

NAME: SANSKRUTI GOVARDHANE

ROLL NO: 852

PRN NO: 202201040219

BATCH: H3

Lab Assignment:

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

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

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

```
import pandas as pd
```

```
from itertools import combinations
```

```
from collections import Counter
```

```
df = pd.read_csv('/content/grainsales (2).csv')
```

```
print(df)
```

```
df['Sales'] = pd.to_numeric(df['Sales'])
```

```
monthly_sales = df.groupby('Months')['Sales'].sum()
best_month = monthly_sales.idxmax()
earnings = monthly_sales.loc[best_month]
print("The best month for sales was", best_month)
print("The earnings for that month were", earnings)
```

```
product_sales = df.groupby('GrainName')['Sales'].sum()
best_product = product_sales.idxmax()
print("The product that sold the most is", best_product)
```

```
city_sales = df.groupby('City')['Sales'].sum()
best_city = city_sales.idxmax()
print("The city that sold the most products is", best_city)
```

```
# Read the CSV file into a DataFrame
```

```
df = pd.read_csv('/content/grainsales (2).csv')
```

```
# Group the data by sales and create a list of products for each sale
```

```
grouped_sales = df.groupby('Sales')['GrainName'].apply(list)
```

```
# Create a list of all product combinations for each sale
```

```
product_combinations = [list(combinations(products, 2)) for products in
grouped_sales]
```

```
# Flatten the list of combinations
```

```

flattened_combinations = [item for sublist in product_combinations for item in
sublist]

# Count the occurrences of each product combination
combination_counts = Counter(flattened_combinations)

# Find the most common product combinations
most_common_combinations = combination_counts.most_common()

# Print the result
print("The most frequently sold product combinations are:")
for combination, count in most_common_combinations:
    print(combination[0], "and", combination[1], "- Sold together", count, "times")

```

OUTPUT:

	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 Bengal	Asansole	JULY	2023	4000000
19	Corn	UP	Kanpur	AUG	2023	4500000
20	Sooji	Tamil Nadu	Madurai	MAY	2023	3000000
21	Brown rice	Telangana	Hyderabad	JUNE	2023	3500000
22	Wheat	West Bengal	Asansole	JULY	2023	4000000
23	Corn	UP	Kanpur	AUG	2023	4500000
24	Ragi	Maharashtra	Nagpur	JAN	2023	1000000
25	Brown rice	Telangana	Hyderabad	JUNE	2023	3500000
26	Wheat	West Bengal	Asansole	JULY	2023	4000000

The best month for sales was JULY

The earnings for that month were 16000000

The product that sold the most is Wheat

The city that sold the most products is Asansole

The most frequently sold product combinations are:

Ragi and Ragi - Sold together 10 times

Bajra and Bajra - Sold together 6 times

Brown rice and Brown rice - Sold together 6 times

Wheat and Wheat - Sold together 6 times

Sooji and Sooji - Sold together 3 times

Corn and Corn - Sold together 3 times

Oats and Oats - Sold together 1 times

Sattu and Sattu - Sold together 1 times