
Question 5. In the code cell below, create a visualization that will help us determine if there is an association between birth rate and death rate during this time interval. It may be helpful to create an intermediate table containing the birth and death rates for each state. **(4 Points)**

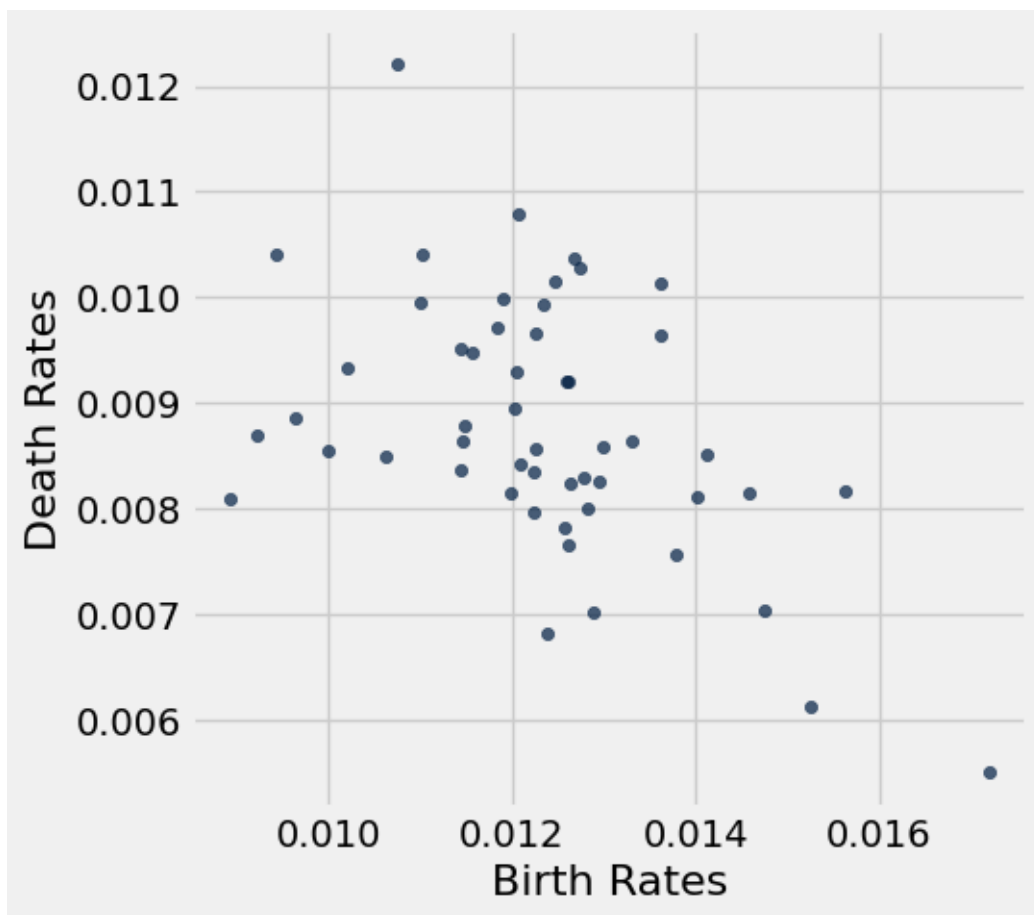
Things to consider:

- What type of chart will help us illustrate an association between 2 variables?
- How can you manipulate a certain table to help generate your chart?
- Check out the [Recommended Reading](#) for this homework!

```
In [28]: # In this cell, use birth_rates and death_rates to generate your visualization
        birth_rates_2015 = pop.column('BIRTHS') / pop.column('2015')
        death_rates_2015 = pop.column('DEATHS') / pop.column('2015')

        vis_pop_table = Table().with_columns(
            "Birth Rates", birth_rates_2015,
            "Death Rates", death_rates_2015,
        )

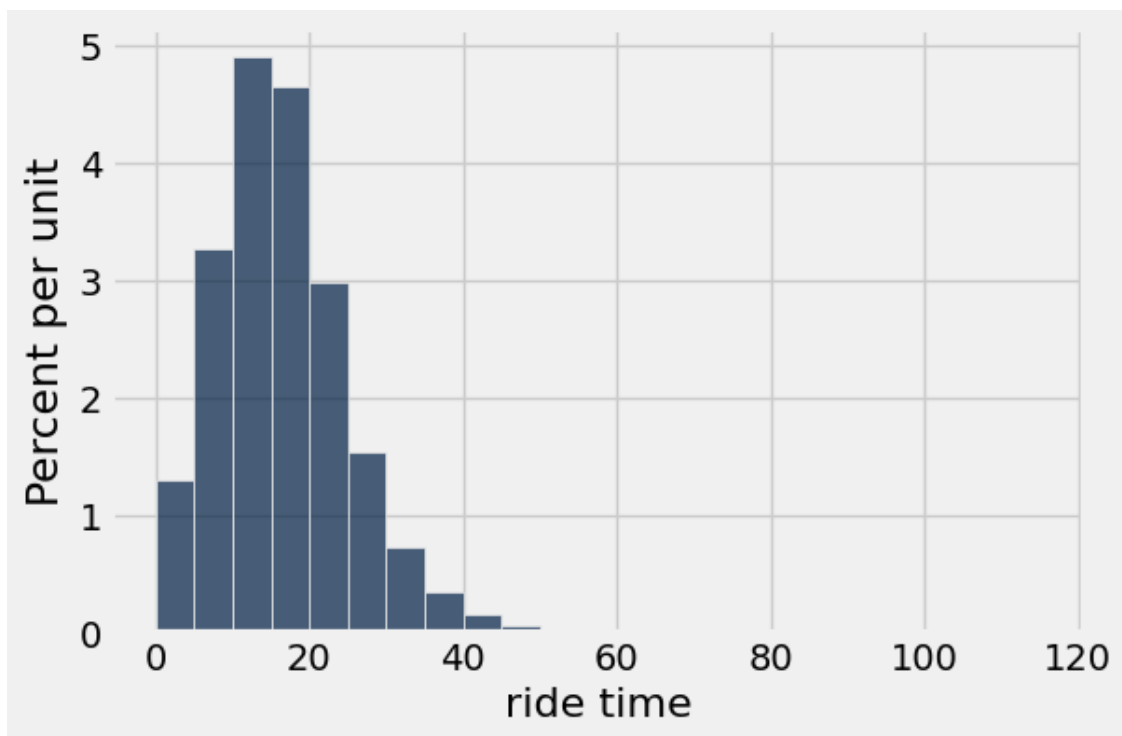
        vis_pop_table.scatter("Birth Rates", "Death Rates")
```



Question 1. Produce a histogram that visualizes the distributions of all ride times in Boston using the given bins in `equal_bins`. (4 Points)

Hint: See [Chapter 7.2](#) if you're stuck on how to specify bins.

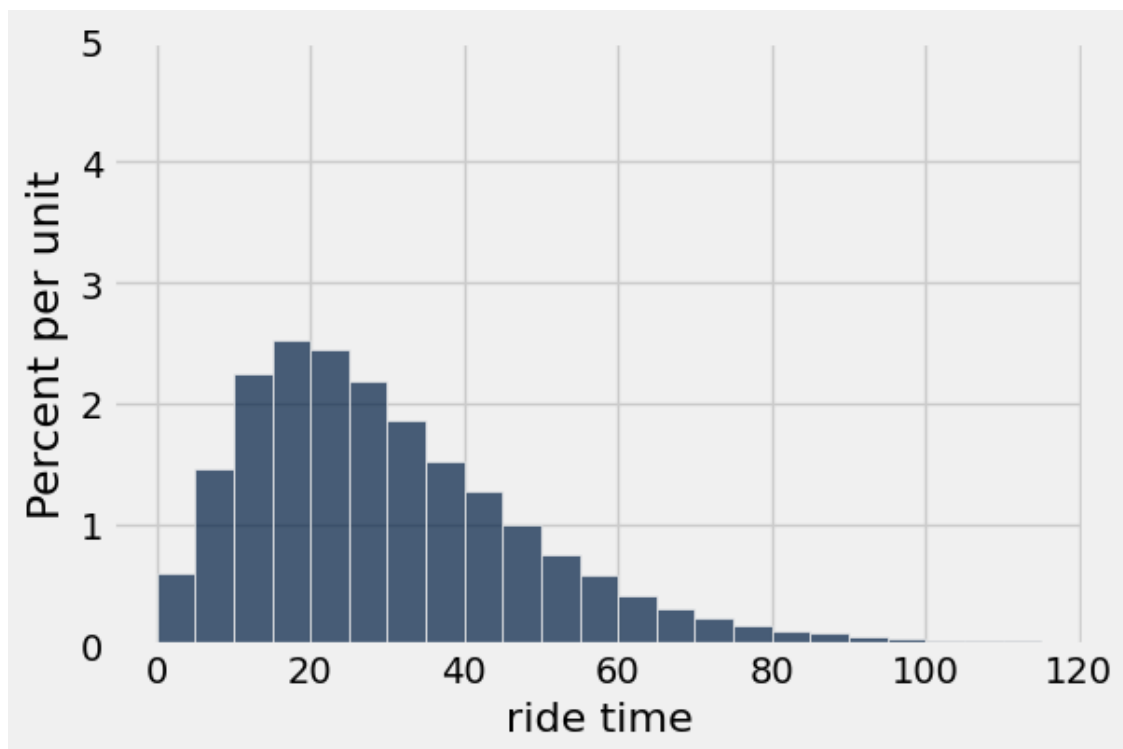
```
In [32]: equal_bins = np.arange(0, 120, 5)
        boston.hist('ride time', bins=equal_bins)
```



Question 2. Now, produce a histogram that visualizes the distribution of all ride times in Manila using the given bins. (4 Points)

```
In [33]: equal_bins = np.arange(0, 120, 5)
         manila.hist('ride time', bins=equal_bins)

         # Don't delete the following line!
         plt.ylim(0, 0.05);
```



Question 6. Identify one difference between the histograms, in terms of the statistical properties. > *Hint:* Without performing any calculations, can you comment on the average or skew of each histogram? (4 Points)

Boston seems to be skewed towards left-hand side, meaning less average ride time, while Manila shows longer ride time

Question 7. Why is your solution in Question 6 the case? Based on one of the following two readings, why are the distributions for Boston and Manila different? **(4 Points)**

- [Boston reading](#)
- [Manila reading](#)

Hint: Try thinking about external factors of the two cities that may be causing the difference! The readings provide some potential factors – try to connect them to the ride time data.

Type your answer here, replacing this text.

Question 2. State at least one reason why you chose the histogram from Question 1. **Make sure to clearly indicate which histogram you selected** (ex: “I chose histogram A because ...”). **(5 Points)**

Type your answer here, replacing this text.

Question 4. State at least one reason why you chose the histogram from Question 3. **Make sure to clearly indicate which histogram you selected** (ex: “I chose histogram A because ...”). **(5 Points)**

I chose histogram B because according to the scatter plot, y doesn't have any data points around the value 0

