

# STAT 408: Midterm

## Due: October 24 at 5:00 PM

### Name:

Please turn in the exam to D2L and include both the R Markdown code *and either* a Word or PDF file. Please verify that all of the code has compiled and the graphics look like you think they should on your Word or PDF file. If you are having issues with the Word file contorting your images, you may submit the image files directly to D2L as well.

While the exam is open book, meaning you are free to use any resources from class, this is strictly an individual endeavor. **Discussing the problems with anyone outside the course instructor is a violation of the MSU student conduct code.** The instructor will answer questions related to expectations or understanding of the exam, but will not fix or troubleshoot broken code.

## 1. (18 points)

The Rubik's cube was a popular game in the 1980's which players attempt to match colors on the six sides of a cube. To win the game, a player must have a single color on each side of the cube. For additional details see the wikipedia page: [https://en.wikipedia.org/wiki/Rubik's\\_Cube](https://en.wikipedia.org/wiki/Rubik's_Cube).

Write a function that:

- Takes a 3-by-3-by-6 array of character values of colors as input (note the 3-by-3 submatrices represent a side of the cube and there should be a total of 9 elements for each of the following colors: red, green, yellow, blue, white, and orange),
- Assuming a proper cube has been entered, the function returns either 'This cube is a completed Rubik's cube', or 'Try Again: cube is not a completed Rubik's cube',
- include all necessary documentation and notation for your function and also include errors for incorrect inputs.

Verify your function by testing it on the following arrays. Each call should either return 'This cube is a completed Rubik's cube', 'Try Again: cube is not a completed Rubik's cube', or an error. For full credit errors must be returned for all improper inputs. Even if you cannot complete the entire problem, include the code you have for consideration of partial credit.

```
array1 <- array('red', dim=c(3,3,6))
array2 <- array(1:6, dim = c(3,3,6))
array3 <- array(c('red', 'blue', 'green', 'yellow', 'white', 'orange'), dim=c(3,3,6))
array4 <- array(c('red', 'blue', 'green', 'yellow', 'white', 'orange'), dim=c(3,3,3))
array5 <- array(rep(c('red', 'blue', 'green', 'yellow', 'white', 'orange'), each=9), dim=c(3,3,6))
array6 <- array(rep(c('duck', 'goose', 'eagle', 'crane', 'white', 'orange'), each=9), dim=c(3,3,6))
```

## 2. (20 points)

### a. (4 points)

Find and upload one image that presents data in a misleading way. You may need to revisit the R Markdown cheat sheet to upload the image.

Describe the problem with this particular graph.

**b. (4 points)**

Find and upload an image that you find compelling. Describe what you like about the image.

**c. (12 points)**

Select a data set and create a series of three or more graphs illustrating interesting observations from the data. At least two different types of graphs need to be constructed with a minimum of one graph created using ggplot2. The data set can be one we have used in class or a data set you obtain from elsewhere. However, you should not use the same graphics as a homework or lab.

These figures should adhere to the principles we have stated in class and make sure that your graphics can “stand alone”. You can add captions via R Markdown if you’d like more than just the title to describe the figures. For full credit your data storytelling results should be *compelling*.

**3. (22 points)**

For this question use the Baltimore Towing data set:

<http://www.math.montana.edu/ahoegh/teaching/stat408/datasets/BaltimoreTowing.csv>.

**a. (4 points)**

Describe the data set. What does each row / column represent?

**b. (4 points)**

Compute how many vehicles are towed each month.

**c. (4 points)**

How many vehicles with the vehicleType of SUV were towed (have a receivingDateTime) between 10 PM and 7 AM?

**d. (4 points)**

Compute the average totalPaid for vehicleType of Van by year. Create a table to summarize these results.

**e. (6 points)**

Using this data set create a graphic to highlight an interesting story. Include a 3-4 sentence description of your figure.