

In [21]: *#1.a. Dataframe creation and basic operation*

```
import pandas as pd
data = {
    "Employee": ["John", "Alice", "Bob", "Emma"],
    "Department": ["IT", "HR", "Finance", "IT"],
    "Salary": [60000, 55000, 70000, 72000],
    "Age": [30, 28, 35, 32]
}

df = pd.DataFrame(data)
print(df.head(2))
df["Experience"] = [5, 3, 7, 6]

average_salary = df["Salary"].mean()
print(df)
print(f"Average Salary: {average_salary}")
```

	Employee	Department	Salary	Age
0	John	IT	60000	30
1	Alice	HR	55000	28

	Employee	Department	Salary	Age	Experience
0	John	IT	60000	30	5
1	Alice	HR	55000	28	3
2	Bob	Finance	70000	35	7
3	Emma	IT	72000	32	6

Average Salary: 64250.0

In [23]: *#1.b. Create a dataset of students with name and 3 subject*

```
import pandas as pd

students = {
    "Name" :["Amit", "Neha", "Rohan", "Priya", "Vikas"],
    "Math": [85, 78, 92, 88, 76],
    "Science": [80, 89, 95, 91, 82],
    "English": [75, 83, 88, 79, 85]
}

df_students = pd.DataFrame(students)

high_math_scores = df_students[df_students["Math"] > 80]
print(high_math_scores)

sorted_science = df_students.sort_values(by = "Science", ascending = False)
print(sorted_science)

top_english_student = df_students[df_students["English"] == df_students["English"].max()]
print(top_english_student)
```

	Name	Math	Science	English
0	Amit	85	80	75
2	Rohan	92	95	88
3	Priya	88	91	79

	Name	Math	Science	English
2	Rohan	92	95	88
3	Priya	88	91	79
1	Neha	78	89	83
4	vikas	76	82	85
0	Amit	85	80	75

	Name	Math	Science	English
2	Rohan	92	95	88

In [29]: #2. Household expenses for a month(panda series)

```
import pandas as pd

categories = [ 'Groceries', 'Utilities', 'Rent', 'Transportation', 'Enterainment']

expenses = [500, 200, 1200, 300, 150]

expense_series=pd.Series(expenses, index=categories)

print(expense_series)
```

```
Groceries      500
Utilities      200
Rent           1200
Transportation  300
Enterainment   150
dtype: int64
```

In [32]: #3. Monthly energy consumption(panda series)

```
import pandas as pd
import matplotlib.pyplot as plt

months = ['January', 'February', 'March', 'April', 'May', 'June', 'July', 'August',
electricity_usage = [ 350, 320, 310, 330, 340, 370, 380, 360, 350, 330, 320, 330]
gas_usage = [20, 18, 16, 15, 12, 10, 8, 9, 12, 15, 17, 19]

electricity_series = pd.Series(electricity_usage, index=months, name="Electricity u
gas_series = pd.Series(gas_usage, index=months, name="Gas usage(therms)")

print(electricity_series)

print(gas_series)
```

```

January      350
Februaury    320
March        310
April        330
May          340
June         370
July         380
August       360
September    350
October      330
November     320
December     330
Name: Electricity usage(kwh), dtype: int64
January      20
Februaury    18
March        16
April        15
May          12
June         10
July          8
August        9
September    12
October      15
November     17
December     19
Name: Gas usage(therms), dtype: int64

```

```

In [35]: #4. Monthly revenue generated from advertising(panda series)

import pandas as pd
import matplotlib.pyplot as plt

months = ['January', 'February', 'March', 'April', 'May', 'June', 'July', 'August',
'September', 'October', 'November', 'December']

revenue = [5000, 5200, 4800, 5400, 5600, 5800, 6100, 5900, 6200, 6500, 7000, 6900]

revenue_series = pd.Series(revenue, index=months, name="Monthly Revenue Advertising")

print(revenue_series)

#Total revenue for the year
total_revenue= revenue_series.sum()
print(f"Total Revenue for the year: ${total_revenue}")

```

January	5000
February	5200
March	4800
April	5400
May	5600
June	5800
July	6100
August	5900
September	6200
October	6500
November	7000
December	6900

Name: Monthly Revenue Advertising, dtype: int64

Total Revenue for the year: \$70400

In []: