STAT 463 - Assignment 2

Due Date: Saturday February 25th, 2023

Data Description and Background

In 1929, Edwin Hubble investigated the relationship between the distance of a galaxy from the earth and the velocity with which it appears to be receding. Galaxies appear to be moving away from us no matter which direction we look. This is thought to be the result of the "Big Bang". Hubble hoped to provide some knowledge about how the universe was formed and what might happen in the future. The data collected includes distances (in megaparsecs) to 24 galaxies and their recession velocities (in km/sec). Note: 1 parsec = 3.26 light years. In a short write-up address the following:

1) Visually display the data in an appropriate graph and comment on anything that may be of note. Be sure to adequately label the graph, indicating in a title what is being displayed as well as what each axis represent (including units of measurement). – 2 points
2) It is desired to ascertain whether galaxies that are farther away tend to be receding faster from Earth (i.e. they have higher recession velocities). Based on the plot you constructed in part (1), would a linear regression model with recession velocity as the response variable and distance as the explanatory variable be appropriate for this dataset? – 2 points
3) What is the nature of the relationship between galaxies' distances from earth and their recession velocities? Specifically, as the distance increases, how is the recession velocity affected? Quantify it and interpret this quantity. – 2 points
4) Does the data suggest that there is a statistically significant relationship between galaxies' recession velocities and their distances from earth? – 3 points
5) How well do galaxies' distances from earth explain their recession velocities? Quantify this and interpret this quantity. – 2 points
6) Andromeda, the nearest large spiral galaxy to our home galaxy the Milky Way, is 0.77 megaparsecs away. Give an estimate of its recession velocity. Additionally, provide a range of value that will encompass Andromeda's true recession velocity with probability 0.95. – 4 points