Here is a contingency table of college students with their **Favorite Color** (Red or Blue) down the columns and their **School** (Berkeley or Stanford) across the rows.

| | Red | Blue |
|---|---|---|
| Berkeley | 10 | 90 |
| Stanford | 60 | 40 |

1. Find the proportion of all students who attend Berkeley. What type of proportion is this?

2. Find the proportion of all students who attend Berkeley *and* like red best. What type of proportion is this?

3. *Of the students who attend Berkeley*, find the proportion that like red best. What type of proportion is this?

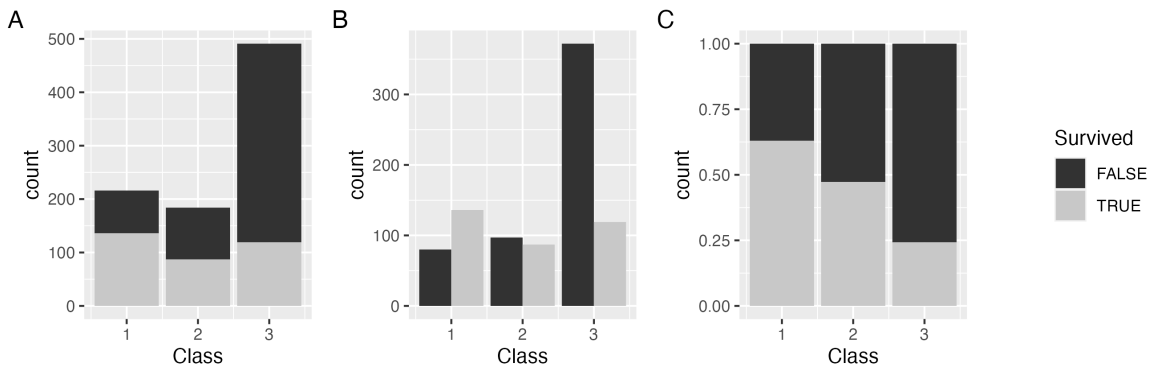
RMS Titanic was an infamous British passenger. It sank in the North Atlantic Ocean on 15 April 1912 after striking an iceberg on her way to New York City. The wreck the deadliest sinking of a single ship at the time, with almost 70% of the passengers and crew dying.

Below are the first five rows of for a data frame called `titanic`, which includes information on some of the passengers.

| Name | Sex | Age | Fare | Class | Survived |
|---|---|---|---|---|---|
| Braund, Mr. Owen Harris | male | 22 | 7.25 | 3 | FALSE |
| Cumings, Mrs. John Bradley (Florence Briggs Thayer) | female | 38 | 71.28 | 1 | TRUE |
| Heikkinen, Miss. Laina | female | 26 | 7.92 | 3 | TRUE |
| Futrelle, Mrs. Jacques Heath (Lily May Peel) | female | 35 | 53.10 | 1 | TRUE |
| Allen, Mr. William Henry | male | 35 | 8.05 | 3 | FALSE |

4. What is the unit of observation for `titanic`?

The three following graphs were generated using `titanic`.




5. Which graph is **best** used to answer the following questions? *Write the letter associated with the graph to the right of each question*

- *How many first class passengers survived the Titanic's sinking?* _____
- *Is there an association between a passenger's class and their survival?* _____
- *What is the mode of the* `Class` *variable?* _____

6. Roughly how many passengers are in the `titanic` data frame, based on the graphs?

The `promote` data frame can be found in the `stat20data` R package.

7. Using `promote`, write `ggplot2` code to make a stacked, normalized bar chart having identified gender on the x axis and the bars filled in by promotion decision. Write the code you used below. *Make sure you load in any libraries where necessary!*

8. Write a one-sentence claim about the association between identified gender and promotion decision.

9. Consider the vector `q9` which was made as follows:

```
nums <- c(1,2,3,4,5,6,NA)
```

Copy this code as-is into your R session. Then, without changing `nums`, write R code to calculate the mean of `nums` so that the result is 3.5 (this is the mean of the numbers 1 through 6). Copy your code in the space below. Again, you don't need to turn in your code file, just write the working code below.