Behavioral Analysis of Investment Decisions on Shark Tank India- Final Report

By *George Paul* (2021121006), a part of
Team GAS
George Paul (2021121006)
Adith John Rajeev (2020114010)
Sankalp Bhat (2020112018)

Note: George marks a contribution from George, Adith marks a contribution from Adith, Sankalp marks a contribution from Sankalp. Any unmarked section was a collective effort.

Introduction to Shark Tank

On Shark Tank India, entrepreneurs pitch their business ideas to a panel of investors known as "Sharks." They pitch their company, its product or service, and the reasons why it's a compelling investment opportunity. Entrepreneurs also state how much money they're seeking and the percentage of equity they're willing to give up. Sharks then question the entrepreneurs rigorously, evaluating the viability of the business and the capability of the founders. Interested Sharks may extend offers, which can be for the original amount or with modified terms. Entrepreneurs have the choice to accept, negotiate, or decline these offers. If an entrepreneur and one or more Sharks reach an agreement on the investment amount and equity, a deal is struck.

Dataset Description

The dataset is a curation of every startup that has appeared on the first three seasons of Shark Tank India. The data describes facets such as details about the presenter(s) and company. The dataset also includes details about the finances - The initial ask, final investments, type of investment made (royalty, equity, conditions etc.) by each shark.

Hypotheses to test

Loss Aversion

Sankalp

Sharks will exhibit loss aversion, demonstrating a decreased likelihood of investing in startups with indicators of financial risk.

Relevant Parameters

- ➔ Gross Margin
- ➔ Net Margin
- ➔ Cash Burn
- ➔ Received Offer (indicator of funding success)

Analysis Methods

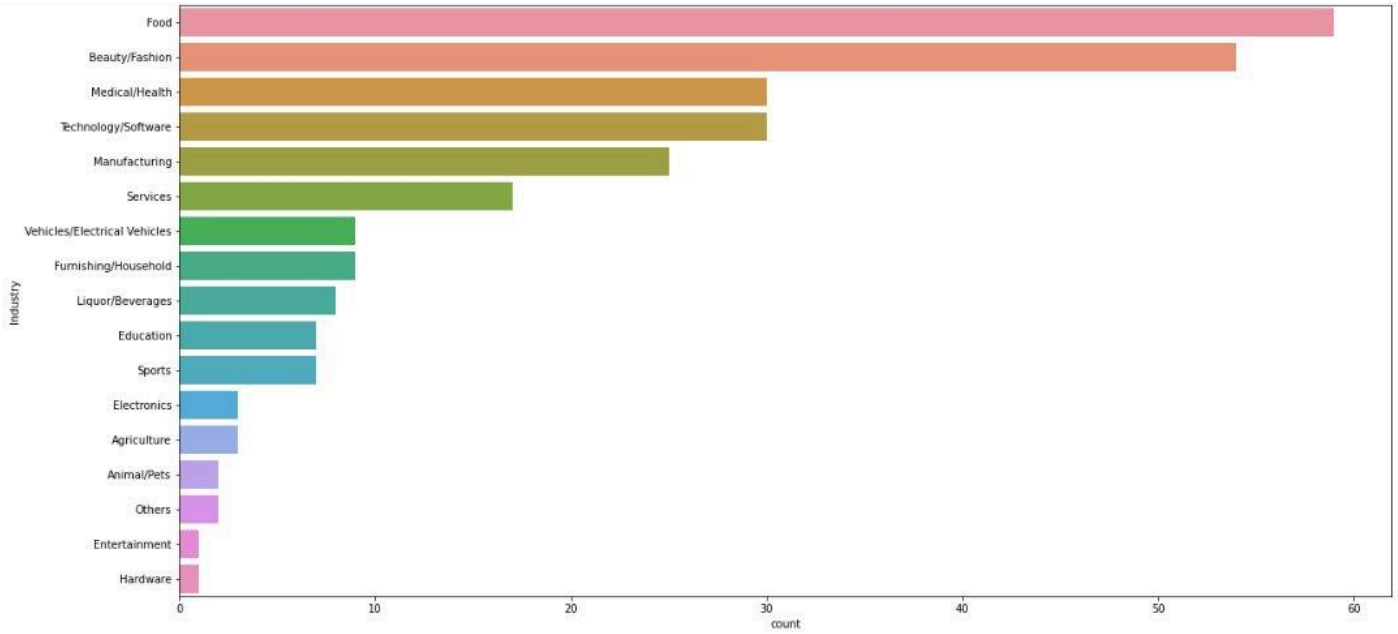
Logistic regression to predict deal success based on financial indicators. Logistic regression was done using Received Offer against Gross Margin, Net Margin, Cash Burn.

Results and Discussion

```
==============================================================================
     coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
  -0.0009      0.009     -0.100      0.920      -0.019       0.017
==============================================================================
```

Resulting p-value for the test was 0.92. This result isn't significant and we can accept the Null Hypothesis.

Industry-wise distribution



George *Fig.1*

Fig.1 shows the number of pitches given in each industry. There seems to be a heavy leaning towards the Manufacturing, Tech, Health, Beauty, and Food industries for startups that feature on Shark Tank India.

Per-Industry Investments by Sharks

George

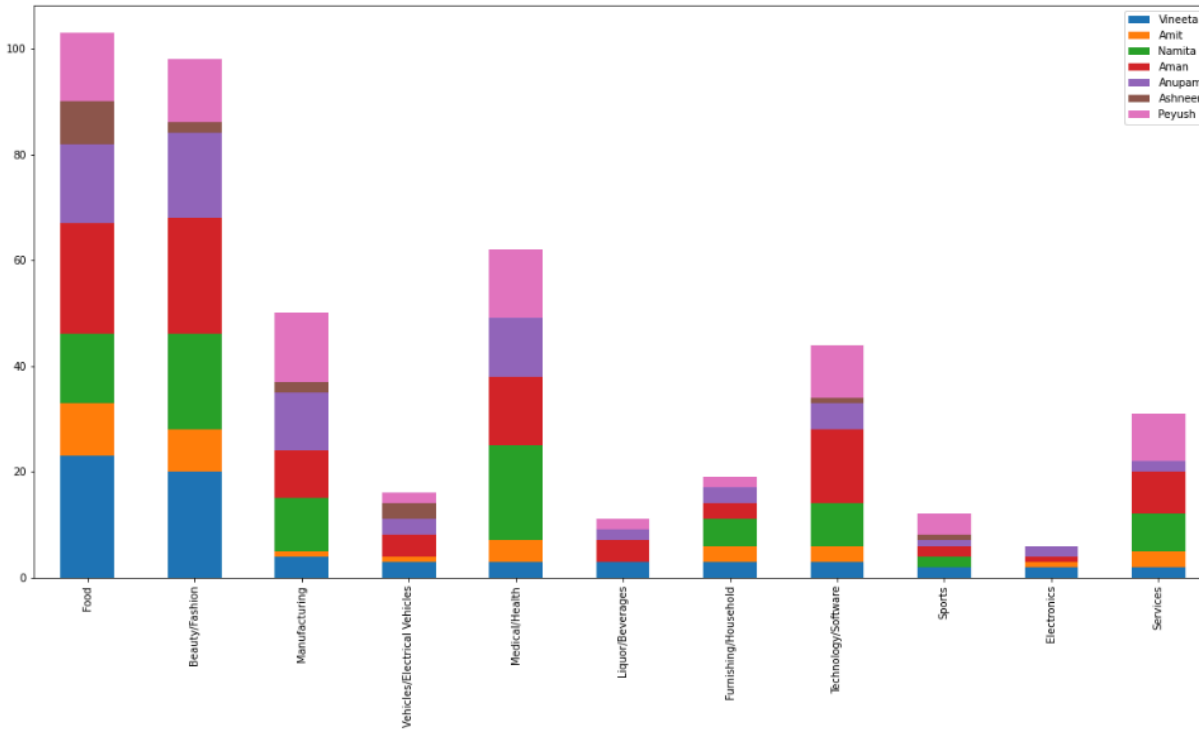


Fig.2

The difference in distribution of investments by each shark in each industry is shown in Fig.2. We hypothesised that certain sharks display propensities towards investing in certain industries versus others. Or even avoid certain industries.

Relevant Parameters

➔ Industry
➔ X invested - A derived value that is true if X invested or gave debt, X being one of the sharks.

Analysis

An expected distribution for investment by each shark was calculated using all the investments made in each season by each shark. That is, a distribution for the investments made by shark X as a fraction of all the investments.

A Chi-Square test was conducted using this expected distribution against the observed distribution for each industry. The tests yielded statistical significance values for whether there was a difference in the distributions.

Since multiple tests (one for each industry) are being conducted across the board for the same hypothesis, we elected to employ a p-value correction method. Initially we went ahead with Benjamini-Hochberg correction but found that even with the relatively stricter Bonferroni correction, our results remained the same so we've ended up with the same.

The corrected α for this hypothesis is `0.00454`

Results and Discussion

We observed statistically significant results in the following industries (paired with the p-values for each):

- ➔ Vehicles/EV: 0.001
- ➔ Health: 1.4e-15
- ➔ Furnishing: 9.6e-9
- ➔ Sports: 1.2e-5
- ➔ Electronics: 4.2e-18
- ➔ Services: 0.0005

This lines up well with the earlier visualisation that shows similar results. A drawback of this analysis method is that we have no concrete way to tell which of the sharks is causing the significant difference. Although one can tell by observing Fig.2 and comparing with the expected distribution.

Indeed, we see that Namita's share of investments in the Medical/Health industry is higher than usual. Conversely, Ashneer's is not present considering he hasn't invested in a single Health startup on the show.

Age distribution of presenters

George

During the EDA stage, we noticed that hardly any presenters are of Old age (classified by the dataset as >55 years of age). This led us to make certain hypotheses of the age distributions of the presenters, notably among the industries as well. We hypothesise that the age distribution of presenters are significantly different for certain industries

Analysis

A Chi-Square test was employed to test the differences present in each industry. An expected distribution was formulated by calculating the age distribution of presenters across the show. Then the distribution in each industry was tested against this.

And since the same hypothesis was tested across 11 industries, similar to the hypothesis made of Per-industry shark investment, we corrected the α value accordingly using a Bonferroni correction. The corrected α being 0.00454.

Results and Discussion

The following industries showed statistically significant differences in the age distribution of presenters (paired with the corresponding p-value):

- ➔ Electronics : 6.3e-16
- ➔ Liquor/Beverages : 0.0001

And again, similar to the hypothesis made of Per-industry shark investment, the drawback of running the Chi-Square test is that we have no concrete pointer to the relevant age that is showing a significant difference. Although, the results align with the below visualisations and we can see the changes much easier since there are only a few sections (Young, Middle-age).

Visualisations

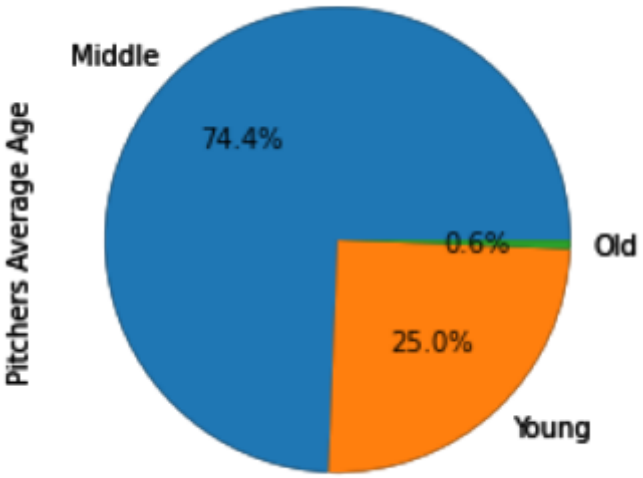


Fig.3 - Age distribution of presenters

The visualisation shows that there are barely any Old presenters, and is dominated by Middle age presenters.

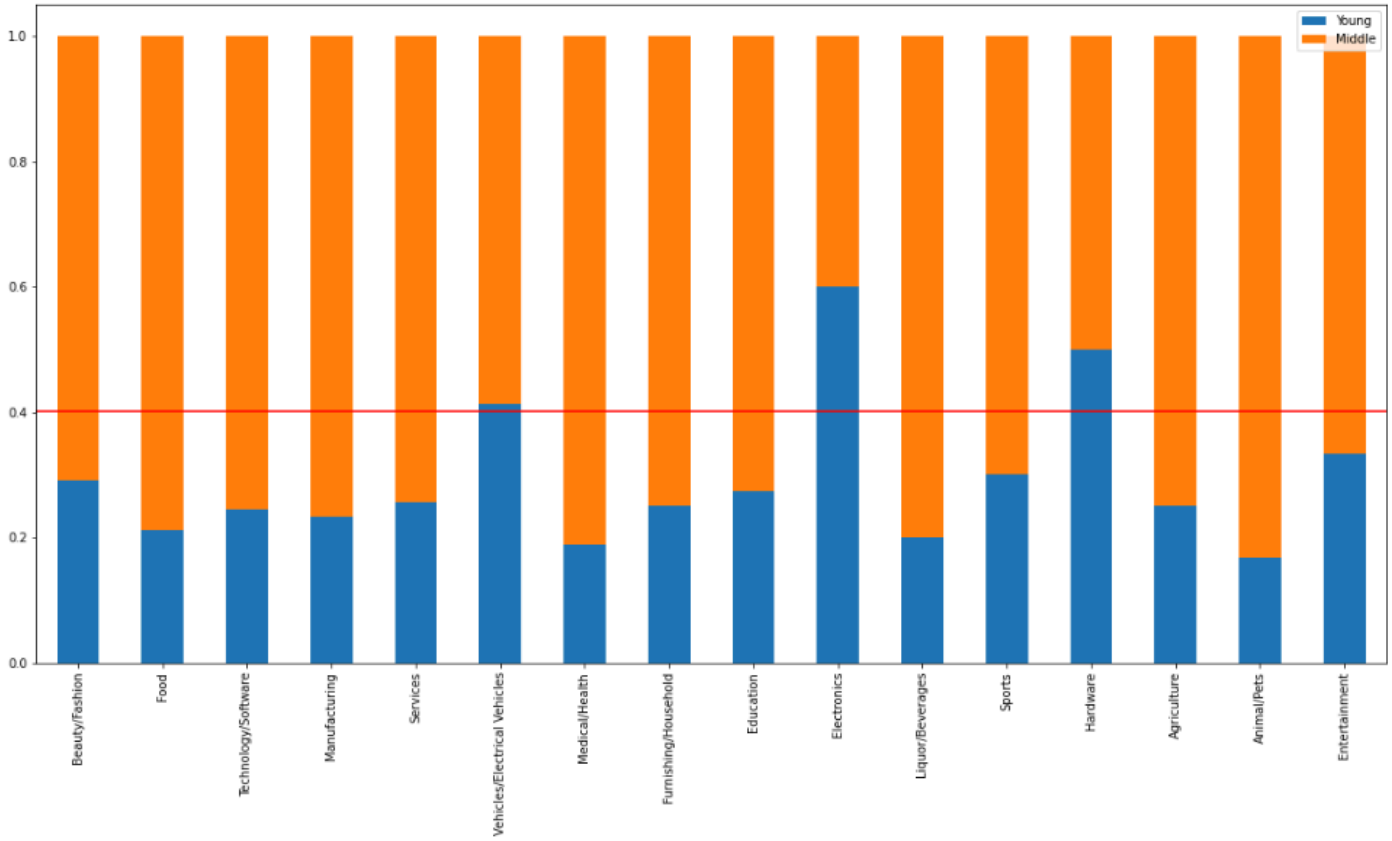


Fig.4 - Age distribution for each industry, Red line shows the expected distribution

As was seen in our analysis, The Electronics industry shows a significant change with Young presenters dominating. And the Liquor/Beverages industry shows a high number of Middle-age presenters.

Risk-taking Propensity

Adith

Individual Sharks will demonstrate consistent risk-taking profiles, with some favouring higher-risk investments than others.

Relevant Parameters

- ➔ SKUs
- ➔ Bootstrapped
- ➔ Has Patents
- ➔ Received Offer

Analysis Methods

Logistic Regression analysis was done with SKUs against Received Offer. Along with this, we also tested the effect of Patents and Bootstrapped startups using a Chi-Square test.

Results

Logistic Regression results:

```
==============================================================
   coef    std err          z      P>|z|      [0.025      0.975]
--------------------------------------------------------------
 0.0014      0.002      0.615      0.539      -0.003       0.006
==============================================================
```

Chi-Square with Bootstrapped vs Received offer:

`p-value: 0.8444523314961103`

Chi-Square with Has Patents vs Received offer:

`p-value: 0.9812035739391029`

Anchoring Effect

Sankalp

Sharks' valuations and investment decisions will be significantly influenced by the startup's initial ask and valuation.

Relevant Parameters

- ➔ Original Ask Amount
- ➔ Original Offered Equity
- ➔ Valuation Requested
- ➔ Total Deal Amount

Analysis Methods

A Linear Regression analysis was conducted with Original Ask Amount, Original Offered Equity, Valuation Requested against Total Deal Amount.

Results and Discussion

	coef	std err	t	P>|t|	[0.025	0.975]
const	22.8522	5.005	4.566	0.000	12.997	32.708
Original Ask Amount	0.6483	0.064	10.151	0.000	0.523	0.774
Original Offered Equity	-1.3391	0.671	-1.997	0.047	-2.660	-0.019
Valuation Requested	0.0004	0.000	1.133	0.258	-0.000	0.001

We observed mostly significant results for this hypothesis but as was pointed out in both our feedback and something we figured in the exploratory analysis stage, we thought that this result would be significant, but decided to test anyway considering that, after an initial analysis, we saw many under-valued propositions compared to the original ask.

Gender and Background Biases

Adith

Systematic differences in investment patterns may exist based on Shark gender or other background factors.

Relevant Parameters

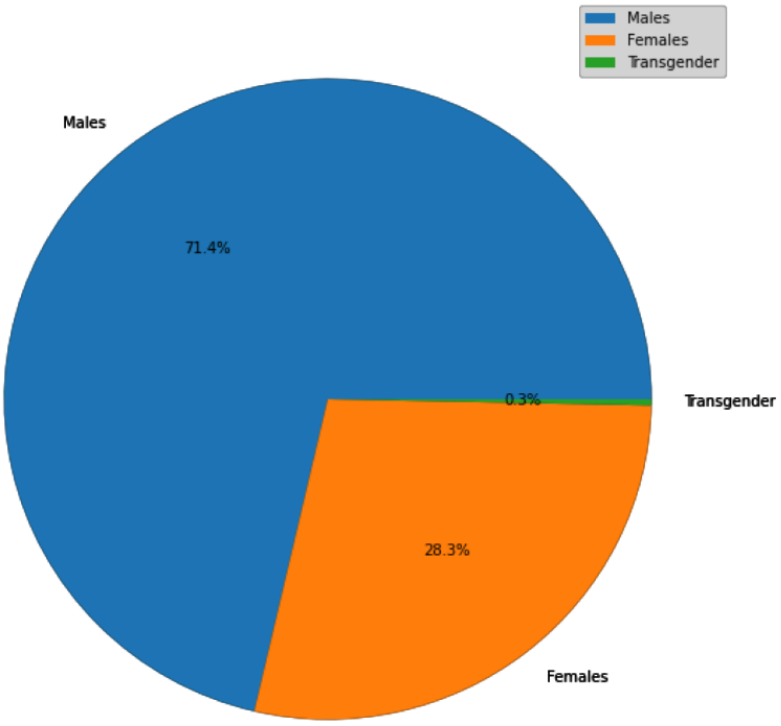
Analysis Methods

An exploratory data analysis should yield any discrepancies followed by a deeper analysis featuring tests. EDA didn't show much in terms of correlations worth testing. We noticed that the shark Amit showed a slight propensity towards Male presenters. We tested this using a correlation analysis between Amit's Investments and the number of Male Presenters in a pitch.

Results

Significant results were observed with a correlation coefficient of the `0.76` and a corresponding p-value of `0.007`.

Visualisations



George *Fig.5 - A pie chart that shows the fraction of presenters of each gender*

The visualisation shows, also, that a large majority of the presenters on Shark Tank India are Males.

Cooperation and Competition

Adith

Initially, we wanted to test whether the presence of Competition among Sharks will lead to more favourable terms for the startup founders. After a few attempts we couldn't accomplish a satisfactory analysis. Instead, we found that we can make an analysis of the effect of cooperation of the sharks on the final deal. We hypothesised that the presence of multiple sharks having a share in the final deal might result in better value for the startup founder.

Relevant Parameters

- ➔ Number of Sharks in Deal
- ➔ A derived value *Favorability* - A ratio between the original valuation versus the final valuation. A value greater than 1 indicates a higher valuation than initially requested.

Analysis Methods

Ordinary Least Squares Linear regression of Favorability against the Number of Sharks in Deal.

Visualisations

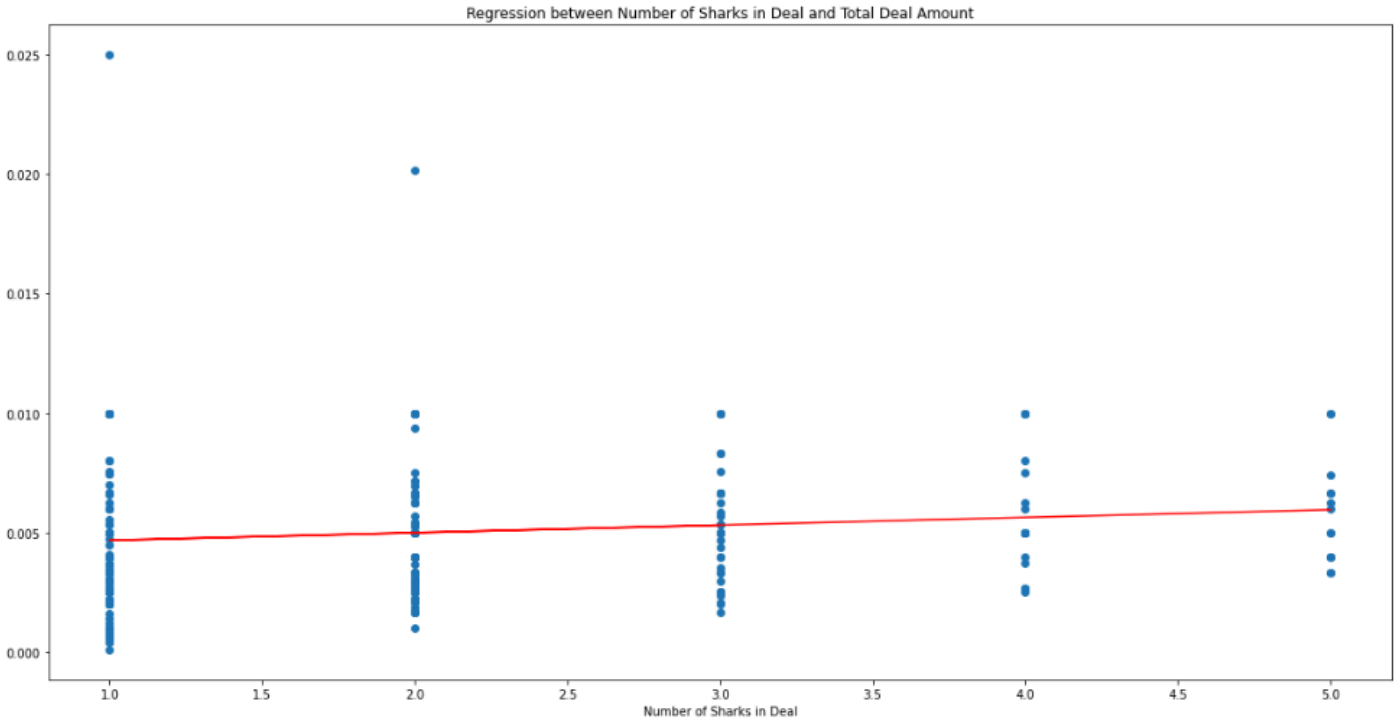


Fig.6 - Regression line produced by this analysis

Results and Discussion

A significant result was observed. With a p-value of `0.043`, We can reject the null hypothesis. The presence of multiple sharks in the deal leads to more favourable terms for the startup.

Other Visualisations

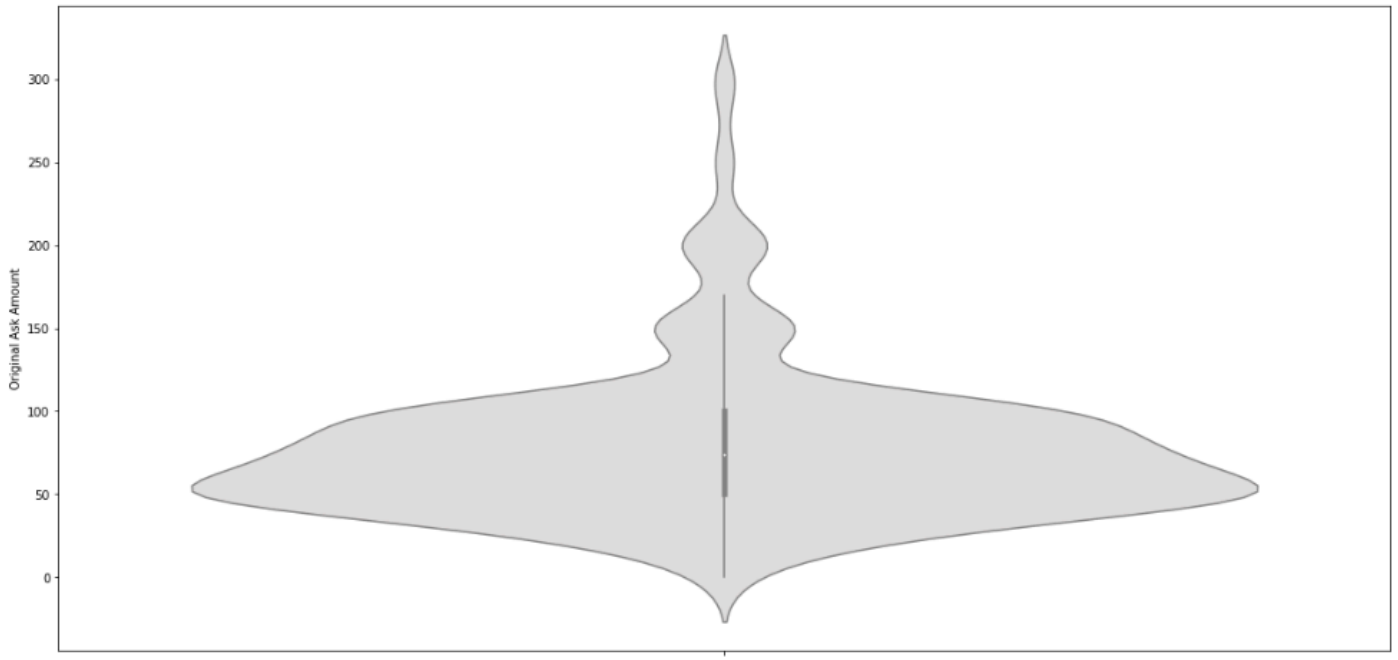


Fig.7 - Violin plot of ask amounts

Some bumps are seen at multiples of 50 lakhs. Also of note is that some presenters entered with very small ask amounts (5rs-100rs) since they valued the sharks' mentorship very highly. These outliers were removed.

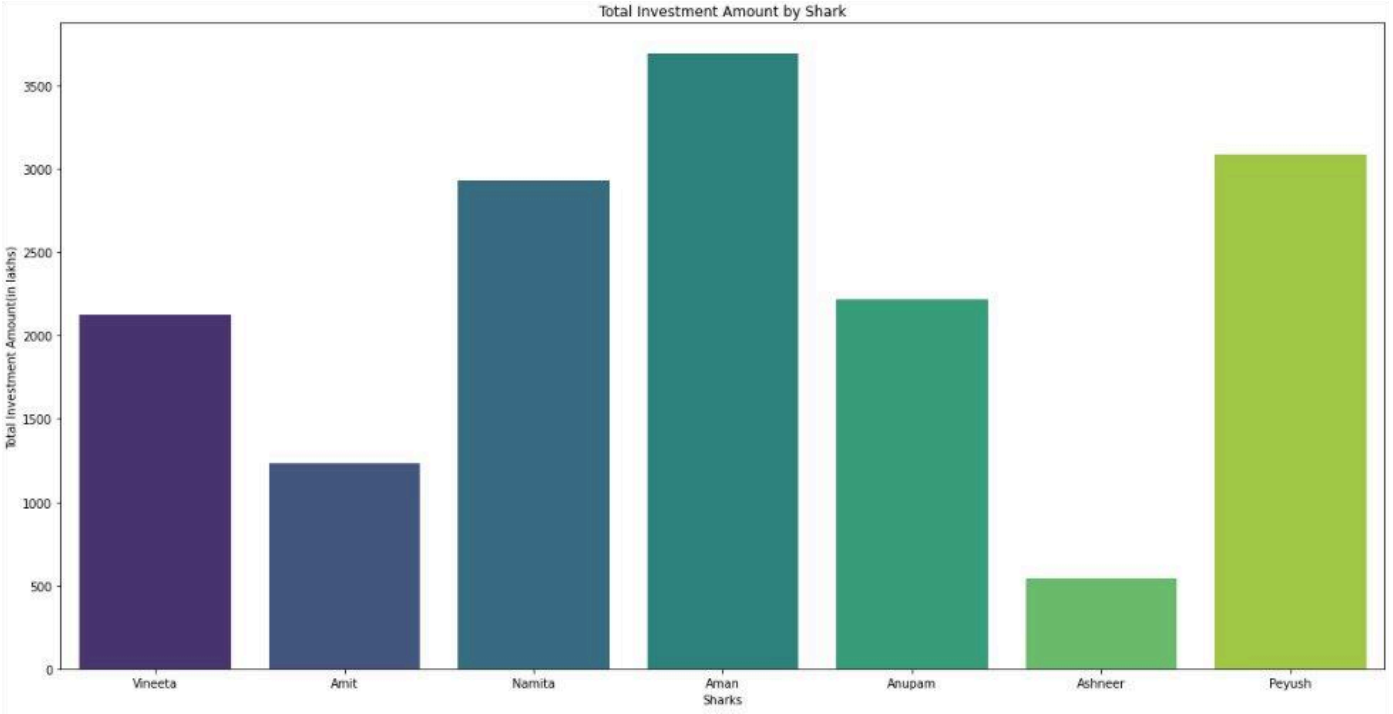


Fig.8 - Total Investment Amounts for each shark

This visualisation is intended to show the general willingness of each shark to invest in startups. Do note, these values were not adjusted for the presence of each shark in each season. Ashneer, in fact, is only present for season 1.

Conclusion

We have taken good care to present all the results of hypotheses that we tested. Importantly because many of the results showed no statistical significance and hence we accept those null hypotheses. What we learned and now hope to show with this analysis is that most sharks invest without a real regard for the cold, hard numbers involved in companies and the pitchers. This aligns with an underlying message of the show - the sharks being business magnates that pride themselves on the sixth sense of whether a company will succeed. All in all, we may not have fully succeeded with our target of helping future startups to make better pitches but have shone light on the nature of short-form pitches towards investors that operate mostly with the intangible aspects of presenters and the company behind them.