

DATA_301_Assignment_02A_Ishaan_Sathaye_and_Sreshta_Talluri

October 23, 2023

This assignment is in two parts. This is Part A; make sure to also complete Part B. Read the notebook from the beginning and answer the questions as you go. You can add as many cells as you want. Submission instructions are at the end. See Canvas for general rules about Assignments and collaboration.

1 Shark Tank

Shark Tank is a reality TV show. Contestants present their idea for a company to a panel of investors (a.k.a. “sharks”), who then decide whether or not to invest in that company. The investors give a certain amount of money in exchange for a percentage stake in the company (“equity”). If you are not familiar with the show, you may want to watch part of an episode [here](#) to get a sense of how it works.

The data that you will examine in this lab contains data about all contestants from the first 6 seasons of the show, including: - the name and industry of the proposed company - whether or not it was funded (i.e., the “Deal” column) - which sharks chose to invest in the venture (N.B. There are 7 regular sharks, not including “Guest”. Each shark has a column in the data set, labeled by their last name.) - if funded, the amount of money the sharks put in and the percentage equity they got in return

To earn full credit on this lab, you should: - use built-in `pandas` methods (like `.sum()` and `.max()`) instead of writing a for loop over a `DataFrame` or `Series` - use the split-apply-combine pattern wherever possible

Of course, if you can’t think of a vectorized solution, a `for` loop is still better than no solution at all!

```
[18]: import pandas as pd
```

1.1 Question 0. Getting and Cleaning the Data

The data is stored in the CSV file <https://dlsun.github.io/pods/data/sharktank.csv>. Read in the data into a `Pandas DataFrame`.

```
[19]: # YOUR CODE HERE
df_shark = pd.read_csv("https://dlsun.github.io/pods/data/sharktank.csv")
df_shark
```

| [19]: | Season | No. in series | Company | Deal |
|-------|--------|---------------|-----------------------------|-------|
| 0 | 1.0 | 1.0 | Ava the Elephant | Yes \ |
| 1 | 1.0 | 1.0 | Mr. Tod's Pie Factory | Yes |
| 2 | 1.0 | 1.0 | Wisspots | No |
| 3 | 1.0 | 1.0 | College Foxes Packing Boxes | No |
| 4 | 1.0 | 1.0 | Ionic Ear | No |
| .. | ... | ... | ... | ... |
| 490 | 6.0 | 28.0 | You Kick Ass | Yes |
| 491 | 6.0 | 29.0 | Shark Wheel | Yes |
| 492 | 6.0 | 29.0 | Gato Cafe | No |
| 493 | 6.0 | 29.0 | Sway Motorsports | Yes |
| 494 | 6.0 | 29.0 | Spikeball | Yes |

| | Industry | Entrepreneur | Gender | Amount | Equity | Corcoran |
|-----|----------------------|--------------|--------|-----------|--------|----------|
| 0 | Healthcare | | Female | \$50,000 | 55% | 1.0 \ |
| 1 | Food and Beverage | | Male | \$460,000 | 50% | 1.0 |
| 2 | Business Services | | Male | NaN | NaN | NaN |
| 3 | Lifestyle / Home | | Male | NaN | NaN | NaN |
| 4 | Uncertain / Other | | Male | NaN | NaN | NaN |
| .. | ... | ... | ... | ... | ... | ... |
| 490 | Children / Education | | Female | \$100,000 | 10% | NaN |
| 491 | Fitness / Sports | | Male | \$225,000 | 8% | NaN |
| 492 | Uncertain / Other | | Female | NaN | NaN | NaN |
| 493 | Green/CleanTech | | Male | \$300,000 | 20% | NaN |
| 494 | Fitness / Sports | | Male | \$500,000 | 20% | NaN |

| | Cuban | Greiner | Herjavec | John | O'Leary | Harrington | Guest |
|-----|-------|---------|----------|------|---------|------------|-------|
| 0 | NaN | NaN | NaN | NaN | NaN | NaN | NaN \ |
| 1 | NaN | NaN | NaN | 1.0 | NaN | NaN | NaN |
| 2 | NaN | NaN | NaN | NaN | NaN | NaN | NaN |
| 3 | NaN | NaN | NaN | NaN | NaN | NaN | NaN |
| 4 | NaN | NaN | NaN | NaN | NaN | NaN | NaN |
| .. | ... | ... | ... | ... | ... | ... | ... |
| 490 | 1.0 | NaN | NaN | NaN | NaN | NaN | NaN |
| 491 | 1.0 | NaN | 1.0 | NaN | NaN | NaN | 1.0 |
| 492 | NaN | NaN | NaN | NaN | NaN | NaN | NaN |
| 493 | 1.0 | NaN | NaN | NaN | NaN | NaN | NaN |
| 494 | NaN | NaN | NaN | 1.0 | NaN | NaN | NaN |

| | Details / Notes |
|-----|-----------------|
| 0 | NaN |
| 1 | NaN |
| 2 | NaN |
| 3 | NaN |
| 4 | NaN |
| .. | ... |
| 490 | NaN |

```

491 10% royalty until $500K; then converts to 5% e...
492                                     NaN
493                                     NaN
494                                     NaN

```

```
[495 rows x 17 columns]
```

There is one column for each of the sharks. A 1 indicates that they chose to invest in that company, while a missing value indicates that they did not choose to invest in that company. Notice that these missing values show up as NaNs when we read in the data. Fill in these missing values with zeros. Other columns may also contain NaNs; be careful not to fill those columns with zeros, or you may end up with strange results down the line.

```

[20]: # YOUR CODE HERE
df_shark[["Corcoran", "Cuban", "Greiner", "Herjavec", "John", "O'Leary",
↪ "Harrington", "Guest"]] = df_shark[["Corcoran", "Cuban", "Greiner",
↪ "Herjavec", "John", "O'Leary", "Harrington", "Guest"]].fillna(0)
df_shark

```

```

[20]:      Season  No. in series      Company Deal
0         1.0         1.0      Ava the Elephant  Yes \
1         1.0         1.0    Mr. Tod's Pie Factory  Yes
2         1.0         1.0      Wispots           No
3         1.0         1.0 College Foxes Packing Boxes  No
4         1.0         1.0      Ionic Ear         No
..      ...           ...
490      6.0         28.0      You Kick Ass       Yes
491      6.0         29.0      Shark Wheel       Yes
492      6.0         29.0      Gato Cafe         No
493      6.0         29.0      Sway Motorsports  Yes
494      6.0         29.0      Spikeball        Yes

      Industry Entrepreneur Gender  Amount Equity  Corcoran
0      Healthcare      Female  $50,000  55%      1.0 \
1  Food and Beverage      Male $460,000  50%      1.0
2  Business Services      Male   NaN     NaN     0.0
3  Lifestyle / Home      Male   NaN     NaN     0.0
4  Uncertain / Other      Male   NaN     NaN     0.0
..      ...
490 Children / Education  Female $100,000  10%     0.0
491  Fitness / Sports      Male $225,000   8%     0.0
492  Uncertain / Other  Female   NaN     NaN     0.0
493  Green/CleanTech      Male $300,000  20%     0.0
494  Fitness / Sports      Male $500,000  20%     0.0

      Cuban  Greiner  Herjavec  John  O'Leary  Harrington  Guest
0      0.0     0.0     0.0   0.0     0.0     0.0     0.0 \

```

| | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|
| 1 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| 2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 3 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 4 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| .. | ... | ... | ... | ... | ... | ... | ... |
| 490 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 491 | 1.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | 1.0 |
| 492 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 493 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 494 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 |

| Details / Notes | |
|-----------------|--|
| 0 | NaN |
| 1 | NaN |
| 2 | NaN |
| 3 | NaN |
| 4 | NaN |
| .. | ... |
| 490 | NaN |
| 491 | 10% royalty until \$500K; then converts to 5% e... |
| 492 | NaN |
| 493 | NaN |
| 494 | NaN |

[495 rows x 17 columns]

Notice that Amount and Equity are currently being treated as categorical variables (dtype: object). Can you figure out why this is? Clean up these columns and cast them to numeric types (i.e., a dtype of int or float) because we'll need to perform mathematical operations on these columns.

```
[21]: # YOUR CODE HERE
df_shark["Amount"] = df_shark["Amount"].str.replace("$", "")
df_shark["Amount"] = df_shark["Amount"].str.replace(",", "")
df_shark["Amount"] = df_shark["Amount"].astype(float)
df_shark["Equity"] = df_shark["Equity"].str.replace("%", "")
df_shark["Equity"] = df_shark["Equity"].astype(float)
df_shark
```

```
[21]:      Season  No. in series      Company Deal
0         1.0           1.0      Ava the Elephant  Yes \
1         1.0           1.0    Mr. Tod's Pie Factory  Yes
2         1.0           1.0           Wispots      No
3         1.0           1.0 College Foxes Packing Boxes  No
4         1.0           1.0         Ionic Ear      No
..         ...           ...
490        6.0          28.0      You Kick Ass  Yes
```

| | | | | |
|-----|-----|------|------------------|-----|
| 491 | 6.0 | 29.0 | Shark Wheel | Yes |
| 492 | 6.0 | 29.0 | Gato Cafe | No |
| 493 | 6.0 | 29.0 | Sway Motorsports | Yes |
| 494 | 6.0 | 29.0 | Spikeball | Yes |

| | Industry | Entrepreneur | Gender | Amount | Equity | Corcoran |
|-----|----------------------|--------------|--------|----------|--------|----------|
| 0 | Healthcare | | Female | 50000.0 | 55.0 | 1.0 \ |
| 1 | Food and Beverage | | Male | 460000.0 | 50.0 | 1.0 |
| 2 | Business Services | | Male | NaN | NaN | 0.0 |
| 3 | Lifestyle / Home | | Male | NaN | NaN | 0.0 |
| 4 | Uncertain / Other | | Male | NaN | NaN | 0.0 |
| .. | ... | | ... | ... | ... | |
| 490 | Children / Education | | Female | 100000.0 | 10.0 | 0.0 |
| 491 | Fitness / Sports | | Male | 225000.0 | 8.0 | 0.0 |
| 492 | Uncertain / Other | | Female | NaN | NaN | 0.0 |
| 493 | Green/CleanTech | | Male | 300000.0 | 20.0 | 0.0 |
| 494 | Fitness / Sports | | Male | 500000.0 | 20.0 | 0.0 |

| | Cuban | Greiner | Herjavec | John | O'Leary | Harrington | Guest |
|-----|-------|---------|----------|------|---------|------------|-------|
| 0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 \ |
| 1 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| 2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 3 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 4 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| .. | ... | ... | ... | ... | ... | ... | |
| 490 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 491 | 1.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | 1.0 |
| 492 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 493 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 494 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 |

| | Details / Notes |
|-----|--|
| 0 | NaN |
| 1 | NaN |
| 2 | NaN |
| 3 | NaN |
| 4 | NaN |
| .. | ... |
| 490 | NaN |
| 491 | 10% royalty until \$500K; then converts to 5% e... |
| 492 | NaN |
| 493 | NaN |
| 494 | NaN |

[495 rows x 17 columns]

1.2 Question 1. Which Company was Worth the Most?

The valuation of a company is how much it is worth. If someone invests \$10,000 for a 40% equity stake in the company, then this means the company must be valued at \$25,000, since 40% of \$25,000 is \$10,000.

Calculate the valuation of each company that was funded. Which company was most valuable? Is it the same as the company that received the largest total investment from the sharks?

```
[22]: # YOUR CODE HERE
df_shark_q1 = df_shark[df_shark["Equity"] != 0]
df_shark_q1["Valuation"] = df_shark_q1["Amount"] / (df_shark_q1["Equity"] / 100)
df_shark_q1
```

```
/var/folders/q8/mqm68gfx7pjfpqftf7y_v6140000gn/T/ipykernel_99181/2936311911.py:3
```

```
: SettingWithCopyWarning:
```

```
A value is trying to be set on a copy of a slice from a DataFrame.
```

```
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

```
df_shark_q1["Valuation"] = df_shark_q1["Amount"] / (df_shark_q1["Equity"] / 100)
```

```
[22]:
```

| | Season | No. in series | Company | Deal |
|-----|--------|---------------|-----------------------------|-------|
| 0 | 1.0 | 1.0 | Ava the Elephant | Yes \ |
| 1 | 1.0 | 1.0 | Mr. Tod's Pie Factory | Yes |
| 2 | 1.0 | 1.0 | Wisspots | No |
| 3 | 1.0 | 1.0 | College Foxes Packing Boxes | No |
| 4 | 1.0 | 1.0 | Ionic Ear | No |
| .. | ... | ... | ... | ... |
| 490 | 6.0 | 28.0 | You Kick Ass | Yes |
| 491 | 6.0 | 29.0 | Shark Wheel | Yes |
| 492 | 6.0 | 29.0 | Gato Cafe | No |
| 493 | 6.0 | 29.0 | Sway Motorsports | Yes |
| 494 | 6.0 | 29.0 | Spikeball | Yes |

| | Industry | Entrepreneur | Gender | Amount | Equity | Corcoran |
|-----|----------------------|--------------|--------|----------|--------|----------|
| 0 | Healthcare | | Female | 50000.0 | 55.0 | 1.0 \ |
| 1 | Food and Beverage | | Male | 460000.0 | 50.0 | 1.0 |
| 2 | Business Services | | Male | NaN | NaN | 0.0 |
| 3 | Lifestyle / Home | | Male | NaN | NaN | 0.0 |
| 4 | Uncertain / Other | | Male | NaN | NaN | 0.0 |
| .. | ... | | ... | ... | ... | ... |
| 490 | Children / Education | | Female | 100000.0 | 10.0 | 0.0 |
| 491 | Fitness / Sports | | Male | 225000.0 | 8.0 | 0.0 |
| 492 | Uncertain / Other | | Female | NaN | NaN | 0.0 |
| 493 | Green/CleanTech | | Male | 300000.0 | 20.0 | 0.0 |
| 494 | Fitness / Sports | | Male | 500000.0 | 20.0 | 0.0 |

| | Cuban | Greiner | Herjavec | John | O'Leary | Harrington | Guest |
|-----|-------|---------|----------|------|---------|------------|-------|
| 0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 \ |
| 1 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| 2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 3 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 4 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| .. | ... | ... | ... | ... | ... | ... | |
| 490 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 491 | 1.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | 1.0 |
| 492 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 493 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 494 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 |

| | Details / Notes | Valuation |
|-----|--|--------------|
| 0 | NaN | 9.090909e+04 |
| 1 | NaN | 9.200000e+05 |
| 2 | NaN | NaN |
| 3 | NaN | NaN |
| 4 | NaN | NaN |
| .. | ... | ... |
| 490 | NaN | 1.000000e+06 |
| 491 | 10% royalty until \$500K; then converts to 5% e... | 2.812500e+06 |
| 492 | NaN | NaN |
| 493 | NaN | 1.500000e+06 |
| 494 | NaN | 2.500000e+06 |

[494 rows x 18 columns]

```
[23]: max_valuation = df_shark_q1["Valuation"].max()
df_shark_q1.loc[df_shark_q1["Valuation"] == max_valuation]
```

```
[23]: Season  No. in series  Company Deal      Industry
421      6.0             11.0    Zipz  Yes  Food and Beverage \

      Entrepreneur Gender    Amount  Equity  Corcoran  Cuban  Greiner
421              Male  2500000.0    10.0      0.0    0.0    0.0 \

      Herjavec  John  O'Leary  Harrington  Guest
421      0.0    0.0      1.0          0.0    0.0 \

      Details / Notes  Valuation
421  with an option for another $2.5 Million for an...  25000000.0
```

```
[24]: df_shark_q1[df_shark_q1["Amount"] == df_shark_q1["Amount"].max()]
```

```
[24]:      Season  No. in series Company Deal      Industry Entrepreneur Gender
483      6.0              27.0 AirCar  Yes  Green/CleanTech              Male  \

      Amount  Equity  Corcoran  Cuban  Greiner  Herjavec  John  O'Leary
483 5000000.0   50.0      0.0   0.0      0.0      1.0   0.0      0.0  \

      Harrington  Guest              Details / Notes
483      0.0      0.0  Contingent on getting deal to bring to contine...  \

      Valuation
483 10000000.0
```

YOUR EXPLANATION HERE

There is one row/company that has an equity of 0% (The Wall DoctorRX). This would mean that their valuation is infinite which seems like a special case. For that reason, we will ignore that observation and consider the other ones. The company that was the most valuable was Zipz and it is not the same as the one that received the largest total in investment.

1.3 Question 2. Which Shark Invested the Most?

Calculate the total amount of money that each shark invested over the 6 seasons. Which shark invested the most total money over the 6 seasons?

Hint: If n sharks funded a given venture, then the amount that each shark invested is the total amount divided by n .

```
[25]: # ENTER CODE HERE.
df_shark["Num_Investors"] = df_shark["Corcoran"] + df_shark["Cuban"] +
    ↪df_shark["Greiner"] + df_shark["Herjavec"] + df_shark["John"] +
    ↪df_shark["O'Leary"] + df_shark["Harrington"] + df_shark["Guest"]
df_shark
```

```
[25]:      Season  No. in series      Company Deal
0          1.0          1.0      Ava the Elephant  Yes  \
1          1.0          1.0  Mr. Tod's Pie Factory  Yes
2          1.0          1.0      Wispots          No
3          1.0          1.0  College Foxes Packing Boxes  No
4          1.0          1.0      Ionic Ear          No
..      ...      ...
490      6.0          28.0      You Kick Ass  Yes
491      6.0          29.0      Shark Wheel  Yes
492      6.0          29.0      Gato Cafe  No
493      6.0          29.0      Sway Motorsports  Yes
494      6.0          29.0      Spikeball  Yes

      Industry Entrepreneur Gender  Amount  Equity  Corcoran
0      Healthcare      Female  50000.0   55.0      1.0  \
1  Food and Beverage      Male  460000.0   50.0      1.0
```


| | | | | | |
|-----|----------------------|--------|----------|------|-----|
| 2 | Business Services | Male | NaN | NaN | 0.0 |
| 3 | Lifestyle / Home | Male | NaN | NaN | 0.0 |
| 4 | Uncertain / Other | Male | NaN | NaN | 0.0 |
| .. | ... | ... | ... | ... | ... |
| 490 | Children / Education | Female | 100000.0 | 10.0 | 0.0 |
| 491 | Fitness / Sports | Male | 225000.0 | 8.0 | 0.0 |
| 492 | Uncertain / Other | Female | NaN | NaN | 0.0 |
| 493 | Green/CleanTech | Male | 300000.0 | 20.0 | 0.0 |
| 494 | Fitness / Sports | Male | 500000.0 | 20.0 | 0.0 |

| | Cuban | Greiner | Herjavec | John | O'Leary | Harrington | Guest | |
|-----|-------|---------|----------|------|---------|------------|-------|---|
| 0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | \ |
| 1 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | |
| 2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | |
| 3 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | |
| 4 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | |
| .. | ... | ... | ... | ... | ... | ... | ... | |
| 490 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | |
| 491 | 1.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | 1.0 | |
| 492 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | |
| 493 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | |
| 494 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | |

| | Details / Notes | Num_Investors |
|-----|--|---------------|
| 0 | | NaN |
| 1 | | NaN |
| 2 | | NaN |
| 3 | | NaN |
| 4 | | NaN |
| .. | ... | ... |
| 490 | | NaN |
| 491 | 10% royalty until \$500K; then converts to 5% e... | 3.0 |
| 492 | | NaN |
| 493 | | NaN |
| 494 | | NaN |

[495 rows x 18 columns]

```
[26]: df_shark["Amount_Per_Investor"] = df_shark["Amount"] / df_shark["Num_Investors"]
```

```
[27]: names = ["Corcoran", "Cuban", "Greiner", "Herjavec", "John", "O'Leary",
             ↪ "Harrington", "Guest"]
total_amounts = []
for x in names:
    total_amounts.append((df_shark[df_shark[x] == 1.0]["Amount_Per_Investor"].
    ↪sum(), x))
```

```
[28]: total_amounts
```

```
[28]: [(4912500.0, 'Corcoran'),
      (17817500.0, 'Cuban'),
      (8170000.0, 'Greiner'),
      (16297500.0, 'Herjavec'),
      (8154000.0, 'John'),
      (7952500.0, "O'Leary"),
      (800000.0, 'Harrington'),
      (400000.0, 'Guest')]
```

```
[29]: max(total_amounts)
```

```
[29]: (17817500.0, 'Cuban')
```

YOUR EXPLANATION HERE

Over the 6 seasons, Mark Cuban invested the most total money.

1.4 Question 3. Do the Sharks Prefer Certain Industries?

Calculate the funding rate (the proportion of companies that were funded) for each industry. Make a visualization showing this information.

```
[30]: # ENTER CODE HERE.
def bin_deal(a):
    if a == "Yes":
        return 1
    else:
        return 0
df_shark['Bin_Deal'] = df_shark['Deal'].map(bin_deal)
df_shark
```

```
[30]:
```

| | Season | No. in series | Company | Deal |
|-----|--------|---------------|-----------------------------|-------|
| 0 | 1.0 | 1.0 | Ava the Elephant | Yes \ |
| 1 | 1.0 | 1.0 | Mr. Tod's Pie Factory | Yes |
| 2 | 1.0 | 1.0 | Wispots | No |
| 3 | 1.0 | 1.0 | College Foxes Packing Boxes | No |
| 4 | 1.0 | 1.0 | Ionic Ear | No |
| .. | ... | ... | ... | ... |
| 490 | 6.0 | 28.0 | You Kick Ass | Yes |
| 491 | 6.0 | 29.0 | Shark Wheel | Yes |
| 492 | 6.0 | 29.0 | Gato Cafe | No |
| 493 | 6.0 | 29.0 | Sway Motorsports | Yes |
| 494 | 6.0 | 29.0 | Spikeball | Yes |

| | Industry | Entrepreneur | Gender | Amount | Equity | Corcoran |
|---|------------|--------------|--------|---------|--------|----------|
| 0 | Healthcare | | Female | 50000.0 | 55.0 | 1.0 \ |

| | | | | | |
|-----|----------------------|--------|----------|------|-----|
| 1 | Food and Beverage | Male | 460000.0 | 50.0 | 1.0 |
| 2 | Business Services | Male | NaN | NaN | 0.0 |
| 3 | Lifestyle / Home | Male | NaN | NaN | 0.0 |
| 4 | Uncertain / Other | Male | NaN | NaN | 0.0 |
| .. | ... | ... | ... | ... | ... |
| 490 | Children / Education | Female | 100000.0 | 10.0 | 0.0 |
| 491 | Fitness / Sports | Male | 225000.0 | 8.0 | 0.0 |
| 492 | Uncertain / Other | Female | NaN | NaN | 0.0 |
| 493 | Green/CleanTech | Male | 300000.0 | 20.0 | 0.0 |
| 494 | Fitness / Sports | Male | 500000.0 | 20.0 | 0.0 |

| | Cuban | Greiner | Herjavec | John | O'Leary | Harrington | Guest | |
|-----|-------|---------|----------|------|---------|------------|-------|---|
| 0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | \ |
| 1 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | |
| 2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | |
| 3 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | |
| 4 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | |
| .. | ... | ... | ... | ... | ... | ... | ... | |
| 490 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | |
| 491 | 1.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | 1.0 | |
| 492 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | |
| 493 | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | |
| 494 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | |

| | Details / Notes | Num_Investors | |
|-----|--|---------------|---|
| 0 | NaN | 1.0 | \ |
| 1 | NaN | 2.0 | |
| 2 | NaN | 0.0 | |
| 3 | NaN | 0.0 | |
| 4 | NaN | 0.0 | |
| .. | ... | ... | |
| 490 | NaN | 1.0 | |
| 491 | 10% royalty until \$500K; then converts to 5% e... | 3.0 | |
| 492 | NaN | 0.0 | |
| 493 | NaN | 1.0 | |
| 494 | NaN | 1.0 | |

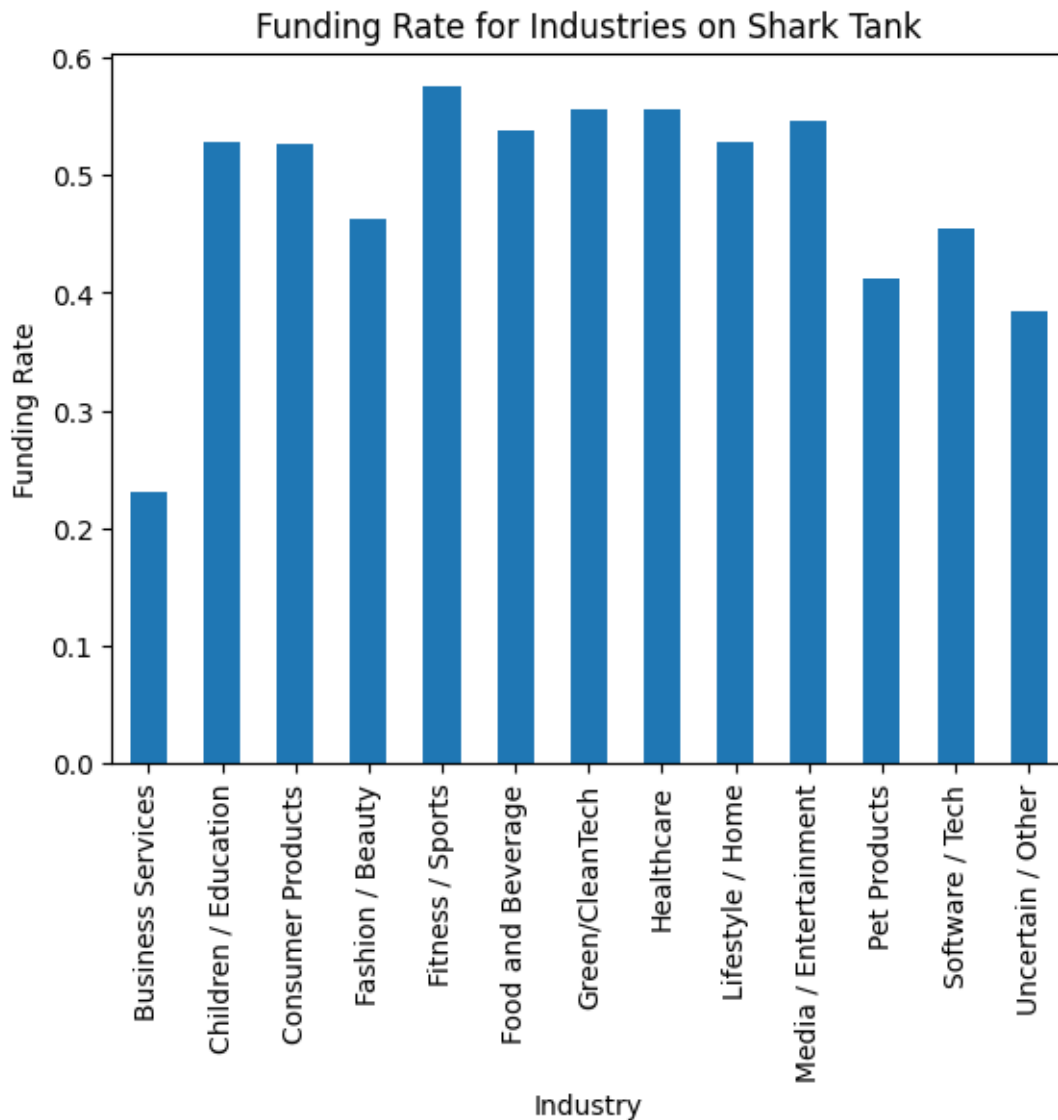
| | Amount_Per_Investor | Bin_Deal |
|-----|---------------------|----------|
| 0 | 50000.0 | 1 |
| 1 | 230000.0 | 1 |
| 2 | NaN | 0 |
| 3 | NaN | 0 |
| 4 | NaN | 0 |
| .. | ... | ... |
| 490 | 100000.0 | 1 |
| 491 | 75000.0 | 1 |
| 492 | NaN | 0 |

```
493          300000.0      1
494          500000.0      1
```

```
[495 rows x 20 columns]
```

```
[31]: df_shark.groupby("Industry")["Bin_Deal"].mean().plot.bar(ylabel="Funding Rate",
↳title="Funding Rate for Industries on Shark Tank")
```

```
[31]: <Axes: title={'center': 'Funding Rate for Industries on Shark Tank'},
xlabel='Industry', ylabel='Funding Rate'>
```



YOUR EXPLANATION HERE

From the bar plot, it can be seen that Business Services were in general least likely to be funded in Shark Tank. On the other hand, Fitness / Sports is the most likely to be funded in Shark Tank.

1.5 Submission Instructions

- After you have completed the notebook, select **Runtime > Run all**
- After the notebook finishes rerunning check to make sure that you have no errors and everything runs properly. Fix any problems and redo this step until it works.
- Rename this notebook by clicking on “DATA 301 Lab 2A - YOUR NAMES HERE” at the very top of this page. Replace “YOUR NAMES HERE” with the first and last names of you and your partner (if you worked with one).
- Expand all cells with View > Expand Sections
- Save a PDF version: File > Print > Save as PDF
 - Under “More Settings” make sure “Background graphics” is checked
 - Printing Colab to PDF doesn’t always work so well and some of your output might get cutoff. That’s ok.
 - It’s not necessary, but if you want a more nicely formatted PDF you can uncomment and run the code in the following cell. (Here’s a [video](#) with other options.)
- Download the notebook: File > Download .ipynb
- Submit the notebook and PDF in Canvas. If you worked in a pair, only one person should submit in Canvas.

```
[32]: # !wget -nc https://raw.githubusercontent.com/brpy/colab-pdf/master/colab_pdf.py
      # from colab_pdf import colab_pdf
      # colab_pdf('DATA 301 Lab1A - YOUR NAMES HERE.ipynb')
```