Assessment 1 - Statistics

Date for submission: To be fixed in class

Using the data provided in the course, please apply the following statistical techniques to questions that you deem appropriate. You can use different data for different techniques.

Present the results in a word document and then interpret them.

Save you working in R in a text file and submit that with your report.

The report should be succinct with the emphasis on showing that
1. You know what each statistical technique is used for – this is demonstrated by your choice of variables etc. in the data.
2. You know how to run the techniques in R, this is demonstrated by your results presented in the report but also by your working in R.
3. You know how to interpret the results – this is demonstrated by your description / interpretations of the results in your report.

I am available for zoom discussion most days if you have problems.

Techniques for the report
1.   Chi2 Test for independence + Pearson Residuals
2.   Multiple Correspondence Analysis
2b.   Cluster Analysis of Results of Correspondence Analysis – Factor Map
3.   Agglomerative Cluster Analysis
3b.   K-Means (Medoid) Cluster Analysis – Silhouette Plot
4.   Binary Fixed-Effect Multiple Logistic Regression
4b.   Regression Diagnostics: Orthogonality; Parsimony; Cross-Validation
        (Bootstrapping); Influential Observations

The marking will be based on a score of 60 (20 for each of the three points made above, 5 Point for each of the techniques). Your mark will then be divided by 3 to give a final mark out of 20.