

Tables are a fundamental way of representing data sets. A table can be viewed in two ways:

- A sequence of named columns that each describe a single attribute of all entries in a data set, or
- A sequence of rows where each row contains all the attribute information about that entry in the data set

Data 8 uses a library consisting of many Table functions that will allow you to manipulate and visualize data. All of the functions we will use in this course are listed on the Data 8 course webpage under Python Reference, and a similar outline will be provided during exams.

1. Rise and Shine

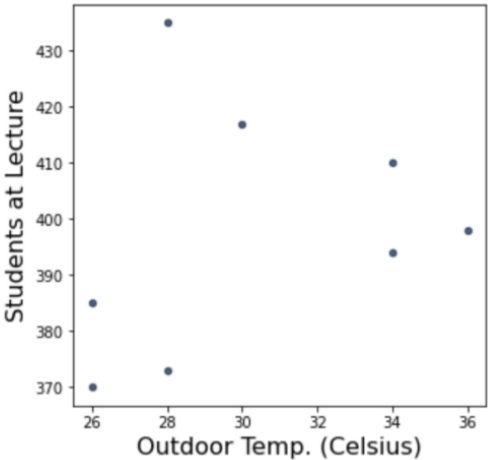
Let’s look at an example table called weather.

Date	Outdoor Temp. (Celsius)	Students at Lecture
June 21	28	435
June 22	30	417
June 23	34	394
June 24	36	398
June 27	34	410
June 28	26	385
June 29	26	370
June 30	28	373

The table has 8 rows, each corresponding to a day of the year. Each row has three attributes: the date, the outdoor temperature in Celsius, and the number of students at lecture that day.

a. Using just the information provided in the weather table, can you generate the following tables and visualizations? If not, what additional information do you need? *Note: You do not need to know/write the specifics of the code that you would use to generate these outputs.*

Outdoor Temp. (Celsius)	Mean # of Students at Lecture
26	377.5
28	404
30	417
34	402
36	398



YES/NO

YES/NO

b. Matthew prefers using units of Fahrenheit instead of Celsius. Fill in the blank line of code to calculate an array of new values that we can add to the weather table.

Recall the following formula:  $F = \frac{9}{5}C + 32$

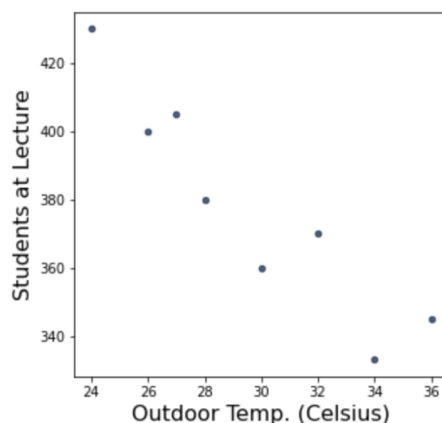
```
temp_in_celsius = weather.column('Outdoor Temp. (Celsius)')
```

```
temp_in_fahrenheit = _____
```

```
new_table = weather.with_columns('Outdoor Temp. (Fahrenheit)', temp_in_fahrenheit)
```

Date	Outdoor Temp. (Celsius)	Students at Lecture	Outdoor Temp. (Fahrenheit)
June 21	28	435	82.4
June 22	30	417	86
June 23	34	394	93.2
June 24	36	398	96.8
June 27	34	410	93.2
June 28	26	385	78.8
June 29	26	370	78.8
June 30	28	373	82.4

c. Matthew wants to compare the data in the weather table with a previous summer's data. He found the following scatterplot:



Is the data collected previously the result of an observational study or a randomized controlled experiment? Why?

d. Using our answer to part c and the visualization, is there a relationship between outdoor temperature and the number of students at lecture – an association, a causal relationship or something else? Why?

## 2. Code made with ♥ and Coffee

Ciara collected the following information about her coworkers' methods of getting to work and their coffee consumption.

Method	Number of Coworkers	Average Cups of Coffee per Day
Take the Bus to Work	12	1.1
Drive to Work	15	1.9

a. Ciara is trying to compute the absolute value of the difference between the total number of cups drunk by driving coworkers per year and the total number of cups drunk by bussing coworkers per year. She will do all of this in a single cell. Identify the errors in the following cell and correct them. *Make sure that the cell outputs (shows the result) a single, positive number. You may need to add another line to achieve this.*

```
number_cups_bus = 12(1.1)
number_cups_drive = 15(1.9)
number_cups_day_difference = ((number_cups_bus - number_cups_drive)
number_cups_week_difference = number_cups_difference * 7
yearly_cups = number_cups_week_difference * 52
```

b. Ciara wants to determine whether she is spending more money on coffee compared to boba over the course of 1 semester (15 weeks long). Assume she purchases 4 coffees and 3 bobas per week. Complete the following lines of code, which should assign `result` to `True` if Ciara spends *strictly more* on coffee and `False` otherwise.

```
cost_coffee = 4.50
```

```
cost_boba = 6.00
```

```
total_coffee = _____
```

```
total_boba = _____
```

```
result = _____
```

### 3. Sp17 Practice Midterm Q3b

A study followed 369 people with cardiovascular disease, randomly selected from hospital patients. A year later, those who owned a dog were four times more likely to be alive than those who didn't. For all following questions, please provide a brief explanation.

a. True or False. This study is a randomized controlled experiment.

b. True or False. This study shows that dog owners live longer than cat owners on average.

c. True or False. This study shows that for someone with cardiovascular disease, adopting a dog causes them to live longer.