CS 541 HW1
Spring 2023

Instructions:

-This assignment is due on **20th February 2023** at 11:59 pm EST.
-Any sign of collaboration or plagiarism will result in a 0 and being reported to the Graduate Academic Integrity Board. This includes all sorts of online resources and AI tools.

Late policy:
90% if late by 1 date, 70 % if late by 2 days, 50 % if late by 5 days and 0 there after.

-Please submit the code only after checking it. If the code does not compile, assignment will be graded as 0.

Submission Pattern:
-You should submit a single zip file containing everything including code, data, README etc.

-Answers to essay type questions should be included in the same folder as the code files.

-For questions involving mathematical calculations, you are free to use Microsoft word or Overleaf. Final submission file should be pdf only.

-The zip folder should be named as CS541_HW1_FirstName _LastName

1 Probability (30 points)

A. Suppose you repeatedly roll a fair six-sided die until you roll a 1 (and then you stop). Every time you roll a 2, you lose A points, and every time you roll a 6, you win B points. You do not win or lose any points if you roll a 3, 4, or a 5. What is the expected number of rolls (as a function of A and B) you will have when you stop?

B. Suppose that we have a collection of customer reviews for two restaurants, i.e., Chipotle and Five Guys. Chipotle gets 200 reviews, where 120 reviews are positive while the other 80 reviews are negative. Five Guys gets 40 positive reviews and 60 negative reviews with a total of 100 reviews. Suppose we first randomly choose a restaurant (thus each restaurant would have an equal probability of being selected), then select one of its customer reviews at random. Given the result that we finally get a positive review, what is the probability that this review is about Chipotle? Solve the problem using **Bayes rule**.

C. Suppose the probability of a coin turning up heads is $0 < p < 1$, and that we flip it 7 times and get H, H, T, H, T, T, H. We know the probability of obtaining this sequence is L(p) = pp(1-p)p(1-p)(1-p)p = $p^4(1-p)^3$. What is the value of p that maximizes L(p)? What is an intuitive interpretation of this value p?
Hint: Consider taking the derivative of log L(p). You can also directly take the derivative of L(p), but its cleaner and more natural to differentiate log L(p).(No need to prove this in the solution)

2 Find the First Word (20 points)

Implement a program findFirstWord.py where a function takes two arguments, 'words' and 'order' from the user, and returns the first word to occur in the user given string('words') according to the order specified by the user. **Implement this without sorting the words.**
For example:

```
Input:
Enter the words separated by a space: apple bear cat
Enter the order: bacdefghi
Output:
bear

Input:
Enter the words separated by a space: zoo xmas yellow water
Enter the order: xyzw
Output:
xmas
```

```
Input:
Enter the words separated by a space: bcab abbca desc ccdd dcab
Enter the order: edcbaxyz
Output:
desc
```

3 Data Summary (25 points)

Write a program ***summary.py*** that computes minimum, maximum, mean and standard deviation of the number of COVID cases over time of a USA state, input by the user.
Find the csv file here.

4 Matrix Manipulation. (25 points)

Write a program ***grayscale.py*** that converts a colored image into grayscale and then saves that image as grayscale_image in the same folder. A color image is a width $\times$ height $\times$ 3 matrix. Use any python package to read the image as a matrix. Convert each pixel (with R,G,B values) to a single grayscale value using the equation below. **NOTE: Direct usage of any inbuilt packages and/or libraries to convert image to grayscale is not allowed. You need to implement this by yourself.**
Value = 0.2989R + 0.5870G + 0.1140B
An example code to read the image is:

```
from matplotlib.image import imread
image\_rgb = imread(infile)
```

An example code to write the image is:

```
from matplotlib import pyplot as plt
\#pyplot requires pixel values to be between 0 and 1
image\_gray = plt.imsave('outfile.jpeg',outfile/255)
```