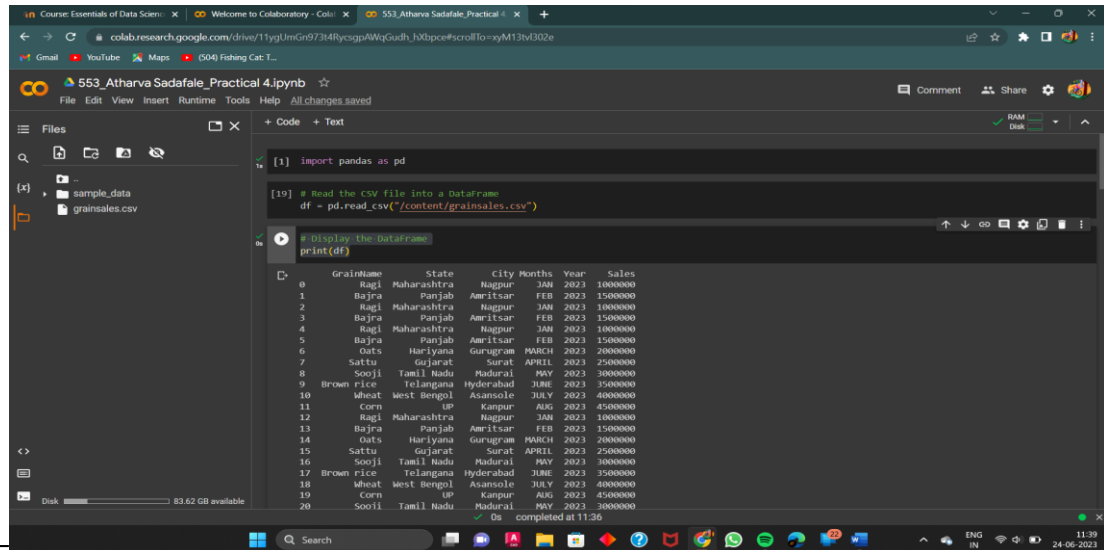


Name: ATHARVA RAHUL SADAFale

Roll no: 553

PRN No: 202201070048

EDS Practical 4



The screenshot shows a Google Colab environment. The notebook is titled "553_Atharva Sadafale_Practical 4.ipynb". The code in the notebook is as follows:

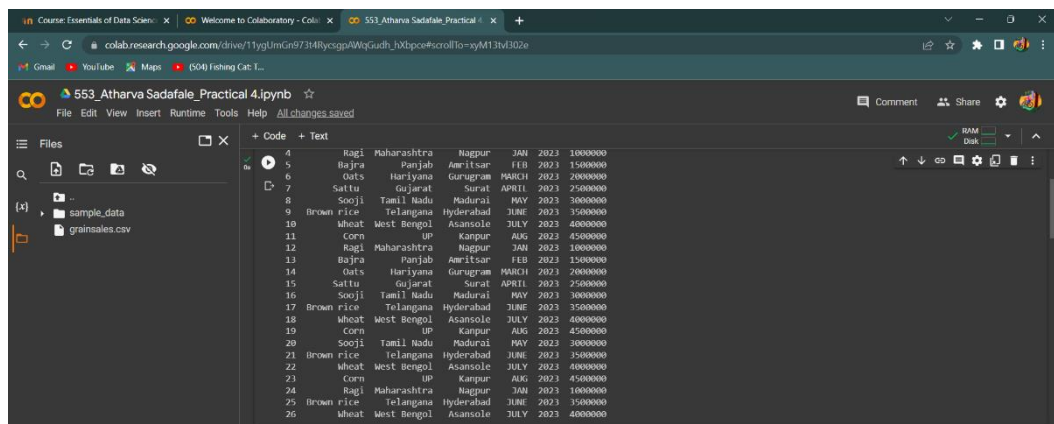
```
[1] import pandas as pd

[19] # Read the CSV file into a DataFrame
df = pd.read_csv("/content/grainsales.csv")

# Display the DataFrame
print(df)
```

The output of the code is a DataFrame with the following columns: GrainName, State, City, Month, Year, and Sales. The data is as follows:

GrainName	State	City	Month	Year	Sales
Ragi	Maharashtra	Nagpur	JAN	2023	1000000
Bajra	Punjab	Amritsar	FEB	2023	1500000
Ragi	Maharashtra	Nagpur	JAN	2023	1000000
Bajra	Punjab	Amritsar	FEB	2023	1500000
Ragi	Maharashtra	Nagpur	JAN	2023	1000000
Bajra	Punjab	Amritsar	FEB	2023	1500000
Oats	Haryana	Gurgaon	MARCH	2023	2000000
Sattu	Gujarat	Surat	APRIL	2023	2500000
Sooji	Tamil Nadu	Madurai	MAY	2023	3000000
Brown rice	Telangana	Hyderabad	JUNE	2023	3500000
Wheat	West Bengal	Asansole	JULY	2023	4000000
Corn	UP	Kanpur	AUG	2023	4500000
Ragi	Maharashtra	Nagpur	JAN	2023	1000000
Bajra	Punjab	Amritsar	FEB	2023	1500000
Oats	Haryana	Gurgaon	MARCH	2023	2000000
Sattu	Gujarat	Surat	APRIL	2023	2500000
Sooji	Tamil Nadu	Madurai	MAY	2023	3000000
Brown rice	Telangana	Hyderabad	JUNE	2023	3500000
Wheat	West Bengal	Asansole	JULY	2023	4000000
Corn	UP	Kanpur	AUG	2023	4500000
Sooji	Tamil Nadu	Madurai	MAY	2023	3000000



The screenshot shows the same Google Colab environment. The code in the notebook is as follows:

```
4 Ragi Maharashtra Nagpur JAN 2023 1000000
5 Bajra Punjab Amritsar FEB 2023 1500000
6 Oats Haryana Gurgaon 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 Bengal Asansole JULY 2023 4000000
11 Corn UP Kanpur AUG 2023 4500000
12 Ragi Maharashtra Nagpur JAN 2023 1000000
13 Bajra Punjab Amritsar FEB 2023 1500000
14 Oats Haryana Gurgaon 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
```

```
1] #Count the number of unique grain names:
num_grains = df['GrainName'].nunique()
print("Number of unique grain names:", num_grains)
```

Result: Number of unique grain names: 8

```
2] #Calculate the total sales amount for each grain:
grain_sales = df.groupby('GrainName')['Sales'].sum()
print(grain_sales)
```

Result :

GrainName	
Bajra	6000000
Brown rice	14000000
Corn	13500000
Oats	4000000
Ragi	5000000
Sattu	5000000
Sooji	9000000
Wheat	16000000

```
3] #The state with the highest total sales:
top_state = df.groupby('State')['Sales'].sum().idxmax()
print("State with the highest total sales:", top_state)
```

Result: State with the highest total sales: West Bengal

```
4] #Calculate the average sales amount per month:
monthly_average = df.groupby('Months')['Sales'].mean()
print(monthly_average)
```

Result:

Months	
APRIL	2500000.0
AUG	4500000.0
FEB	1500000.0
JAN	1000000.0
JULY	4000000.0
JUNE	3500000.0
MARCH	2000000.0
MAY	3000000.0

```
5] #The total sales for each grain in Telangana:
telangana_grain_sales = df[df['State'] ==
'Telangana'].groupby('GrainName')['Sales'].sum()
print(telangana_grain_sales)
```

Result:

GrainName	
Brown rice	14000000

```
6] #the total sales for each state:
state_sales = df.groupby('State')['Sales'].sum()
print(state_sales)
```

Result:

State	
Gujarat	5000000
Hariyana	4000000
Maharashtra	5000000
Panjab	6000000
Tamil Nadu	9000000
Telangana	14000000
UP	13500000
West Bengol	16000000

```
7] #Find the top 3 cities with the highest sales:
top_cities = df.groupby('City')['Sales'].sum().nlargest(3)
print(top_cities)
```

Result:

City	
Asansole	16000000
Hyderabad	14000000
Kanpur	13500000

```
8] #Data to show high data sales of more than 2000000 amount
high_sales_data = df[df['Sales'] > 1500000]
print(high_sales_data)
```

Result:

	GrainName	State	City	Months	Year	Sales
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
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
25	Brown rice	Telangana	Hyderabad	JUNE	2023	3500000
26	Wheat	West Bengal	Asansole	JULY	2023	4000000

```
9] #Sort the data by sales amount in descending order:
sorted_data = df.sort_values(by='Sales', ascending=False)
print(sorted_data)
```

Result:

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

```
10] #Filter the data to include only sales in Tamil Nadu and May:
tamil_nadu_may_sales = df[(df['State'] == 'Tamil Nadu') & (df['Months']
== 'MAY')]
print(tamil_nadu_may_sales)
```

Result:

	GrainName	State	City	Months	Year	Sales
8	Sooji	Tamil Nadu	Madurai	MAY	2023	3000000
16	Sooji	Tamil Nadu	Madurai	MAY	2023	3000000
20	Sooji	Tamil Nadu	Madurai	MAY	2023	3000000