Proposal additions course Regressions II 2022 - University of Groningen (**To be discussed)**

Juan Rivillas. Updated: 5th Aug 2022

| Week | Syllabus RUG | Current, new changes and additions | Learning outcomes |
| --- | --- | --- | --- |
| Week 1 Multiple Regression(5 -9 sep) | * Understand the challenges and opportunities of multiple regression. * Understand how to test multiple. hypotheses using multiple regression * Asses model fit. * Asses violations of regression assumptions. | What is exploratory data analysis? (Week 6)   * Preparing dataset for analysis * Getting data into a usable format * Importing and working with data * Visualising distributions * Dealing with outliers and missing values * Exploring variation and covariation | * **Format** research data ready for analysis * **Formulate** questions about a dataset * **Select** a suitable visualisation for a given question * **Generate** useful visualisations from your data * **Evaluate** the effectiveness of a data visualisation * **Identify** potential sources of bias in the data. |
| Week 2. Challenges in Multiple Regression(12 -16 sep) | * Identify potential pitfalls in multiple regression * Take steps to prevent and/or deal with potential pitfalls * Understand how to model quadratic relationships and interactions | What is correlation analysis and test the significance?   * Steps to test associations between categorical variables * Correlation test between variables using R packages. * Interpretate correlograms or correlation matrix. | * **Understand** the need to check association before assess causation. * **Interpret** graphical correlations. * **Calculate** correlations for sampled data. * **Apply** correlations methods to answer questions about a population. |
| Week 3. Generalised Linear Models(19 -23 sep) | * Interpret the weaknesses of Linear Models for binary and count outcomes. * Understand and use Logistic regression for binary dummy variables. * Understand and use Poisson regression for count data. * Hypothesis testing using GLMs. | Clustering Analysis and Latent Class Analysis (LCA) (Week 5)   * Understand the mechanics and motivation behind CA. * Understand the mechanics and motivation behind LCA. * Perform CA and LCA on data and interpret outcome. * Visualise the outcome of CA and LCA. | * **Understand** advantages and disadvantages of both methods. * **Identify** the best fit model in cluster analysis and Latent Class Analysis (LCA). * **Formulate** comparisons CA and LCA in the data. * **Interpretate** results and selection best fit models. |
| Week 4. Application of Regression(26-30 sep) | * Apply regression analysis in real world datasets. * Understand how to combine exploratory analysis with regression analysis. * Interpret the outcome of the analysis. | Linear Regression (Week 3)   * Directed Acyclic Graph (DAGs) for model accuracy. * Regression - residuals and least squares. * Regression with a single predictor variable * Interpret the weaknesses of Linear Models for binary and count outcomes. | * **Identify** the correlation coefficient as a single measure of linear association. * **Apply** linear regression to model a response variable in terms of a single or multiple variables * **Assess** model validity by checking model assumptions. * **Assess**model fitness by comparing the results produced by the model. |
|  |  | Generalized Linear Models   * Logistic regression for binary and categorical variables. * Generalized Linear Models (GLMs): Poisson regression and Analysis of Variance (ANOVA) models. |  |
| Week 5. Clustering Analysis and Latent Class Analysis (LCA)(3-7 oct) | * Understand the mechanics and motivation behind Cluster Analysis (CA). * Understand the mechanics and motivation behind Latent Class Analysis (LCA). * Perform CA and LCA on datasets and interpret the outcome. * Visualise the outcome of CA and LCA. | Multiple regression (Weeks 1 and 2)   * Challenges and opportunities of multiple regression. * What do we assume? * Steps in multiple regression modelling process * How to test multiple hypotheses using multiple explanatory variables. * Confounding factors. * Interpretation with multiple variables. * Model interactions, model checking, and model fitting. * Asses violations of regression assumptions. * Identify potential pitfalls in multiple regression * Take steps to prevent and/or deal with potential pitfalls * How to model quadratic relationships and interactions | * **Understand** the challenges and opportunities of multiple regression * **Understand** how to test multiple hypotheses using multiple regression * **Assess** model validity by checking model assumptions. * **Assess**model fitness by comparing the results produced by the model. |
| Week 6. Importing, Tidying and Joining Datasets(10-14 oct) | * Import a wide variety of datasets into R. * Understand the concept of joins from relational algebra. * Understand the concept of tidy data. * Join and tidy datasets using R code. | Application of Regression (Week 4)   * Hands-on example. * Apply regression analysis in real world datasets. * Understand how to combine exploratory analysis with regression analysis. * Interpret the outcome of the analysis. | * **Generate** useful visualisations from your data * **Evaluate** the effectiveness of a data visualisation * **Assess**model fitness by comparing the results produced by the model with your data. * **Interpretate** results and selection best fit models |
| Week 7.Open datasets for SDG monitoring(17-21 oct) | * Become familiar with all major open datasets for social science research. * Use said datasets to collect information for their research projects. * Understand the tools used to monitor the achievement of Sustainable Development Growth. | Open datasets for SDG monitoring   * R as a data-handling environment. * The tidyverse packages. * Reading tabular data with readr. * Tidying data with tidyr. * Transforming data with dplyr. * Data visualisation with ggplot2. * Applying statistical tests to tabular data. | * **Use** R to import and tidy tabular data. * **Combine** data from multiple sources. * **Transform** data to generate derived variables. * **Apply** statistical tests to data in R. * **Create** high-quality data visualisations. |
| Week 8.Putting it all together (24-28 oct) | * Reflect on the material covered in the course. * Use feedback from the lecturer and each other to come full circle on what they have learned and how to use it * Have a chance to ask questions about their final project. | Putting it all together | * **Reflect** on the material covered in the course. * **Use** feedback from the lecturer. * **Apply** new data science skills on final project. |
| Week 9. Your Turn!  (1-4 nov) | * Present the results of their final project thus gaining exposure to presenting data driven reports | Your Turn!   * Present research report. * Give and receive constructive feedback. | * Know your research report. * Know your role and part in the presentation. * Practice the presentation. |