Assignment NO.4

Name: AISHU SINHA

Rollno:306

PRN:202201060068

```
import pandas as pd

# Read the CSV file into a DataFrame df =
pd.read_csv('/content/grainsales.csv')

# Display the DataFrame print(df)
```

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 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

```
# Identify 10 grains from the dataset grains =
df['GrainName'].unique()[:10] print(grains)
```

OUTPUT:

```
['Ragi' 'Bajra' 'Oats' 'Sattu ' 'Sooji' 'Brown rice ' 'Wheat' 'Corn']
```

```
# Group data by month and calculate total sales
monthly_sales = df.groupby('Months')['Sales'].sum()

# The month with the highest sales best_month
= monthly_sales.idxmax()

# Get the earnings for the best month earnings
= monthly_sales.loc[best_month]
 print("Best month for sales:", best_month)
print("Earnings for the best month:", earnings)
```

OUTPUT:

```
Best month for sales: JULY
Earnings for the best month:
16000000
```

```
# Group data by product and calculate
total sales product_sales =
df.groupby('GrainName')['Sales'].sum()

# The product with the
highest sales
best_product =
product_sales.idxmax()
 print("Product that
sold the most:",
best_product)
```

OUTPUT:

```
Product that sold the most: Wheat
```