Introduction Draft/Notes

Methodology Notes

(Support this with cited texted book or article reference. Any book on statistics or R statistics should do.)

Linear Regression

https://www.statisticssolutions.com/free-resources/directory-of-statistical-analyses/assumptions-of-linear-regression/

1. Assumptions.

Linear regression has 5 key assumptions.

1st is a linear relationship, that is, the relationship between the independent and dependent variables are linear. From my initial scatter plots, the relationships I’ve been looking at, pop\_density ~ distance\_from Sydney, the relationship doesn’t seem to be linear.

Our correlation coefficient is around 0.2 which is quite low.

2nd all variables are multivariate normal.

Checked with histogram of Q-Q plot.

Log-transformation give the data a more linear distribution.

Check normality with a goodness of fit test (Kolmogorov-Sminov test)

3rd no or little multicollinearity.

Distance Between Two latitude and longitude points

1. Haversine