practical 4

Name – Kamlesh Pawar

Roll number - 379

PRN no - 202201070138

[75]:

[76]:
```
import pandas as pd
```

[77]:
```
df=pd.read_csv("grainsales.csv")
```

[78]:
```
df
```

[78]:
[ ]:

	GrainName	State	City	Months	Year	Sales
0	Ragi	Maharashtra	Nagpur	JAN	2023	1000000
1	Bajra	Panjab	Amritsar	FEB	2023	1500000
2	Ragi	Maharashtra	Nagpur	JAN	2023	1000000
3	Bajra	Panjab	Amritsar	FEB	2023	1500000
4	Ragi	Maharashtra	Nagpur	JAN	2023	1000000
5	Bajra	Panjab	Amritsar	FEB	2023	1500000
6	Oats	Hariyana	Gurugram	MARCH	2023	2000000
7	Sattu	Gujarat	Surat	APRIL	2023	2500000
8	Sooji	Tamil Nadu	Madurai	MAY	2023	3000000
9	Brown rice	Telangana	Hyderabad	JUNE	2023	3500000
10	Wheat	West Bengol	Asansole	JULY	2023	4000000
11	Corn	UP	Kanpur	AUG	2023	4500000
12	Ragi	Maharashtra	Nagpur	JAN	2023	1000000
13	Bajra	Panjab	Amritsar	FEB	2023	1500000
14						

```
          Oats Hariyana Gurugram MARCH 2023 2000000
15 Sattu Gujarat Surat APRIL 2023 2500000 16 Sooji Tamil
Nadu Madurai MAY 2023 3500000
   17 Brown rice Telangana Hyderabad JUNE 2023
3500000 18 Wheat West Bengol      Asansole JULY 2023
                        4000000
19        Corn            UP Kanpur AUG 2023 4500000
```

[ ]:

1 Q1. Which was the best month for sales? How much was earned that month?

[79]:
```
df1=df groupby([ Months ]) max( Sales )
df1
```

[79]:
```
        Year Sales
Months
APRIL 2023 2500000
 AUG  2023 4500000
 FEB  2023 1500000
 JAN  2023 1000000
JULY  2023 4000000
 JUNE   2023 3500000
 MARCH  2023 2000000
 MAY    2023  3500000
```

[ ]:

[80]:
```
df1=df groupby(["Months"])[["Sales"]].sum()
df1
```

[80]:
```
          Sales
 Months
 APRIL  5000000
 AUG  9000000
 FEB  6000000
 JAN  4000000
JULY  8000000
JUNE  7000000
MARCH 4000000
 MAY 6500000
```

[81]:
```
df1=df groupby([ Months'],sort  False)[["Sales"]].sum()
max1   df1["Sales"].max()
df1[df1["Sales"]==max1]
```

[81]:

Months	Sales
AUG	9000000

2 Q2. Which product sold the most? Why do you think it did?

[82]:
```
df2=df groupby(['GrainName'],sort  False)[["Sales"]].sum()
max1   df2["Sales"].max()
df2[df2["Sales"]==max1]
```

[82]:

GrainName	Sales
Corn	9000000

3 Q3. Which city sold the most products?

[83]:
```
df2=df groupby(['City'],sort  False)[["Sales"]].sum()
max1   df2["Sales"].max()
df2[df2["Sales"]==max1]
```

[83]:

City	Sales
Kanpur	9000000

4 Q4. What Products are most often sold together?

[84]:
```
import pandas as pd
from itertools import combinations
from collections import Counter
```

[85]:
```
product_combinations   df.groupby( Months )['GrainName ].apply(lambda x:␣
 ↪list(combinations(x, 2))).tolist()
all_combinations   [item for sublist in product_combinations for item in␣
 ↪sublist]
```

[86]:
```
combination_counts   Counter(all_combinations)
sorted_combinations   sorted(combination_counts.items(), key=lambda x: x[1] ␣
 ↪reverse=True)
```

[87]:
```
print("Most often sold together products: )
for combination, count in sorted_combinations:
    print(combination, "-"  count)
```

```
Most often sold together products:
 ('Bajra', 'Bajra') - 6
 ('Ragi', 'Ragi') - 6
('Sattu ', 'Sattu ') - 1
('Corn', 'Corn') - 1
('Wheat', 'Wheat') - 1
('Brown rice ', 'Brown rice ') - 1
('Oats', 'Oats') - 1
('Sooji', 'Sooji') - 1
```