

Mobile Phone Data Analysis

```
import pandas as pd
```

```
df = pd.read_csv('/content/mobile.csv')
```

```
df.head()
```

	Name	Spec Score	rating	price	img	tag	sim	processor	storage	battery	display	camera	memoryE
0	OPPO Reno 14 Pro 5G	89	4.65	41990	https://cdn1.smartprix.com/rx-is822PXo3-w280-h...	UPCOMING	Dual Sim, 3G, 4G, 5G, VoLTE, Wi-Fi, NFC, IR Bl...	Dimensity 8450, Octa Core, 3.25 GHz Processor	12 GB RAM, 256 GB inbuilt	6200 mAh Battery with 80W Fast Charging	6.83 inches, 1272 x 2800 px, 120 Hz Display wi...	50 MP + 50 MP + 50 MP Triple Rear & 50 MP Fron...	
1	OPPO Reno 14 5G	87	4.75	32990	https://cdn1.smartprix.com/rx-iRGgfcGDH-w280-h...	UPCOMING	Dual Sim, 3G, 4G, 5G, VoLTE, Wi-Fi, NFC, IR Bl...	Dimensity 8350, Octa Core, 3.35 GHz Processor	8 GB RAM, 256 GB inbuilt	6000 mAh Battery with 80W Fast Charging	6.59 inches, 1256 x 2760 px, 120 Hz Display wi...	50 MP + 50 MP + 8 MP Triple Rear & 50 MP Front...	
2	Poco F7 5G	83	4.75	31999	https://cdn1.smartprix.com/rx-icmgBU9Q2-w280-h...	LAUNCHED	Dual Sim, 3G, 4G, 5G, VoLTE, Wi-Fi, NFC, IR Bl...	Snapdragon 8s Gen4, Octa Core, 3.2 GHz Processor	12 GB RAM, 256 GB inbuilt	7550 mAh Battery with 90W Fast Charging	6.83 inches, 1280 x 2772 px, 120 Hz Display wi...	50 MP + 8 MP Dual Rear & 20 MP Front Camera	Mem Not S

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1019 entries, 0 to 1018
Data columns (total 15 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Name                   1019 non-null  object
1   Spec Score             1019 non-null  int64
2   rating                 1019 non-null  float64
3   price                  1019 non-null  int64
4   img                    1019 non-null  object
5   tag                    1019 non-null  object
6   sim                    1019 non-null  object
7   processor              993 non-null   object
8   storage                999 non-null   object
9   battery                1013 non-null  object
10  display                1018 non-null  object
11  camera                 1008 non-null  object
12  memoryExternal         668 non-null   object
13  version                971 non-null   object
14  fm                     294 non-null   object
dtypes: float64(1), int64(2), object(12)
memory usage: 119.5+ KB
```

Data Cleaning & Pre-Processing

In this stage, we will clean the data set by handling missing values, converting data types and trimming extra spaces from textual data.

```
print("Missing values for each column:")
print(df.isnull().sum())
```

```
Missing values for each column:
Name                0
Spec Score          0
rating              0
price               0
img                 0
```

```

tag          0
sim          0
processor    26
storage      20
battery      6
display      1
camera      11
memoryExternal 351
version      48
fm           725
dtype: int64

```

The dataset shows there are some missing values in the following column:

1.Processor 2.Storage 3.Battery 4.Display 5.Camera 6.memoryExternal 7.Version 8.fm

#for the column processor, drop missing values because the column has string data

```

# Drop rows where 'processor' is NaN
df = df.dropna(subset=['processor'])

```

```

# Display the cleaned DataFrame
print(df)

```

```

↔

```

	Name	Spec	Score	rating	price \
2	Poco F7 5G		83	4.75	31999
3	Vivo X200 FE		89	4.65	49990
4	OPPO K13x 5G		73	4.20	11999
7	Vivo T4 Lite 5G		73	4.10	9999
8	Vivo S30 Pro Mini		89	4.65	40990
...
1012	Apple iPhone 16e (512GB)		85	4.35	80999
1013	Vivo X300 Pro Mini		90	4.35	59990
1014	Realme GT 7 Pro Racing Edition		85	4.05	36999
1017	Samsung Galaxy A14 4G (6GB RAM + 128 GB)		74	4.00	9699
1018	Tecno Spark 30C 5G (8GB RAM + 128GB)		75	4.65	12999

	img	tag \
2	https://cdn1.smartprix.com/rx-icmgBU9Q2-w280-h...	LAUNCHED
3	https://cdn1.smartprix.com/rx-iHI7IaOgQ-w280-h...	UPCOMING
4	https://cdn1.smartprix.com/rx-iXUulomIY-w280-h...	LAUNCHED
7	https://cdn1.smartprix.com/rx-ic3ZJSusM-w280-h...	LAUNCHED
8	https://cdn1.smartprix.com/rx-ihxl2mWnr-w280-h...	UPCOMING
...
1012	https://cdn1.smartprix.com/rx-iA5rc7mM3-w280-h...	LAUNCHED
1013	https://cdn1.smartprix.com/rx-inekH0PgW-w280-h...	RUMORED
1014	https://cdn1.smartprix.com/rx-iA2dbnJIw-w280-h...	UPCOMING
1017	https://cdn1.smartprix.com/rx-igcSDy9DV-w280-h...	LAUNCHED
1018	https://cdn1.smartprix.com/rx-ilU0Yr0Ho-w280-h...	LAUNCHED

	sim \
2	Dual Sim, 3G, 4G, 5G, VoLTE, Wi-Fi, NFC, IR Bl...
3	Dual Sim, 3G, 4G, 5G, VoLTE, Wi-Fi, NFC, IR Bl...
4	Dual Sim, 3G, 4G, 5G, VoLTE, Wi-Fi
7	Dual Sim, 3G, 4G, 5G, VoLTE, Wi-Fi
8	Dual Sim, 3G, 4G, 5G, VoLTE, Vo5G, Wi-Fi, NFC,...
...	...
1012	Dual Sim, 3G, 4G, 5G, VoLTE, Wi-Fi, NFC
1013	Dual Sim, 3G, 4G, 5G, VoLTE, Wi-Fi, NFC, IR Bl...
1014	Dual Sim, 3G, 4G, 5G, VoLTE, Wi-Fi, NFC, IR Bl...
1017	Dual Sim, 3G, 4G, VoLTE, Wi-Fi
1018	Dual Sim, 3G, 4G, 5G, VoLTE, Wi-Fi, NFC, IR Bl...

	processor \
2	Snapdragon 8s Gen4, Octa Core, 3.2 GHz Processor
3	Dimensity 9300 Plus, Octa Core, 3.25 GHz Proce...
4	Dimensity 6300, Octa Core, 2.4 GHz Processor
7	Dimensity 6300, Octa Core, 2.4 GHz Processor
8	Dimensity 9300 Plus, Octa Core, 3.4 GHz Processor
...	...
1012	Bionic A18, Hexa Core, 4.04 GHz Processor
1013	Dimensity 9500, Octa Core, 3.63 GHz Processor
1014	Snapdragon 8 Elite, Octa Core, 4.32 GHz Processor
1017	Exynos 850, Octa Core, 2 GHz Processor
1018	Dimensity 6300, Octa Core, 2.4 GHz Processor

	storage	battery \
2	12 GB RAM, 256 GB inbuilt	7550 mAh Battery with 90W Fast Charging
3	12 GB RAM, 256 GB inbuilt	6500 mAh Battery with 90W Fast Charging
4	4 GB RAM, 128 GB inbuilt	6000 mAh Battery with 45W Fast Charging
7	4 GB RAM, 128 GB inbuilt	6000 mAh Battery with 15W Fast Charging
8	12 GB RAM, 256 GB inbuilt	6500 mAh Battery with 90W Fast Charging

```
# for externalMemory
df = df.dropna(subset=['memoryExternal'])
print(df)
```

	Name	Spec	Score	rating	price
2	Poco F7 5G		83	4.75	31999
3	Vivo X200 FE		89	4.65	49990
4	OPPO K13x 5G		73	4.20	11999
7	Vivo T4 Lite 5G		73	4.10	9999
8	Vivo S30 Pro Mini		89	4.65	40990
...
1012	Apple iPhone 16e (512GB)		85	4.35	80999
1013	Vivo X300 Pro Mini		90	4.35	59990
1014	Realme GT 7 Pro Racing Edition		85	4.05	36999
1017	Samsung Galaxy A14 4G (6GB RAM + 128 GB)		74	4.00	9699
1018	Tecno Spark 30C 5G (8GB RAM + 128GB)		75	4.65	12999

	img	tag
2	https://cdn1.smartprix.com/rx-icmgBU9Q2-w280-h...	LAUNCHED
3	https://cdn1.smartprix.com/rx-iHI7Ia0gQ-w280-h...	UPCOMING
4	https://cdn1.smartprix.com/rx-iXUulomTY-w280-h...	LAUNCHED
7	https://cdn1.smartprix.com/rx-ic3ZJ5usM-w280-h...	LAUNCHED
8	https://cdn1.smartprix.com/rx-ihx12mWnr-w280-h...	UPCOMING
...
1012	https://cdn1.smartprix.com/rx-iA5rc7mM3-w280-h...	LAUNCHED
1013	https://cdn1.smartprix.com/rx-inekH0PgW-w280-h...	RUMORED
1014	https://cdn1.smartprix.com/rx-iA2dbnJIw-w280-h...	UPCOMING
1017	https://cdn1.smartprix.com/rx-igcSDy9DV-w280-h...	LAUNCHED
1018	https://cdn1.smartprix.com/rx-ilU0Yr0Ho-w280-h...	LAUNCHED

	sim
2	Dual Sim, 3G, 4G, 5G, VoLTE, Wi-Fi, NFC, IR Bl...
3	Dual Sim, 3G, 4G, 5G, VoLTE, Wi-Fi, NFC, IR Bl...
4	Dual Sim, 3G, 4G, 5G, VoLTE, Wi-Fi
7	Dual Sim, 3G, 4G, 5G, VoLTE, Wi-Fi
8	Dual Sim, 3G, 4G, 5G, VoLTE, Vo5G, Wi-Fi, NFC,...
...	...
1012	Dual Sim, 3G, 4G, 5G, VoLTE, Wi-Fi, NFC
1013	Dual Sim, 3G, 4G, 5G, VoLTE, Wi-Fi, NFC, IR Bl...
1014	Dual Sim, 3G, 4G, 5G, VoLTE, Wi-Fi, NFC, IR Bl...
1017	Dual Sim, 3G, 4G, VoLTE, Wi-Fi
1018	Dual Sim, 3G, 4G, 5G, VoLTE, Wi-Fi, NFC, IR Bl...

	processor
2	Snapdragon 8s Gen4, Octa Core, 3.2 GHz Processor
3	Dimensity 9300 Plus, Octa Core, 3.25 GHz Proce...
4	Dimensity 6300, Octa Core, 2.4 GHz Processor
7	Dimensity 6300, Octa Core, 2.4 GHz Processor
8	Dimensity 9300 Plus, Octa Core, 3.4 GHz Processor
...	...
1012	Bionic A18, Hexa Core, 4.04 GHz Processor
1013	Dimensity 9500, Octa Core, 3.63 GHz Processor
1014	Snapdragon 8 Elite, Octa Core, 4.32 GHz Processor
1017	Exynos 850, Octa Core, 2 GHz Processor
1018	Dimensity 6300, Octa Core, 2.4 GHz Processor

	storage	battery
2	12 GB RAM, 256 GB inbuilt	7550 mAh Battery with 90W Fast Charging
3	12 GB RAM, 256 GB inbuilt	6500 mAh Battery with 90W Fast Charging
4	4 GB RAM, 128 GB inbuilt	6000 mAh Battery with 45W Fast Charging
7	4 GB RAM, 128 GB inbuilt	6000 mAh Battery with 15W Fast Charging
8	12 GB RAM, 256 GB inbuilt	6500 mAh Battery with 90W Fast Charging

```
# Strip whitespace from string columns
str_cols = df.select_dtypes(include=['object']).columns
for col in str_cols:
    df[col] = df[col].str.strip()
```

```
# Drop rows where critical numeric data might be missing
df.dropna(subset=['Spec Score', 'rating', 'price'], inplace=True)
```

```
print(df[['Spec Score', 'rating', 'price']].describe())
```

	Spec Score	rating	price
count	648.000000	648.000000	648.000000
mean	78.327160	4.376620	31161.012346
std	10.274339	0.237039	39498.842233
min	11.000000	3.450000	947.000000
25%	74.000000	4.150000	11999.000000
50%	80.000000	4.400000	18994.500000
75%	84.000000	4.600000	33999.000000
max	96.000000	4.750000	48000.000000

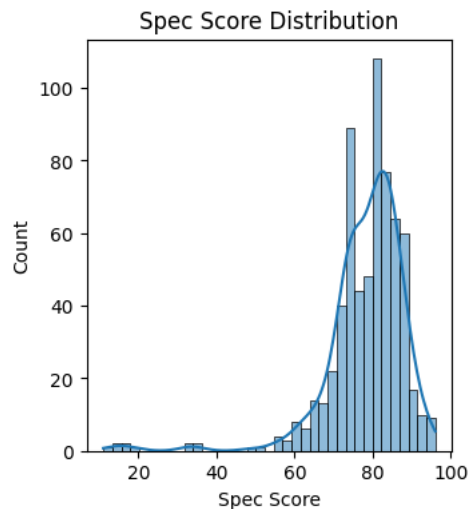
✧ Exploratory Data Analysis

After cleaning the dataset we are now heading towards visualisaion. The following diagrams/figures are representing distribution of numerical data and the frequency of categorical feature.

```
import matplotlib.pyplot as plt
%matplotlib inline
```

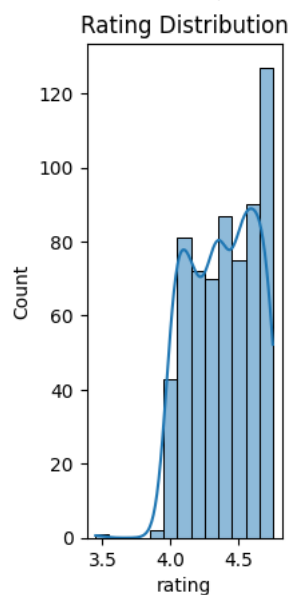
```
import seaborn as sns
# Visualizing distributions with histograms
plt.figure(figsize=(12, 4))
plt.subplot(1, 3, 1)
sns.histplot(df['Spec Score'], kde=True)
plt.title('Spec Score Distribution')
```

```
Text(0.5, 1.0, 'Spec Score Distribution')
```



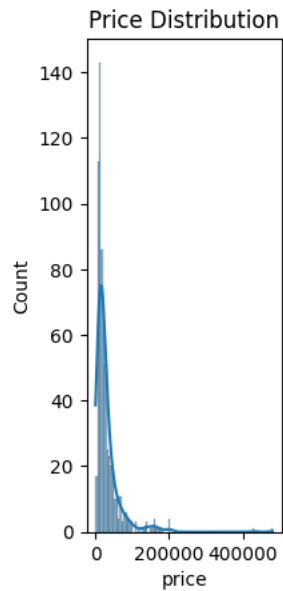
```
plt.subplot(1, 3, 2)
sns.histplot(df['rating'], kde=True)
plt.title('Rating Distribution')
```

```
Text(0.5, 1.0, 'Rating Distribution')
```

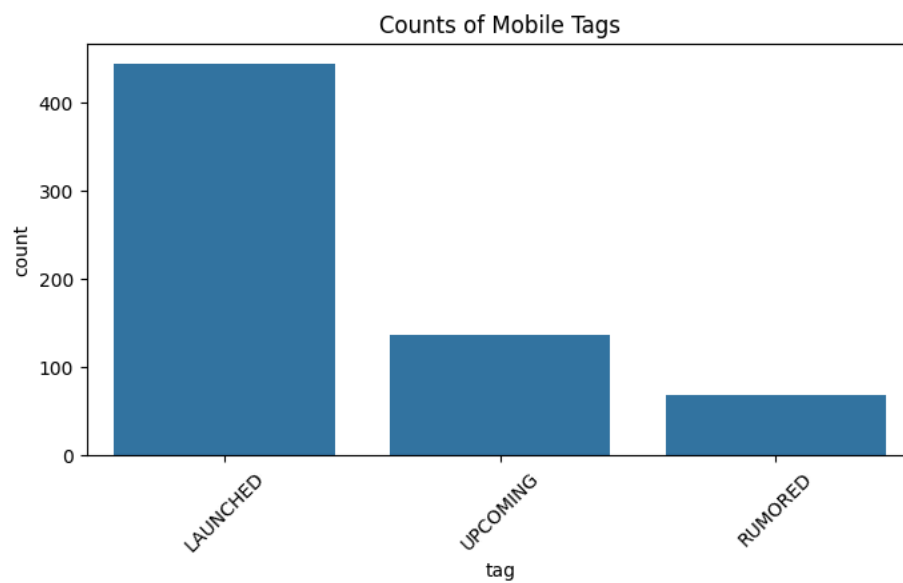


```
plt.subplot(1, 3, 3)
sns.histplot(df['price'], kde=True)
plt.title('Price Distribution')
```

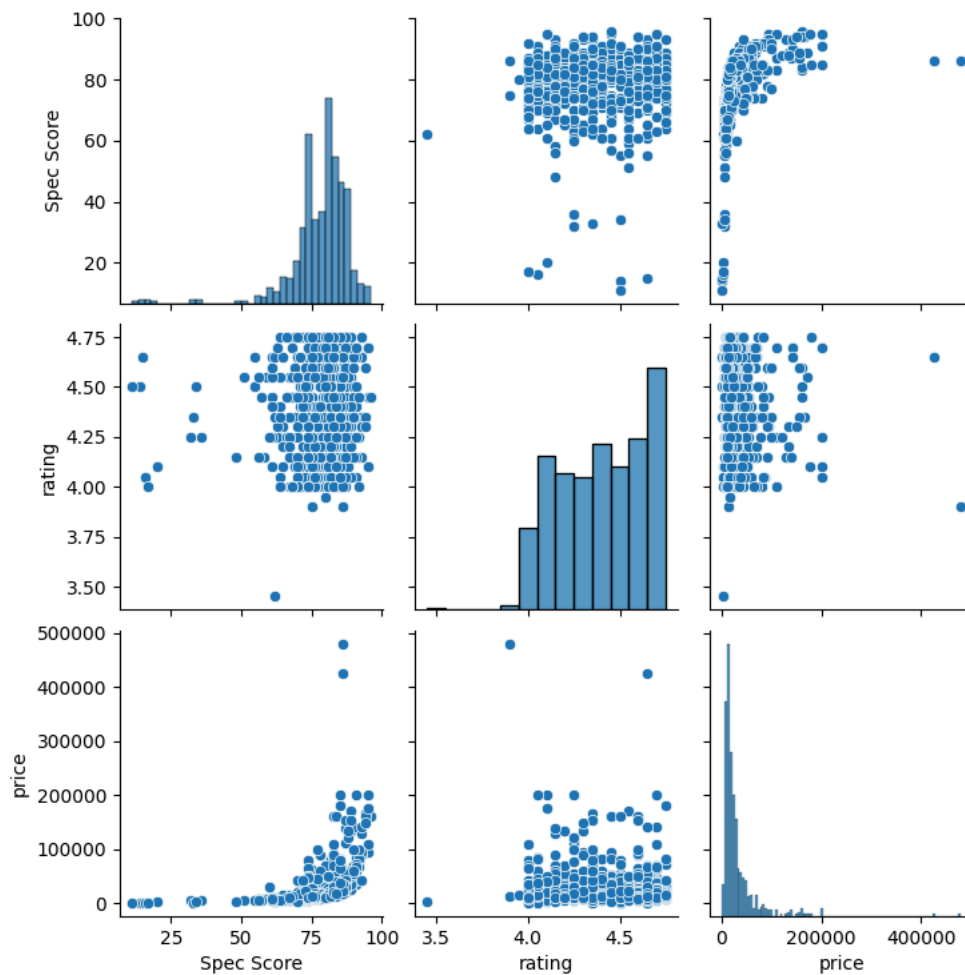
```
Text(0.5, 1.0, 'Price Distribution')
```



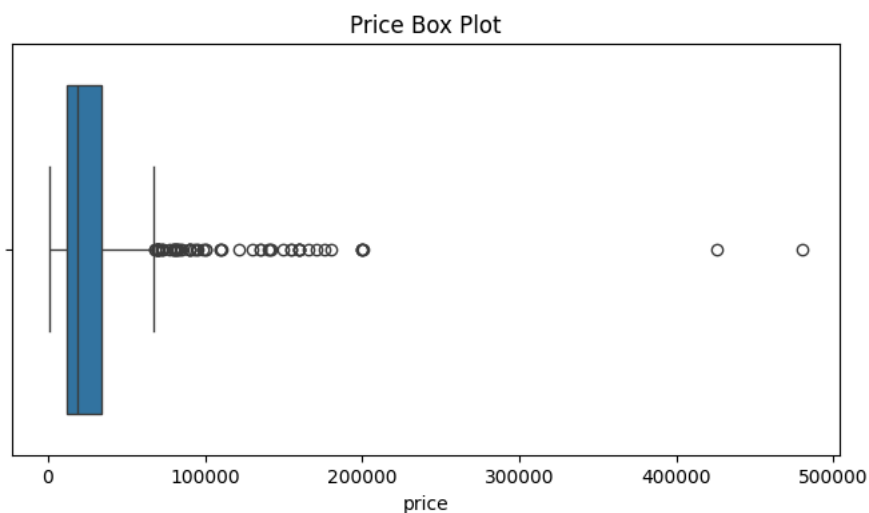
```
# Bar plot for categorical features such as 'tag'
plt.figure(figsize=(8, 4))
sns.countplot(x='tag', data=df)
plt.title('Counts of Mobile Tags')
plt.xticks(rotation=45)
plt.show()
```



```
# Pair Plot for numerical relationships
sns.pairplot(df[['Spec Score', 'rating', 'price']])
plt.show()
```



```
# Box Plot for Price to check for outliers
plt.figure(figsize=(8, 4))
sns.boxplot(x=df['price'])
plt.title('Price Box Plot')
plt.show()
```



```
#for heat map
numeric_cols = ['Spec Score', 'rating', 'price']
df_subset = df[numeric_cols]

# Compute the correlation matrix
correlation_matrix = df_subset.corr()

# Plot the heatmap
sns.heatmap(correlation_matrix, annot=True, fmt=".2f", cmap='coolwarm', linewidths=0.5)

# Customize plot
plt.title('Correlation Heatmap: Spec Score, Rating, and Price')
plt.xticks(rotation=45)
```

```
plt.yticks(rotation=0)
```

```
plt.show()
```

