```
#Name= Darshan Ramagade
#Roll No.= 548
#PRN = 202201070044
#Batch = E3

import pandas as pd
data= pd.read_csv('/content/grainsales.csv')
print(data)
```

```
       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
20        Sooji   Tamil Nadu    Madurai    MAY  2023  3000000
21   Brown rice    Telangana  Hyderabad   JUNE  2023  3500000
22        Wheat  West Bengol   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 Bengol   Asansole   JULY  2023  4000000
```

```
all_data=pd.read_csv("/content/grainsales.csv")
```

```
all_data.head()
```

```
GrainNameStateCityMonthsYearSales0RagiMaharashtraNagpurJAN202310000001BajraPanjabAmritsarFEB202315000002RagiMaharashtraNagpurJAN2
02310000003BajraPanjabAmritsarFEB202315000004RagiMaharashtraNagpurJAN20231000000
```

```
# 1 Which was the best month for sales? How much was earned that month?
```

```
import pandas as pd

df = pd.read_csv('/content/grainsales.csv')
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 months were", earnings)
```

```
O/P= The best month for sales was JULY
        The earnings for that months were 16000000
```

```
# 2 Which product sold the most? Why do you think it did?
```

```
import pandas as pd

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

```
O/P= The product that sold the most is Wheat
```

```
# 3 Which city sold the most products?
```

```
import pandas as pd
```

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

```
O/P = The city that sold the most products is Asansole
```

```
# 4 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.csv')
grouped_sales = df.groupby('Sales')['GrainName'].apply(list)
product_combinations = [list(combinations(products, 2)) for products in grouped_sales]
flattened_combinations = [item for sublist in product_combinations for item in sublist]
combination_counts = Counter(flattened_combinations)
most_common_combinations = combination_counts.most_common()
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")
```

```
O/P= 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 S
       attu - Sold together 1 times
```

```
#Extra Operations
```

```
nan_df = all_data[all_data.isna().any(axis=1)]
```

```
display(nan_df.head())
```

```
all_data = all_data.dropna(how='all')
all_data.head()
```

GrainName State City Months Year Sales

	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

```
prices = all_data.groupby('City').mean()['Sales']
```

```
print(prices)
```

```
O/P = City Amritsar 1500000.0 Asansole 4000000.0 Gurugram 2000000.0 Hyderabad 3500000.0 Kanpur
4500000.0 Madurai 3000000.0 Nagpur 1000000.0 Surat 2500000.0 Name: Sales, dtype: float64
```