report

Student ID: 11484265

2024-09-19

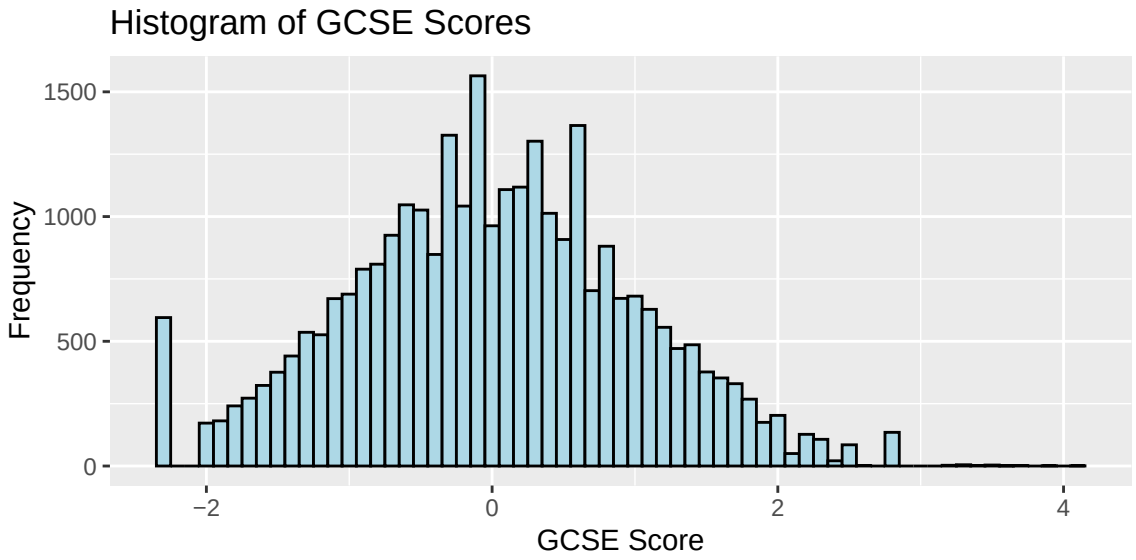
Question 1: Descriptive Analysis (10%)

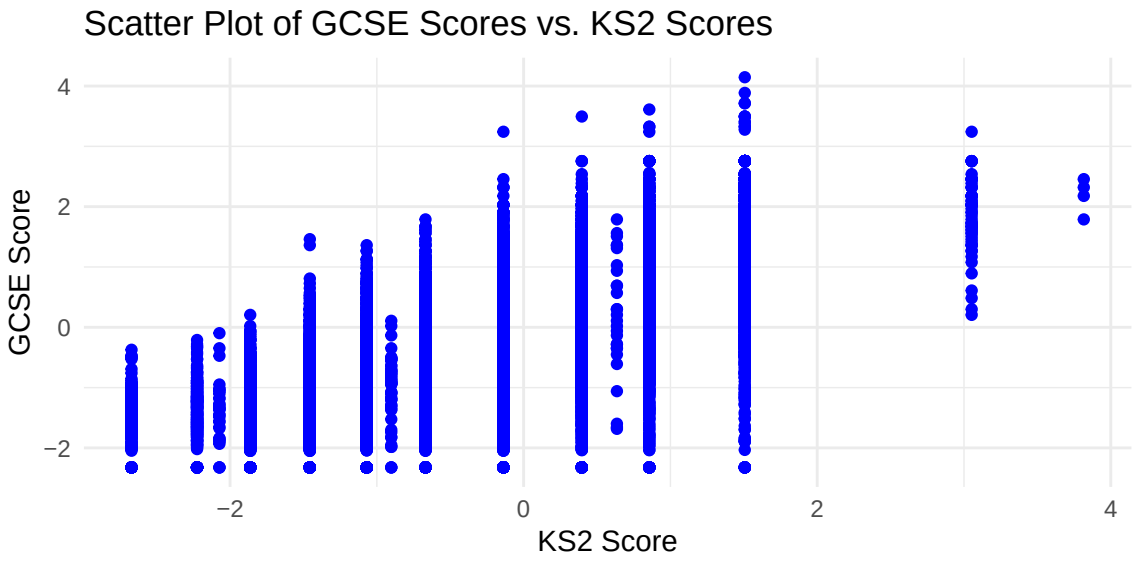
Describe the dataset: -How many pupils and schools are there in the dataset?

There are 54 schools and 29506 pupils.

-What is the average number of pupils per school? 546

-Create histograms for “gcsescore” and “ks2score”.





While there are pupils who score low on the GCSE on a wide range of KS2 scores, the general trend is that higher GCSE scorers generally have higher KS2. Those who scored extremely high on the KS2 score may not always have the highest GCSE, but are all above average.

Question 2: Model Estimation (30%)

Estimate the following multilevel models and summarise the results in a table:

-Model 0 (m0): An empty model with "gcsescore" as the dependent variable and a random intercept at the school level.

-Model 1 (m1): m0 + a fixed effect for "ks2score".

-Model 2 (m2): m1 + "ks2score" squared to capture non-linear effects.

-Model 3 (m3): m2 + a random slope for "ks2score" to allow the relationship between "ks2score" and "gcsescore" to vary by school.