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

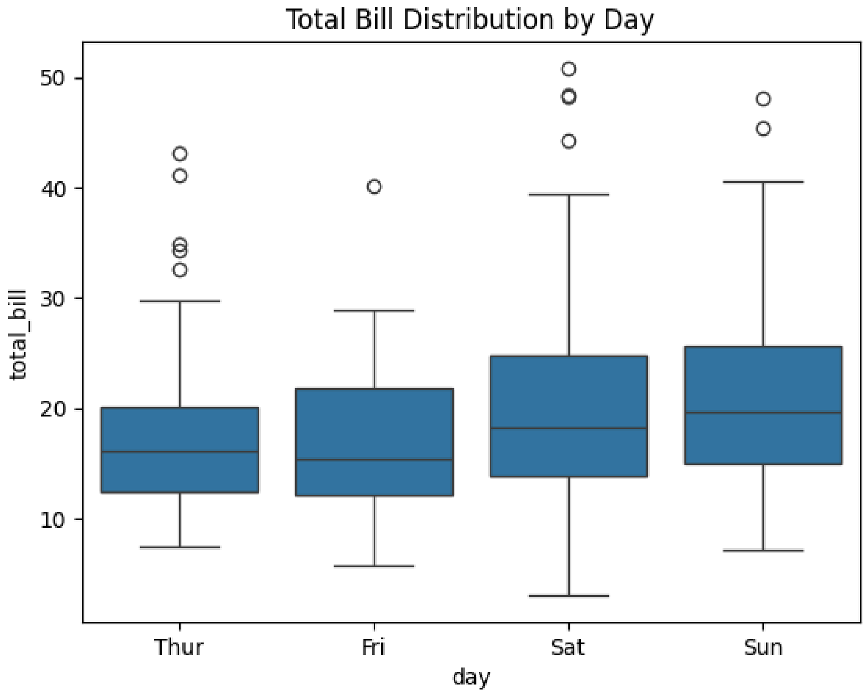
```
df = sns .load_dataset('tips')
```

```
df
```

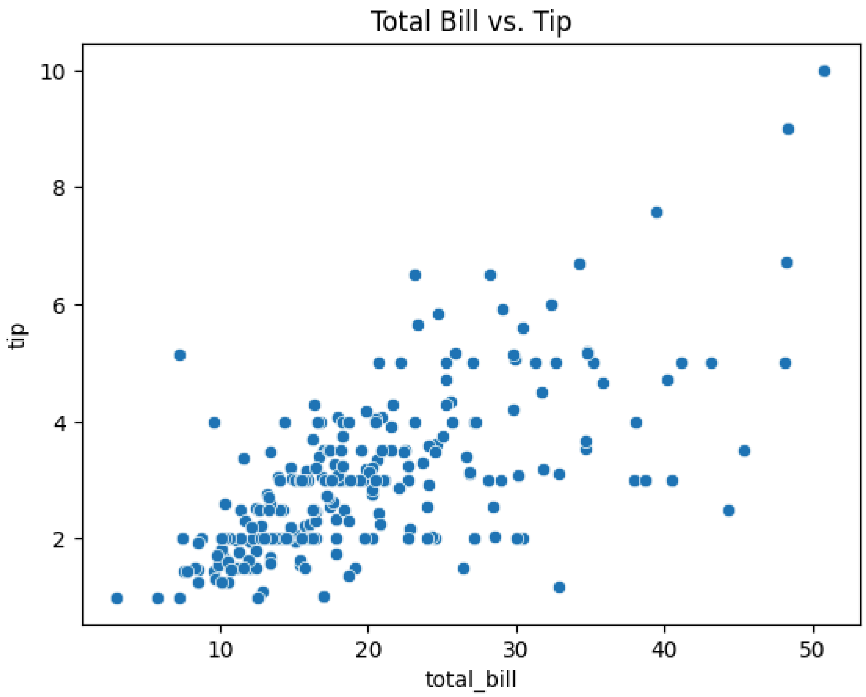
	total_bill	tip	sex	smoker	day	time	size
0	16.99	1.01	Female	No	Sun	Dinner	2
1	10.34	1.66	Male	No	Sun	Dinner	3
2	21.01	3.50	Male	No	Sun	Dinner	3
3	23.68	3.31	Male	No	Sun	Dinner	2
4	24.59	3.61	Female	No	Sun	Dinner	4
...	...	...	...	...	...	...	...
239	29.03	5.92	Male	No	Sat	Dinner	3
240	27.18	2.00	Female	Yes	Sat	Dinner	2
241	22.67	2.00	Male	Yes	Sat	Dinner	2
242	17.82	1.75	Male	No	Sat	Dinner	2
243	18.78	3.00	Female	No	Thur	Dinner	2

244 rows × 7 columns

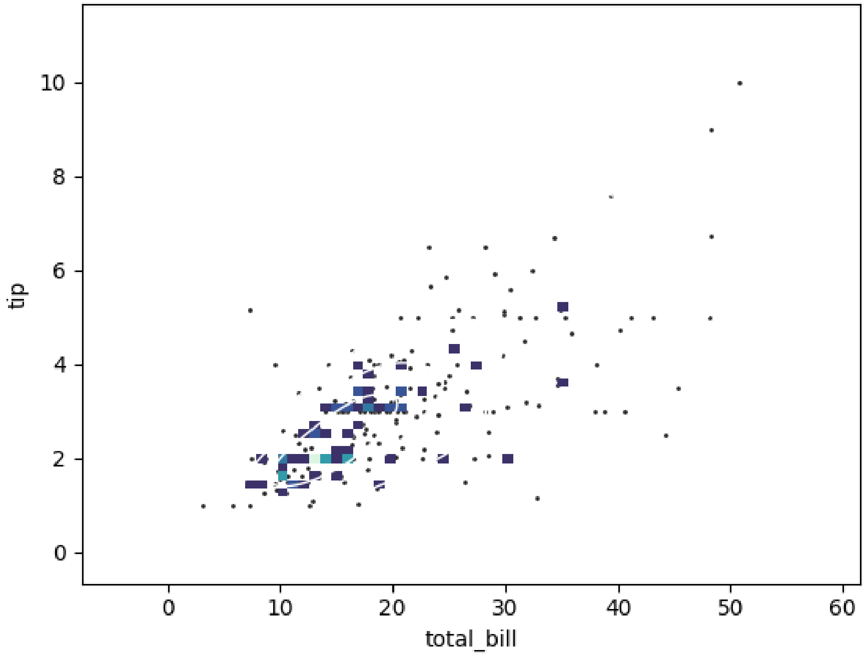
```
sns.boxplot(data=df, x='day', y='total_bill')
plt.title('Total Bill Distribution by Day')
plt.show()
```



```
sns.scatterplot(data=df, x='total_bill', y='ti
plt.title('Total Bill vs. Tip')
plt.show()
```

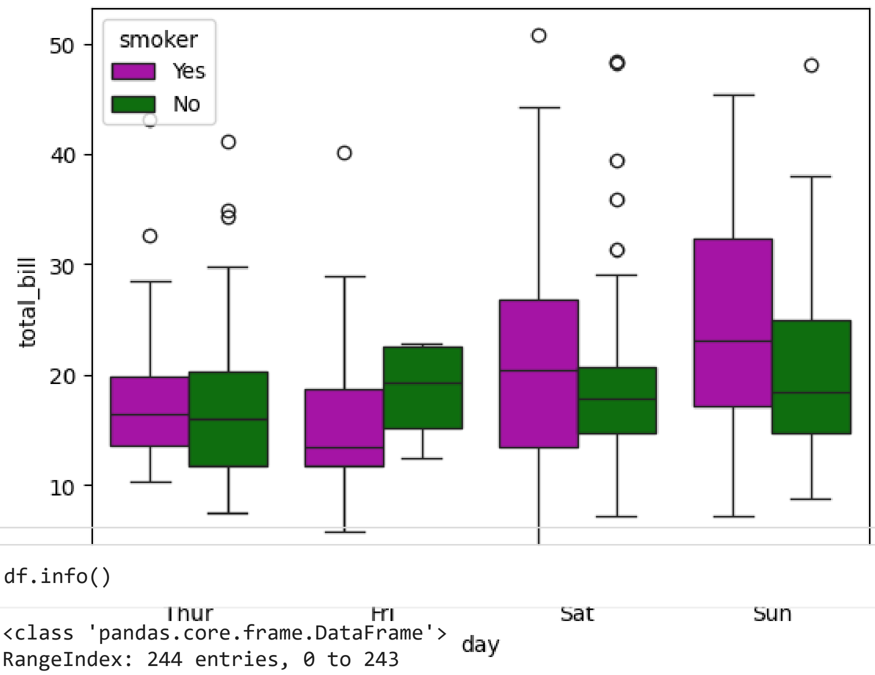


```
sns.scatterplot(data=df, x='total_bill', y='tip', s=5, color=".15")
sns.histplot(data=df, x='total_bill', y='tip', bins=50, pthresh=.1, cmap="mako")
sns.kdeplot(data=df, x='total_bill', y='tip', levels=5, color="w", linewidths=1)
plt.show()
```



```
sns.boxplot(x="day", y="total_bill",
            hue="smoker", palette=["m", "g"],
            data=df)
```

```
<Axes: xlabel='day', ylabel='total_bill'>
```



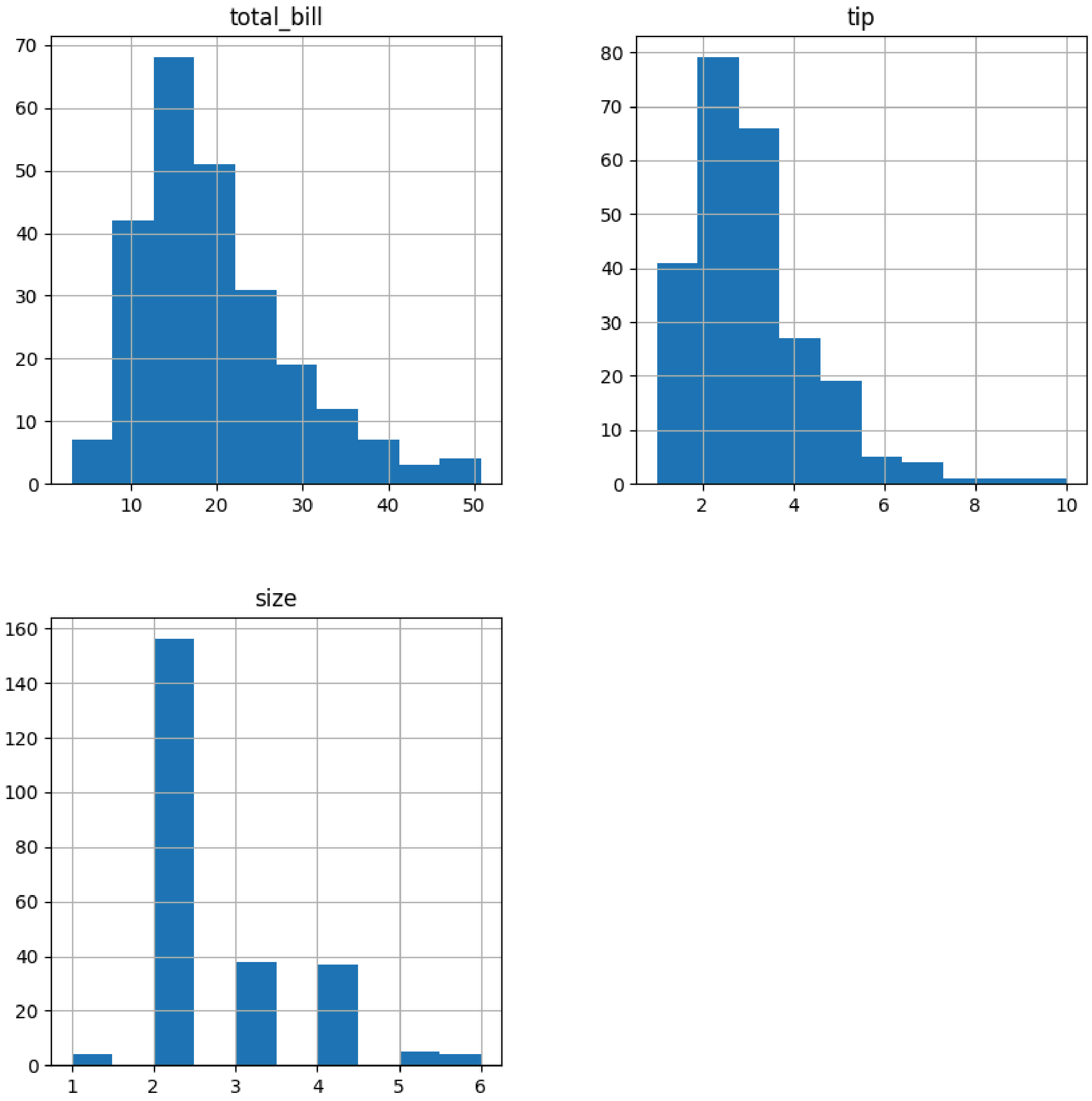
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 244 entries, 0 to 243
Data columns (total 7 columns):
 #   Column      Non-Null Count  Dtype   
---  ------      --------------  -----   
 0   total_bill  244 non-null    float64 
 1   tip         244 non-null    float64 
 2   sex         244 non-null    category
 3   smoker      244 non-null    category
 4   day         244 non-null    category
 5   time        244 non-null    category
 6   size        244 non-null    int64   
dtypes: category(4), float64(2), int64(1)
memory usage: 7.4 KB
```

```
df.describe()
```

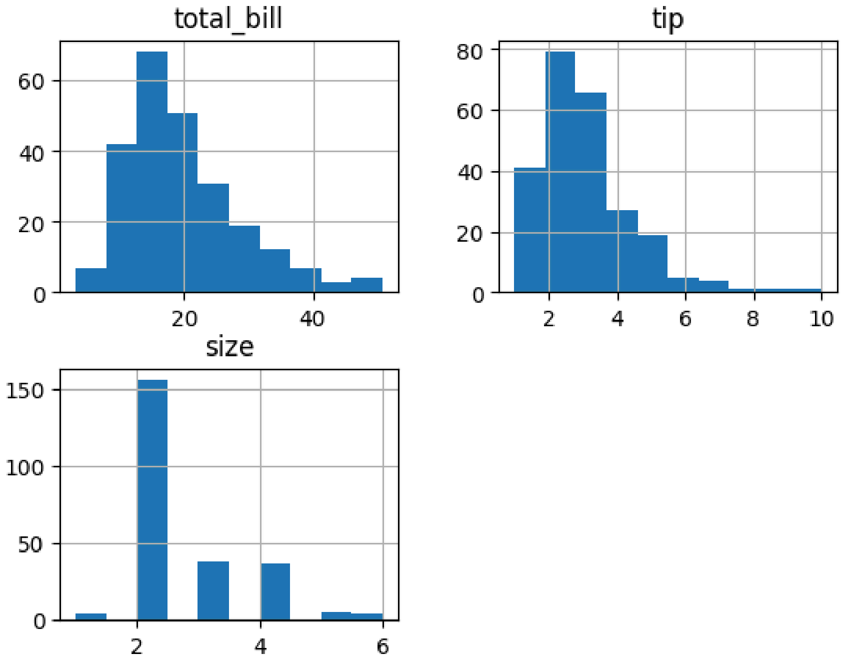
	total_bill	tip	size
count	244.000000	244.000000	244.000000
mean	19.785943	2.998279	2.569672
std	8.902412	1.383638	0.951100
min	3.070000	1.000000	1.000000
25%	13.347500	2.000000	2.000000
50%	17.795000	2.900000	2.000000
75%	24.127500	3.562500	3.000000
max	50.810000	10.000000	6.000000

```
df.hist(figsize=(10,10))
plt.show()
```


total_bill
70
60
50
40
30
20
10
0
10
20
30
40
50
tip
80
70
60
50
40
30
20
10
0
2
4
6
8
10
size
160
140
120
100
80
60
40
20
0
1
2
3
4
5
6

```
df.hist()
```

```
array([[<Axes: title={'center': 'total_bill'}>,
        <Axes: title={'center': 'tip'}>],
       [<Axes: title={'center': 'size'}>, <Axes: >]], dtype=object)
```



```
df.isnull()
```

	total_bill	tip	sex	smoker	day	time	size
0	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False
...	...	...	...	...	...	...	...
239	False	False	False	False	False	False	False
240	False	False	False	False	False	False	False
241	False	False	False	False	False	False	False
242	False	False	False	False	False	False	False
243	False	False	False	False	False	False	False

244 rows × 7 columns

```
df.notnull()
```

	total_bill	tip	sex	smoker	day	time	size
0	True	True	True	True	True	True	True
1	True	True	True	True	True	True	True
2	True	True	True	True	True	True	True
3	True	True	True	True	True	True	True
4	True	True	True	True	True	True	True
...	...	...	...	...	...	...	...
239	True	True	True	True	True	True	True
240	True	True	True	True	True	True	True
241	True	True	True	True	True	True	True
242	True	True	True	True	True	True	True
243	True	True	True	True	True	True	True

244 rows × 7 columns

```
df.dropna()
```

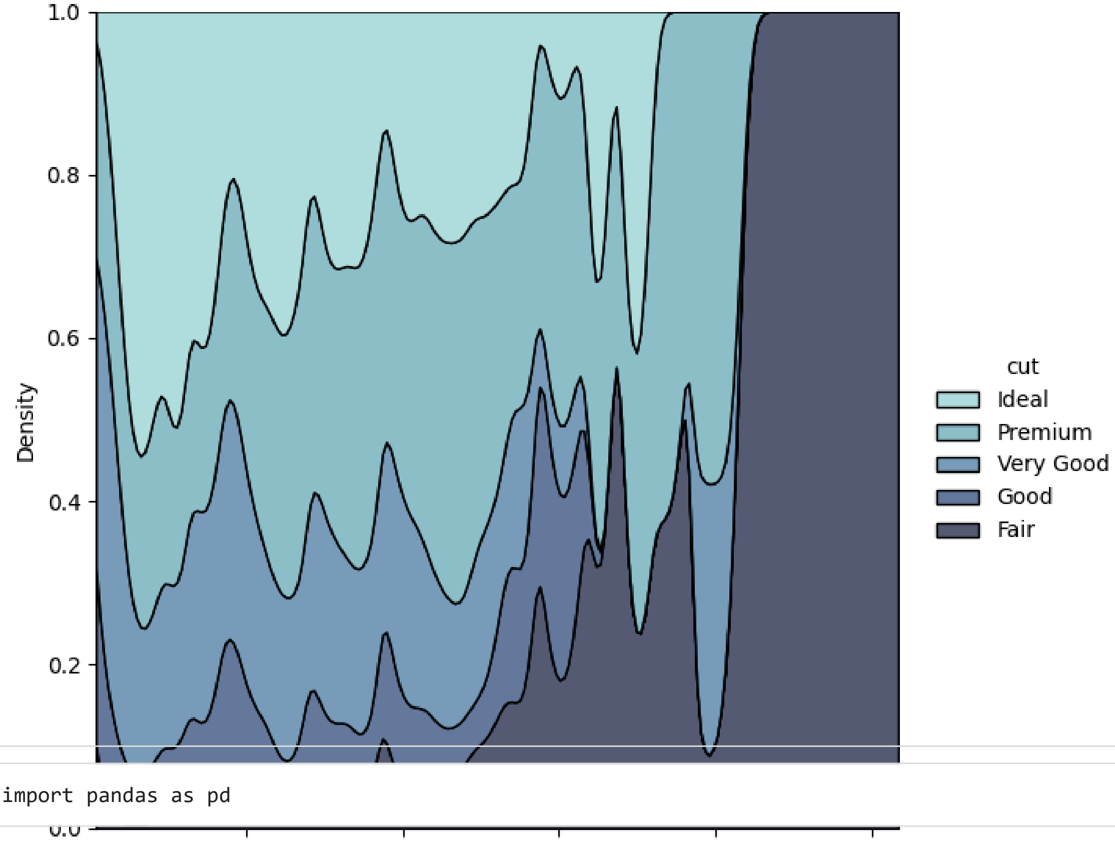
	total_bill	tip	sex	smoker	day	time	size
0	16.99	1.01	Female	No	Sun	Dinner	2
1	10.34	1.66	Male	No	Sun	Dinner	3
2	21.01	3.50	Male	No	Sun	Dinner	3
3	23.68	3.31	Male	No	Sun	Dinner	2
4	24.59	3.61	Female	No	Sun	Dinner	4
...	...	...	...	...	...	...	...
239	29.03	5.92	Male	No	Sat	Dinner	3
240	27.18	2.00	Female	Yes	Sat	Dinner	2
241	22.67	2.00	Male	Yes	Sat	Dinner	2
242	17.82	1.75	Male	No	Sat	Dinner	2
243	18.78	3.00	Female	No	Thur	Dinner	2

244 rows × 7 columns

```
diamonds = sns.load_dataset('diamonds')
```

```
sns.displot(
    data=diamonds,
    x="carat", hue="cut",
    kind="kde", height=6,
    multiple="fill", clip=(0, None),
    palette="ch:rot=-.25,hue=1,light=.75",
)
```

```
<seaborn.axisgrid.FacetGrid at 0x7a22c52543e0>
```



```
import pandas as pd
```

```
data = {'Country': ['Belgium', 'India', 'Brazil'],
'Capital': ['Brussels' , 'New delhi' , 'brazil'],
'Population': [11190846, 130317105, 207847528]}
```

```
df = pd.DataFrame(data,columns=['Country', 'Capital', 'Population'])
```

```
df
```

	Country	Capital	Population
0	Belgium	Brussels	11190846
1	India	New delhi	130317105
2	Brazil	brazil	207847528

```
import pandas as pd
```

```
df = pd.read_csv('car_resale_prices.csv.zip')
```

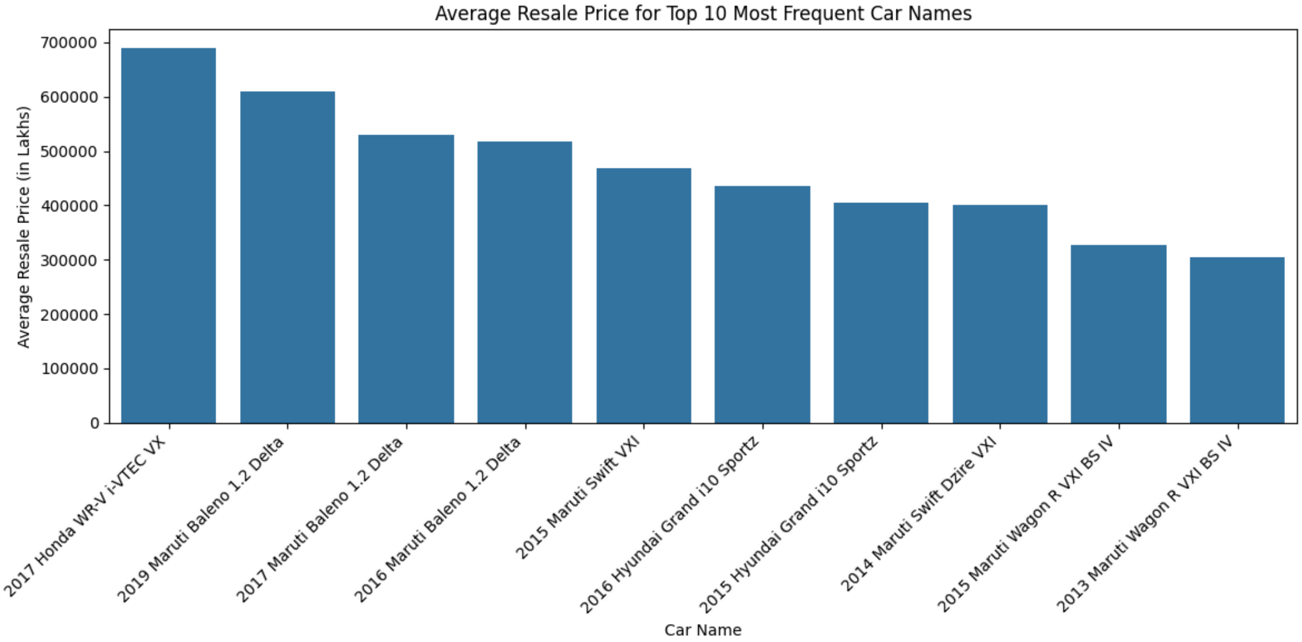
```
df.head(5)
```

	Unnamed: 0	full_name	resale_price	registered_year	engine_capacity	insurance	transmission_type	kms_driv
0	0	2017 Maruti Baleno 1.2 Alpha	₹ 5.45 Lakh	2017	1197 cc	Third Party insurance	Manual	40,000 K
1	1	2018 Tata Hexa XTA	₹ 10 Lakh	2018	2179 cc	Third Party insurance	Automatic	70,000 K
2	2	2015 Maruti Swift Dzire VXI	₹ 4.50 Lakh	2015	1197 cc	Third Party insurance	Manual	70,000 K
3	3	2015 Maruti Swift Dzire VXI	₹ 4.50 Lakh	2015	1197 cc	Third Party insurance	Manual	70,000 K
4	4	2009 Hyundai i10 Magna 1.1	₹ 1.60 Lakh	2009	1086 cc	Third Party insurance	Manual	80,000 K

```
df.groupby('column').name()
```

```
---------------------------------------------------------------------------
KeyError                                  Traceback (most recent call last)
/tmp/ipython-input-2210427764.py in <cell line: 0>()
----> 1 df.groupby('column').name()

                                  2 frames
/usr/local/lib/python3.12/dist-packages/pandas/core/groupby/grouper.py in get_grouper(obj, key, axis, level,
sort, observed, validate, dropna)
   1041                 in_axis, level, gpr = False, gpr, None
   1042             else:
-> 1043                 raise KeyError(gpr)
   1044         elif isinstance(gpr, Grouper) and gpr.key is not None:
   1045             # Add key to exclusions

KeyError: 'column'
```

```
df['resale_price_numeric'] = df['resale_price'].str.replace('₹', '').str.replace('Lakh', 'e5').str.replace('
name_counts = df['full_name'].value_counts()
top_names = name_counts.head(10).index.tolist()
df_top_names = df[df['full_name'].isin(top_names)]


average_price_by_name = df_top_names.groupby('full_name')['resale_price_numeric'].mean().sort_values(ascendi

plt.figure(figsize=(12, 6))
sns.barplot(x=average_price_by_name.index, y=average_price_by_name.values)
plt.title('Average Resale Price for Top 10 Most Frequent Car Names')
plt.xlabel('Car Name')
plt.ylabel('Average Resale Price (in Lakhs)')
plt.xticks(rotation=45, ha='right')
plt.tight_layout()
plt.show()
```



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

```
data=pd.read_csv("bigbasket_products.csv")
```

```
data
```

	Unnamed: 0	product	category	sub_category	brand	sale_price	market_price	
0	0	Original Disinfectant Toilet Cleaner Liquid	Cleaning & Household	All Purpose Cleaners	Harpic	489.00	534.0	https://www.bigbasket.cc
1	1	Disinfectant Surface & Floor Cleaner Liquid - ...	Cleaning & Household	All Purpose Cleaners	Lizol	302.00	380.0	https://www.bigbasket.cc
2	2	Surface Disinfectant Spray	Cleaning & Household	All Purpose Cleaners	Savlon	256.76	318.0	https://www.bigbasket.cc
3	3	Harpic Disinfectant Toilet Cleaner Original200...	Cleaning & Household	All Purpose Cleaners	bb Combo	74.48	76.0	https://www.bigbasket.cc
4	4	Harpic Toilet Cleaner Liquid - Original 1 L + ...	Cleaning & Household	All Purpose Cleaners	bb Combo	462.00	558.0	https://www.bigbasket.cc
...	...	...	...	...	...	...	...	
27550	27550	Canned Sardine - in Tomato Sauce	Eggs, Meat & Fish	Fish & Seafood	Golden Prize	75.05	79.0	https://www.bigbasket.con
27551	27551	Silver Belly	Eggs, Meat & Fish	Fish & Seafood	H.S. Dry Fish	45.00	50.0	https://www.bigbasket.con
27552	27552	Prawns - Pd Xtra Large	Eggs, Meat & Fish	Fish & Seafood	Seastar	246.50	290.0	https://www.bigbasket.con
27553	27553	Ready to Cook - Large Prawns (Peeled & Deveined)	Eggs, Meat & Fish	Fish & Seafood	Buffet	274.50	305.0	https://www.bigbasket.con
27554	27554	Tuna Spread In Mayonnaise	Eggs, Meat & Fish	Fish & Seafood	Golden Prize	185.25	195.0	https://www.bigbasket.con

27555 rows × 13 columns

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 27555 entries, 0 to 27554
Data columns (total 13 columns):
 #   Column        Non-Null Count  Dtype 
---  ------        --------------  ----- 
 0   Unnamed: 0    27555 non-null  int64 
 1   product       27554 non-null  object
 2   category      27555 non-null  object
 3   sub_category  27555 non-null  object
 4   brand         27554 non-null  object 
 5   sale_price    27555 non-null  float64
 6   market_price  27555 non-null  float64
 7   image_url     27555 non-null  object 
 8   p_url         27555 non-null  object 
 9   type          27555 non-null  object 
 10  eancode       26991 non-null  object 
 11  rating        18892 non-null  float64
 12  description   27440 non-null  object 
dtypes: float64(3), int64(1), object(9)
memory usage: 2.7+ MB
```

```
data.head()
```

	Unnamed: 0	product	category	sub_category	brand	sale_price	market_price	
0	0	Original Disinfectant Toilet Cleaner Liquid	Cleaning & Household	All Purpose Cleaners	Harpic	489.00	534.0	https://www.bigbasket.com/med
1	1	Disinfectant Surface & Floor Cleaner Liquid - ...	Cleaning & Household	All Purpose Cleaners	Lizol	302.00	380.0	https://www.bigbasket.com/med
2	2	Surface Disinfectant Spray	Cleaning & Household	All Purpose Cleaners	Savlon	256.76	318.0	https://www.bigbasket.com/med
3	3	Harpic Disinfectant Toilet Cleaner Original200...	Cleaning & Household	All Purpose Cleaners	bb Combo	74.48	76.0	https://www.bigbasket.com/med
4	4	Harpic Toilet Cleaner Liquid - Original 1 L + ...	Cleaning & Household	All Purpose Cleaners	bb Combo	462.00	558.0	https://www.bigbasket.com/med

```
data.isnull()
```

```
   Unnamed:  product  category  sub_category  brand  sale_price  market_price  image_url  p_url  type  ean
          0
0     False    False     False         False  False       False         False      False  False  False
```

```
data.describe().sum()
```

```
                          0
Unnamed: 0    118171.587670
sale_price     41510.227923
market_price   41766.787382
```

```
2     False    False     False         False  False       False         False      False  False  False
3     False    False     False         False  False       False         False      False  False  False
                         False         False  False       False         False      False  False  False
...     ...      ...       ...           ...    ...         ...           ...        ...    ...    ...
                         False         False  False       False         False      False  False  False
```