Name - Mohit Rajpurohit Registration No. - 12415859 Roll No. - 48

Questions

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

df1= pd.read_csv("accounts.csv")
df2= pd.read_csv("data_dictionary.csv")
df3= pd.read_csv("products.csv")
df4= pd.read_csv("sales_pipeline.csv")
df5= pd.read_csv("sales_teams.csv")

df4.head()
```

```
  opportunity_id      sales_agent         product  account deal_stage
\
0       1C1I7A6R      Moses Frase  GTX Plus Basic  Cancity        Won

1       Z063OYW0  Darcel Schlecht          GTXPro    Isdom        Won

2       EC4QE1BX  Darcel Schlecht      MG Special  Cancity        Won

3       MV1LWRNH      Moses Frase       GTX Basic  Codehow        Won

4       PE84CX4O        Zane Levy       GTX Basic   Hatfan        Won


  engage_date  close_date  close_value
0  2016-10-20  2017-03-01       1054.0
1  2016-10-25  2017-03-11       4514.0
2  2016-10-25  2017-03-07         50.0
3  2016-10-25  2017-03-09        588.0
4  2016-10-25  2017-03-02        517.0
```

```
df6=pd.merge(df1,df4 )
df7=pd.merge(df6,df1)
df8=pd.merge(df7,df5)
df=pd.merge(df8,df4)
df
```

```
               account     sector  year_established  revenue
employees  \
0     Acme Corporation  technolgy              1996  1100.04
2822
1     Acme Corporation  technolgy              1996  1100.04
2822
2     Acme Corporation  technolgy              1996  1100.04
2822
3     Acme Corporation  technolgy              1996  1100.04
```

```
2822
4     Acme Corporation  technolgy              1996  1100.04
2822
...                ...        ...               ...      ...         ..
.
7370          Zumgoity    medical              1984   441.08
1210
7371          Zumgoity    medical              1984   441.08
1210
7372          Zumgoity    medical              1984   441.08
1210
7373          Zumgoity    medical              1984   441.08
1210
7374          Zumgoity    medical              1984   441.08
1210

     office_location subsidiary_of opportunity_id
sales_agent  \
0      United States           NaN       N4SD17JR       Reed Clapper

1      United States           NaN       TBZMXKH4     Wilburn Farren

2      United States           NaN       S3W6Q07M      Kami Bicknell

3      United States           NaN       V4324HG4   Violet Mclelland

4      United States           NaN       TK9T01QM       Reed Clapper

...              ...           ...            ...                ...

7370   United States           NaN       RWR0W3U8    Darcel Schlecht

7371   United States           NaN       1PXQZ40I  Versie Hillebrand

7372   United States           NaN       YL29KX3L      Anna Snelling

7373   United States           NaN       D01U08PR    Darcel Schlecht

7374   United States           NaN       YMI68O7S  Marty Freudenburg


              product   deal_stage engage_date  close_date  close_value
\
0          GTX Basic          Won  2016-11-12  2017-03-01        556.0

1        MG Advanced          Won  2017-01-15  2017-03-29       3833.0

2       GTX Plus Pro     Engaging  2017-01-16         NaN          NaN

3         MG Special         Lost  2017-01-23  2017-05-06          0.0
```

```
4              GTXPro         Lost  2017-01-25  2017-05-07          0.0

...               ...          ...         ...         ...          ...

7370           GTXPro         Lost  2017-10-26  2017-10-27          0.0

7371      MG Advanced          Won  2017-12-13  2017-12-27       3735.0

7372  GTX Plus Basic  Prospecting         NaN         NaN          NaN

7373           GTXPro  Prospecting         NaN         NaN          NaN

7374  GTX Plus Basic  Prospecting         NaN         NaN          NaN


                manager regional_office
0         Rocco Neubert            East
1            Cara Losch            East
2         Summer Sewald            West
3            Cara Losch            East
4         Rocco Neubert            East
...                  ...             ...
7370      Melvin Marxen         Central
7371   Dustin Brinkmann         Central
7372   Dustin Brinkmann         Central
7373      Melvin Marxen         Central
7374      Melvin Marxen         Central

[7375 rows x 16 columns]
```

```
df.dtypes
```

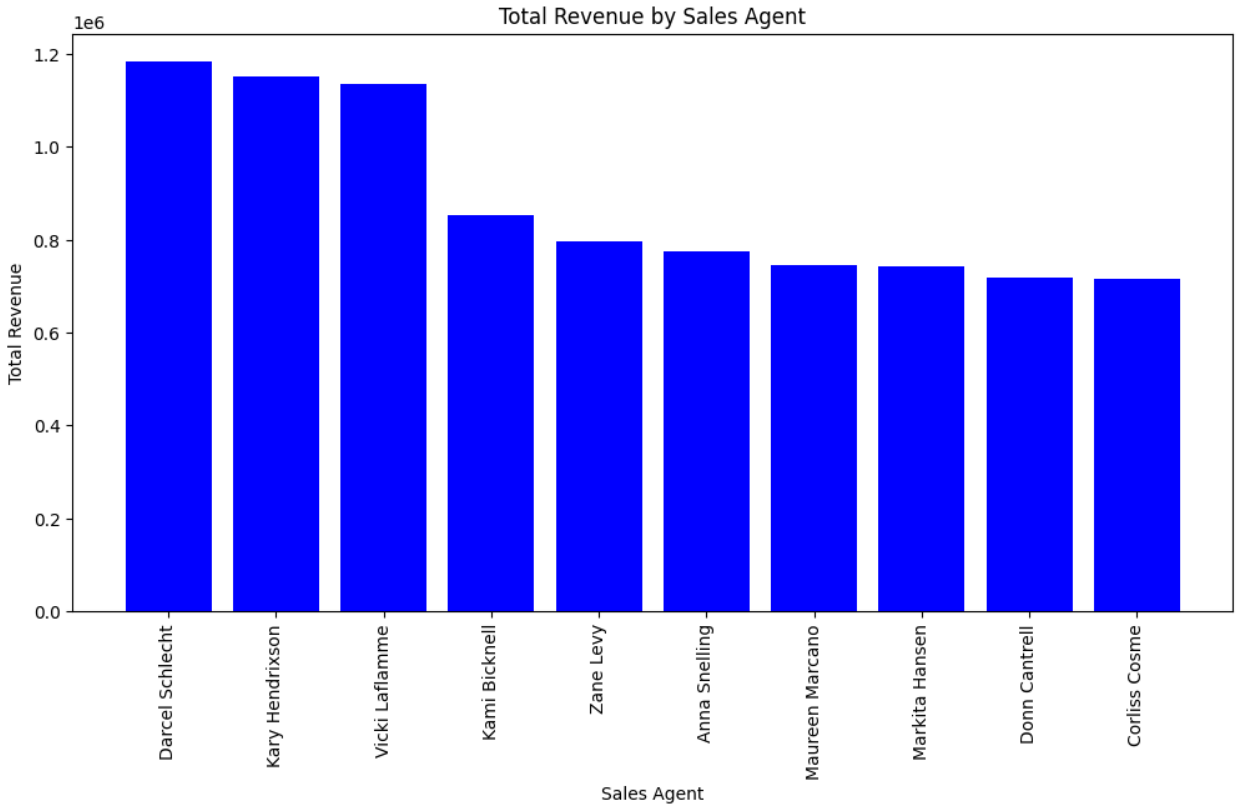
```
account              object
sector               object
year_established      int64
revenue             float64
employees             int64
office_location      object
subsidiary_of        object
opportunity_id       object
sales_agent          object
product              object
deal_stage           object
engage_date          object
close_date           object
close_value         float64
manager              object
regional_office      object
dtype: object
```

1. Sales Trends Over Time

```
df.head()
```

```
            account     sector              year_established  revenue
\
0  Acme Corporation  technolgy 1970-01-01 00:00:00.000001996  1100.04

1  Acme Corporation  technolgy 1970-01-01 00:00:00.000001996  1100.04

2  Acme Corporation  technolgy 1970-01-01 00:00:00.000001996  1100.04

3  Acme Corporation  technolgy 1970-01-01 00:00:00.000001996  1100.04

4  Acme Corporation  technolgy 1970-01-01 00:00:00.000001996  1100.04


   employees office_location subsidiary_of opportunity_id
sales_agent  \
0       2822   United States           NaN       N4SD17JR      Reed
Clapper
1       2822   United States           NaN       TBZMXKH4    Wilburn
Farren
2       2822   United States           NaN       S3W6Q07M     Kami
Bicknell
3       2822   United States           NaN       V4324HG4  Violet
Mclelland
4       2822   United States           NaN       TK9T01QM      Reed
Clapper

        product deal_stage engage_date close_date  close_value
manager  \
0     GTX Basic        Won  2016-11-12 2017-03-01        556.0  Rocco
Neubert
1   MG Advanced        Won  2017-01-15 2017-03-29       3833.0
Cara Losch
2  GTX Plus Pro   Engaging  2017-01-16 2017-05-22          NaN  Summer
Sewald
3    MG Special       Lost  2017-01-23 2017-05-06          0.0
Cara Losch
4        GTXPro       Lost  2017-01-25 2017-05-07          0.0  Rocco
Neubert

  regional_office
0            East
1            East
2            West
3            East
4            East
```

```
df['engage_date'] =
pd.to_datetime(df['engage_date'].fillna(df['engage_date'].mode()[0]))
df['close_date'] =
pd.to_datetime(df['close_date'].fillna(df['close_date'].mode()[0]))
df['year_established'] = pd.to_datetime(df['year_established'])
df1 = df.groupby("sales_agent")
["revenue"].sum().reset_index().sort_values(by="revenue",
ascending=False).head(10)
df1

# Plot the graph
plt.figure(figsize=(12, 6))
plt.bar(df1['sales_agent'], df1["revenue"], color='blue')
plt.xlabel("Sales Agent")
plt.ylabel("Total Revenue")
plt.title("Total Revenue by Sales Agent")
plt.xticks(rotation=90)
plt.show()
```



3.Sales by Region

```
df3 = df3.groupby("product")
["sales_price"].sum().reset_index().sort_values(by="sales_price",
ascending=False)
```

```
# Plot the bar chart
plt.figure(figsize=(10, 4))
plt.bar(df3['product'], df3['sales_price'], color='green')
plt.xlabel("Product")
plt.ylabel("Total Sales Price")
plt.title("Total Sales Price by Product")
plt.xticks(rotation=45)
plt.show()
```