

# Assignment #3

April 12, 2022

## 1 Assignment #3: Statistical Methods

In this assignment, you'll demonstrate your understanding of the statistical methods we've covered in class. These include basic statistical methods such as the mean and median, creating a simulation by repeatedly replicating an event, and performing a linear regression to predict future data points based on prior data points.

You may receive partial credit for incomplete answers. Consult our class notebooks if you're feeling stuck.

### 1.1 Simulating Coin Flips

In these questions, you'll work toward creating a simulation. The simulated event will be flipping a coin ten (10) times. You will replicate the event 100000 times, and then plot the distribution of outcomes. Each question below will build toward running the simulation and plotting the results.

```
[2]: # Import the libraries you'll need for the simulation
import pandas
import random
```

Write a line of code that randomly selects a number between 1 and 52

```
[ ]:
```

Write a line of code that randomly generates a number that is either 1 or 2

```
[ ]:
```

Write a loop that prints out the word "hello" ten times. (You will likely want to use the "range" function.)

```
[ ]:
```

Write two lines of code. In the first line, create an empty list and assign it to a variable. In the second line, add something to that list.

```
[ ]:
```

In the code below, a list is defined. Write two more lines of code, a line of code to transform it to a Pandas series and assign to a variable, and another to get the mean using the `.mean()` method.

```
[20]: budget = [50, 55, 40, 70, 50]
```

Write a function that flips 10 coins. If the coin is tails, score it as a 1. If the coin is heads, score it as a two. Add up all the flips and return the result. (This is like our roll dice function, but with two possible results instead of 6. You won't need the iterations part, just use the number 10 instead of the iterations variable.) I've written the first line for you.

If you're having trouble with this question, consult our [class notebook on simulations](#). Try to understand what each piece of code is doing, not just copy and paste.

```
[66]: def flip_10_coins():  
      pass # Delete this line and write your code here
```

Write a function to run the `flip_10_coins` function 100000 times. (If you're consulting the class notebook, you don't need the iterations variable, just use the number 100000.)

```
[64]: def run_simulation():  
      pass # Delete this line and write your code here
```

Run your simulation (by running the `run_simulation` function) and create a histogram using the `.plot(kind='hist')` method.

```
[ ]:
```

## 2 Linear Regression

```
[12]: # Import some data  
df = pandas.read_csv('https://raw.githubusercontent.com/smythp/  
↳foundations-data-science/master/data/ice_cream.csv')
```

Print out the columns of the data frame.

```
[ ]:
```

Create a scatterplot of revenue and temperature.

```
[ ]:
```

Describe the relationship between temperature and revenue that you see in the plot in a sentence or short paragraph.

What is a linear regression? In your own words, describe how a linear regression works and why it might be useful in examples such as the above about ice cream and temperature, or in cases like our electric car example from class. Give your response in the form of a short paragraph.

### 2.1 Bonus

Use the rest of the notebook to calculate a linear regression and plot it on a scatter plot. It may be easier to use `matplotlib.pyplot` to plot two items at once on the same graph. Feel free to consult

our class notebook on linear regression.

This is a bonus question. You will receive additional credit for completing it, but you will not lose any points for not attempting it, or if you provide an incomplete answer.

```
[13]: # You might be able to do it without pyplot, but it can help  
import matplotlib.pyplot as plt  
from sklearn.linear_model import LinearRegression
```

```
[ ]:
```

```
[ ]:
```