

MTH208a: Midsem Exam - Part II

Instructions

1. Follow the instructions for each question **exactly**.
2. If your code is not properly commented or horizontal/vertical spacing is missing, points will be deducted.
3. There are three questions and this Part II is overall 70 marks.

Questions

1. Instructions:

- Copy and paste your final code for this whole question in the **ans1.R** file in your exam repository.
- Make sure you DO NOT have command **rm(list = ls())** in the R Script
- Do not call any external libraries in your R script.

Birthday problem: Suppose there are $n = 25$ people in a room, and every person's birthday can be on any of the 365 days with equal probability (ignoring leap years).

- a. (15 points) Write an R function, call it **birthday()**, to determine whether any two people of the n people in the room share a birthday. This function should have one argument **n**, the number of people in the room.

Note: if you use the function **sample()**, be sure to choose option **replace = TRUE**.

Hint: you may also find the function **unique()** helpful.

- b. (10 points) Run the function **birthday()** 1000 times and calculate the proportion of times a shared birthday was found. Save this proportion in **ans25**.
- c. (5 points) Repeat the above for $n = 50$ and save the final proportion in **ans50**.

2. Instructions:

- Copy and paste your final function in the **ans2.R** file in your exam repository.
- The code you submit should contain your function and also any libraries you require. NOTHING ELSE.
- Make sure you DO NOT have command **rm(list = ls())** in the R Script.

Image problem: (20 points) Write a function, **flip()**, that takes an **imager** image as an argument, and returns an **imager** image that is mirrored across the the vertical axis. For example, if input image

is shown on the left below, and the output image should be the one on the right.



Note: The input argument for the function MUST be an image loaded with `imager` and NOT the location or file name of the image.

3. Instructions:

- Copy all relevant code and paste in the file `ans3.R`. Make sure to include all lines of code that are required to produce the plot exactly.
- Your final plot should be saved in an object called `finalP` and not actually plotted.
- Make sure you DO NOT have command `rm(list = ls())` in the R Script.

Titanic Dataset: (20 points) Load the dataset `titanic.csv` provided in your repository. This dataset contains information regarding the passengers on the Titanic. The columns are:

- `PassengerId` - unique ID for each passenger
- `Survived` - whether the passenger survived or not
- `Pclass` - the class of their ticket: 1st class, 2nd class, 3rd class
- `Sex` - male or female
- `Age` - age
- `SibSp` - number of siblings/spouse traveling with them
- `Parch` - number of parents/children traveling with them
- `Fare` - cost of their ticket
- `Embarked` - the Port from which they boarded the ship. C = Cherbourg, Q = Queenstown, S = Southampton.

Write R code to recreate the plot below, **exactly**. The final `ggplot` plot should be saved in the object `finalP`. That is, your last line of code should look like

```
finalP <- ggplot( ... ) ....
```

Hint: You may need to use `color = Sex` at some point.

Hint: You may find it useful to convert the dataframe to a tibble using `as_tibble()` in library `tibble`.

Fare vs Survival

Irrespective of Sex, richer people survived

