```python
import pandas as pd
import plotly.express as px
import plotly.graph_objects as go
import plotly.io as pio
import plotly.colors as colors

pio.templates.default = "plotly_white"

data = pd.read_csv("G:\OneDrive\Desktop\Superstore.csv",
encoding='latin-1')

data.head()
```

```
    Row ID        Order ID  Order Date    Ship Date       Ship Mode
Customer ID  \
0        1  CA-2016-152156   11/8/2016   11/11/2016     Second Class
CG-12520
1        2  CA-2016-152156   11/8/2016   11/11/2016     Second Class
CG-12520
2        3  CA-2016-138688    6/12/2016    6/16/2016     Second Class
DV-13045
3        4  US-2015-108966  10/11/2015   10/18/2015   Standard Class
SO-20335
4        5  US-2015-108966  10/11/2015   10/18/2015   Standard Class
SO-20335

      Customer Name      Segment         Country             City  ...  \
0       Claire Gute     Consumer   United States        Henderson  ...
1       Claire Gute     Consumer   United States        Henderson  ...
2   Darrin Van Huff    Corporate   United States      Los Angeles  ...
3   Sean O'Donnell     Consumer    United States  Fort Lauderdale  ...
4   Sean O'Donnell     Consumer    United States  Fort Lauderdale  ...

   Postal Code  Region      Product ID         Category Sub-
Category  \
0        42420   South  FUR-BO-10001798        Furniture     Bookcases

1        42420   South  FUR-CH-10000454        Furniture        Chairs

2        90036    West  OFF-LA-10000240  Office Supplies        Labels

3        33311   South  FUR-TA-10000577        Furniture        Tables

4        33311   South  OFF-ST-10000760  Office Supplies       Storage


                                    Product Name     Sales
Quantity  \
0                 Bush Somerset Collection Bookcase   261.9600
2
1  Hon Deluxe Fabric Upholstered Stacking Chairs,...   731.9400
```

```
3
2   Self-Adhesive Address Labels for Typewriters b...    14.6200
2
3        Bretford CR4500 Series Slim Rectangular Table  957.5775
5
4                       Eldon Fold 'N Roll Cart System    22.3680
2

   Discount      Profit
0      0.00     41.9136
1      0.00    219.5820
2      0.00      6.8714
3      0.45   -383.0310
4      0.20      2.5164

[5 rows x 21 columns]
```

```python
df_shape = data.shape
```

```python
df_shape
```

```
(9994, 21)
```

```python
missing_values  = data.isnull().sum()
```

```python
missing_values
```

```
Row ID            0
Order ID          0
Order Date        0
Ship Date         0
Ship Mode         0
Customer ID       0
Customer Name     0
Segment           0
Country           0
City              0
State             0
Postal Code       0
Region            0
Product ID        0
Category          0
Sub-Category      0
Product Name      0
Sales             0
Quantity          0
Discount          0
Profit            0
dtype: int64
```

```python
data_cleaned = data.dropna()
```

```
data_cleaned

       Row ID          Order ID  Order Date   Ship Date          Ship
Mode  \
0             1  CA-2016-152156   11/8/2016   11/11/2016    Second Class

1             2  CA-2016-152156   11/8/2016   11/11/2016    Second Class

2             3  CA-2016-138688   6/12/2016    6/16/2016    Second Class

3             4  US-2015-108966  10/11/2015   10/18/2015  Standard Class

4             5  US-2015-108966  10/11/2015   10/18/2015  Standard Class

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

9989       9990  CA-2014-110422   1/21/2014    1/23/2014    Second Class

9990       9991  CA-2017-121258   2/26/2017     3/3/2017  Standard Class

9991       9992  CA-2017-121258   2/26/2017     3/3/2017  Standard Class

9992       9993  CA-2017-121258   2/26/2017     3/3/2017  Standard Class

9993       9994  CA-2017-119914    5/4/2017     5/9/2017    Second Class


      Customer ID     Customer Name     Segment          Country
City  \
0       CG-12520       Claire Gute     Consumer   United States
Henderson
1       CG-12520       Claire Gute     Consumer   United States
Henderson
2       DV-13045    Darrin Van Huff   Corporate   United States        Los
Angeles
3       SO-20335     Sean O'Donnell    Consumer   United States   Fort
Lauderdale
4       SO-20335     Sean O'Donnell    Consumer   United States   Fort
Lauderdale
...          ...               ...         ...             ...         ...
...
9989    TB-21400   Tom Boeckenhauer    Consumer   United States
Miami
9990    DB-13060        Dave Brooks    Consumer   United States
Costa Mesa
9991    DB-13060        Dave Brooks    Consumer   United States
Costa Mesa
9992    DB-13060        Dave Brooks    Consumer   United States
Costa Mesa
9993    CC-12220       Chris Cortes    Consumer   United States
Westminster
```

```
      ... Postal Code  Region      Product ID          Category Sub-
Category  \
0     ...       42420   South  FUR-BO-10001798          Furniture
Bookcases
1     ...       42420   South  FUR-CH-10000454          Furniture
Chairs
2     ...       90036    West  OFF-LA-10000240  Office Supplies
Labels
3     ...       33311   South  FUR-TA-10000577          Furniture
Tables
4     ...       33311   South  OFF-ST-10000760  Office Supplies
Storage
...   ...         ...     ...              ...              ...
...
9989  ...       33180   South  FUR-FU-10001889          Furniture
Furnishings
9990  ...       92627    West  FUR-FU-10000747          Furniture
Furnishings
9991  ...       92627    West  TEC-PH-10003645         Technology
Phones
9992  ...       92627    West  OFF-PA-10004041  Office Supplies
Paper
9993  ...       92683    West  OFF-AP-10002684  Office Supplies
Appliances

                                           Product Name      Sales
Quantity  \
0                    Bush Somerset Collection Bookcase   261.9600
2
1       Hon Deluxe Fabric Upholstered Stacking Chairs,...   731.9400
3
2       Self-Adhesive Address Labels for Typewriters b...    14.6200
2
3           Bretford CR4500 Series Slim Rectangular Table   957.5775
5
4                          Eldon Fold 'N Roll Cart System    22.3680
2
...                                                   ...        ...
...
9989                               Ultra Door Pull Handle    25.2480
3
9990  Tenex B1-RE Series Chair Mats for Low Pile Car...    91.9600
2
9991                                 Aastra 57i VoIP phone   258.5760
2
9992  It's Hot Message Books with Stickers, 2 3/4" x 5"    29.6000
4
9993  Acco 7-Outlet Masterpiece Power Center, Wihtou...   243.1600
```

```
2

      Discount      Profit
0         0.00     41.9136
1         0.00    219.5820
2         0.00      6.8714
3         0.45   -383.0310
4         0.20      2.5164
...        ...        ...
9989      0.20      4.1028
9990      0.00     15.6332
9991      0.20     19.3932
9992      0.00     13.3200
9993      0.00     72.9480

[9994 rows x 21 columns]

descriptive_numerical_stats = data_cleaned.describe()

descriptive_numerical_stats

             Row ID    Postal Code           Sales      Quantity
Discount  \
count   9994.000000    9994.000000     9994.000000   9994.000000
9994.000000
mean    4997.500000   55190.379428      229.858001      3.789574
0.156203
std     2885.163629   32063.693350      623.245101      2.225110
0.206452
min        1.000000    1040.000000        0.444000      1.000000
0.000000
25%     2499.250000   23223.000000       17.280000      2.000000
0.000000
50%     4997.500000   56430.500000       54.490000      3.000000
0.200000
75%     7495.750000   90008.000000      209.940000      5.000000
0.200000
max     9994.000000   99301.000000    22638.480000     14.000000
0.800000

             Profit
count   9994.000000
mean      28.656896
std      234.260108
min    -6599.978000
25%        1.728750
50%        8.666500
75%       29.364000
max     8399.976000
```

```python
numerical_column = 'Sales'

mean_value = data_cleaned[numerical_column].mean()

mean_value
```

229.8580008304938

```python
data_cleaned['Order Date'] = pd.to_datetime(data_cleaned['Order Date'])

data_cleaned['Month'] = data_cleaned['Order Date'].dt.to_period('M').astype(str)

data_cleaned.head()
```

```
   Row ID        Order ID Order Date   Ship Date      Ship Mode Customer ID  \
0       1  CA-2016-152156 2016-11-08  11/11/2016    Second Class CG-12520
1       2  CA-2016-152156 2016-11-08  11/11/2016    Second Class CG-12520
2       3  CA-2016-138688 2016-06-12   6/16/2016    Second Class DV-13045
3       4  US-2015-108966 2015-10-11  10/18/2015  Standard Class SO-20335
4       5  US-2015-108966 2015-10-11  10/18/2015  Standard Class SO-20335

      Customer Name     Segment        Country            City  ... Region  \
0       Claire Gute    Consumer  United States        Henderson  ... South
1       Claire Gute    Consumer  United States        Henderson  ... South
2   Darrin Van Huff   Corporate  United States      Los Angeles  ... West
3    Sean O'Donnell    Consumer  United States  Fort Lauderdale  ... South
4    Sean O'Donnell    Consumer  United States  Fort Lauderdale  ... South

        Product ID         Category Sub-Category  \
0  FUR-BO-10001798        Furniture    Bookcases
1  FUR-CH-10000454        Furniture       Chairs
2  OFF-LA-10000240  Office Supplies       Labels
3  FUR-TA-10000577        Furniture       Tables
4  OFF-ST-10000760  Office Supplies      Storage

                                     Product Name     Sales  Quantity  \
```

```
0                    Bush Somerset Collection Bookcase   261.9600
2
1   Hon Deluxe Fabric Upholstered Stacking Chairs,...   731.9400
3
2   Self-Adhesive Address Labels for Typewriters b...    14.6200
2
3       Bretford CR4500 Series Slim Rectangular Table   957.5775
5
4                    Eldon Fold 'N Roll Cart System    22.3680
2

    Discount      Profit      Month
0       0.00     41.9136    2016-11
1       0.00    219.5820    2016-11
2       0.00      6.8714    2016-06
3       0.45   -383.0310    2015-10
4       0.20      2.5164    2015-10

[5 rows x 22 columns]
```

```python
monthly_sales = data_cleaned.groupby('Month')
['Sales'].sum().reset_index()

monthly_sales.head()
```

```
      Month         Sales
0   2014-01   14236.895
1   2014-02    4519.892
2   2014-03   55691.009
3   2014-04   28295.345
4   2014-05   23648.287
```

```python
monthly_sales['Month'] = monthly_sales['Month'].astype(str)

Monthly_Sales = px.line(monthly_sales, x='Month', y='Sales',
title='Monthly Sales')

Monthly_Sales.update_layout(height=500, width=800)
```

Monthly Sales
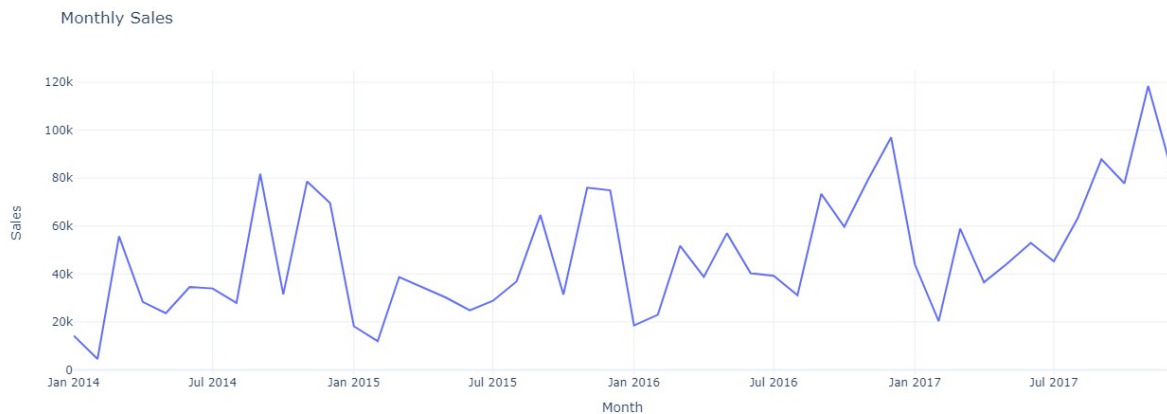

```
sales_by_category = data_cleaned.groupby(['Month', 'Category'])
['Sales'].sum().reset_index()

sales_by_category

        Month         Category        Sales
0     2014-01         Furniture     6242.5250
1     2014-01   Office Supplies     4851.0800
2     2014-01        Technology     3143.2900
3     2014-02         Furniture     1839.6580
4     2014-02   Office Supplies     1071.7240
..       ...               ...           ...
139   2017-11   Office Supplies    31472.3370
140   2017-11        Technology    49918.7730
141   2017-12         Furniture    31407.4668
142   2017-12   Office Supplies    30436.9420
143   2017-12        Technology    21984.9100

[144 rows x 3 columns]

Sales_by_Category = px.bar(sales_by_category, x='Month', y='Sales',
color='Category', title='Sales by Category')

Sales_by_Category.update_layout(height=500, width=800)
```
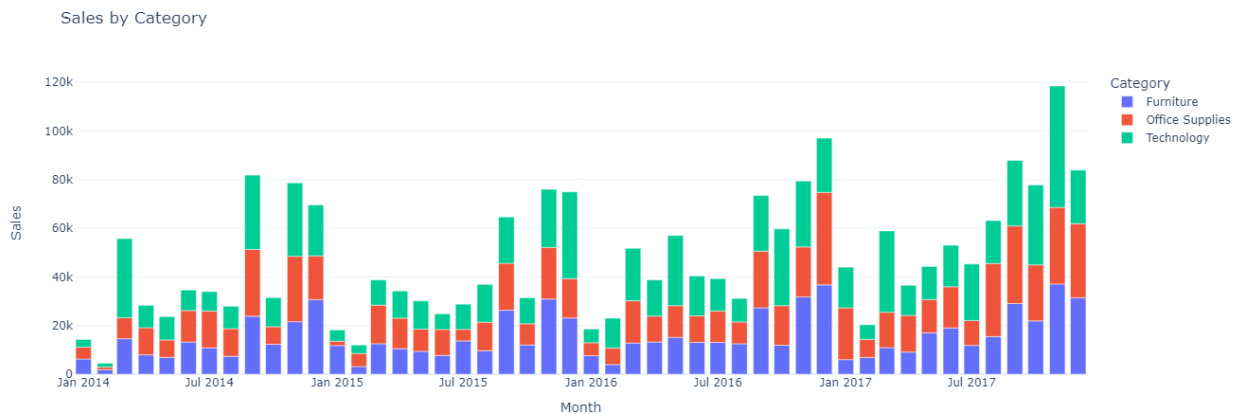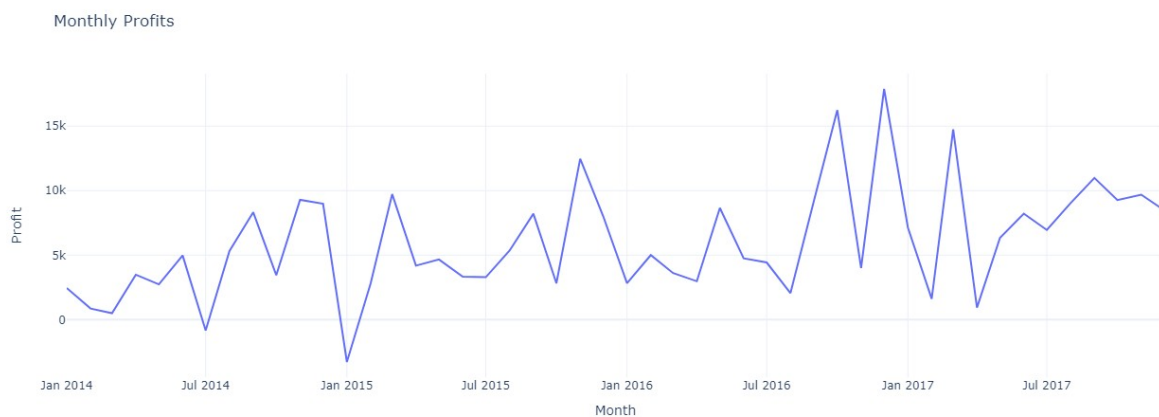
Sales by Category


```
data_cleaned['Profit'] =
data_cleaned['Profit'].astype(str).str.replace(',', '').astype(float)

monthly_profits = data_cleaned.groupby('Month')
['Profit'].sum().reset_index()

Monthly_Profits = px.line(monthly_profits, x='Month', y='Profit',
title='Monthly Profits')

Monthly_Profits.update_layout(height=500, width=800)
```

Monthly Profits


```
data_cleaned['Profit'] =
data_cleaned['Profit'].astype(str).str.replace(',', '').astype(float)

profit_by_category1 = data_cleaned.groupby('Category')
['Profit'].sum().reset_index()

profit_by_category1
```
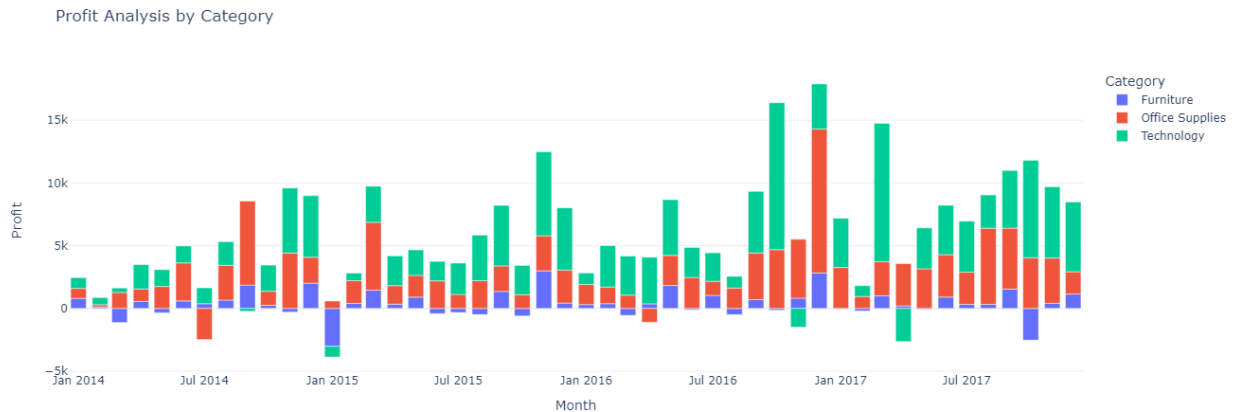
	Category	Profit
0	Furniture	18451.2728

```
1  Office Supplies   122490.8008
2       Technology   145454.9481
```

```python
profit_by_category = data_cleaned.groupby(['Month', 'Category'])
['Profit'].sum().reset_index()

Profit_Analysis_by_Category = px.bar(profit_by_category, x='Month',
y='Profit', color='Category', title='Profit Analysis by Category')

Profit_Analysis_by_Category.update_layout(height=500, width=800)
```


Profit Analysis by Category

```python
Pie_Chart = px.pie(profit_by_category, names='Category',
values='Profit', title='Profit Analysis by Category ')

Pie_Chart.update_layout(height=500, width=800)
```


Profit Analysis by Category