Shark Tank Indian Company Data Analysis

Importing Libraries

```
In [67]:  import numpy as np
          import pandas as pd
          import seaborn as sns
          import matplotlib.pyplot as plt
```

Finding file path

```
In [3]:   import os
          os.listdir()
```

Out[3]:   ['.ipynb_checkpoints', 'ShartankIndiaAllPitches.csv', 'Untitled.ipynb']

Reading Data

```
In [5]:   df = pd.read_csv('ShartankIndiaAllPitches.csv')
          df.head()
```

Out[5]:

	Episode Number	Brand	Investment Amount (In Lakhs INR)	Debt (In lakhs INR)	Equity	Anupam	Ashneer	Namita	Aman	Peyush	Vin
0	1	BluePine Industries	75	0	18%	N	Y	N	Y	N	
1	1	Booz scooters	40	0	50%	N	Y	N	N	N	
2	1	Heart up my Sleeves	25	0	30%	Y	N	N	N	N	
3	2	Tagz Foods	70	0	3%	N	Y	N	N	N	
4	2	Head and Heart	0	0	0%	N	N	N	N	N	

checking null values

In [6]:
```python
df_null = df.isna().sum()
print('Number of null values: - \n',df_null)
```

```
Number of null values: -
 Episode Number                      0
Brand                                0
Investment Amount (In Lakhs INR)     0
Debt (In lakhs INR)                  0
Equity                               0
Anupam                               0
Ashneer                              0
Namita                               0
Aman                                 0
Peyush                               0
Vineeta                              0
Ghazal                               0
Season                               0
dtype: int64
```

Statistical/Discriptive Analysis

In [7]:
```python
df.describe()
```

Out[7]:

	Episode Number	Investment Amount (In Lakhs INR)	Debt (In lakhs INR)	Season
count	117.000000	117.000000	117.000000	117.0
mean	18.735043	21.683761	2.572650	1.0
std	10.070778	26.067766	11.544753	0.0
min	1.000000	0.000000	0.000000	1.0
25%	10.000000	0.000000	0.000000	1.0
50%	19.000000	10.000000	0.000000	1.0
75%	27.000000	50.000000	0.000000	1.0
max	35.000000	80.000000	99.000000	1.0

In [10]:
```python
df.describe().transpose()
```

Out[10]:

	count	mean	std	min	25%	50%	75%	max
Episode Number	117.0	18.735043	10.070778	1.0	10.0	19.0	27.0	35.0
Investment Amount (In Lakhs INR)	117.0	21.683761	26.067766	0.0	0.0	10.0	50.0	80.0
Debt (In lakhs INR)	117.0	2.572650	11.544753	0.0	0.0	0.0	0.0	99.0
Season	117.0	1.000000	0.000000	1.0	1.0	1.0	1.0	1.0

Anupam Invest/Debt Summary

In [27]:
```python
df_count_Anupam = df[df["Anupam"]=="Y"]["Anupam"].value_counts()
print("Anupam Investing Total Companies: - ", df_count_Anupam)
print()

df_Invest_Anupam = df[df['Anupam']=="Y"]['Investment Amount (In Lakhs INR)'].su
print('Anupam Invest Amount in Lakhs: - ',df_Invest_Anupam)
print()

df_Debt_Amount = df[df['Anupam']=="Y"]['Debt (In lakhs INR)'].sum()
print('Anupam Debt Amount in Lakhs: - ',df_Debt_Amount)
```

```
Anupam Investing Total Companies: -  Y    24
Name: Anupam, dtype: int64

Anupam Invest Amount in Lakhs: -  816

Anupam Debt Amount in Lakhs: -  30
```

Finding all investor deatils by using For loop

In [46]:
```python
df.columns
```

Out[46]:
```
Index(['Episode Number', 'Brand', 'Investment Amount (In Lakhs INR)',
       'Debt (In lakhs INR)', 'Equity', 'Anupam', 'Ashneer', 'Namita', 'Ama
n',
       'Peyush', 'Vineeta', 'Ghazal', 'Season'],
      dtype='object')
```

In [47]:
```python
# Creating Investor List
Investor_List = ['Anupam', 'Ashneer', 'Namita', 'Aman',
        'Peyush', 'Vineeta', 'Ghazal']
```

In [64]:
```python
for investor in Investor_List:

    df_Invset = df[df[investor]=='Y']['Investment Amount (In Lakhs INR)'].sum(
    df_Debt = df[df[investor]=="Y"]['Debt (In lakhs INR)'].sum()
    df_count = df[df[investor]=="Y"][investor].value_counts()

    print(f'{investor} Invest Amount in Lakhs: - ',df_Invset)
    print(f'{investor} Debt in Lakhs: - ',df_Debt)
    print(f'{investor} Investing Total Companies: - ',df_count)

    print()
```

```
Anupam Invest Amount in Lakhs: -  816
Anupam Debt in Lakhs: -  30
Anupam Investing Total Companies: -  Y    24
Name: Anupam, dtype: int64

Ashneer Invest Amount in Lakhs: -  731
Ashneer Debt in Lakhs: -  129
Ashneer Investing Total Companies: -  Y    21
Name: Ashneer, dtype: int64

Namita Invest Amount in Lakhs: -  870
Namita Debt in Lakhs: -  0
Namita Investing Total Companies: -  Y    22
Name: Namita, dtype: int64

Aman Invest Amount in Lakhs: -  1060
Aman Debt in Lakhs: -  50
Aman Investing Total Companies: -  Y    28
Name: Aman, dtype: int64

Peyush Invest Amount in Lakhs: -  886
Peyush Debt in Lakhs: -  92
Peyush Investing Total Companies: -  Y    27
Name: Peyush, dtype: int64

Vineeta Invest Amount in Lakhs: -  570
Vineeta Debt in Lakhs: -  30
Vineeta Investing Total Companies: -  Y    15
Name: Vineeta, dtype: int64

Ghazal Invest Amount in Lakhs: -  245
Ghazal Debt in Lakhs: -  0
Ghazal Investing Total Companies: -  Y    7
Name: Ghazal, dtype: int64
```

In [227]:
```python
total_invested = 0
for investor in investor_list:
    invested_amount = df[df[investor] == 'Y']['Investment Amount (In Lakhs INR
    total_invested += invested_amount

print('Total Invested Amount by all Investors:', total_invested)
```

Total Invested Amount by all Investors: 5178

In [231]:
```python
debt_amounts = {}
total_debt = 0

for investor in investor_list:
    debt_amount = df[df[investor] == 'Y']['Debt (In lakhs INR)'].sum()
    debt_amounts[investor] = debt_amount
    total_debt += debt_amount

print('Total Debt Amount:', total_debt)
```

Total Debt Amount: 331

Data Visualization

In [87]:
```python
# Invesrment for each Brand
plt.figure(figsize = (15,6))
sns.set(font_scale=1.2)
sns.barplot(x='Brand', y='Investment Amount (In Lakhs INR)', data=df)
plt.xticks(rotation=90,fontsize=8)
plt.yticks(fontsize=10)
plt.title('Investment for each Brand',fontsize=20)
plt.show()
```
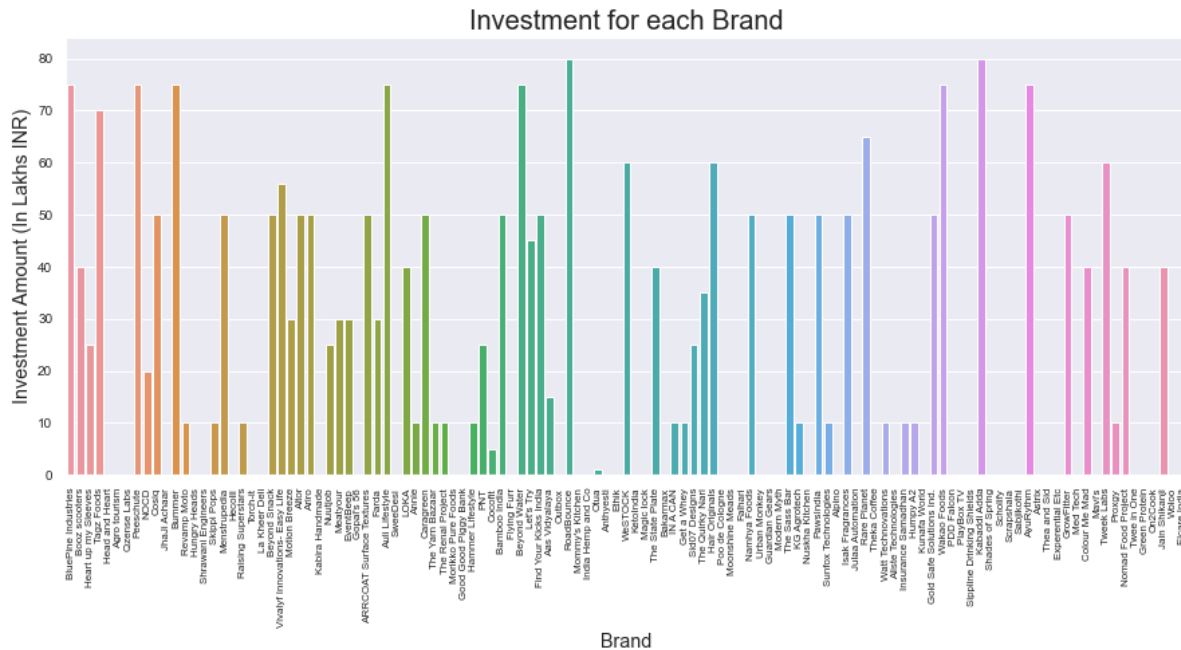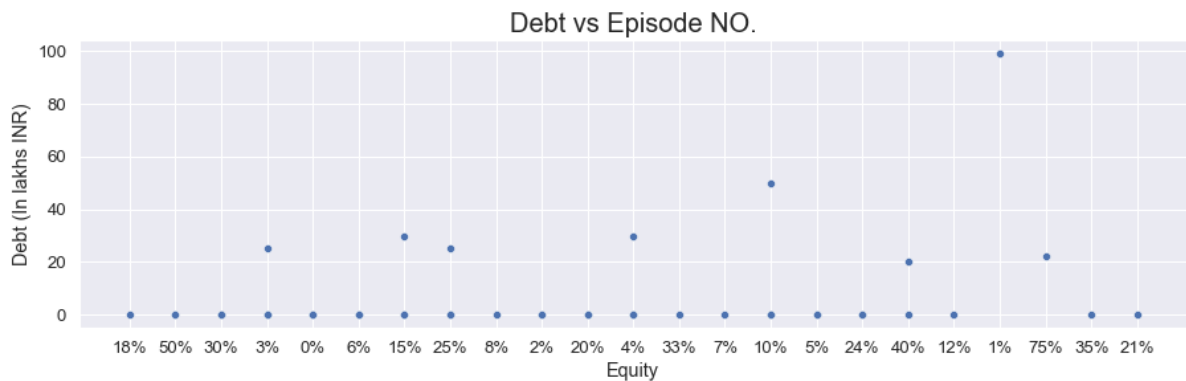
In [91]:
```python
# Investment for Episode
plt.figure(figsize=(15,6))
sns.set(font_scale=1.2)
sns.lineplot(x='Episode Number',y='Investment Amount (In Lakhs INR)',data=df)
plt.title('Investment by Episode NO.',fontsize=20)
plt.show()
```


In [100]:
```python
# Correlation between Equity and Investment
plt.figure(figsize=(15,6))
sns.scatterplot(x='Equity',y='Investment Amount (In Lakhs INR)',data=df)
plt.title('Investment vs Episode NO.',fontsize=20)
plt.show()
```

In [235]:
```python
# Correlation between Equity and Debt
plt.figure(figsize=(15,4))
sns.scatterplot(x='Equity',y='Debt (In lakhs INR)',data=df)
plt.title('Debt vs Episode NO.',fontsize=20)
plt.show()
```


Thank You

In [ ]: