Analysis of Shark Tank Data

Catherine Rauch

Abstract

This report examines 10 seasons of “Shark Tank” to examine the breakdown of pitches, and how these changed over time, as well as how the Sharks spend their money. Methods used in this report include: grouping, aggregating with pandas, graphs and linear regression in Altair. Exploratory data analysis showed that a majority of deals had an equity stake less than 40% and were valued at less than $400,000. Some key results include: 56 percent pitches got a deal, and the most popular pitch categories were Lifestyle/Home, Fashion/Beauty, and Food/Beverage. The overall value of companies increased over the seasons, starting at just $200,000 and increasing to $1.8 million. A total of approximately $195 million dollars was invested over the 10 seasons in 497 deals.

Introduction

The primary goal of this report is to explore and analyze the dataset, looking specifically at the composition of deals, how these changed over the seasons, how the sharks have invested their money and how likely you are to get a shark to invest with your valuation. The show "Shark Tank" is a TV series on NBC; it follows a group of investors (the sharks) and those pitching their business (the entrepreneurs). The idea is to convince the other side to accept their valuation of the business and negotiate a deal.

There a 6 “Sharks”: Mark Cuban, Robert Herjavec, Barbara Corcoran, Lori Greiner, Daymond John and Kevin O’Leary. Additional data about the sharks is included in the appendix [1]. An entrepreneur’s pitch revolves around a valuation, which is what the company is currently valued at in the market. Investopedia [2] explains that valuation is usually determined through factors such as revenue, earnings, and the value of other companies within the same sector. If an entrepreneur asks for a 10% equity stake in exchange for $100,000 they are valuing their company at a $1 million valuation. Most entrepreneurs come in with high valuations, while the sharks try to undercut with a counter at much lower valuations. A recent analysis of the shark tank data was published on the Science of People website by Vanessa Van Edwards [3], she analyzed 495 pitches and found 10 specific tactics successful entrepreneurs can use to get their own yes. Successful pitches were more effective in 10 ways: Credible, Agreeable, Interactive, Captivating, Relevant, Entertaining, Confident, Powerful, Funny, and Inspirational. Another analysis from 2018 by Mithun Desai [4], worked to develop a predictive model to predict deal or no deal using Text Mining with a CART Model, Random Forest, and with logistic regression.

The dataset used in this project contains hand-compiled data from season 1 to season 10 (2009 to 2019). It includes information about company name, deal status, category of pitch, gender of team, valuation of company, and what shark invested.

According to Wikipedia [5], over $150 million has been invested to date and the businesses that appear on the show get the unique opportunity to be broadcast to 8 million viewers. Examining the data from this show can gain insights into how to successful seek an investment in any kind of company and can shed light on what these investors are looking for in a business presentation.

Questions of Interest

1. What is the breakdown of pitches? What percent got deals?

2. How have pitches and deals changed over the seasons?

3. How do the sharks invest their money, ie do they favor certain categories? Who has the most money invested?

4. How likely is an entrepreneur to get the valuation originally asked from a shark?

Methods and Data

Data frames will be used for easy data manipulation as calculation will need the data to be aggregatated and sorted. Plots such as bar, line, scatter, density and a heat map will be used to illustrate results. Multiple linear regression lines will be fit individually to Shark data to understand how likely an entrepreneur is to get a deal from a shark at the asking valuation.

The data set was compiled by Halle Tecco, a self proclaimed shark tank fanatic, who is an adjunct professor at Columbia Business School, and an investor at Techammer. The data was copied to an excel workbook from a google docs link [6].

Each data row is one pitch on the show, it includes company name, gender of entrepreneurs, and category of business. From first glance, it appears than men led pitches dominate the data, with few women and mixed teams. Additionally, a large number of businesses appear to be either Food/Beverage or Fashion/Beauty based. This is not a feature of the data collection but rather because the data was drawn from a recorded show, the makeup of contestants was chosen through application via show producers.

As the data was collected by hand, by a single person, it's possible that incorrect or incomplete information was included in the data this could indicate that the "principle measurement" of precision was violated. As Shark Tank is a nationally televised show and watching the recording is how the data was collected, it is unlikely, another principle, distortion would be relevant.

The analysis of this data is also unlikely to cause any ethical issues; contestants on the show must sign a contract and have released the rights for their name/company and any interactions are nationally televised.

From the initial overview of the data, it became apparent that missing values were recorded as NAN. These values were replaced with N/As. Columns were renamed to be more easily used and then preliminary scatter plots were examined for any possible relationships between variables. It was also discovered that data from Seasons 1-4 was missing in the equity/amount/valuation asked for columns, but did contain data for equity/valuation/amount received in the case of a deal.

Figure 0

From figure 0, some preliminary relationships can be seen. The greater the valuation the lower the equity taken by the shark, that graph appears to be very similar to equity asked and valuation given in the deal. Valuation has increased in more recent seasons. It also can be seen that valuation and amount asked for are loosely linearly positive related, while amount given and valuation given are less so. To deeper explore the variables Equity and Dollar Amount, density graphs were gen

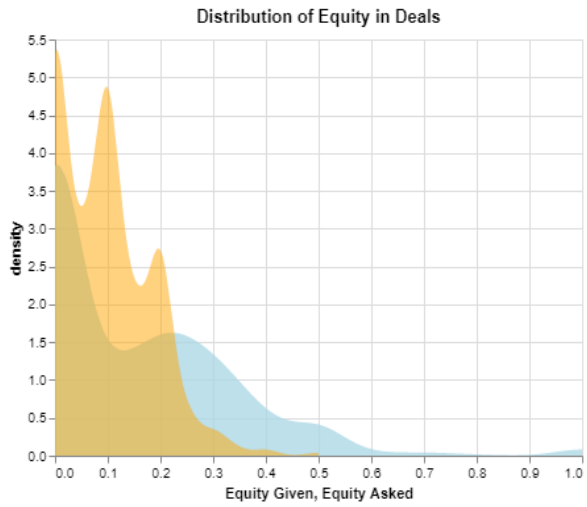


Figure 1.1

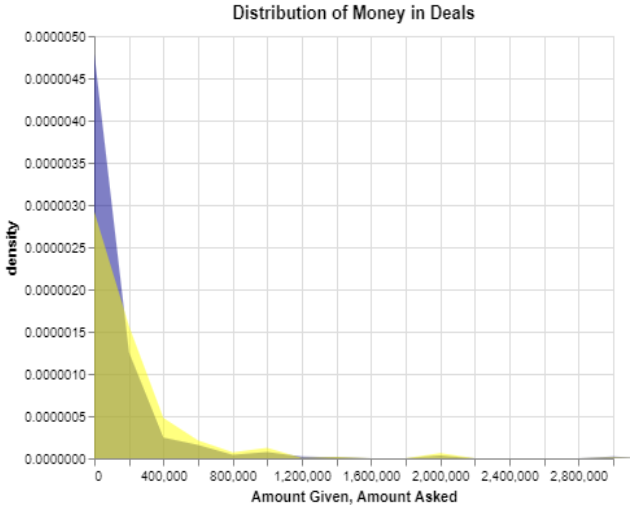


Figure 1.2

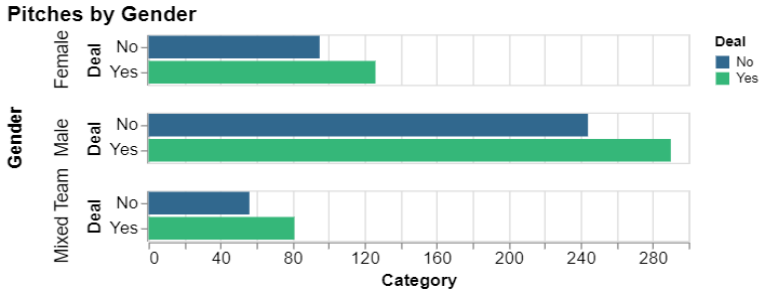
From figure 1.1 it can be seen that for most deals, equity asked was below 20% and none were above 50%. With the equity actually given being more spread out, the majority being below 40%. From figure 1.2 it's clear that the amount of money asked for and the amount of money given has more overlap, with a more concentrated density of amount given below $200,000.

Analysis, Results and Interpretation

Question 1

To determine the composition of pitches and calculate what percent got deals, the data was grouped by deal status and then gender of the pitching team; those results are graphed and discussed below.

Figure 2



About 56 percent of entrepreneurs that pitched got a deal. There was a total of 892 pitches ,of which 534 were made by teams of men, that's about 60% of all pitches, 25% were women and the remaining 15% were mixed teams. As seen in Figure 2, although significantly more men made it on the show to pitch, they are slightly less likely to get a deal. 54% percent of all men teams successful while for all women teams 57% struck a deal. Mixed teams had the fewest total number of pitches, but they had the best luck convincing the sharks to make a deal at 59% successful.

Breaking down the pitches by category, Food and Beverage is the most popular pitch category with 181 total pitches. The next two popular categories, as seen in figure 3, are Fashion/Beauty at 167 pitches and Lifestyle/Home with 139. These 3 categories make up more than half of all pitches seen on the show.

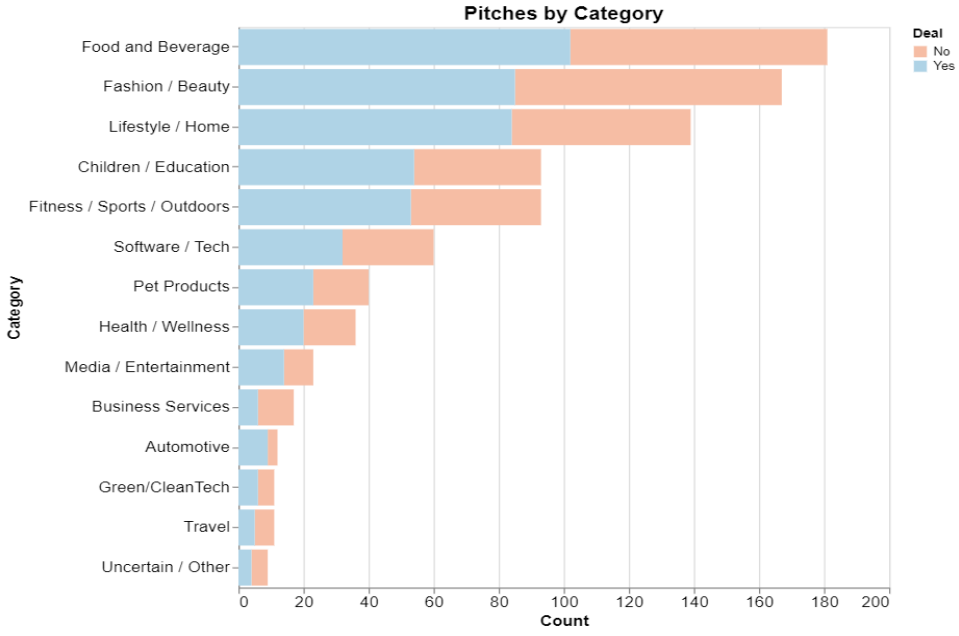


Figure 3

In these categories, while the automotive category has less than 20 total pitches, it has the highest chances of getting a deal at 75%. The second highest at 60% is Fashion/Beauty, usually clothing

companies and Lifestyle/Home, often products. While the lowest category with 35% was Business Services which includes service event planning companies. The only other categories below a 50% closing percentage were in Travel such as a luggage shipping company.

Question 2

To illustrate how pitches and deals have changed over the seasons the data was grouped by season and the mean calculated on the aggregated data frame.

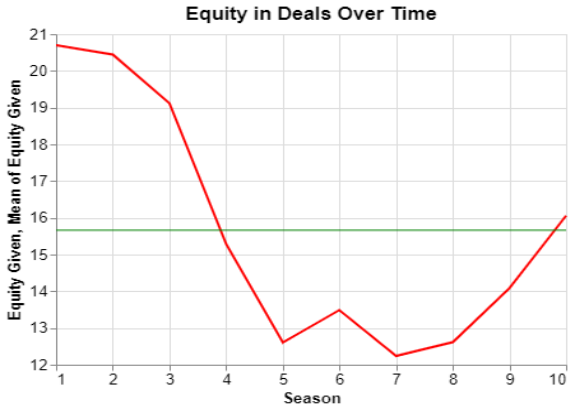


Figure 5.1

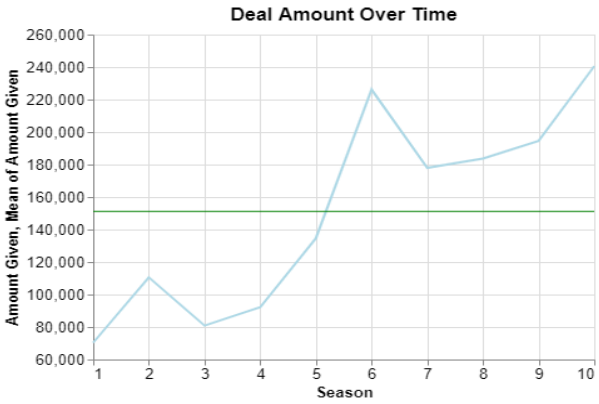


Figure 5.2

During season 1, entrepreneurs gave up over 20% of their company for a deal which is well above the average 15.7%. It can be seen in figure 5.1, as time went on, average equity demanded dropped considerably. This coincided with average money given in a deal increasing; this could be as the show gained in popularity more successful companies were pitched on the show. The bigger a company is before appearing on the show, the less equity they need to give up at a higher valuation, this can be confirmed in figure 5.3 where its demonstrated where valuation has in general been increasing.

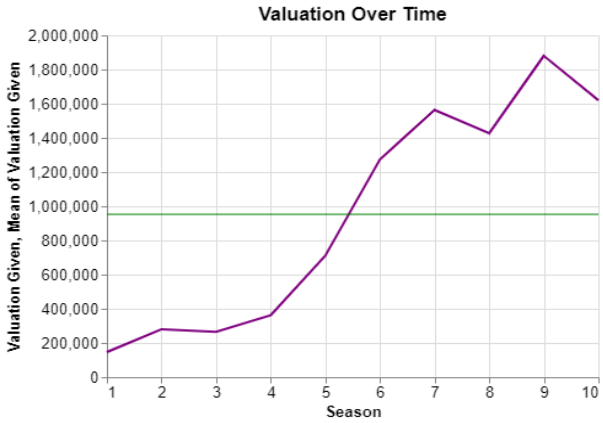


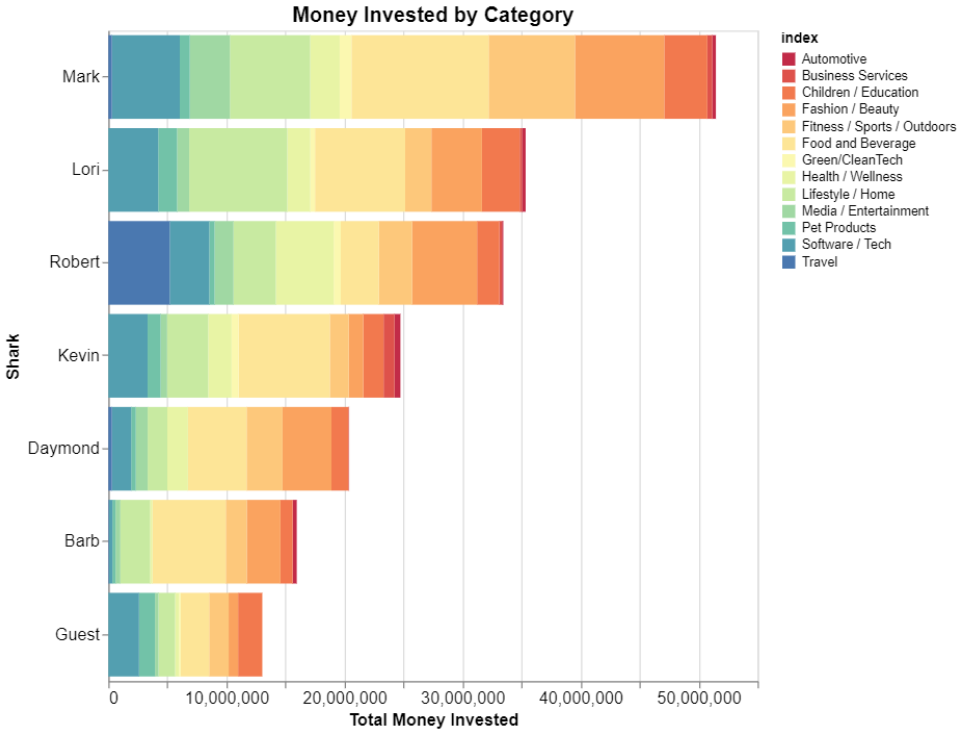
Figure 5.3

In the first season, the average worth of a deal was $70,317, this is below the total average for all ten seasons, $ 151,163, and this can be seen in figure 5.2. With the total amount of money in deals increasing and the sharks taking less equity, valuation has gone up commensurately as can be seen in figure 5.3

Question 3

To explore how the shark panel has invested their money, separate data frames were created, each containing data about the deals each shark made individually. It was found that there is a total of $194,409,666 dollars invested over the 10 seasons in 497 deals. Mark Cuban has the most money invested at $52,715,000, in a total of 151 deals, which is over 30% of all deals made in the shark tank. This might be expected as he also has the highest net worth of the 6 sharks. Lori, although having significantly less money invested at $35,350,000 in 118 deals, her total deals account for about 25 % of all contracts made on the show.

Figure 6



To better visualize which shark invests in which categories the most, a heat map was graphed below. In figure 7 it can be seen that Mark Cuban has the most deals in Food and Beverage while Lori has the most in Lifestyle/Home.

Figure 7

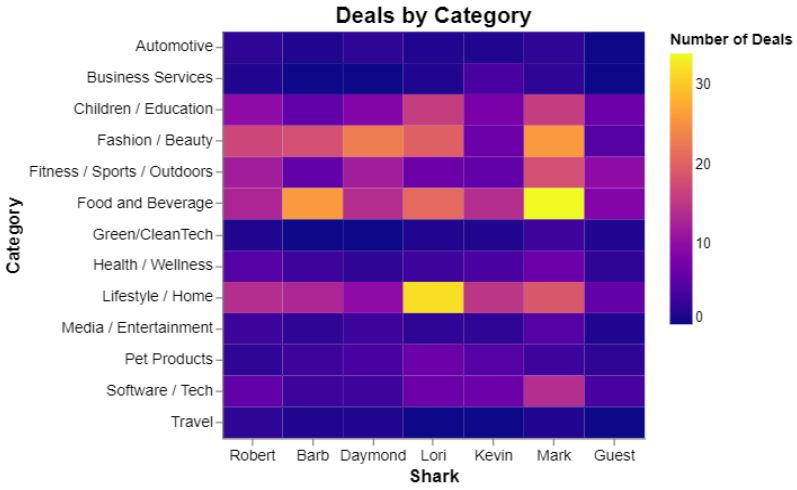


Figure 7 also confirms what was found earlier, the sharks all appear to strike deals in Food and Beverage, Fashion/Beauty and lifestyle/Home the most. Additionally all seem to avoid Green/CleanTech, which could be a result of fewer total pitches. Both Barbara and Daymond stay away from deals in Pet products, Media/Entertainment and Software/Tech. It's also worth noting that in a rather difficult category to understand, Software/Tech, Mark who made his fortune selling software companies has the most deals. Daymond specializes in Fashion / Beauty deals, which can be expected as he made in fortune with a fashion brand. Mark and Robert appear to have the most well rounded categories having at least 15 deals in more than 5 categories.

Question 4

To explore how sharks differ when accepting a valuation, a plot of valuation given vs valuation asked was generated.

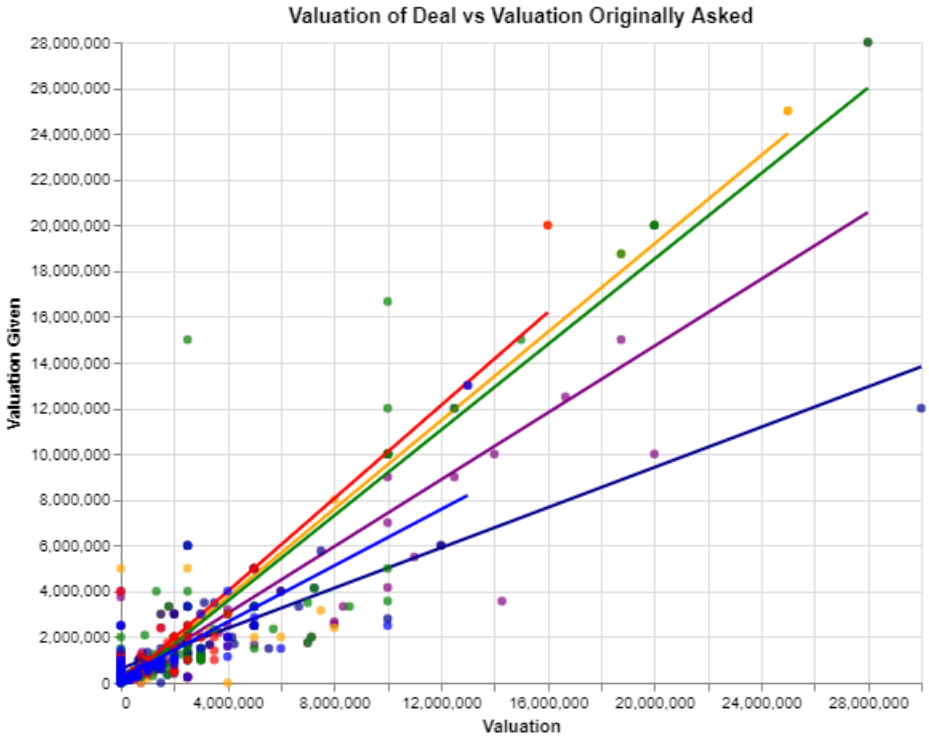


Figure 8

Orange - Kevin
Blue - Daymond
Dark Blue – Robert
Green – Lori
Red - Barbara
Purple - Mark

A flatter slope means a shark demands a lower valuation than what was pitched by the entrepreneur. As can be seen in figure 8, Lori and Kevin have a slope of close to 1; this means they are most likely to give an entrepreneur a deal at the valuation they asked for. Barbara has a steeper slope, which implies she is more likely to give you a higher valuation then you asked for. This could be a result of either kindness or could be a strategic edge, offering either more money or a smaller equity stake to edge out another shark. Mark and Daymond both had flatter slopes. This indicates they were less likely to give an entrepreneur the valuation they were asking for. They are more likely to suggest a lower valuation; this usually means a larger amount of equity for the same amount of money of the original pitch. And Robert has the flattest slope indicating that he is the most likely to undercut an entrepreneurs valuation, this could be due to the categories he often invests in, as certain entrepreneurs tend to overestimate their own companies. It can be concluded that different shark invest differently. Some invest more than

others. But they all commonly invest in their interests and where their expertise lay. Certain sharks are more likely to agree with your valuation while others are prone to undervaluing.

Conclusions and Future Work

The primary goal of this report was to explore and analyze the dataset. It was successful in analyzing the composition of deals, how these changed over the seasons, and how the sharks have invested their money. However due to missing information on asking price and asking equity graphs for the first 4 seasons, calculations requiring this information was omitted. From the data that was complete, it was found that more than half all pitches were successful in striking a deal.

Men led teams were the most represented, but mixed gender teams were the most likely to close a deal. Categories with the most pitches made up more than half of the total pitches, these categories were Lifestyle/Home, Fashion/Beauty, and Food and Beverage had. While these had the most, the most successful were those in the Automotive category while business services had the lowest percentage. In the early days of the show (circa 2009), equity demanded from the entrepreneur's was high, but after the first season it started to fall exponentially. On the other side, the amount of money provided by the sharks in a deal started low and increased linearly. In general, the overall worth of companies increased substantially over the seasons. There is close to $195 million invested in all the deals on shark tank with Mark Cuban having the most money invested at just over $50 million.

Together, Lori Greiner and Mark Cuban alone, account for over half the deals done in the shark tank. Generally, the Sharks invest differently. Some invest in all kinds of business; others invest in just their interests. Many avoid categories that don't align with their expertise. Certain sharks are more likely to agree with an entrepreneur's original valuation while others are prone to disagreement, undervaluing and counteroffers; they're called sharks for a reason after all. From this data, future work could include exploring a prediction model for a deal using asking valuation and category as predictors.

Appendix

1. Lois Crouse and Martha Hurwitz. "Shark Tank Cast - Who Are The Sharks & Guest Sharks?" allsharktankproducts.com/shark-tank-cast/

Mark Cuban, owner of the Dallas Mavericks, he made his money selling companies to Yahoo ($5.7B) and CompuServe ($6M); Robert Herjavec, the CEO of the Herjavec Group (company specializing in IT Security), made his money selling to AT&T Canada ($30.2M) and Nokia ($225M); Barbara Corcoran, specializes in real estate, sold The Corcoran Group for $66 million; Lori Greiner, known as the "The Queen of QVC" (a free-to-air television network), is an entrepreneur who has created more than 600 products and holds over 120 patents; Daymond John, a fashion expert who founded the FUBU apparel company; Kevin O'Leary who made his money through venture capital investments and owns O'Leary Mortgages, O'Leary books, O'Shares Investments and O'Leary Fine Wines.

2. Jea Yu. "How Is a Business Valued on 'Shark Tank?"

www.investopedia.com/articles/company-insights/092116/how-business-valued-shark-tank.asp.

3. Vanessa Van Edwards "Learn the Secrets Behind the Best Shark Tank Pitches of All Time."

www.scienceofpeople.com/shark-tank-pitch/

4. Mithun Desai. "Text Mining and Predictive Modelling on Shark Tank."

www.tabvizexplorer.com/text-mining-and-predictive-modelling-on-shark-tank/.

5. Wikipedia, main page for Shark Tank: includes information about ratings and viewers

https://en.wikipedia.org/wiki/Shark_Tank

6. Dataset source
https://docs.google.com/spreadsheets/u/1/d/1Lr0gi_QJB_JU0lBMjJ7WiBRxA0loml1FlM-KlmKsaEY/htmlview?pli=1#