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

```
df = pd.read_csv("Sample - Superstore.csv", encoding="latin1")
df.head()
```

```
df.shape
df.info()
```

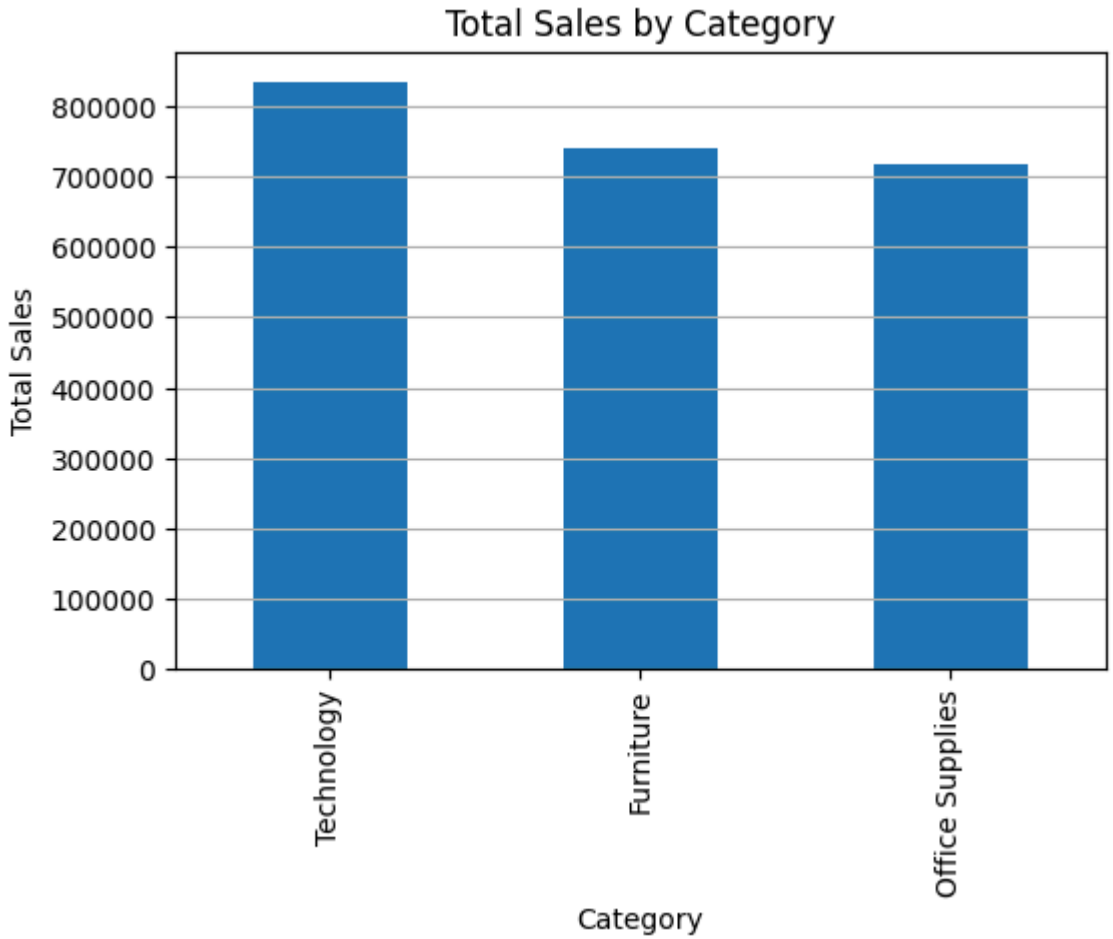
```
(9994, 21)
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 9994 entries, 0 to 9993
Data columns (total 21 columns):
 #   Column         Non-Null Count  Dtype  
---  ------         --------------  -----  
 0   Row ID         9994 non-null   int64  
 1   Order ID       9994 non-null   object 
 2   Order Date     9994 non-null   object 
 3   Ship Date      9994 non-null   object 
 4   Ship Mode      9994 non-null   object 
 5   Customer ID    9994 non-null   object 
 6   Customer Name  9994 non-null   object 
 7   Segment        9994 non-null   object 
 8   Country        9994 non-null   object 
 9   City           9994 non-null   object 
 10  State          9994 non-null   object 
 11  Postal Code    9994 non-null   int64  
 12  Region         9994 non-null   object 
 13  Product ID     9994 non-null   object 
 14  Category       9994 non-null   object 
 15  Sub-Category   9994 non-null   object 
 16  Product Name   9994 non-null   object 
 17  Sales          9994 non-null   float64
 18  Quantity       9994 non-null   int64  
 19  Discount       9994 non-null   float64
 20  Profit         9994 non-null   float64
dtypes: float64(3), int64(3), object(15)
memory usage: 1.6+ MB
```

```
top_categories = (
    df.groupby("Category")["Sales"]
    .sum()
    .sort_values(ascending=False)
```

```
)

top_categories.plot(
    kind="bar",
    title="Total Sales by Category",
    figsize=(6,4)
)

plt.xlabel("Category")
plt.ylabel("Total Sales")
plt.grid(axis="y")
plt.show()
```



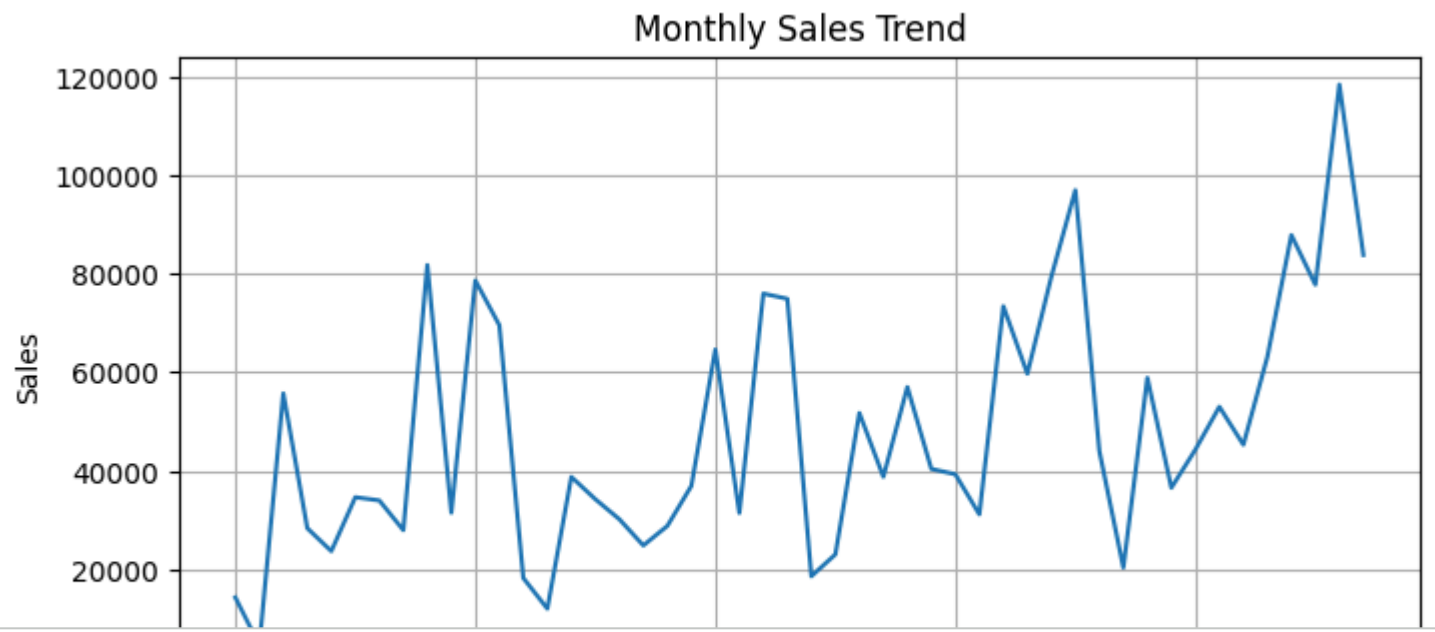

```
df["Order Date"] = pd.to_datetime(df["Order Date"])
```

```
sales_trend = (
    df.groupby(df["Order Date"].dt.to_period("M"))["Sales"]
    .sum()
)

sales_trend.index = sales_trend.index.astype(str)
```

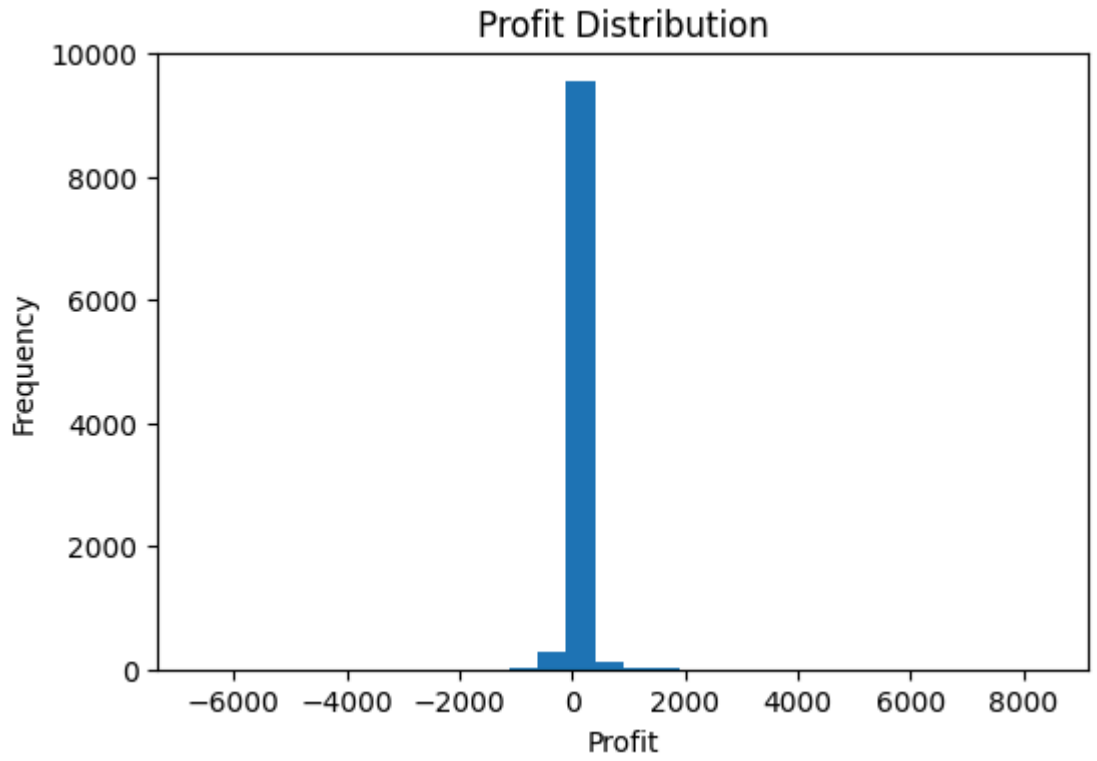
```
sales_trend.plot(
    kind="line",
    title="Monthly Sales Trend",
    figsize=(8,4)
)

plt.xlabel("Month")
plt.ylabel("Sales")
plt.grid()
plt.show()
```

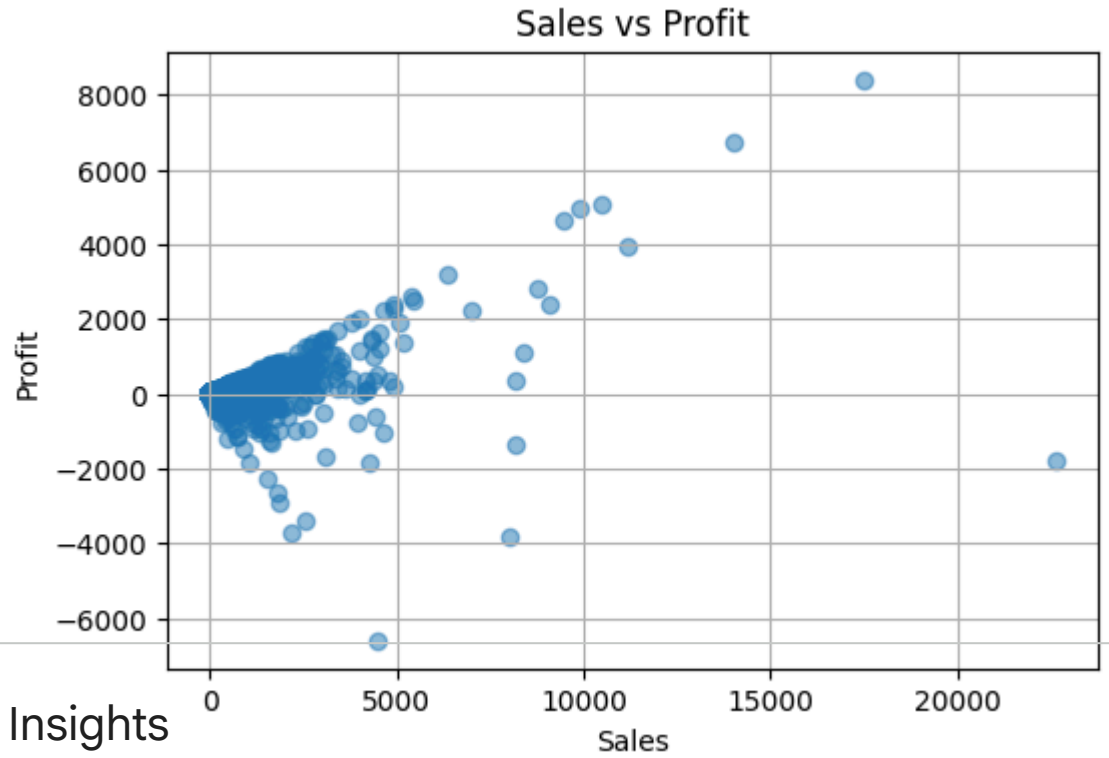



```
df["Profit"].plot(
    kind="hist",
    bins=30,
    title="Profit Distribution",
    figsize=(6,4)
)
```

```
plt.xlabel("Profit")
plt.show()
```




```
plt.figure(figsize=(6,4))
plt.scatter(df["Sales"], df["Profit"], alpha=0.5)
plt.title("Sales vs Profit")
plt.xlabel("Sales")
plt.ylabel("Profit")
plt.grid()
plt.show()
```



Key Insights

1. Technology category generates the highest total sales among all categories.
2. Monthly sales show seasonal spikes, indicating high-demand periods.