

# Does Industry Type Affect Making a Deal on Shark Tank?

STAT 213 Elements of Prob & Math Stat

Joshua Jones, Johanna Lee, Lan Nguyen

## **Abstract:**

It is proposed that certain industry sectors are more lucrative to invest in than others, and this idea has been gaining popularity after the rise of start-up-pitching shows, such as Shark Tank, The Profit, and Billion Dollar Buyer. Thus, there is a focus on investing in certain industries, such as retail, construction, and wholesale. With start-ups gaining popularity quicker than ever before, in this project, we aim to explore if there is a relationship between the industry type and choice to be invested in by examining observed and expected values and determining whether the categories and deal acceptances are dependent. After collecting data about Shark Tanks from Kaggle and summarizing the statistics of industry types and whether the investment choices succeed or fail, a Pearson's Chi-squared test showed that the p-value is greater than the significance level and thus, there was insufficient evidence to support that the industry choice would impact the possibility of its investment acceptance or rejection.

## Background:

Shark Tank is a well-known American business reality television series that gives entrepreneurs the chance for their business dreams to become a profitable empire. Our research focuses on determining if there is some categorical bias in the industry that sharks invest in. In the American business world, the most profitable industries are mineral products, aerospace, machinery & mechanical appliances, automotive, and software according to Forbes. On the other hand, according to Investopedia, food, healthcare, and technology firms are considered to be the safest industries because the products are essential for human life. The business world is driven by demand, and entrepreneurs strive to create products to fulfill this need. To support these entrepreneurs, investors like the Sharks put in their own money hoping to help that product succeed so that they can also profit. The question that we want to address is whether there is a relationship between sharks saying yes to a start-up based on their product's industry. The null hypothesis is that the industry type and whether sharks say yes to invest are independent. The alternative hypothesis is that type of industry and the deal's acceptance are dependent on each other.

## Methods:

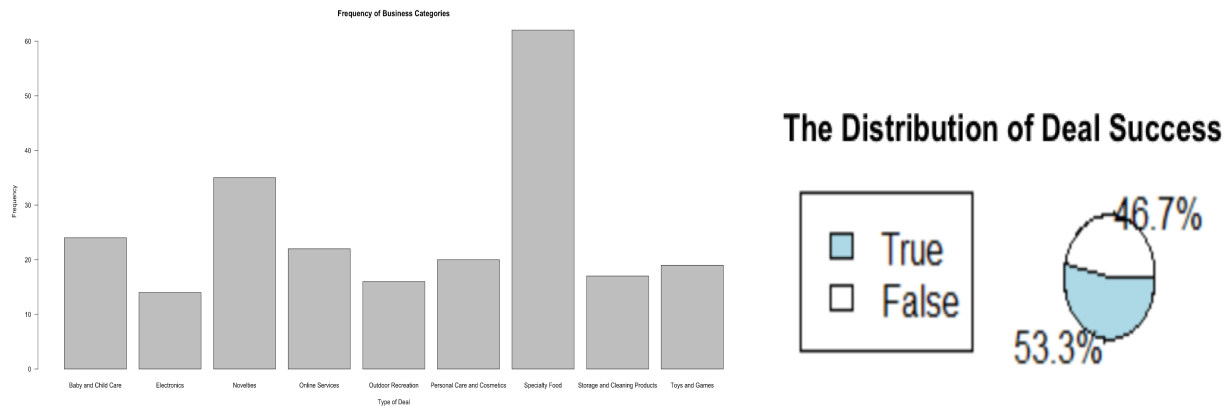
Through Kaggle, we utilized a dataset called "Shark Tank Pitches" which contained 495 different company pitches and their corresponding industry type, deal success, entrepreneurs, etc. This information was collected based on the first six seasons of the show. We do not expect our data to contain any bias in collection since this is directly from the show. However, the Sharks may be skewed in their decisions as some are more specialized in specific fields. Additionally, there may be bias in the type of companies that are able to be pitched in front of the sharks.

Our two main variables were industry category type (categorical) and whether the company was able to make a deal (categorical). There was a total of 32 different categories, hence we cleaned and simplified the data to only include only those categories that appeared more than 10 times, resulting in a total of 9 unique categories (Baby and Child Care, Electronics, Novelties, Online Services, Outdoor Recreation, Personal Care and Cosmetics, Specialty Food, Storage and Cleaning Products, Toys and Games). This simplified data had 229 unique variables which were run for a univariate analysis for our two main variables: a pie chart for deal acceptance and a bar chart for business category type. Our study represents a Case CC situation, therefore we ran a bivariate summary analysis and created a stacked bar chart. Finally to test our hypothesis, we ran our inferential statistics by completing a chi-square test.

## Results:

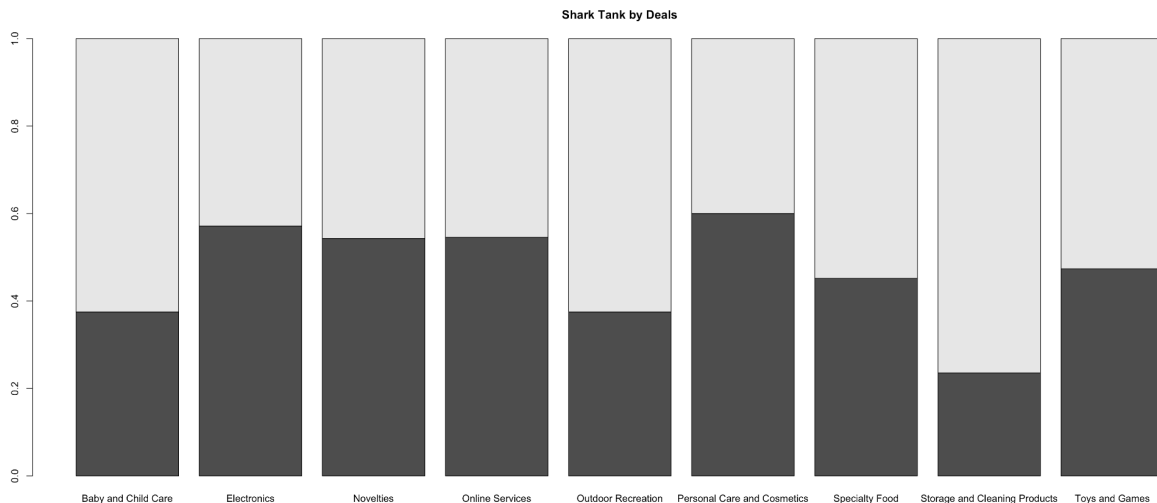
With a total sample size of 229, 9 different categories of industries and 2 choices of True or False, 53.3% of the deals receive a True result, being accepted by the sharks, while the other 46.7% is False (Fig. 2). As one can see from Fig. 1 below, specialty food is the most common business category invested in by far at a frequency of 62, followed by novelties at around 35 occurrences. Electronics is the least common category to be invested in at a frequency of 14, but not significantly lower than other less common choices such as Storage and Cleaning Products and Personal Care and Cosmetics, respectively at frequencies of 17 and 20.

**Figures 1 and 2: The Respective Distribution of Industry Choice of Investment and Investment Decision (False/True) Results in the Sample**



After cleaning up the data to include only the 9 most commonly invested industry, the most likely industry to be invested in (i.e. the one with the highest probability of being invested in), based on Figure 3 as shown below, is Storage and Cleaning Products (13 True's and 4 False's), followed by Baby and Child Care (15 True's and 9 False's). Interestingly, these are some of the two least chosen industries to be invested in as mentioned from other sources. Novelties and online services both have almost equivalent chances of being rejected or accepted, as seen in Figure 3. The industry category most likely to be rejected is Personal Care and Cosmetics, but even then, the difference between being rejected and accepted are not significant.

**Figure 3: Stacked Bar Chart of the Possibility of Being Accepted (White) or Refused (Black) for Each Category of Investment.**



One can use the Pearson Chi-square test with significance level of 0.05, given the observed counts from the graph above and the expected counts under the assumption that the null hypothesis is true, to test if the two categorical variables are dependent on each other, or alternatively, if one categorical variable has any impact on the other. After finding the test statistic, which is 8.4765, one could determine the probability of obtaining the sample statistics available while assuming the null hypothesis was true by taking the p-value of the test statistic

on a t-distributed curve with a 8 degrees of freedom. This p-value was found to be 0.3884, which is considerably higher than the alpha level, hence we fail to reject the null hypothesis. There is insufficient evidence to suggest that the industry choice of investment will impact the possibility of being accepted or rejected by the sharks. All calculations were done in RStudio and can be found in the Appendix.

### **Discussion:**

To conclude, this study was conducted to find if there is a correlation between the industry choice and the possibility of being accepted by entrepreneurs as this has been seen as a trend in the entertainment industry on shows such as Shark Tank. This research helps start-ups and founders see the landscape of the investing decision. The null hypothesis is that the industry choice and whether entrepreneurs say yes to the product are independent of one another. On the other hand, the alternative hypothesis is that the product's industry and whether entrepreneurs choose to invest in the product are dependent on each other. We concluded that the test statistic is 8.4765 and the p-value is 0.3884. Since the p-value of 0.3884 is larger than the significance level of 0.05, we fail to reject the null hypothesis. There is insufficient evidence to support the claim that there is a correlation between the industry of the product and whether entrepreneurs are going to accept and invest in that product. This doesn't suggest investors don't have a preference. Some may be more likely to invest in something because as seen in the chart, there are a lot more Food Industries (62) products that sharks wanted to invest in than others such as Outdoor Recreation (16), but rather the chance of being invested is not dependent on the industry type. This is different from what the research from Investopedia suggests that certain industries are more profitable than others, but this study shows that there is no relationship between Sharks on Shark Tank investing in a product and the industry of the product. There are limitations to our study as Sharks are experienced investors and know how to decipher which investments will be profitable. Moreover, on a show like Shark Tank and as an entrepreneur, whether they get that investment is also dependent on their pitch and what they are able to bring to the table. Our shortcomings are that we cleaned up the table, so not all data in the various industries were presented. We only used the data of the nine most popular industries. Also, there are different investors as the seasons go on and different investors may result in different answers. Finally, Shark Tank is a televised TV show, hence, it is possibly not as accurate or objective as the reality of the investing world. Therefore, there is a possibility that there is a relationship between the industry of the product and the possibility of getting invested in by the sharks. A future study could be performed about how as time goes on, how different economic scenarios and different investor choice and preferences according to the market affect the relationship between a product's industry and whether an investor will invest in it and do this using data from non-televised deals from different investors.

**References:**

DePietro, A. (2018b, July 11). *The Most Profitable Industry In All 50 U.S. States*. Forbes.  
<https://www.forbes.com/sites/andrewdepietro/2018/07/11/most-profitable-industry-state/?sh=46fbf0a26c9>

*The Safest Industries to Invest In*. (2022, January 31). Investopedia.  
<https://www.investopedia.com/articles/investing/052815/safest-industries-invest.asp>

*Shark Tank Pitches*. (2017, December 19). Kaggle.  
<https://www.kaggle.com/datasets/rahulsathyajit/shark-tank-pitches?resource=download>

**Appendix:***Univariate Summary of Deal Distribution*

```
> deals <- table(Sharktank$deal)
> deals
FALSE TRUE
107 122
> pie(deals,main="The Distribution of Deal Success")
```

*Univariate Summary of Industry Distribution*

```
> category <- table(Sharktank$category)
> category
Baby and Child Care    Electronics    Novelties    Online Services    Outdoor Recreation
24          14          35          22          16
Personal Care and Cosmetics    Specialty Food    Storage and Cleaning Products    Toys and Games
20          62          17          19
> barplot(category,xlab="Type of Deal",ylab="Frequency",main="Frequency of Business Categories")
```

*Bivariate Summary Statistics*

```
> t1 <- table(Sharktank$deal, Sharktank$category)
> barplot(prop.table(t1,margin =2), main = "Shark Tank by Deals")
```

*Inference:*

```
> t2 <- table(Sharktank$category, Sharktank$deal)
> t2
```

	FALSE	TRUE
Baby and Child Care	9	15
Electronics	8	6
Novelties	19	16
Online Services	12	10
Outdoor Recreation	6	10
Personal Care and Cosmetics	12	8
Specialty Food	28	34
Storage and Cleaning Products	4	13
Toys and Games	9	10

```
> dealtest <- chisq.test(t2, correct = FALSE)
> dealtest
Pearson's Chi-squared test
data: t2
X-squared = 8.4765, df = 8, p-value = 0.3884
```