

Shark Tank EDA



```
In [1]: import numpy as np  
import pandas as pd  
import matplotlib.pyplot as plt  
import seaborn as sns  
%matplotlib inline
```

```
In [2]: shark_data=pd.read_csv("C:/Users/em/Downloads/Shark Tank India Dataset.csv")
```

In [3]: `shark_data # check the data`

Out[3]:

	episode_number	pitch_number	brand_name	idea	deal	pitcher_ask_amount	ask_equity	ask_valuation	deal_amount
0	1	1	BluePine Industries	Frozen Momos	1	50.0	5.0	1000.00	75.0
1	1	2	Booz scooters	Renting e-bike for mobility in private spaces	1	40.0	15.0	266.67	40.0
2	1	3	Heart up my Sleeves	Detachable Sleeves	1	25.0	10.0	250.00	25.0
3	2	4	Tagz Foods	Healthy Potato Chips	1	70.0	1.0	7000.00	70.0
4	2	5	Head and Heart	Brain Development Course	0	50.0	5.0	1000.00	0.0
...
112	34	113	Green Protein	Plant-Based Protein	0	60.0	2.0	3000.00	0.0
113	34	114	On2Cook	Fastest Cooking Device	0	100.0	1.0	10000.00	0.0
114	35	115	Jain Shikanji	Lemonade	1	40.0	8.0	500.00	40.0
115	35	116	Woloo	Washroom Finder	0	50.0	4.0	1250.00	0.0
116	35	117	Elcare India	Carenting for Elders	0	100.0	2.5	4000.00	0.0

117 rows × 28 columns



In [4]: `shark_data.head(5) # to check top 5 record`

Out[4]:

	episode_number	pitch_number	brand_name	idea	deal	pitcher_ask_amount	ask_equity	ask_valuation	deal_amount
0	1	1	BluePine Industries	Frozen Momos	1	50.0	5.0	1000.00	75.0
1	1	2	Booz scooters	Renting e-bike for mobility in private spaces	1	40.0	15.0	266.67	40.0
2	1	3	Heart up my Sleeves	Detachable Sleeves	1	25.0	10.0	250.00	25.0
3	2	4	Tagz Foods	Healthy Potato Chips	1	70.0	1.0	7000.00	70.0
4	2	5	Head and Heart	Brain Development Course	0	50.0	5.0	1000.00	0.0

5 rows × 28 columns



In [5]: `shark_data.tail(5) # to check bottom 5 record`

Out[5]:

	episode_number	pitch_number	brand_name	idea	deal	pitcher_ask_amount	ask_equity	ask_valuation	deal_amount
112	34	113	Green Protein	Plant-Based Protein	0	60.0	2.0	3000.0	0.0
113	34	114	On2Cook	Fastest Cooking Device	0	100.0	1.0	10000.0	0.0
114	35	115	Jain Shikanji	Lemonade	1	40.0	8.0	500.0	40.0
115	35	116	Woloo	Washroom Finder	0	50.0	4.0	1250.0	0.0
116	35	117	Elcare India	Caring for Elders	0	100.0	2.5	4000.0	0.0

5 rows × 28 columns

In [6]: `shark_data.shape # to check the number of rows and column in data`

Out[6]: (117, 28)

In [7]: `# basic information of data
shark_data.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 117 entries, 0 to 116
Data columns (total 28 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   episode_number    117 non-null    int64  
 1   pitch_number      117 non-null    int64  
 2   brand_name        117 non-null    object  
 3   idea              117 non-null    object  
 4   deal              117 non-null    int64  
 5   pitcher_ask_amount 117 non-null    float64 
 6   ask_equity        117 non-null    float64 
 7   ask_valuation     117 non-null    float64 
 8   deal_amount       117 non-null    float64 
 9   deal_equity       117 non-null    float64 
 10  deal_evaluation   117 non-null    float64 
 11  ashneer_present   117 non-null    int64  
 12  anupam_present    117 non-null    int64  
 13  aman_present      117 non-null    int64  
 14  namita_present    117 non-null    int64  
 15  vineeta_present   117 non-null    int64  
 16  peyush_present    117 non-null    int64  
 17  ghazal_present    117 non-null    int64  
 18  ashneer_deal      117 non-null    int64  
 19  anupam_deal       117 non-null    int64  
 20  aman_deal         117 non-null    int64  
 21  namita_deal       117 non-null    int64  
 22  vineeta_deal      117 non-null    int64  
 23  peyush_deal       117 non-null    int64  
 24  ghazal_deal       117 non-null    int64  
 25  total_sharks_invested 117 non-null    int64  
 26  amount_per_shark  117 non-null    float64 
 27  equity_per_shark  117 non-null    float64 

dtypes: float64(8), int64(18), object(2)
memory usage: 25.7+ KB
```

In [8]: `shark_data.isnull().sum() # to check the null value present in our dataset`

Out[8]:

episode_number	0
pitch_number	0
brand_name	0
idea	0
deal	0
pitcher_ask_amount	0
ask_equity	0
ask_valuation	0
deal_amount	0
deal_equity	0
deal_valuation	0
ashneer_present	0
anupam_present	0
aman_present	0
namita_present	0
vineeta_present	0
peyush_present	0
ghazal_present	0
ashneer_deal	0
anupam_deal	0
aman_deal	0
namita_deal	0
vineeta_deal	0
peyush_deal	0
ghazal_deal	0
total_sharks_invested	0
amount_per_shark	0
equity_per_shark	0
dtype:	int64

In [9]: `shark_data.describe()`

Out[9]:

	episode_number	pitch_number	deal	pitcher_ask_amount	ask_equity	ask_valuation	deal_amount	deal_equity	de...
count	117.000000	117.000000	117.000000	117.000000	117.000000	117.000000	117.000000	117.000000	117.000000
mean	18.735043	59.000000	0.555556	319.854709	5.188034	3852.462479	31.982915	8.963504	8.963504
std	10.070778	33.919021	0.499041	2767.842777	3.892121	11931.601957	36.687391	13.106769	13.106769
min	1.000000	1.000000	0.000000	0.001010	0.250000	0.010000	0.000000	0.000000	0.000000
25%	10.000000	30.000000	0.000000	45.000000	2.500000	666.670000	0.000000	0.000000	0.000000
50%	19.000000	59.000000	1.000000	50.000000	5.000000	1250.000000	25.000000	3.000000	3.000000
75%	27.000000	88.000000	1.000000	80.000000	7.500000	2857.140000	50.000000	15.000000	15.000000
max	35.000000	117.000000	1.000000	30000.000000	25.000000	120000.000000	150.000000	75.000000	75.000000

8 rows × 26 columns

In [10]: `shark_data.columns # to check all the columns present in our dataset`

Out[10]:

```
Index(['episode_number', 'pitch_number', 'brand_name', 'idea', 'deal',
       'pitcher_ask_amount', 'ask_equity', 'ask_valuation', 'deal_amount',
       'deal_equity', 'deal_valuation', 'ashneer_present', 'anupam_present',
       'aman_present', 'namita_present', 'vineeta_present', 'peyush_present',
       'ghazal_present', 'ashneer_deal', 'anupam_deal', 'aman_deal',
       'namita_deal', 'vineeta_deal', 'peyush_deal', 'ghazal_deal',
       'total_sharks_invested', 'amount_per_shark', 'equity_per_shark'],
      dtype='object')
```

In [11]: `shark_data.select_dtypes(include=['object'])`

Out[11]:

	brand_name	idea
0	BluePine Industries	Frozen Momos
1	Booz scooters	Renting e-bike for mobility in private spaces
2	Heart up my Sleeves	Detachable Sleeves
3	Tagz Foods	Healthy Potato Chips
4	Head and Heart	Brain Development Course
...
112	Green Protein	Plant-Based Protein
113	On2Cook	Fastest Cooking Device
114	Jain Shikanji	Lemonade
115	Woloo	Washroom Finder
116	Elcare India	Carenting for Elders

117 rows × 2 columns

There are 2 columns which is in categorial value

In [12]: `shark_data.select_dtypes(include=['int', 'float'])`

Out[12]:

	episode_number	pitch_number	deal	pitcher_ask_amount	ask_equity	ask_valuation	deal_amount	deal_equity	deal_valuation
0	1	1	1	50.0	5.0	1000.00	75.0	16.00	468.0
1	1	2	1	40.0	15.0	266.67	40.0	50.00	80.0
2	1	3	1	25.0	10.0	250.00	25.0	30.00	83.0
3	2	4	1	70.0	1.0	7000.00	70.0	2.75	2545.0
4	2	5	0	50.0	5.0	1000.00	0.0	0.00	0.0
...
112	34	113	0	60.0	2.0	3000.00	0.0	0.00	0.0
113	34	114	0	100.0	1.0	10000.00	0.0	0.00	0.0
114	35	115	1	40.0	8.0	500.00	40.0	30.00	133.0
115	35	116	0	50.0	4.0	1250.00	0.0	0.00	0.0
116	35	117	0	100.0	2.5	4000.00	0.0	0.00	0.0

117 rows × 26 columns

In [13]: # check the data of first 10 columns
shark_data.iloc[0:5,0:10]

Out[13]:

	episode_number	pitch_number	brand_name	idea	deal	pitcher_ask_amount	ask_equity	ask_valuation	deal_amount
0	1	1	BluePine Industries	Frozen Momos	1	50.0	5.0	1000.00	75.0
1	1	2	Booz scooters	Renting e-bike for mobility in private spaces	1	40.0	15.0	266.67	40.0
2	1	3	Heart up my Sleeves	Detachable Sleeves	1	25.0	10.0	250.00	25.0
3	2	4	Tagz Foods	Healthy Potato Chips	1	70.0	1.0	7000.00	70.0
4	2	5	Head and Heart	Brain Development Course	0	50.0	5.0	1000.00	0.0

In [14]: # check data 10 to 20 columns
shark_data.iloc[0:5,10:21]

Out[14]:

	deal_valuation	ashneer_present	anupam_present	aman_present	namita_present	vineeta_present	peyush_present	ghazal_pres
0	468.75	1	1	1	1	1	1	0
1	80.00	1	1	1	1	1	1	0
2	83.33	1	1	1	1	1	1	0
3	2545.45	1	1	1	1	1	1	0
4	0.00	1	1	1	1	1	1	0

2: Cleaning the data

Checking missing values if yes then handle them accordingly

- we see that from above analysis there is no missing and null values

In [15]: # to check duplicate record
shark_data.duplicated()

Out[15]:

```
0      False
1      False
2      False
3      False
4      False
...
112    False
113    False
114    False
115    False
116    False
Length: 117, dtype: bool
```

- we can see that there is no duplicate record present so no need to drop the columns

INSIGHTS

In [16]: `shark_data.head(2)`

Out[16]:

	episode_number	pitch_number	brand_name	idea	deal	pitcher_ask_amount	ask_equity	ask_valuation	deal_amount	deal
0	1	1	BluePine Industries	Frozen Momos	1	50.0	5.0	1000.00	75.0	
1	1	2	Booz scooters	Renting e-bike for mobility in private spaces	1	40.0	15.0	266.67	40.0	

2 rows × 28 columns



How many Duplicated rows present

In [19]: `shark_data[shark_data.duplicated()]`

Out[19]:

	episode_number	pitch_number	brand_name	idea	deal	pitcher_ask_amount	ask_equity	ask_valuation	deal_amount	deal_equ
0	0	0								

0 rows × 28 columns



- There is no duplicate rows present in our dataset

Find the number of episodes

In [20]: `shark_data['episode_number'].nunique()`

Out[20]: 35

- Total number of 35 episode in shark tank

In [21]: `# to check the count of number of deal done or not
shark_data['deal'].value_counts()`

Out[21]:

1	65
0	52
Name: deal, dtype: int64	

- Total number of deal done is 1 is 65
- Total number of deal not done is 0 is 52

In [22]: `# to check the maximum deal amount done on the episode
shark_data[shark_data['deal_amount']==shark_data['deal_amount'].max()]`

Out[22]:

	episode_number	pitch_number	brand_name	idea	deal	pitcher_ask_amount	ask_equity	ask_valuation	deal_amount	dea
50	17	51	Aas Vidyalaya	EdTech App	1	150.0	3.0	5000.0	150.0	

1 rows × 28 columns



In [23]: # to check the maximum ask valuation in the episode

```
shark_data[shark_data['ask_valuation']==shark_data['ask_valuation'].max()]
```

Out[23]:

episode_number	pitch_number	brand_name	idea	deal	pitcher_ask_amount	ask_equity	ask_valuation	deal_amount	deal
30	11	31 Gopal's 56	Fiber Ice Cream	0	30000.0	25.0	120000.0	0.0	

1 rows × 28 columns

- Maximum ask valuation is 120000

In [24]: # to check the minimum equity ask in the episode

```
shark_data[shark_data['ask_equity']==shark_data['ask_equity'].min()]
```

Out[24]:

episode_number	pitch_number	brand_name	idea	deal	pitcher_ask_amount	ask_equity	ask_valuation	deal_amount	de
6	3	7 Qzense Labs	Food Freshness Detector	0	100.0	0.25	40000.0	0.0	

1 rows × 28 columns

- Minimum asked equity is 0.25

In [25]: # to check maximum ask valuation

```
shark_data[shark_data['ask_valuation']==shark_data['ask_valuation'].max()]
```

Out[25]:

episode_number	pitch_number	brand_name	idea	deal	pitcher_ask_amount	ask_equity	ask_valuation	deal_amount	deal
30	11	31 Gopal's 56	Fiber Ice Cream	0	30000.0	25.0	120000.0	0.0	

1 rows × 28 columns

Max, min , mean of asked amount,asked equity, asked valuation ?

In [26]: shark_data[['pitcher_ask_amount', 'ask_equity', 'ask_valuation']].agg([max,min,'mean'])

Out[26]:

	pitcher_ask_amount	ask_equity	ask_valuation
max	30000.000000	25.000000	120000.000000
min	0.001010	0.250000	0.010000
mean	319.854709	5.188034	3852.462479

Max & Min asked- equity,asked-valuation and asked amount episode-wise ?

```
In [27]: shark_data.head(2)
```

Out[27]:

	episode_number	pitch_number	brand_name	idea	deal	pitcher_ask_amount	ask_equity	ask_valuation	deal_amount	deal
0	1	1	BluePine Industries	Frozen Momos	1	50.0	5.0	1000.00	75.0	
1	1	2	Booz scooters	Renting e-bike for mobility in private spaces	1	40.0	15.0	266.67	40.0	

2 rows × 28 columns



In [28]: `shark_data.groupby(['episode_number'])[['pitcher_ask_amount','ask_equity','ask_valuation']].agg([max,min])`

```
C:\Users\em\AppData\Local\Temp\ipykernel_4552\3623356012.py:1: FutureWarning: Indexing with multiple keys (implicitly converted to a tuple of keys) will be deprecated, use a list instead.
    shark_data.groupby(['episode_number'])[['pitcher_ask_amount','ask_equity','ask_valuation']].agg([max,
min])
```

Out[28]:

episode_number	pitcher_ask_amount		ask_equity		ask_valuation	
	max	min	max	min	max	min
1	50.0	25.00000	15.0	5.00	1000.00	250.00
2	70.0	50.00000	5.0	1.00	7000.00	1000.00
3	100.0	50.00000	4.0	0.25	40000.00	1875.00
4	75.0	50.00000	10.0	4.00	1875.00	500.00
5	100.0	10.00000	20.0	1.00	10000.00	50.00
6	100.0	45.00000	10.0	1.00	10000.00	500.00
7	100.0	50.00000	7.5	1.00	7500.00	666.67
8	56.0	30.00000	7.5	2.50	2000.00	746.67
9	100.0	50.00000	5.0	2.50	2000.00	1000.00
10	30.0	25.00000	5.0	2.00	1500.00	500.00
11	30000.0	30.00000	25.0	5.00	120000.00	300.00
12	75.0	40.00000	5.0	3.00	1875.00	800.00
13	50.0	30.00000	10.0	2.00	2500.00	500.00
14	100.0	45.00000	5.0	3.00	3333.33	900.00
15	50.0	5.00000	5.0	3.00	1250.00	100.00
16	80.0	45.00000	7.0	2.00	2250.00	1071.43
17	150.0	50.00000	10.0	3.00	5000.00	500.00
18	100.0	50.00000	4.0	1.00	10000.00	1250.00
19	125.0	15.00000	8.0	1.25	10000.00	300.00
20	65.0	35.00000	5.0	2.00	3250.00	700.00
21	100.0	35.00000	10.0	5.00	1250.00	470.00
22	80.0	50.00000	5.0	2.00	3000.00	1500.00
23	100.0	30.00000	5.0	1.00	10000.00	600.00
24	40.0	20.00000	10.0	8.00	500.00	200.00
25	150.0	50.00000	4.0	2.00	7500.00	1250.00
26	65.0	50.00000	10.0	1.00	6500.00	500.00
27	100.0	0.00101	10.0	1.00	10000.00	0.01
28	90.0	50.00000	5.0	4.00	1875.00	1000.00
29	100.0	75.00000	5.0	3.00	2857.14	1500.00
30	300.0	50.00000	15.0	1.00	30000.00	500.00
31	75.0	50.00000	10.0	2.00	3750.00	500.00
32	200.0	35.00000	7.0	1.00	5000.00	583.33
33	40.0	35.00000	10.0	1.00	3500.00	400.00
34	100.0	30.00000	10.0	1.00	10000.00	400.00
35	100.0	40.00000	8.0	2.50	4000.00	500.00

Brand names in which 2 ,3 or 4 sharks are invested ?

In [29]: # dataframe shows that those brand get invested 2 to 4
shark_data[(shark_data['total_sharks_invested']>=2) & (shark_data['total_sharks_invested']<=4)][['brand_name', 'total_sharks_invested']]

	brand_name	total_sharks_invested
0	BluePine Industries	3
1	Booz scooters	2
2	Heart up my Sleeves	2
9	Cosiq	2
11	Bummer	2
12	Revamp Moto	2
18	Raising Superstars	2
21	Beyond Snack	2
22	Vivalyf Innovations- Easy Life	2
24	Altor	2
25	Ariro	2
27	Nuutjob	3
28	Meatyour	3
29	EventBeep	3
32	Farda	2
35	LOKA	3
36	Annie	3
37	Caragreen	2
38	The Yarn Bazaar	4
39	The Renal Project	2
44	Cocofit	3
45	Bamboo India	2
47	Beyond Water	2
48	Let's Try	2
50	Aas Vidyalaya	3
58	WeSTOCK	4
64	Get a Whey	3
66	The Quirky Nari	2
67	Hair Originals	3
75	The Sass Bar	2
85	Watt Technovations	4
88	Humpy A2	3
90	Gold Safe Solutions Ind.	3
91	Wakao Foods	3
95	Kabaddi Adda	2
108	Tweek Labs	3
109	Proxgy	2
110	Nomad Food Project	4
114	Jain Shikanji	4

Episode wise minimum and maximum of sharks invested ?

```
In [30]: shark_data.head(2)
```

Out[30]:

	episode_number	pitch_number	brand_name	idea	deal	pitcher_ask_amount	ask_equity	ask_valuation	deal_amount	deal
0	1	1	BluePine Industries	Frozen Momos	1	50.0	5.0	1000.00	75.0	
1	1	2	Booz scooters	Renting e-bike for mobility in private spaces	1	40.0	15.0	266.67	40.0	

2 rows × 28 columns



```
In [31]: #maximum and minimum invested by sharks in every episode  
shark_data.groupby(['episode_number'])['total_sharks_invested'].agg([max,min])
```

Out[31]:

episode_number	max	min
1	3	2
2	1	0
3	1	0
4	2	0
5	2	0
6	5	0
7	2	0
8	2	1
9	2	0
10	3	3
11	2	0
12	3	0
13	4	2
14	2	0
15	3	1
16	2	0
17	5	0
18	1	0
19	4	0
20	5	0
21	3	1
22	3	0
23	1	0
24	2	0
25	5	0
26	1	0
27	4	0
28	3	0
29	3	0
30	2	0
31	1	0
32	1	0
33	3	0
34	4	0
35	4	0

Observation

- We can see that in episode 6,17,20,25 all the sharks invested
- minimum is in so many episodes 0

Minimum and Maximum sharks invested in each brand ?

In [32]: `shark_data.groupby(['brand_name'])['total_sharks_invested'].agg([max,min])`

Out[32]:

brand_name	max	min
ARRCOAT Surface Textures	1	1
Aas Vidyalaya	3	3
Agro tourism	0	0
Aliste Technologies	0	0
Alpino	0	0
...
Vivalyf Innovations- Easy Life	2	2
Wakao Foods	3	3
Watt Technovations	4	4
WeSTOCK	4	4
Woloo	0	0

117 rows × 2 columns

Brand names who ask for 1 crore and got a deal ?

In [33]: `shark_data.head(2)`

Out[33]:

episode_number	pitch_number	brand_name	idea	deal	pitcher_ask_amount	ask_equity	ask_valuation	deal_amount	deal
0	1	1	BluePine Industries	Frozen Momos	1	50.0	5.0	1000.00	75.0
1	1	2	Booz scooters	Renting e-bike for mobility in private spaces	1	40.0	15.0	266.67	40.0

2 rows × 28 columns

In [34]: `shark_data[(shark_data['pitcher_ask_amount']==100)&(shark_data['deal_amount']==100)][['brand_name','pitcher_ask_amount','deal_amount']]`

Out[34]:

	brand_name	pitcher_ask_amount	deal_amount
12	Revamp Moto	100.0	100.0
18	Raising Superstars	100.0	100.0
39	The Renal Project	100.0	100.0
64	Get a Whey	100.0	100.0
79	Sunfox Technologies	100.0	100.0
87	Insurance Samadhan	100.0	100.0

Brand names who asked equity between 0 to 50 ?

In [35]: `shark_data.head(2)`

	episode_number	pitch_number	brand_name	idea	deal	pitcher_ask_amount	ask_equity	ask_valuation	deal_amount	deal
0	1	1	BluePine Industries	Frozen Momos	1		50.0	5.0	1000.00	75.0
1	1	2	Booz scooters	Renting e-bike for mobility in private spaces	1		40.0	15.0	266.67	40.0

2 rows × 28 columns

In [36]: `shark_data[(shark_data['ask_equity'] >= 0) & (shark_data['ask_equity'] <= 50)][['brand_name', 'ask_equity']]`

	brand_name	ask_equity
0	BluePine Industries	5.0
1	Booz scooters	15.0
2	Heart up my Sleeves	10.0
3	Tagz Foods	1.0
4	Head and Heart	5.0
...
112	Green Protein	2.0
113	On2Cook	1.0
114	Jain Shikanji	8.0
115	Woloo	4.0
116	Elcare India	2.5

117 rows × 2 columns

Brand names where deal equity is between 0 to 50 ?

In [37]: `# showing which brand name in which deal equity is between 0 to 50
shark_data[(shark_data['deal_equity'] >= 0) & (shark_data['deal_equity'] <= 50)][['brand_name', 'deal_equity']]`

	brand_name	deal_equity
0	BluePine Industries	16.00
1	Booz scooters	50.00
2	Heart up my Sleeves	30.00
3	Tagz Foods	2.75
4	Head and Heart	0.00
...
112	Green Protein	0.00
113	On2Cook	0.00
114	Jain Shikanji	30.00
115	Woloo	0.00
116	Elcare India	0.00

116 rows × 2 columns

Find the number of brands participated in each episode ?

```
In [38]: shark_data['episode_number'].value_counts()
```

```
Out[38]: 18    4
30    4
17    4
16    4
22    4
23    4
27    4
31    4
32    4
33    4
34    4
19    4
29    3
28    3
20    3
26    3
25    3
24    3
21    3
1     3
2     3
15    3
14    3
13    3
12    3
11    3
10    3
9     3
8     3
7     3
6     3
5     3
4     3
3     3
35    3
Name: episode_number, dtype: int64
```

Insights

- maximum number of brands participated in per episode is 4
- minimum number of brands participated in per episode is 3

How many sharks participated in this show and What were their names ?

In [39]: `shark_data.head(2)`

	episode_number	pitch_number	brand_name	idea	deal	pitcher_ask_amount	ask_equity	ask_valuation	deal_amount	deal
0	1	1	BluePine Industries	Frozen Momos	1	50.0	5.0	1000.00	75.0	
1	1	2	Booz scooters	Renting e-bike for mobility in private spaces	1	40.0	15.0	266.67	40.0	

2 rows × 28 columns

In [40]: `# sharks name which is participated in ths show
shark_data.columns[11:18]`

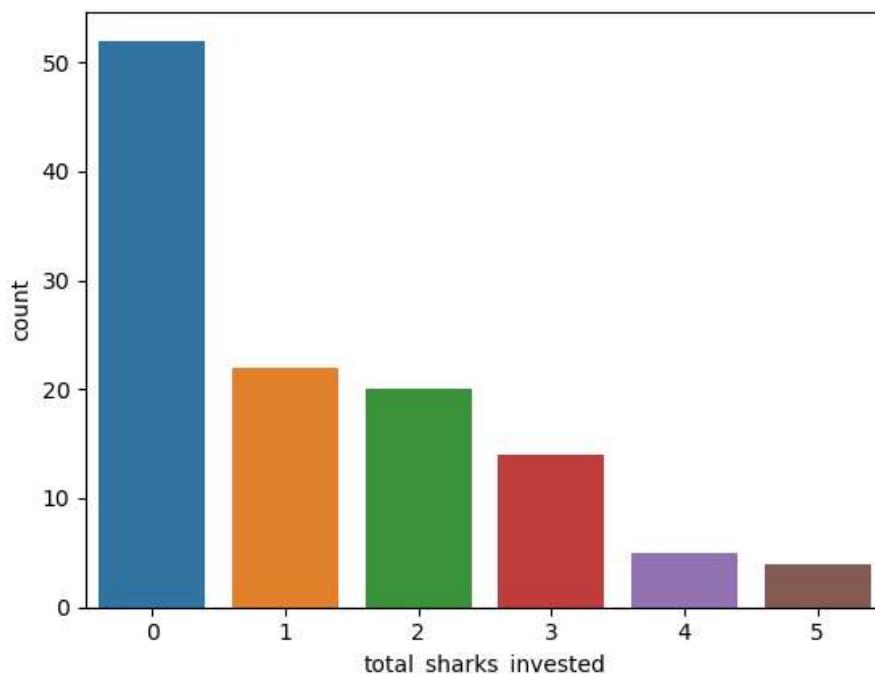
Out[40]: `Index(['ashneer_present', 'anupam_present', 'aman_present', 'namita_present', 'vineeta_present', 'peyush_present', 'ghazal_present'],
dtype='object')`

In [41]: `# count of sharks which are participated in show
len(shark_data.columns[11:18])`

Out[41]: 7

In [42]: `sns.countplot(x='total_sharks_invested', data=shark_data)`

Out[42]: <AxesSubplot:xlabel='total_sharks_invested', ylabel='count'>

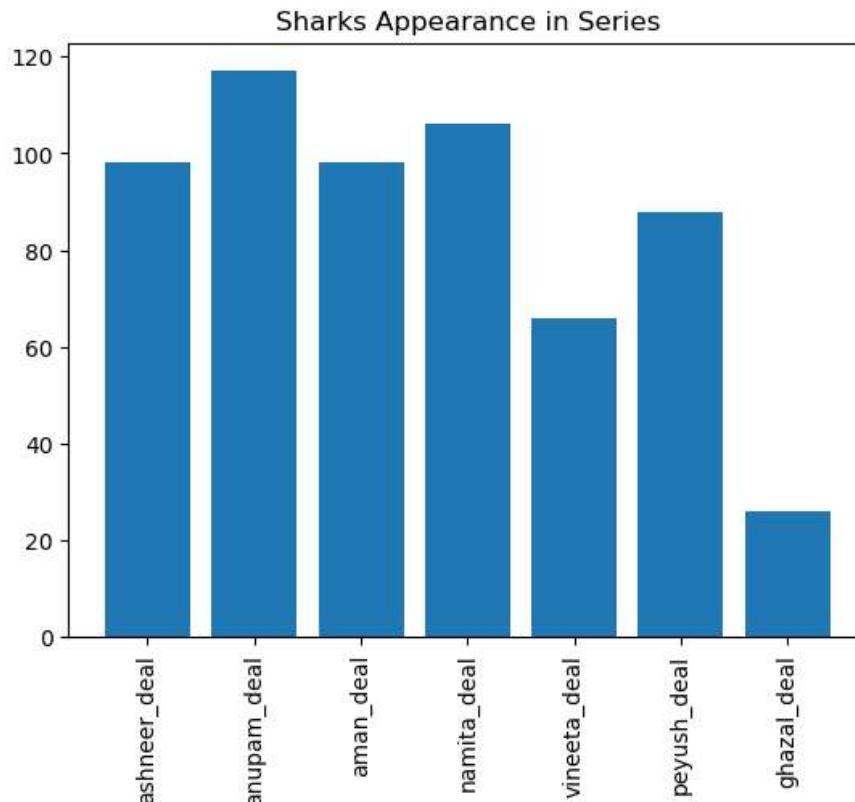


Find appearance of each sharks?

```
In [43]: l = []
for i in shark_data.columns[11:18]:
    l.append(shark_data[i].sum())

# Create a DataFrame for shark appearances
sharks_appearance_data = {'sharks': ['ashneer_deal', 'anupam_deal', 'aman_deal', 'namita_deal', 'vineeta_deal'],
                           'sharks_appearance_data': 1}
sharks_appearance = pd.DataFrame(sharks_appearance_data)

# Plot the bar chart
plt.bar(sharks_appearance['sharks'], sharks_appearance['sharks_appearance_data'])
plt.xticks(rotation=90)
plt.title('Sharks Appearance in Series')
plt.show()
```



Observations

- Anupam appeared maximum in episode in shark tank
- Ghazal appeared minimum in episode in shark tank

How many entrepreneurs were present ?

In [44]: `shark_data.head(2)`

Out[44]:

	episode_number	pitch_number	brand_name	idea	deal	pitcher_ask_amount	ask_equity	ask_valuation	deal_amount	deal
0	1	1	BluePine Industries	Frozen Momos	1	50.0	5.0	1000.00	75.0	
1	1	2	Booz scooters	Renting e-bike for mobility in private spaces	1	40.0	15.0	266.67	40.0	

2 rows × 28 columns

In [46]: `shark_data['brand_name'].nunique()`

Out[46]: 117

Observation

- Total 117 number of enterpenure is present

How many times each investors invested the deal ?

```
In [47]: for i in shark_data.columns[18:25]:
    s=shark_data[i].sum()
    print(i,'deal count',s)
```

```
ashneer_deal deal count 21
anupam_deal deal count 24
aman_deal deal count 28
namita_deal deal count 22
vineeta_deal deal count 15
peyush_deal deal count 27
ghazal_deal deal count 7
```

Observation

- Maximum number of investor is aman_deal
- Minimum number of investor is ghazal_deal

Calculating the total amount invested by the sharks

```
In [51]: l=[]
for i in shark_data.columns[18:25]:
    x=shark_data[shark_data[i]==1]
    s=x['amount_per_shark'].sum()
    l.append(s)
    print(i,'invested',s,'amount')
amount=sorted(l,reverse=False)
amount
```

```
ashneer_deal invested 494.33333333 amount
anupam_deal invested 533.83360253 amount
aman_deal invested 887.500016693 amount
namita_deal invested 648.333602533 amount
vineeta_deal invested 328.333333330001 amount
peyush_deal invested 719.6669191630001 amount
ghazal_deal invested 130.0002525 amount
```

Observations

- Aman invested maximum amount i.e 887.50001 amount
- Lowest amount invested by ghazal i.e 130 amount

Find the equity percent that each sharks gets ?

```
In [52]: shark_data.head(2)
```

	episode_number	pitch_number	brand_name	idea	deal	pitcher_ask_amount	ask_equity	ask_valuation	deal_amount	deal
0	1	1	BluePine Industries	Frozen Momos	1	50.0	5.0	1000.00	75.0	
1	1	2	Booz scooters	Renting e-bike for mobility in private spaces	1	40.0	15.0	266.67	40.0	

2 rows × 28 columns

```
In [54]: l=[]
for i in shark_data.columns[18:25]:
    x=shark_data[shark_data[i]==1]
    s=x['ask_equity'].sum()
    l.append(s)
    print(i,'gets total',s,'equity')
amount=sorted(l,reverse=True)
amount
```

```
ashneer_deal gets total 93.5 equity
anupam_deal gets total 127.0 equity
aman_deal gets total 118.0 equity
namita_deal gets total 119.0 equity
vineeta_deal gets total 92.5 equity
peyush_deal gets total 135.0 equity
ghazal_deal gets total 44.0 equity
```

Observation

- Highest percentage of equity owned by peyush
- Lowest percentage of equity owned by ghazal

Find the total number of amount invested in this show?

In [55]: `shark_data.head(2)`

	episode_number	pitch_number	brand_name	idea	deal	pitcher_ask_amount	ask_equity	ask_valuation	deal_amount	deal
0	1	1	BluePine Industries	Frozen Momos	1	50.0	5.0	1000.00	75.0	
1	1	2	Booz scooters	Renting e-bike for mobility in private spaces	1	40.0	15.0	266.67	40.0	

2 rows × 28 columns

In [56]: `shark_data['deal_amount'].sum()`

Out[56]: 3742.00106

- Total 3742.00106 amount invested in the show

Find the highest equity a shark gets ?

In [60]: `shark_data['equity_per_shark'].max()`

Out[60]: 75.0

Find the most attracted ideas accepted by sharks ?

In [62]: `shark_data['total_sharks_invested'].unique()`

Out[62]: array([3, 2, 1, 0, 5, 4], dtype=int64)

In [63]: `# calculating total_sharks_invested greater than 3 with brand name and idea
shark_data[shark_data['total_sharks_invested']>3][['brand_name','idea']]`

	brand_name	idea
15	Skippi Pops	Ice-Pops
38	The Yarn Bazaar	Yarn-Trading App
49	Find Your Kicks India	Sneaker Resale
58	WeSTOCK	Livestock health monitoring AI
63	IN A CAN	Can Cocktails
79	Sunfox Technologies	Portable ECG Device
85	Watt Technovations	Ventilated PPE Kits
110	Nomad Food Project	Bacon Jams
114	Jain Shikanji	Lemonade

```
In [64]: shark_data[shark_data['ask_valuation']==shark_data['ask_valuation'].max()]
```

episode_number	pitch_number	brand_name	idea	deal	pitcher_ask_amount	ask_equity	ask_valuation	deal_amount	deal
30	11	31 Gopal's 56	Fiber Ice Cream	0	30000.0	25.0	120000.0	0.0	

1 rows × 28 columns

How many companies in front Owner speak ?

```
In [65]: shark_data.head(2)
```

episode_number	pitch_number	brand_name	idea	deal	pitcher_ask_amount	ask_equity	ask_valuation	deal_amount	deal
0	1	1 BluePine Industries	Frozen Momos	1	50.0	5.0	1000.00	75.0	
1	1	2 Booz scooters	Renting e-bike for mobility in private spaces	1	40.0	15.0	266.67	40.0	

2 rows × 28 columns

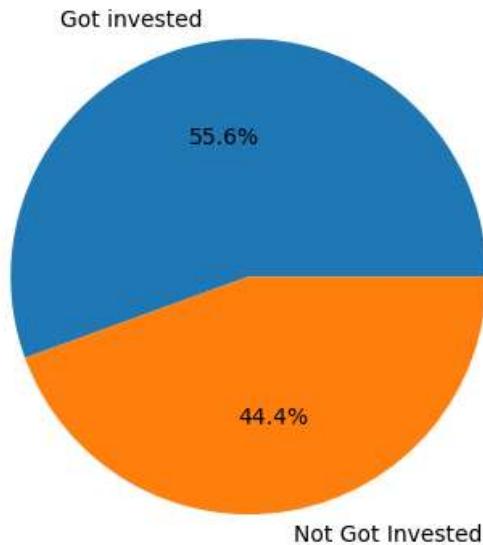
```
In [66]: list=['ashneer_deal','anupam_deal','aman_deal','namita_deal','vineeta_deal','peyush_deal','ghazal_deal']
for i in list:
    pres=shark_data[i].sum()
    print(i[:-8],"present in front of",pres,'companies')
```

ashn present in front of 21 companies
 anu present in front of 24 companies
 a present in front of 28 companies
 nam present in front of 22 companies
 vine present in front of 15 companies
 pey present in front of 27 companies
 gha present in front of 7 companies

```
In [73]: got_invest=shark_data[shark_data['deal_amount']>0.0].shape  
got_not_invest=shark_data[shark_data['deal_amount']==0.0].shape  
print('total number of companies got the investment',got_invest[0])  
print('total number of companies not got the investment',got_not_invest[0])  
list1=['Got invested','Not Got Invested']  
plt.pie([got_invest[0],got_not_invest[0]],labels=list1,autopct=".1f%%")  
plt.title("Numbers of Company which got invested or not")  
plt.show()
```

total number of companies got the investment 65
total number of companies not got the investment 52

Numbers of Company which got invested or not



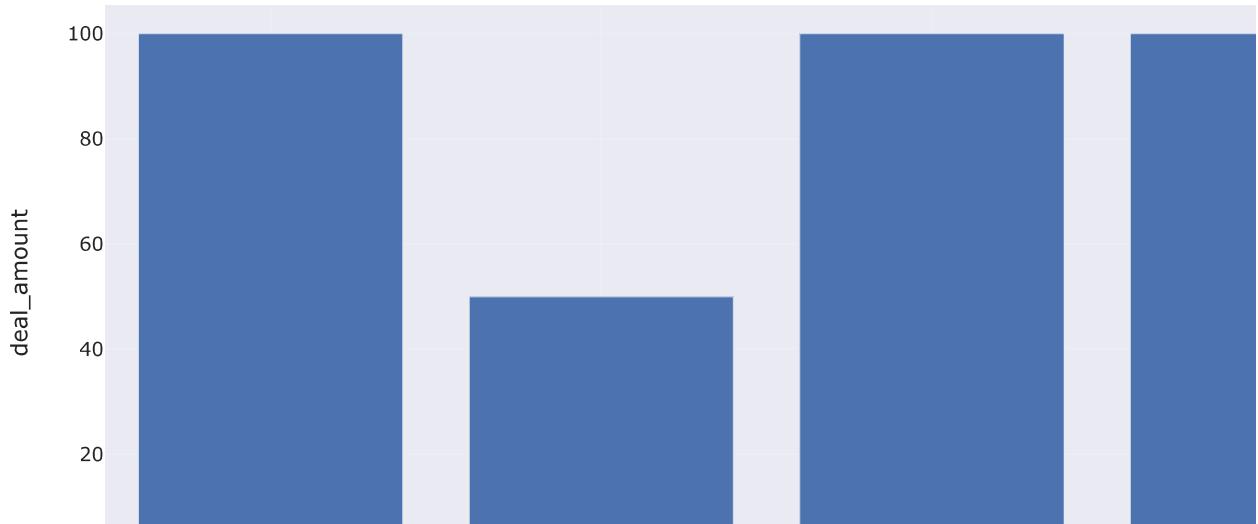
```
In [76]: shark_data.hist(figsize=(22,22),bins=10)
plt.show()
```



```
In [3]: import plotly.express as px
```

```
In [14]: total_shark = shark_data[shark_data["total_sharks_invested"] == 5]
figure = px.bar(total_shark, x='brand_name', y='deal_amount', title="Five Shark Deal ", template="seaborn")
figure.show()
```

Five Shark Deal

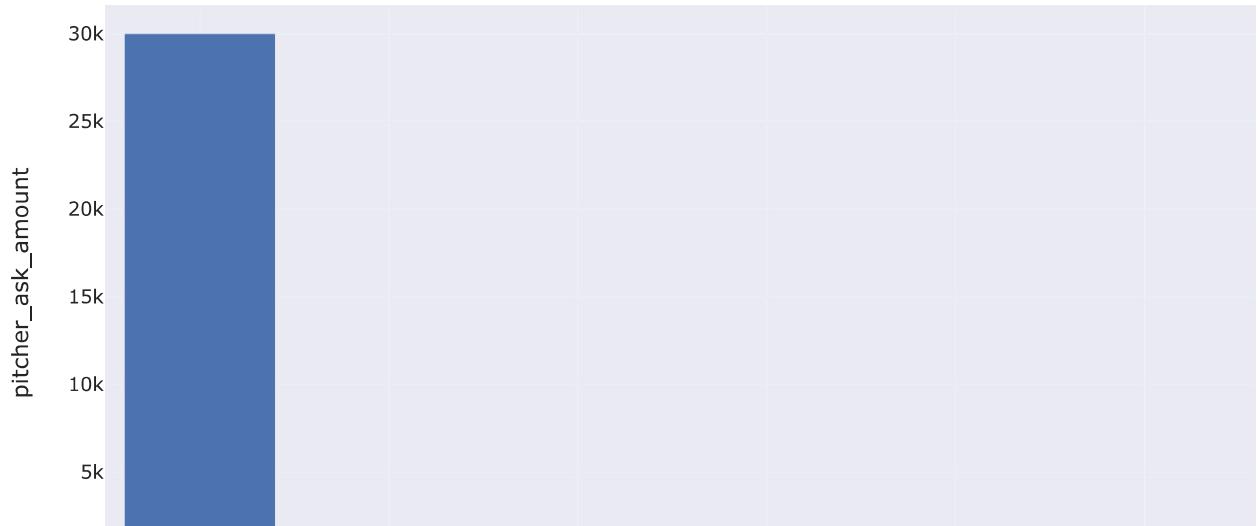


From above plot, we get to know that there are all total 4 five shark deal brands and also with the pitcher deal amount and the deal amount of the brand.

Highest Pitch Ask Amount:

```
In [15]: high=shark_data[shark_data['pitcher_ask_amount']>100]
figure=px.bar(high,x='brand_name',y='pitcher_ask_amount',title='Highest pitched asked',template='seaborn'
figure.show()
```

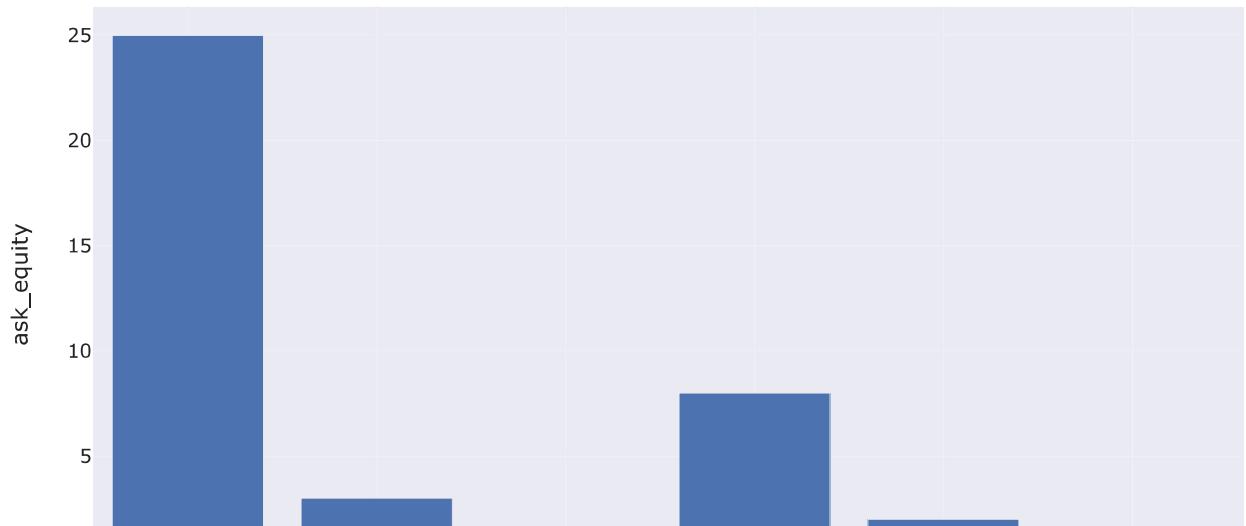
Highest pitched asked



Ask Equity and Deal Equity:

```
In [16]: figure=px.bar(high,x='brand_name',y='ask_equity',title='Ask equity and deal equity' ,template='seaborn')
figure.show()
```

Ask equity and deal equity



```
In [7]: shark_data.head(2)
```

```
Out[7]:
```

	episode_number	pitch_number	brand_name	idea	deal	pitcher_ask_amount	ask_equity	ask_valuation	deal_amount	deal
0	1	1	BluePine Industries	Frozen Momos	1	50.0	5.0	1000.00	75.0	
1	1	2	Booz scooters	Renting e-bike for mobility in private spaces	1	40.0	15.0	266.67	40.0	

2 rows × 28 columns

```
In [17]: #No. of brands the sharks have invested:
```

```
num_deal_shark=[shark_data.ashneer_deal.sum(), shark_data.anupam_deal.sum(), shark_data.aman_deal.sum(),  
all_sharks=["Ashneer", "Anupam", "Aman", "Namita", "Vineeta", "Peyush", "Ghazal"]  
figure=px.bar(total_shark, x=all_sharks, y=num_deal_shark,title="Number of deals done",template="seaborn  
figure.show()
```

Number of deals done



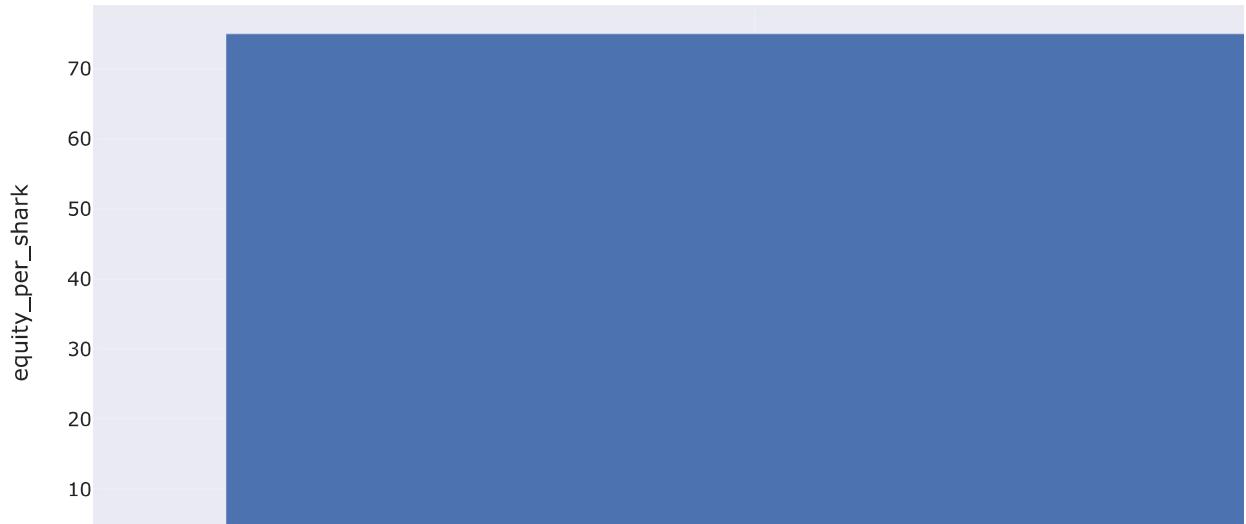
insights

- Aman invested maximum times and Ghazal invested minimum times

Above 50% Equity taken by the shark in which brand:

```
In [18]: shark_equity=shark_data[shark_data['equity_per_shark']>50]
figure=px.bar(shark_equity,x='brand_name',y='equity_per_shark',title='Above 50% equity',template='seaborn')
figure.show()
```

Above 50% equity



Observation

- only 1 brand name got above 50% equity (Side07 Designs)

```
In [ ]:
```