

Seniority DataFrame:
   ID  Designation_Level
A001                  2
A002                  2
A003                  3
A004                  2
A005                  3

Project DataFrame (with missing values):
   ID    Project    Cost   Status
A001  Project 1 1002000 Finished
A002  Project 2 2000000  Ongoing
A003  Project 3 4500000 Finished
A004  Project 4 5500000  Ongoing
A005  Project 5     NaN Finished
A002  Project 6  680000   Failed
A005  Project 7  400000 Finished
A003  Project 8  350000   Failed
A001  Project 9     NaN  Ongoing
A003 Project 10  300000 Finished
A001 Project 11 2000000   Failed
A004 Project 12 1000000  Ongoing
A004 Project 13 3000000 Finished
A005 Project 14  200000 Finished
```

In [85]:
```
# Task-2 Finding the Missing values for the cost column in the dataframe " pro
```

In [134…
```
import numpy as np   # importing numpy to use numerical values


project = pd.read_csv("project.csv")


running_sum = 0
count = 0
```

In [135…
```
for i in range(len(project)):
    if pd.isna(project.loc[i, "Cost"]):
        if count > 0:
            project.loc[i, "Cost"] = running_sum / count   # calculate the run
            running_sum += project.loc[i, "Cost"]
            count += 1
    else:
        running_sum += project.loc[i, "Cost"]  # if the 'Cost' is not missing,
        count += 1  # increment the count for the number of valid costs encoun
```

In [136…
```
project.to_csv("project_updated.csv", index=False) # saving the updated file a


pd.set_option("display.max_rows", None) # displaying all the rows
pd.set_option("display.max_columns", None) # displaying all the columns
pd.set_option("display.float_format", '{:.1f}'.format)

print("\n Updated project Dataframe:\n")
print(project.to_string(index=False)) # dispaying the project dataframes with
```

```
Updated project Dataframe:

  ID    Project      Cost   Status
A001  Project 1 1002000.0 Finished
A002  Project 2 2000000.0  Ongoing
A003  Project 3 4500000.0 Finished
A004  Project 4 5500000.0  Ongoing
A005  Project 5 3250500.0 Finished
A002  Project 6  680000.0   Failed
A005  Project 7  400000.0 Finished
A003  Project 8  350000.0   Failed
A001  Project 9 2210312.5  Ongoing
A003 Project 10  300000.0 Finished
A001 Project 11 2000000.0   Failed
A004 Project 12 1000000.0  Ongoing
A004 Project 13 3000000.0 Finished
A005 Project 14  200000.0 Finished
```

In [89]:
```
# Task-3 splitting the name column in Employe dataframe into two individual co
```

In [138…
```
employee = pd.read_csv("employee.csv")


employee[["First Name", "Last Name"]] = employee["Name"].str.split(" ", n=1, e


employee = employee.drop(columns=["Name"]) # dropping the name column


employee.to_csv("employee_updated.csv", index=False)

print("\n Employee table updated:\n")
print(employee.to_string(index=False)) # displaying employee table with  "Firs
```

```
Employee table updated:

  ID Gender    City  Age First Name Last Name
A001      M   Paris   25       John     Alter
A002      F  London   27      Alice Luxumberg
A003      M  Berlin   29        Tom Sabestine
A004      F Newyork   31       Nina     Adgra
A005      F  Madrid   30        Amy     Johny
```

In [139…
```
# Task-4 Naming the table with "Final" and joining all three dataframes in a s
```

In [140…
```
employee = pd.read_csv("employee_updated.csv")
seniority = pd.read_csv("seniority.csv")
project = pd.read_csv("project_updated.csv")

merged1 = pd.merge(employee, seniority, on="ID", how="inner") # merging employ
Final = pd.merge(merged1, project, on="ID", how="inner") # merging project dat

Final.to_csv("final.csv", index=False) # saving merged file as "Final"

pd.set_option("display.max_rows", None)
pd.set_option("display.max_columns", None)
pd.set_option("display.float_format", '{:.0f}'.format)

print("\nFinal:\n")
print(Final.to_string(index=False)) # displaying Combined one dataframe
```

```
Final:

  ID Gender    City  Age First Name Last Name  Designation_Level    Project
Cost   Status
A001      M   Paris   25       John     Alter                  2  Project 1 100
2000 Finished
A001      M   Paris   25       John     Alter                  2  Project 9 221
0312  Ongoing
A001      M   Paris   25       John     Alter                  2 Project 11 200
0000   Failed
A002      F  London   27      Alice Luxumberg                  2  Project 2 200
0000  Ongoing
A002      F  London   27      Alice Luxumberg                  2  Project 6  68
0000   Failed
A003      M  Berlin   29        Tom Sabestine                  3  Project 3 450
0000 Finished
A003      M  Berlin   29        Tom Sabestine                  3  Project 8  35
0000   Failed
A003      M  Berlin   29        Tom Sabestine                  3 Project 10  30
0000 Finished
A004      F Newyork   31       Nina     Adgra                  2  Project 4 550
0000  Ongoing
A004      F Newyork   31       Nina     Adgra                  2 Project 12 100
0000  Ongoing
A004      F Newyork   31       Nina     Adgra                  2 Project 13 300
0000 Finished
A005      F  Madrid   30        Amy     Johny                  3  Project 5 325
0500 Finished
A005      F  Madrid   30        Amy     Johny                  3  Project 7  40
0000 Finished
A005      F  Madrid   30        Amy     Johny                  3 Project 14  20
0000 Finished
```

In [113…
```
# Task-5 Adding the 5% bonus to the employees who finished the project and add
```

In [141…
```
Final = pd.read_csv("final.csv")
```

```
Final["Bonus"] = 0

for i in range(len(Final)):
    if Final.loc[i, "Status"] == "Finished": # finding out the status finished
        Final.loc[i, "Bonus"] = Final.loc[i, "Cost"] * 0.05   # adding the 5%

Final.to_csv("final_updated.csv", index=False)

pd.set_option("display.max_rows", None)
pd.set_option("display.max_columns", None)
pd.set_option("display.float_format", '{:.0f}'.format)

print("\n Final DataFrame with Bonus column:\n")
print(Final.to_string(index=False))  # displaying the updated dataframe with b
```

```
Final DataFrame with Bonus column:

  ID Gender    City  Age First Name Last Name  Designation_Level    Project
Cost   Status  Bonus
A001      M   Paris   25       John     Alter                  2  Project 1 100
2000 Finished  50100
A001      M   Paris   25       John     Alter                  2  Project 9 221
0312  Ongoing      0
A001      M   Paris   25       John     Alter                  2 Project 11 200
0000   Failed      0
A002      F  London   27      Alice Luxumberg                  2  Project 2 200
0000  Ongoing      0
A002      F  London   27      Alice Luxumberg                  2  Project 6  68
0000   Failed      0
A003      M  Berlin   29        Tom Sabestine                  3  Project 3 450
0000 Finished 225000
A003      M  Berlin   29        Tom Sabestine                  3  Project 8  35
0000   Failed      0
A003      M  Berlin   29        Tom Sabestine                  3 Project 10  30
0000 Finished  15000
A004      F Newyork   31       Nina     Adgra                  2  Project 4 550
0000  Ongoing      0
A004      F Newyork   31       Nina     Adgra                  2 Project 12 100
0000  Ongoing      0
A004      F Newyork   31       Nina     Adgra                  2 Project 13 300
0000 Finished 150000
A005      F  Madrid   30        Amy     Johny                  3  Project 5 325
0500 Finished 162525
A005      F  Madrid   30        Amy     Johny                  3  Project 7  40
0000 Finished  20000
A005      F  Madrid   30        Amy     Johny                  3 Project 14  20
0000 Finished  10000
```

In [95]:
```
# Task-6 Decreasing the designation level by 1, whose projects have " Failed"
```

In [142…
```
Final = pd.read_csv("final_updated.csv")


for i in range(len(Final)):
```

```
    if Final.loc[i, "Status"] == "Failed":  # finding out the status failed
        if Final.loc[i, "Designation_Level"] > 1:   # demoting the designation
            Final.loc[i, "Designation_Level"] -= 1

Final = Final[Final["Designation_Level"] <= 4] # deleting the designation leve


Final.to_csv("final_updated_task6.csv", index=False)


pd.set_option("display.max_rows", None)
pd.set_option("display.max_columns", None)
pd.set_option("display.float_format", '{:.0f}'.format)

print("\n Final DataFrame after demotions:\n")
print(Final.to_string(index=False)) # displaying the updated dataframe with de
```

```
Final DataFrame after demotions:

  ID Gender    City  Age First Name Last Name  Designation_Level    Project
Cost   Status  Bonus
A001      M   Paris   25       John     Alter                  2  Project 1 100
2000 Finished  50100
A001      M   Paris   25       John     Alter                  2  Project 9 221
0312  Ongoing      0
A001      M   Paris   25       John     Alter                  1 Project 11 200
0000   Failed      0
A002      F  London   27      Alice Luxumberg                  2  Project 2 200
0000  Ongoing      0
A002      F  London   27      Alice Luxumberg                  1  Project 6  68
0000   Failed      0
A003      M  Berlin   29        Tom Sabestine                  3  Project 3 450
0000 Finished 225000
A003      M  Berlin   29        Tom Sabestine                  2  Project 8  35
0000   Failed      0
A003      M  Berlin   29        Tom Sabestine                  3 Project 10  30
0000 Finished  15000
A004      F Newyork   31       Nina     Adgra                  2  Project 4 550
0000  Ongoing      0
A004      F Newyork   31       Nina     Adgra                  2 Project 12 100
0000  Ongoing      0
A004      F Newyork   31       Nina     Adgra                  2 Project 13 300
0000 Finished 150000
A005      F  Madrid   30        Amy     Johny                  3  Project 5 325
0500 Finished 162525
A005      F  Madrid   30        Amy     Johny                  3  Project 7  40
0000 Finished  20000
A005      F  Madrid   30        Amy     Johny                  3 Project 14  20
0000 Finished  10000
```

In [115…
```
# Task-7  Drop the Gender column and adding "Mr" and "Mrs" to first name.
```

In [143…
```
Final = pd.read_csv("final_updated_task6.csv")
```

```python
for i in range(len(Final)):
    if Final.loc[i, "Gender"] == "M":
        Final.loc[i, "First Name"] = "Mr. " + Final.loc[i, "First Name"] # add
    elif Final.loc[i, "Gender"] == "F":
        Final.loc[i, "First Name"] = "Mrs. " + Final.loc[i, "First Name"] # ad


Final = Final.drop(columns=["Gender"])  # dropping the gender column


Final.to_csv("final_updated_task7.csv", index=False)


pd.set_option("display.max_rows", None)
pd.set_option("display.max_columns", None)
pd.set_option("display.float_format", '{:.0f}'.format)

print("\n Final DataFrame :\n")
print(Final.to_string(index=False)) # displaying the updated dataframe removin
```

```
Final DataFrame :

  ID    City  Age First Name Last Name  Designation_Level    Project    Cost
Status  Bonus
A001   Paris   25   Mr. John     Alter                  2  Project 1 1002000 Fi
nished  50100
A001   Paris   25   Mr. John     Alter                  2  Project 9 2210312  O
ngoing      0
A001   Paris   25   Mr. John     Alter                  1 Project 11 2000000
Failed      0
A002  London   27 Mrs. Alice Luxumberg                  2  Project 2 2000000  O
ngoing      0
A002  London   27 Mrs. Alice Luxumberg                  1  Project 6  680000
Failed      0
A003  Berlin   29    Mr. Tom Sabestine                  3  Project 3 4500000 Fi
nished 225000
A003  Berlin   29    Mr. Tom Sabestine                  2  Project 8  350000
Failed      0
A003  Berlin   29    Mr. Tom Sabestine                  3 Project 10  300000 Fi
nished  15000
A004 Newyork   31  Mrs. Nina     Adgra                  2  Project 4 5500000  O
ngoing      0
A004 Newyork   31  Mrs. Nina     Adgra                  2 Project 12 1000000  O
ngoing      0
A004 Newyork   31  Mrs. Nina     Adgra                  2 Project 13 3000000 Fi
nished 150000
A005  Madrid   30   Mrs. Amy     Johny                  3  Project 5 3250500 Fi
nished 162525
A005  Madrid   30   Mrs. Amy     Johny                  3  Project 7  400000 Fi
nished  20000
A005  Madrid   30   Mrs. Amy     Johny                  3 Project 14  200000 Fi
nished  10000
```

In [119…
```
# Task-8  promoting the designation level by 1 for the employees who age is ab
```

In [144…
```
Final = pd.read_csv("final_updated_task7.csv")


for i in range(len(Final)):

    if Final.loc[i, "Age"] > 29:       # finding out the employees whose age i

        Final.loc[i, "Designation_Level"] += 1  # adding the designation level

Final.to_csv("final_updated_task8.csv", index=False)

pd.set_option("display.max_rows", None)
pd.set_option("display.max_columns", None)
pd.set_option("display.float_format", '{:.0f}'.format)

print("\nFinal DataFrame after promotions:\n")
print(Final.to_string(index=False)) # displaying the updated dataframe
```

```
Final DataFrame after promotions:

  ID    City  Age First Name Last Name  Designation_Level    Project    Cost
Status  Bonus
A001   Paris   25   Mr. John     Alter                  2  Project 1 1002000 Fi
nished  50100
A001   Paris   25   Mr. John     Alter                  2  Project 9 2210312  O
ngoing      0
A001   Paris   25   Mr. John     Alter                  1 Project 11 2000000
Failed      0
A002  London   27 Mrs. Alice Luxumberg                  2  Project 2 2000000  O
ngoing      0
A002  London   27 Mrs. Alice Luxumberg                  1  Project 6  680000
Failed      0
A003  Berlin   29    Mr. Tom Sabestine                  3  Project 3 4500000 Fi
nished 225000
A003  Berlin   29    Mr. Tom Sabestine                  2  Project 8  350000
Failed      0
A003  Berlin   29    Mr. Tom Sabestine                  3 Project 10  300000 Fi
nished  15000
A004 Newyork   31  Mrs. Nina     Adgra                  3  Project 4 5500000  O
ngoing      0
A004 Newyork   31  Mrs. Nina     Adgra                  3 Project 12 1000000  O
ngoing      0
A004 Newyork   31  Mrs. Nina     Adgra                  3 Project 13 3000000 Fi
nished 150000
A005  Madrid   30   Mrs. Amy     Johny                  4  Project 5 3250500 Fi
nished 162525
A005  Madrid   30   Mrs. Amy     Johny                  4  Project 7  400000 Fi
nished  20000
A005  Madrid   30   Mrs. Amy     Johny                  4 Project 14  200000 Fi
nished  10000
```

In [101…
```
# Task-9 Adding the project cost for all employees and nameing it has "TotalPr
```

In [145…
```
Final = pd.read_csv("final_updated_task8.csv")


TotalProjCost = Final.groupby(["ID", "First Name"])["Cost"].sum().reset_index(


TotalProjCost = TotalProjCost.rename(columns={"Cost": "Total Cost"})


TotalProjCost.to_csv("total_project_cost.csv", index=False)


pd.set_option("display.max_rows", None)
pd.set_option("display.max_columns", None)
pd.set_option("display.float_format", '{:.0f}'.format)

print("\n Total project cost per employee:\n")
print(TotalProjCost.to_string(index=False)) # displacing the new dataframe
```

```
Total project cost per employee:

  ID First Name  Total Cost
A001   Mr. John     5212312
A002 Mrs. Alice     2680000
A003    Mr. Tom     5150000
A004  Mrs. Nina     9500000
A005   Mrs. Amy     3850500
```

In [103…
```
#Task-10 Printing the employee details whose city name have "o".
```

In [146…
```
Final = pd.read_csv("final_updated_task8.csv")


filtered = Final[Final["City"].str.contains("o", case=False, na=False)]  # fil


pd.set_option("display.max_rows", None)
pd.set_option("display.max_columns", None)
pd.set_option("display.float_format", '{:.0f}'.format)

print("\nEmployees from cities containing 'o':\n")
print(filtered.to_string(index=False)) # displaying all the employee details w
```

```
Employees from cities containing 'o':

  ID    City  Age First Name Last Name  Designation_Level    Project    Cost
Status  Bonus
A002  London   27 Mrs. Alice Luxumberg                  2  Project 2 2000000  O
ngoing      0
A002  London   27 Mrs. Alice Luxumberg                  1  Project 6  680000
Failed      0
A004 Newyork   31  Mrs. Nina     Adgra                  3  Project 4 5500000  O
ngoing      0
A004 Newyork   31  Mrs. Nina     Adgra                  3 Project 12 1000000  O
ngoing      0
A004 Newyork   31  Mrs. Nina     Adgra                  3 Project 13 3000000 Fi
nished 150000
```

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]: