

EDS PPRACTICAL 4

Name: Shreenivas Shivajirao Dudhate

Roll No: 511

Batch: E1

```
import pandas as pd

df = pd.read_csv('/content/sales_data.csv')

best_month = df.sort_values(by='Sales', ascending=False).iloc[0]
best_month_name = best_month['Month']
best_month_sales = best_month['Sales']
print("The best month for sales was", best_month_name, "with earnings of $", best_month_sales)

best_product = df['Product'].value_counts().idxmax()
print("The product that sold the most was", best_product)
product_sales = df.groupby('Product')['Sales'].sum()
print("Product Sales Distribution:")
print(product_sales)

best_city = df['City'].value_counts().idxmax()
print("The city that sold the most products was", best_city)

products_together = df.groupby('Month')['Product'].agg(lambda x: ', '.join(set(x))).reset_index()
print("Products Most Often Sold Together:")
print(products_together)
```

```
import pandas as pd

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

# Question 1: Which was the best month for sales? How much was earned that month?
best_month = df.sort_values('sales', ascending=False).iloc[0]
best_month_name = best_month['month']
best_month_sales = best_month['sales']
print("The best month for sales was", best_month_name, "with earnings of $", best_month_sales)

# Question 2: Which product sold the most? Why do you think it did?
best_product = df['product'].value_counts().idxmax()
print("The product that sold the most was", best_product)
product_sales = df.groupby('product')['sales'].sum()
print("Product sales distribution")
print(product_sales)

# Question 3: Which city sold the most products?
best_city = df['city'].value_counts().idxmax()
print("The city that sold the most products was", best_city)

# Question 4: What products are most often sold together?
products_together = df.groupby('month')['product'].agg(lambda x: ", ".join(set(x)).reset_index())
print("Products most often sold together:")
print(products_together)
```

The best month for sales was December with earnings of \$ 35000
The product that sold the most was A
Product sales Distribution:
Product
A 35000
B 15000
C 11000
Name: sales, dtype: int64
The city that sold the most products was New York
Products most often sold together:
Month Product
0 April A
1 August B
2 December C
3 February B
4 January A
5 July A
6 June C
7 March C
8 May B
9 November B
10 October A
11 September C