**CS544 Module 2 Assignment**

Abraham Cepeda Oseguera

U75425818

Boston University

MET CS 544: Foundations of Analytics and Data Visualization

Professor Hong Pan

October 21, 2022

**Part 1**

1. **What is the probability that a randomly selected person in this survey will have a BMI of above 30?**
   1. 36.17%
2. **If a randomly selected person had a BMI of above 30, what is the probability of that person being in the age group 18-34 years?**
   1. 29.36%
3. **If a randomly selected person had a BMI of above 30, what is the probability of that person being in the age group 35-49 years?** 
   1. 47.28%
4. **If a randomly selected person had a BMI of above 30, what is the probability of that person being in the age group 50-64 years?** 
   1. 18.14%
5. **If a randomly selected person had a BMI of above 30, what is the probability of that person being in the 65 years & over?**
   1. 5.23%

**Part 2**

1. **The sum of the rolls is greater than 6 but less than 10**

**A picture containing background pattern

Description automatically generatedA picture containing background pattern

Description automatically generated**

**Text

Description automatically generated**

1. **All the three rolls are identical.**

**Text

Description automatically generated**

1. **Only two of the three rolls are identical.**

**Background pattern

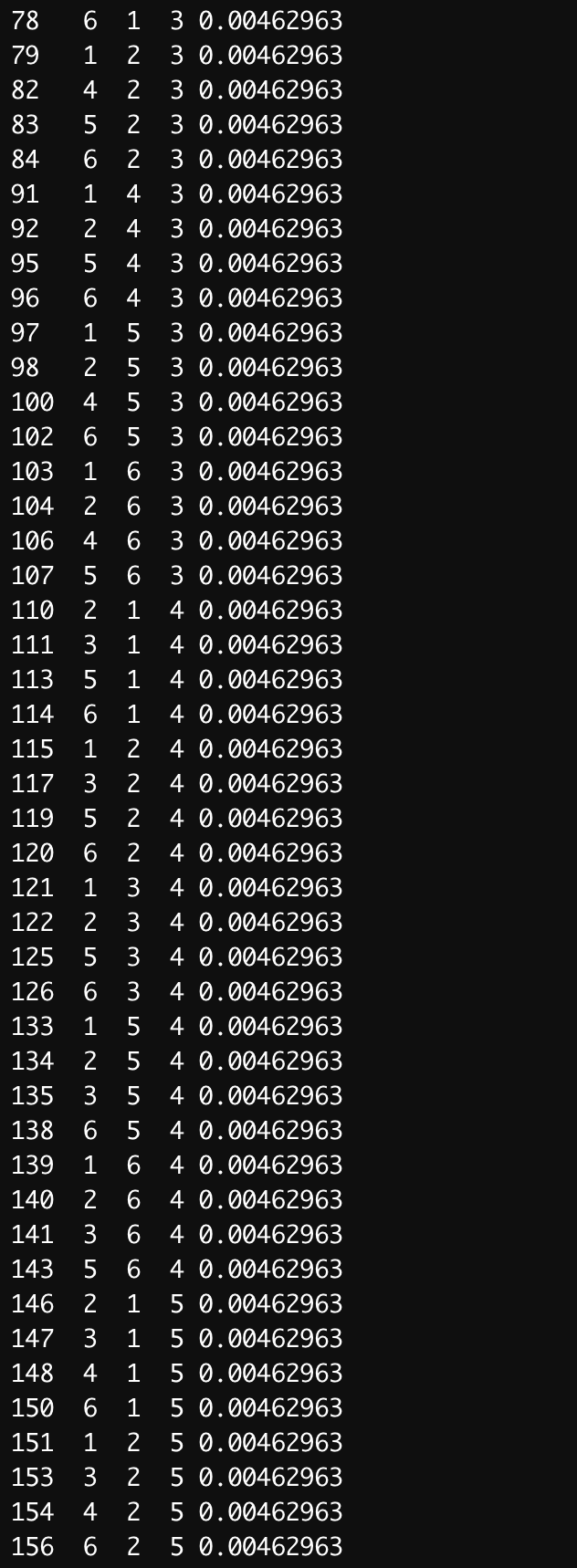
Description automatically generated with medium confidence**

**Text

Description automatically generated with medium confidence**

1. **None of the three rolls are identical.**

**Background pattern

Description automatically generated** **A picture containing background pattern

Description automatically generated**

1. **Only two of the three rolls are identical given that the sum of the rolls is greater than 9.**

**Text

Description automatically generated**

**Part 3**

1. **Using a for loop or a while loop, write your own R function, sum\_of\_first\_N\_odd\_squares (n)**

**Text

Description automatically generated**

1. **Without using a for loop or a while loop, write your own R function, sum\_of\_first\_N\_odd\_squares\_V2 (n)**

**Text

Description automatically generated**

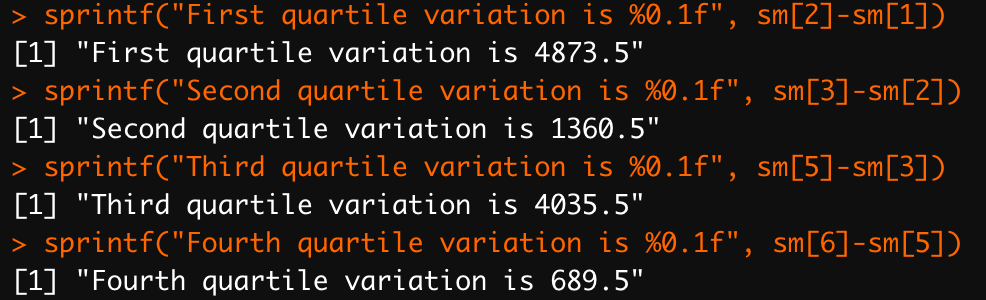
**Part 4**

1. **Store the result of the summary function for the Close attribute as the variable sm. Change the names of this variable so that the output appears as shown below.**

**A picture containing text

Description automatically generated**

**a.2 ) Using the above data, show the quartile variations for the 4 quartiles as shown below**

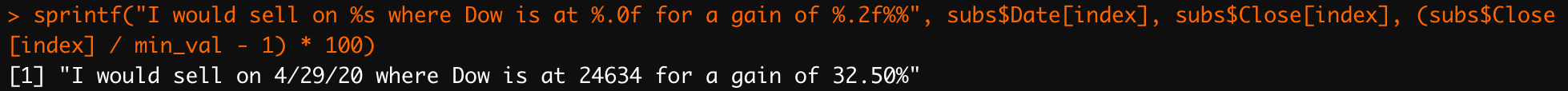
****

1. **Produce the output for the minimum of the Dow closing value in the dataset as shown below**

**Text

Description automatically generated**

1. **Suppose you have an index fund tied to the Dow closing value. If you have invested on the minimum date, what date from the dataset you would have sold to gain the maximum percentage gain. The output is as shown below. Note that the code should be generic so that it works on any such dataset.**

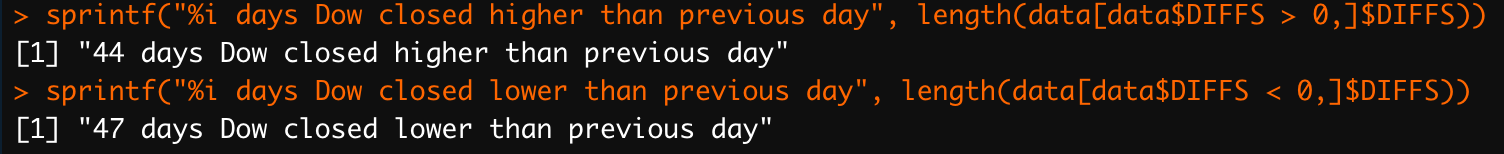
****

1. **Use the diff function to calculate the differences between consecutive closing values in the dataset. Insert the value 0 at the beginning of these differences. Add this result as the DIFFS column of the data frame. The result is as shown below.**

**Text

Description automatically generated**

1. **How many days did the Dow close higher than its previous day value? How many days did the Dow close lower than its previous day value?**

****

1. **Show the subset of the data where there was a gain of at least 1000 points from its previous day value.**

**Text

Description automatically generated**