```
from google.colab import files
uploaded = files.upload()
```

Choose Files flipkart_pro…0250405.csv

- **flipkart_products_20250405.csv**(text/csv) - 903362 bytes, last modified: 4/6/2025 - 100% done

```
Saving flipkart_products_20250405.csv to flipkart_products_20250405 (1).csv
```

```
import pandas as pd

# Replace the file name with your actual uploaded file name
df = pd.read_csv('flipkart_products_20250405.csv')
df.head()
```

	Product Name	Price (₹)	Rating (★)	Number of Buyers	Total Sold	Available Stock	Main Category	Sub Category	Discount (%)	Seller	Return Policy	
0	Krishnamurthy-Devan Laboriosam Ultra Smartphon...	142247.04	3.2	7348	4812	364	Electronics	Smartphones	45	RetailNet	False	https://www.flipkart.com
1	Nanda-Mahal Dignissimos Lite Laptops 1	186922.43	4.1	2342	881	145	Electronics	Laptops	55	Flipkart Assured	False	https://www.flipkart.cor
2	Choudhury LLC Amet Plus Decor 15	11843.41	5.0	739	2580	206	Home	Decor	58	SuperComNet	True	https://www.flipkart.c

Next steps: Generate code with df View recommended plots New interactive sheet

```
# Extracting a Series object
# Check for typos in the column name - It might be 'Product Name' instead of 'product_name'
title_series = df['Product Name']
print(type(title_series))
print(title_series.head())
```

```
<class 'pandas.core.series.Series'>
0    Krishnamurthy-Devan Laboriosam Ultra Smartphon...
1               Nanda-Mahal Dignissimos Lite Laptops 1
2                     Choudhury LLC Amet Plus Decor 15
3             Borah LLC Accusantium Lite Smartphones 9
4                 Murty Inc Placeat Pro Smartwatches 8
Name: Product Name, dtype: object
```

```
# DataFrame info
print(type(df))
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5000 entries, 0 to 4999
Data columns (total 12 columns):
 #   Column            Non-Null Count  Dtype  
---  ------            --------------  -----  
 0   Product Name      5000 non-null   object 
 1   Price (₹)         5000 non-null   float64
 2   Rating (★)        5000 non-null   float64
 3   Number of Buyers  5000 non-null   int64  
 4   Total Sold        5000 non-null   int64  
 5   Available Stock   5000 non-null   int64  
 6   Main Category     5000 non-null   object 
 7   Sub Category      5000 non-null   object 
 8   Discount (%)      5000 non-null   int64  
 9   Seller            5000 non-null   object 
 10  Return Policy     5000 non-null   bool   
 11  Product URL       5000 non-null   object 
dtypes: bool(1), float64(2), int64(4), object(5)
memory usage: 434.7+ KB
```

```
# Selecting column (Series)
# Check if 'price' column exists, if not try 'Price'
print(df['Price']) if 'Price' in df.columns else print(df['price']) if 'price' in df.columns else print("Column 'price' or 'Price' not f

# Selecting multiple columns
# Ensure column names used here match the actual names in your DataFrame
print(df[['product_name', 'Price']].head()) if 'Price' in df.columns else print(df[['product_name', 'price']].head()) if 'price' in df.c


# Row selection using iloc
print(df.iloc[0])  # First row

# Row selection using loc
print(df.loc[0])  # Based on index label
```

```
Column 'price' or 'Price' not found in DataFrame
Column 'price' or 'Price' not found in DataFrame
Product Name        Krishnamurthy-Devan Laboriosam Ultra Smartphon...
Price (₹)                                                   142247.04
Rating (★)                                                        3.2
Number of Buyers                                                 7348
Total Sold                                                       4812
Available Stock                                                   364
Main Category                                             Electronics
Sub Category                                              Smartphones
Discount (%)                                                       45
Seller                                                      RetailNet
Return Policy                                                   False
Product URL         https://www.flipkart.com/Krishnamurthy-Devan-L...
Name: 0, dtype: object
Product Name        Krishnamurthy-Devan Laboriosam Ultra Smartphon...
Price (₹)                                                   142247.04
Rating (★)                                                        3.2
Number of Buyers                                                 7348
Total Sold                                                       4812
Available Stock                                                   364
Main Category                                             Electronics
Sub Category                                              Smartphones
Discount (%)                                                       45
Seller                                                      RetailNet
Return Policy                                                   False
Product URL         https://www.flipkart.com/Krishnamurthy-Devan-L...
Name: 0, dtype: object
```

```
import pandas as pd
import numpy as np

# ... (Your existing code to load the DataFrame) ...

# Check if 'price' or 'Price' column exists in the DataFrame
price_column = 'price' if 'price' in df.columns else 'Price' if 'Price' in df.columns else None

if price_column:
    # Apply np.log1p on prices (safe for 0 values)
    df['log_price'] = np.log1p(df[price_column])
    df[[price_column, 'log_price']].head()
else:
    print("Column 'price' or 'Price' not found in DataFrame")
```

```
Column 'price' or 'Price' not found in DataFrame
```

```
# Check for nulls
df.isnull().sum()

# Drop rows with missing values
df_cleaned = df.dropna()

# Fill missing values
df_filled = df.fillna({'price': 0})
```

```
# Check which rows have null prices, handling potential column name variations
if 'price' in df.columns:
    null_price_rows = df[df['price'].isnull()]
elif 'Price' in df.columns:
    null_price_rows = df[df['Price'].isnull()]
else:
    print("No 'price' or 'Price' column found for null check")
    null_price_rows = pd.DataFrame()  # Empty DataFrame if column not found

print(null_price_rows)
# Replace null product names with "Unknown"
# Check if 'product_name' or 'Product Name' column exists before filling NaNs
product_name_col = 'product_name' if 'product_name' in df.columns else 'Product Name' if 'Product Name' in df.columns else None

if product_name_col:
    df[product_name_col] = df[product_name_col].fillna("Unknown")
else:
    print("Column 'product_name' or 'Product Name' not found in DataFrame")
```

```
No 'price' or 'Price' column found for null check
Empty DataFrame
Columns: []
Index: []
```

```
# Create a MultiIndex using category and brand
if 'category' in df.columns and 'brand' in df.columns:
    df_multi = df.set_index(['category', 'brand'])
    print(df_multi.head())
```