

The collected data has been stored in the Comma Separated Value file "Shark_Tank_India_S1.csv". Each startup is uniquely identified by its startup_number. Every startup contains the following variables:-

- episode_number: Episode number out of 36 episodes
- startup_number: Startup Number out of 121 startups
- brand_name: Name of the startup
- description: Brief introduction of the startup
- startup_ask_amount_lakhs: Startup's demand amount from the sharks
- startup_ask_percentage:
- startup_ask_valuation: Valuation calculated by the startup
- deal_amount_lakhs: Amount agreed by the sharks to invest
- deal_equity: Equity agreed by both the startup and the sharks for the deal
- deal_valuation: Valuation by the company in future
- loan_element_present: If loan is taken or not
- loan_amount: Loan amount if taken
- rannvijay_present: Ranvijay is present in that episode or not
- abish_present: Abish is present in that episode or not
- aman_present: Aman is present in that episode or not
- aman_invested: Aman has invested or not
- anupam_present: Anupam is present in that episode or not
- anupam_invested: Anupam has invested or not
- ashneer_present: Ashneer is present in that episode or not
- ashneer_invested: Ashneer has invested or not
- ghazal_present: Ghazal is present in that episode or not
- ghazal_invested: Ghazal has invested or not
- namita_present: Namita is present in that episode or not
- namita_invested: Namita has invested or not
- peyush_present: Peyush is present in that episode or not
- peyush_invested: Peyush has invested or not
- vineeta_present: Vineeta is present in that episode or not
- vineeta_invested: Vineeta has invested or not
- sharks_offering: Total Sharks present in a single deal
- amount_per_shark: Deal amount divided per sharks
- equity_per_shark: Equity divided per sharks

In [1]:

```
import pandas as pd
import numpy as np
import plotly as py
from plotly.offline import download_plotlyjs, init_notebook_mode, plot,
iplot
import plotly.graph_objects as go
init_notebook_mode ( connected= True)
```

In [2]:

shark_tank_data = pd.read_csv('Shark_Tank_India_S1.csv')

In [3]:

shark_tank_data

Out[3]:

	episode_number	startup_number	brand_name	description	deal_offered	startup_ask_amount_
0	1	1	BluePine Industries	Frozen Momos	1	
1	1	2	Booz scooters	Renting e-bike for mobility in private spaces	1	
2	1	3	Heart up my Sleeves	Detachable Sleeves	1	
3	2	4	Tagz Foods	Healthy Potato Chips	1	
4	2	5	Head and Heart	Brain Development Course	0	
...	
116	35	117	Elcare India	Carenting for Elders	0	
117	36	118	Sneakare	Shoe care and storage solutions	1	
118	36	119	French Crown	Clothing Industry	0	
119	36	120	Store My Goods	Storage solutions	1	
120	36	121	Devnagri	Translation platform	0	

121 rows × 32 columns

In [4]:

#getting the list of columns present in the data set
shark_tank_data.columns

Out[4]:

Index(['episode_number', 'startup_number', 'brand_name', 'description', 'deal_offered', 'startup_ask_amount_lakhs', 'startup_ask_percentage', 'startup_ask_valuation', 'deal_amount_lakhs', 'deal_equity', 'deal_valuation', 'loan_element_present', 'loan_amount', 'rannvijay_present', 'abish_present', 'aman_present', 'aman_invested', 'anupam_present', 'anupam_invested', 'ashneer_present', 'ashneer_invested', 'ghazal_present', 'ghazal_invested', 'namita_present', 'namita_invested', 'peyush_present', 'peyush_invested', 'vineeta_present', 'vineeta_invested', 'sharks_offering', 'amount_per_shark', 'equity_per_shark'], dtype='object')

```
In [5]: shark_tank_data.drop(['rannvijay_present','abish_present'], axis = 1,
        inplace= True)
        # since rannvijay and abish were not among the sharks
```

```
In [6]: shark_tank_data.sample(5)
```

Out[6]:

	episode_number	startup_number	brand_name	description	deal_offered	startup_ask_amount_la
84	27	85	Theka Coffee	Coffee Products	0	
46	16	47	Flying Furr	Dog Hygiene	0	
18	7	19	Raising Superstars	Child Development App	1	1
38	13	39	The Yarn Bazaar	Yarn-Trading App	1	
20	7	21	La Kheer Deli	Kheer in variety of flavors	0	

5 rows × 30 columns

```
In [7]: #checking for the basis info from the data
        shark_tank_data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 121 entries, 0 to 120
Data columns (total 30 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   episode_number                        121 non-null    int64
1   startup_number                        121 non-null    int64
2   brand_name                            121 non-null    object
3   description                            121 non-null    object
4   deal_offered                          121 non-null    int64
5   startup_ask_amount_lakhs              121 non-null    float64
6   startup_ask_percentage                 121 non-null    float64
7   startup_ask_valuation                  121 non-null    float64
8   deal_amount_lakhs                     121 non-null    float64
9   deal_equity                           121 non-null    float64
10  deal_valuation                         121 non-null    float64
11  loan_element_present                   121 non-null    int64
12  loan_amount                           121 non-null    int64
13  aman_present                           121 non-null    int64
14  aman_invested                         121 non-null    int64
15  anupam_present                         121 non-null    int64
16  anupam_invested                       121 non-null    int64
17  ashneer_present                       121 non-null    int64
18  ashneer_invested                      121 non-null    int64
19  ghazal_present                        121 non-null    int64
```

```

20 ghazal_invested      121 non-null    int64
21 namita_present       121 non-null    int64
22 namita_invested      121 non-null    int64
23 peyush_present       121 non-null    int64
24 peyush_invested      121 non-null    int64
25 vineeta_present      121 non-null    int64
26 vineeta_invested     121 non-null    int64
27 sharks_offering      121 non-null    int64
28 amount_per_shark     121 non-null    float64
29 equity_per_shark     121 non-null    float64
dtypes: float64(8), int64(20), object(2)
memory usage: 28.5+ KB

```

In [8]: `shark_tank_data.shape`

Out[8]: (121, 30)

1. Finding the over all statistics of the data

In [9]: `shark_tank_data.describe(include='all')`

Out[9]:

	episode_number	startup_number	brand_name	description	deal_offered	startup_ask_amoun
count	121.000000	121.000000	121	121	121.000000	121.
unique	NaN	NaN	121	121	NaN	
top	NaN	NaN	BluePine Industries	Frozen Momos	NaN	
freq	NaN	NaN	1	1	NaN	
mean	19.305785	61.000000	NaN	NaN	0.561983	312.
std	10.375326	35.073732	NaN	NaN	0.498206	2721.
min	1.000000	1.000000	NaN	NaN	0.000000	0.
25%	11.000000	31.000000	NaN	NaN	0.000000	45.
50%	19.000000	61.000000	NaN	NaN	1.000000	50.
75%	28.000000	91.000000	NaN	NaN	1.000000	80.
max	36.000000	121.000000	NaN	NaN	1.000000	30000.

11 rows × 30 columns

2. Checking for the null values present in the dataframe

In [10]: `shark_tank_data.isnull().sum()`

Out[10]:

episode_number	0
startup_number	0
brand_name	0

```

description          0
deal_offered         0
startup_ask_amount_lakhs 0
startup_ask_percentage 0
startup_ask_valuation 0
deal_amount_lakhs    0
deal_equity          0
deal_valuation       0
loan_element_present 0
loan_amount          0
aman_present         0
aman_invested        0
anupam_present       0
anupam_invested      0
ashneer_present      0
ashneer_invested     0
ghazal_present       0
ghazal_invested      0
namita_present       0
namita_invested      0
peyush_present       0
peyush_invested      0
vineeta_present      0
vineeta_invested     0
sharks_offering      0
amount_per_shark     0
equity_per_shark     0
dtype: int64

```

3.Counting the number of days each shark was present in the show

```
In [11]: shark_tank_data['aman_present'].value_counts()
```

```
Out[11]: 1    102
         0     19
         Name: aman_present, dtype: int64
```

```
In [12]: shark_tank_data['anupam_present'].value_counts()
```

```
Out[12]: 1    121
         Name: anupam_present, dtype: int64
```

```
In [13]: shark_tank_data['peyush_present'].value_counts()
```

```
Out[13]: 1     92
         0     29
         Name: peyush_present, dtype: int64
```

```
In [14]: shark_tank_data['ghazal_present'].value_counts()
```

```
Out[14]: 0     95
         1     26
```

Name: ghazal_present, dtype: int64

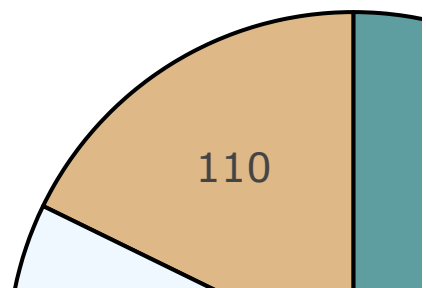
```
In [15]: shark_tank_data['namita_present'].value_counts()
```

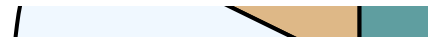
```
Out[15]: 1    110  
        0     11  
        Name: namita_present, dtype: int64
```

```
In [16]: shark_tank_data['ashneer_present'].value_counts()
```

```
Out[16]: 1     98  
        0     23  
        Name: ashneer_present, dtype: int64
```

```
In [17]: #present the same info using the plotly pie chart  
labels = ['aman_present', 'peyush_present', 'namita_present',  
          'anupam_present', 'ghazal_present', 'ashneer_present',  
          'vineeta_present']  
values = []  
for i in labels:  
    x = shark_tank_data[i].value_counts()  
    values.append(x[1])  
  
colors = ['aliceblue', 'azure', 'burlywood', 'cadetblue', 'chartreuse',  
          'cornflowerblue', 'cornsilk']  
  
fig = go.Figure(data = [go.Pie(labels= labels,  
                                values= values)])  
  
fig.update_traces(hoverinfo = 'label+percent', textinfo = 'value',  
                  textfont_size = 20,  
                  marker = dict(colors = colors, line = dict(color  
= '#000000', width = 2)))  
  
fig.show()
```





4. Counting the investments by each shark

```
In [18]: aman = shark_tank_data['aman_invested'].value_counts()
```

```
In [19]: peyush = shark_tank_data['peyush_invested'].value_counts()
```

```
In [20]: anupam = shark_tank_data['anupam_invested'].value_counts()
```

```
In [21]: namita = shark_tank_data['namita_invested'].value_counts()
```

```
In [22]: ashneer = shark_tank_data['ashneer_invested'].value_counts()
```

```
In [23]: vineeta = shark_tank_data['vineeta_invested'].value_counts()
```

```
In [24]: ghazal = shark_tank_data['ghazal_invested'].value_counts()
```

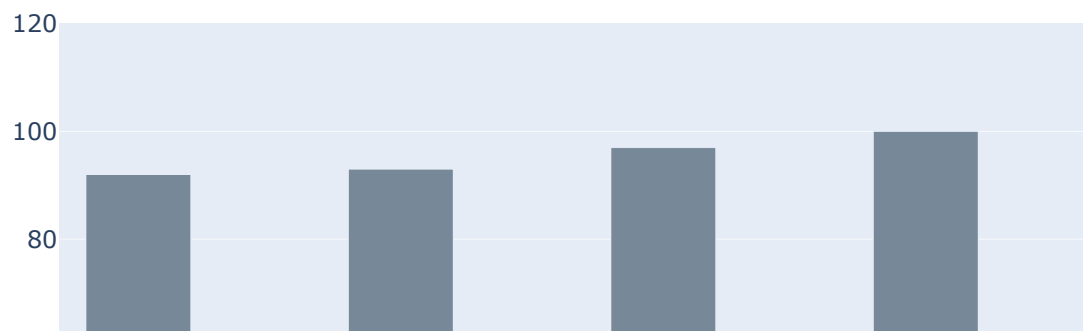
```
In [25]: #plotting the bar graph for the above investments
```

```
In [26]: x = ['aman_invested', 'peyush_invested', 'ashneer_invested',  
            'anupam_invested', 'namita_invested', 'ghazal_invested',  
            'vineeta_invested']  
invested = []  
not_invested = []  
  
sharks = [aman, peyush, anupam, ashneer, ghazal, namita, vineeta]  
  
for i in sharks:
```

```
invested.append(i[0])  
not_invested.append(i[1])
```

In [27]:

```
# bar graph for each shark invested and not invested.  
fig = go.Figure()  
  
fig.add_trace(go.Bar(x = x, y = invested, name = "StartUps_Invested",  
                    marker_color = 'lightslategrey' ))  
  
fig.add_trace(go.Bar(x = x , y = not_invested, name =  
"StartUps_not_Invested",  
                    marker_color = 'lightsalmon'))  
  
fig.update_layout(barmode = 'group', xaxis_tickangle = -45)  
  
fig.show()
```



In [28]:

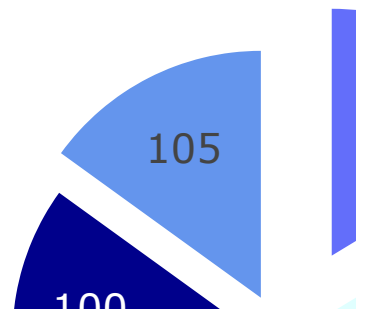
```
# pie chart for the investment made by sharks throughout the seasons.  
colors = ['lightcyan', 'cyan', 'royalblue', 'darkblue', 'navyblue',  
          'turquoise', 'cornflowerblue' ]
```



```
fig = go.Figure(data = [go.Pie(labels = x , values = invested,
                                pull = [0, 0, 0, 0.1, 0.4, 0, 0.2 ])] )

fig.update_traces(hoverinfo = 'label+percent', textinfo = 'value',
                  textfont_size = 20,
                  marker = dict(colors = colors ))

fig.show()
```



the above pie chart shows that namita invested in most of the startups, followed by vineeta and anupam respectively.

5. Finding the top 10 deals offered by the sharks.

In [29]:

```
#based on deal amount.

top_deals = shark_tank_data.sort_values('deal_amount_lakhs', ascending =
False).head(10)
top_deals
```

Out[29]:

	episode_number	startup_number	brand_name	description	deal_offered	startup_ask_amount_lakhs
50	17	51	Aas Vidyalyaya	EdTech App	1	1
36	13	37	Annie	Braille Literary Device	1	
18	7	19	Raising Superstars	Child Development App	1	1
87	27	88	Insurance Samadhan	Insurance Solutions	1	1
15	6	16	Skippi Pops	Ice-Pops	1	
64	21	65	Get a Whey	Sugar-Free Icecream	1	1
12	5	13	Revamp Moto	E-Bike	1	1
39	14	40	The Renal Project	Home Dialysis Treatment	1	1
79	25	80	Sunfox Technologies	Portable ECG Device	1	1
42	15	43	Hammer Lifestyle	Smart Audio Products	1	

10 rows × 30 columns

In [30]:

```
brands_top10 = top_deals['brand_name'].tolist()
asked_amount = top_deals['startup_ask_amount_lakhs'].tolist()
deal_amount = top_deals['deal_amount_lakhs'].tolist()
```

In [31]:

```
# plotting a graph to see the variation in asked amount and deal amount
of top 10 startup.

fig = go.Figure()

fig.add_traces(go.Bar(x = brands_top10, y = asked_amount,
                        name = 'Amount Asked by the StartUps in Lakhks',
                        marker_color = 'crimson'))

fig.add_traces(go.Bar(x = brands_top10, y = deal_amount,
                        name = 'Deal Amount in Lakhs',
                        marker_color = 'cornflowerblue'))

fig.update_layout(
    title = 'Asked Amount v/s Deal Amount',
```

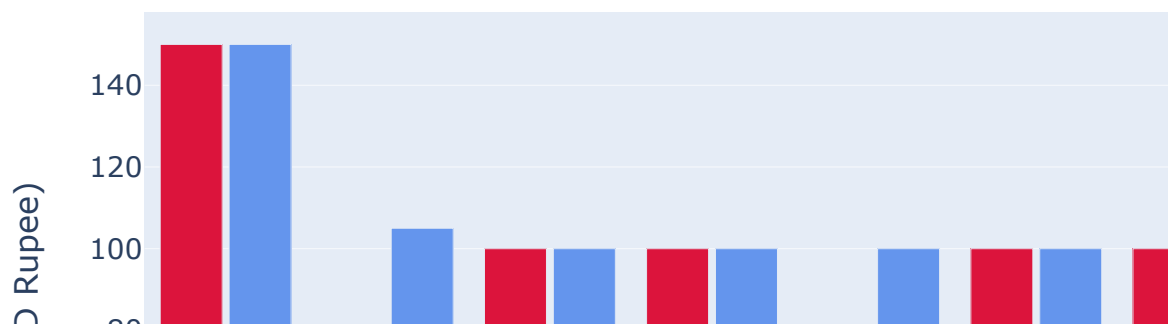
```

axis_tickfont_size = 14,
yaxis = dict(
    title = 'Lakhs (IND Rupee)',
    titlefont_size = 16,
    tickfont_size = 14),
barmode = 'group',
bargap = 0.15,
bargroupgap = 0.1)

fig.show()

```

Asked Amount v/s Deal Amount



Aas Vidyalyaya (EdTech App) was the top deal in the first season of the shark tank with 150 lakhs deal amount

6. Total deals taking place out of 121.

```

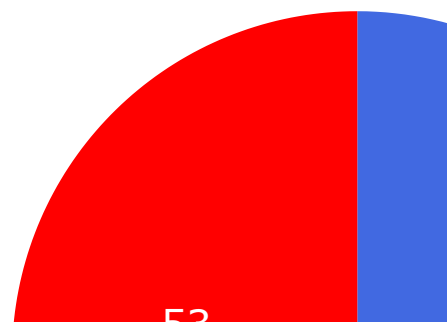
In [32]: deals_taking_place = shark_tank_data[shark_tank_data['deal_offered'] ==
1].shape[0]
deals_not_taking_place = shark_tank_data[shark_tank_data['deal_offered']
== 0].shape[0]

```

```
In [33]: fig = go.Figure(data = [go.Pie(labels = ['Deals taking place', 'Deals not taking place'],
                                         values = [deals_taking_place,
                                         deals_not_taking_place],
                                         )])

fig.update_traces(hoverinfo = 'label+percent', textinfo = 'value',
                  textfont_size = 20,
                  marker = dict(colors = ['royalblue', 'red']))

fig.show()
```



56.2% of 121 deals took place in season 1 of shark tanks.

In []: