STA130 Capstone Project

Poster

Presentation

Riyad | Nicolas | Henry | Chloe

3 Research Questions Includes Data,

Method, Visulization & Conclusion

RESEARCH QUESTION 3

How well can a unary linear regression model predict the galaxy's apparent brightness from redshift?

Authors
Henry & Chloe

The research is the capstone project poster presentation in the course STA130

INTRODUCTION

In GalaxyZoo Tabular Data content, there are two variables

VISUALIZATION
N

`Fitted line plot & Histogram

DATA

`Galaxy Zoo Tabular Data Contents'

INTRODUCTION

In GalaxyZoo Tabular Data content, there are two variables that are related which are redshift and elpetro_absmag_r.
Therefore, we decide to investigate how well can a unary linear regression model predict the galaxy's apparent brightness from redshift?

VISUALIZATION N

`Fitted line plot & Histogram

METHODOLOGY

The method we used in research question 3 is simple linear regression. Simple linear regression is a statistical method that allows us to summarize and study relationships between two continuous or quantitative variables.
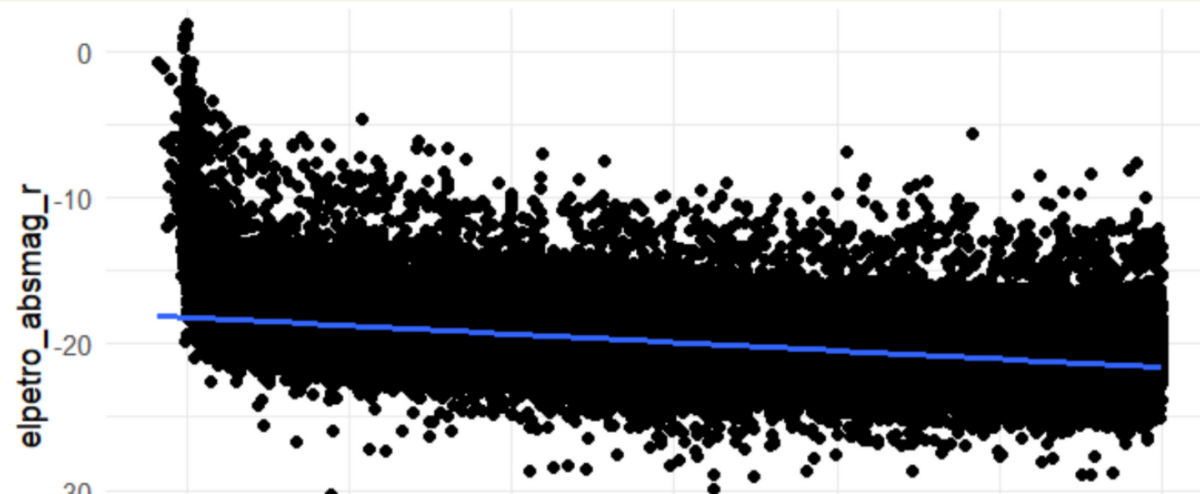
DATA

`Galaxy Zoo Tabular Data Contents'
redshift - related to how far away that galaxy is from us;
elpetro_absmag_r - An estimate of the galaxy's total luminosity brightness or intrinsic brightness measured in absolute magnitude

ANALYSIS

Firstly, we update the data by loading libraries ('tidyverse' and 'arrow') and remove all the 'NA' values.
Secondly, we create a linear regression plot , which 'redshift' is the x-axis and 'elpetro_absmag_r' is the y-axis.
Thirdly, we create a 'mod' the analysis the relationship between 'redshif' and 'elpetro absmag r'. We apply 'summary$coefficients' to find the B1 , B0 and p-value.
Fourly, we can do hypothesis test . H0 will be

which 'redshift' is the x-axis and 'elpetro_absmag_r' is the y-axis.

Thirdly, we create a 'mod' the analysis the relationship between 'redshif' and 'elpetro absmag r'. We apply 'summary$coefficients' to find the B1, B0 and p-value.

Fourly, we can do hypothesis test. H0 will be There is a linear relationship between brightness and redshift. H1 will be There is not a linear relationship between brightness and redshift. The p-value we got form summary will tell us whether to reject H0.
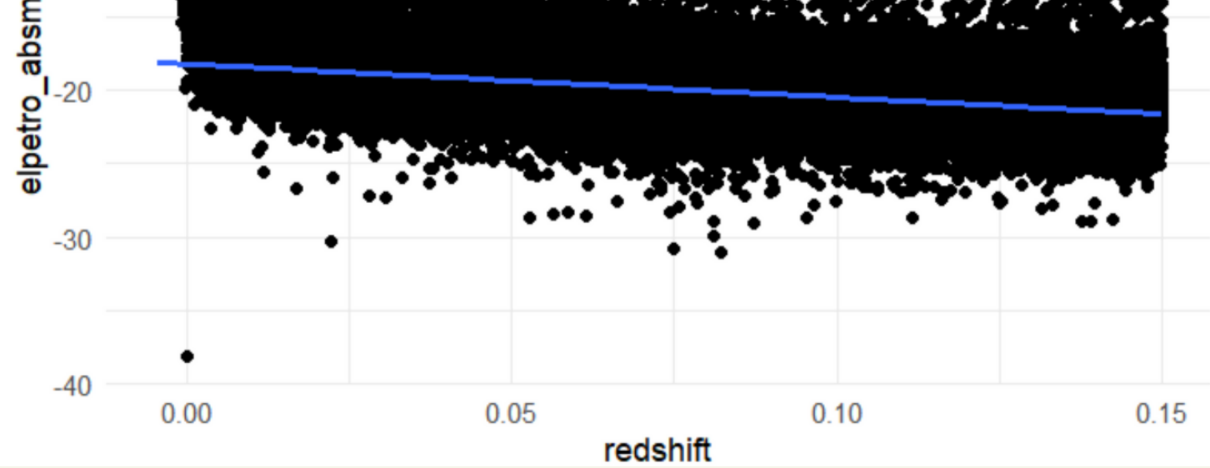
CONCLUSION

The B0 is the brightness when redshift equal to zero, which is -18. B1 is the average change in brightness for 1 unit change in redshift which is - 13.3

Form here, we can get the linear regression model:

$$y_i = \beta_0 + \beta_1 x_i$$
$$y_i = -18 - 13.3 x_i$$

$\alpha = 0.05$. The p-value is 0.11 which is bigger than α, we will reject H1, support H0.

There is a linear relationship between brightness and redshift.