Assignment no : 4

Name : Yukti Wagare

Div : G (G4)

Roll No : 768

PRN No : 202201050035

Problem Statement :-

   Read any real-life dataset. Store the data in Data Frames. Identify 10 grains for the given dataset.

Implement all 20 grains using Pandas methods. The Sample Grains for the Sales Dataset are as:

1. Which was the best month for sales? How much was earned that month?
2. Which product sold the most? Why do you think it did?
3. Which city sold the most products?
4. What Products are most often sold together?

Solution :-

```
import pandas as pd
df=pd.read_csv("/content/sample_data/Grainsales.csv")
df
```

	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	3000000
17	Brown rice	Telangana	Hyderabad	JUNE	2023	3500000
18	Wheat	West Bengol	Asansole	JULY	2023	4000000
19	Corn	UP	Kanpur	AUG	2023	4500000

```python
# Best month for sales? How much was earned that month?
df1=df.groupby(['Months'],sort=False)[["Sales"]].sum()
max1=df1["Sales"].max()
df1[df1["Sales"]==max1]
```

	Sales
Months	
JULY	16000000

```python
# Product sold the most?
df1=df.groupby(['GrainName'],sort=False)[["Sales"]].sum()
max1=df1["Sales"].max()
df1[df1["Sales"]==max1]
# Why do you think it did?
Answer : Wheat is the staple food of India. It is sold mostly in the
months of June and July .
```

	Sales
GrainName	
Wheat	16000000

```
#City sold the most products?
df1=df.groupby(['City'],sort=False)[["GrainName"]].sum()
max1=df1["GrainName"].max()
df1[df1["GrainName"]==max1]
```

City	
Asansole	WheatWheatWheatWheat

```
#Products are most often sold together?
order_products = df[['Sales', 'GrainName']]
order_products =
order_products.groupby('Sales')['GrainName'].apply(list)
product_combinations = order_products.apply(lambda x:
pd.Series(x).value_counts()).fillna(0)
most_common_combinations = product_combinations.sum().nlargest(10)
print("The most often sold products together are:")
print(most_common_combinations)
```

```
The most often sold products together are:
Ragi          5.0
Bajra         4.0
Brown rice    4.0
Wheat         4.0
Sooji         3.0
Corn          3.0
Oats          2.0
Sattu         2.0
dtype: float64
```