

ALY_6000 Project 6 Report

Junchen Yi

College Professional Studies in Analytics, Northeastern University

ALY_6000: Introduction to Analytics

Zhi He

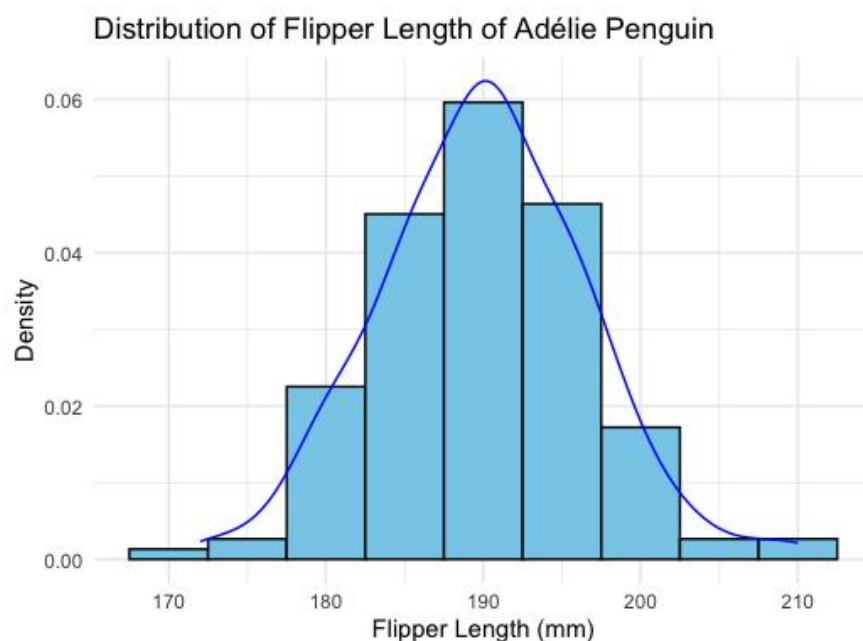
Submission Date: 27/10/2023

Introduction:

It is important to analyze the morphological characteristics of Antarctic penguins to facilitate the understanding of their ecological and adaptive responses. This report analyzes the distributional characteristics of flipper length in Adélie penguins and the relationship between flipper length and beak depth in Antarctic penguins. By analyzing and comparing these two characteristics, I hope to reveal a possible link between them.

Key Findings:

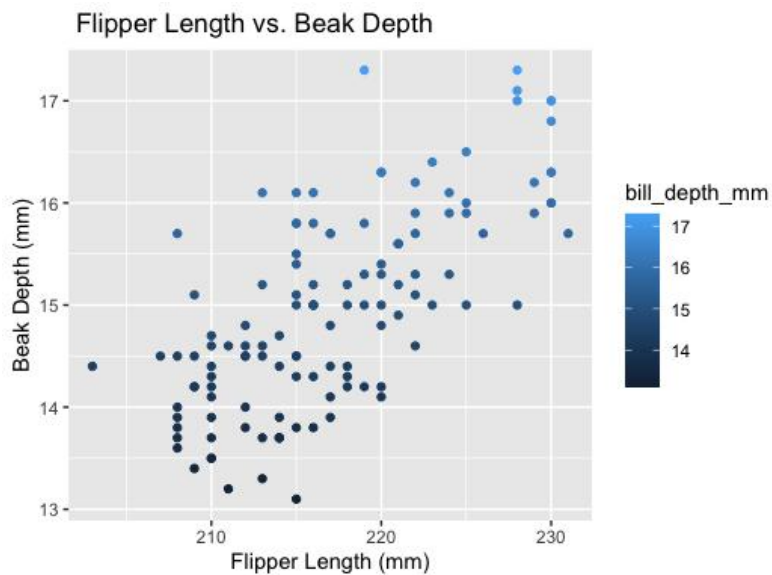
Question 28:



In this question, I first extracted the flipper length data of Adélie penguins from the 'palmerpenguins' dataset, and then drew a histogram and a density plot using the functions of the 'ggplot' package.

I think using histogram to determine the distribution of the data is an appropriate choice, because it can visualize the distribution of the data and help us understand the shape of the data. From the histogram, we can see that the shape of the data is close to a bell curve, which is typical of a normal distribution. We will then perform the Shapiro-Wilk normality test and check for skewness and kurtosis. I got the result of the Shapiro-Wilk normality test of 0.7200466 from the result of running the code, which shows that we don't have enough evidence to reject the original hypothesis, so we can assume that the data may come from a normal distribution. Secondly, the skewness value is equal to 0.08560929 which indicates that the data is almost symmetrical and the positive value indicates that the data is skewed to the right. The value of kurtosis value is equal to 0.2382734 which indicates that the data is similar to the normal distribution. Positive values indicate that the tails of the data are longer than the normal distribution.

To summarize, we can approximate that the data in this dataset is normally distributed.

Question 29:

Since the requirement was to analyze the relationship between two variables, I chose to draw a scatter plot. In addition, I used the Pearson's correlation coefficient to measure the linear relationship between the two data sets. By calculating this, I got a result of 0.7065634, which shows that the length of the flippers and the depth of the beak have a positive linear relationship, which means that both variables increase or decrease at the same time in most cases.

Conclusion:

In conclusion, after data analysis and statistical tests, we found that the flipper length data of Adeline penguins is close to a normal distribution. In addition, we observed a positive linear relationship between flipper length and beak depth in Antarctic penguins, with a Pearson's correlation coefficient of 0.7065634, suggesting that when flipper length increases, beak depth also tends to increase.