Data Analysis with R for Social Scientists

Jakob Tures & Jasper Tjaden

2023-08-04

Contents

Intro **7**

1 Introduction to Seminar **9**

1.1 A section 9

2 Exloratory Data Analysis - I **11**

2.1 Load data 11

2.2 Introduce WVS 11

2.3 glimpse() 11

2.4 skim() 11

2.5 Types of Variables/ Skalen 12

2.6 Univaraite statistics (means, SDs, min, max) 12

2.7 ggplot 12

2.8 gtsummary 12

3 Exloratory Data Analysis - II **13**

3.1 Markdown Introduction 13

3.2 Applying EDA(WVS/own data) 13

4 Linear Regression - Theory **15**

4.1 What is it? 15

4.2 When and for what it can used? 15

4.3 Formula (short) 15

4.4 Assumptions (short) 15

4.5 Mediation 15
4.6 Interactions? 16
4.7 Multiple outcomes 16

5 Linear Regression - Applied **17**
5.1 Incl. Short Theory and DAG 17
5.2 Interpreation of regression tables in practice 17
5.3 Mediation 17
5.4 Regressional Diagnostics 17

6 Linear Regression - Exercises **19**
6.1 Application of Linear Regression 19

7 Logistic Regression - Theory **21**
7.1 What is it? 21
7.2 When and for what it can used? 21
7.3 Formula (short) 21
7.4 Assumptions (short) 21
7.5 Mediation 21
7.6 Multiple outcomes 22
7.7 Multinomial 22

8 Logistic Regression - Applied **23**
8.1 Incl. Short Theory and DAG 23
8.2 Interpreation of regression tables in practice 23
8.3 Mediation 23
8.4 Regressional Diagnostics 24

9 Logistic Regression - Exercises **25**
9.1 Application of Logistic Regression 25

10 Prediction or Margins - Theory **27**
10.1 Predicted probabilities 27
10.2 Marginal Effects 27

11 Prediction or Margins - Exercises **29**

11.1 Application of Regression 29

12 Reporting and Visualization **31**

12.1 Formatted regression tables 31

12.2 Publication-ready formatting/ labelling of visuals 31

12.3 coefficient plots 31

13 Discussion of ideas and term papers **33**

14 Outlook **35**

14.1 Machine Learning 35

Intro

This course offers an accessible and easy introduction to one of the fastest growing statistical packages used in social science and data science more generally.

Please download the data used in the course here. To find more about me, have a look at my website. Also, feel free to watch me as I walk you through each lesson here.

Overview over the Course :

- **Week 1: Introduction to Seminar**
- **Week 2: Exploratory Data Analysis-I**
- **Week 3: Exploratory Data Analysis-II**
- **Week 4: Linear Regression-Theory**
- **Week 5: Linear Regression-Applied**
- **Week 6: Linear Regression-Exercises**
- **Week 7: Logistic Regression-Theory**
- **Week 8: Logistic Regression-Applied**
- **Week 9: Logistic Regression-Exercises**
- **Week 10: Prediction or Margins-Theory**
- **Week 11: Prediction or Margins-Exercises**
- **Week 12: Reporting and Visualizing**
- **Week 13: Discussion of Ideas and term papers**
- **Week 14: Outlook**

Chapter 1

Introduction to Seminar

All chapters start with a first-level heading followed by your chapter title, like the line above. There should be only one first-level heading (`#`) per .Rmd file.

1.1 A section

All chapter sections start with a second-level (`##`) or higher heading followed by your section title, like the sections above and below here. You can have as many as you want within a chapter.

An unnumbered section

Chapters and sections are numbered by default. To un-number a heading, add a `{.unnumbered}` or the shorter `{-}` at the end of the heading, like in this section.

Chapter 2

Exloratory Data Analysis - I

Here goes some texts.

2.1 Load data

Here goes some texts.

2.2 Introduce WVS

Here goes some texts.

2.3 glimpse()

Here goes some texts.

2.4 skim()

Here goes some texts.

2.5 Types of Variables/ Skalen

Here goes some texts.

2.6 Univaraite statistics (means, SDs, min, max)

Here goes some texts.

2.7 ggplot

Here goes some texts.

2.7.1 Histograms

Here goes some texts.

2.7.2 Boxplots

Here goes some texts.

2.7.3 Bar graphs

Here goes some texts.

2.7.4 Scatterplots

Here goes some texts.

2.8 gtsummary

Here goes some texts.

2.8.1 Kreuztabellen

Here goes some texts.

Chapter 3

Exloratory Data Analysis - II

Here goes some texts.

3.1 Markdown Introduction

Here goes some texts.

3.2 Applying EDA(WVS/own data)

Here goes some texts.

Chapter 4

Linear Regression - Theory

4.1 What is it?

Here goes some texts.

4.2 When and for what it can used?

Here goes some texts.

4.3 Formula (short)

Here goes some texts.

4.4 Assumptions (short)

Here goes some texts.

##Interpretation of results Here goes some texts.

4.5 Mediation

Maybe theory into DAG session and example into application?

4.6 Interactions?

Here goes some texts.

4.7 Multiple outcomes

Here goes some texts.

Chapter 5

Linear Regression - Applied

5.1 Incl. Short Theory and DAG

Here goes some texts.

5.1.1 Application with WVS data

Here goes some texts.

5.2 Interpreation of regression tables in practice

Here goes some texts.

5.3 Mediation

Here goes some texts.

5.3.1 Total and Direct effect

Here goes some texts.

5.4 Regressional Diagnostics

(Maybe)

Chapter 6

Linear Regression - Exercises

Here goes some texts.

6.1 Application of Linear Regression

With WVS/own data: Students apply linear regression.

Chapter 7

Logistic Regression - Theory

7.1 What is it?

Here goes some texts.

7.2 When and for what it can used?

Here goes some texts.

7.3 Formula (short)

Here goes some texts.

7.4 Assumptions (short)

Here goes some texts.

##Interpretation of results Here goes some texts.

7.5 Mediation

Maybe theory into DAG session and example into application?

7.6 Multiple outcomes

Here goes some texts.

7.7 Multinomial

Chapter 8

Logistic Regression - Applied

8.1 Incl. Short Theory and DAG

Here goes some texts.

8.1.1 Application with WVS data

Here goes some texts.

8.2 Interpreation of regression tables in practice

Here goes some texts.

8.3 Mediation

Here goes some texts.

8.3.1 Total and Direct effect

Here goes some texts.

8.4 Regressional Diagnostics

(Maybe)

Chapter 9

Logistic Regression - Exercises

Here goes some texts.

9.1 Application of Logistic Regression

With WVS/own data: Students apply linear regression.

Chapter 10

Prediction or Margins - Theory

Here goes some texts.

10.1 Predicted probabilities

At various co-variate levels

10.2 Marginal Effects

Chapter 11

Prediction or Margins - Exercises

Here goes some texts.

11.1 Application of Regression

With WVS/own data: Students apply linear+logistic regression from previous exercises.

Chapter 12

Reporting and Visualization

Here goes some texts.

12.1 Formatted regression tables

Here goes some texts.

12.2 Publication-ready formatting/ labelling of visuals

Here goes some texts.

12.3 coefficient plots

Here goes some texts.

Chapter 13

Discussion of ideas and term papers

Here goes some texts.

Chapter 14

Outlook

Here goes some texts.

14.1 Machine Learning

Here goes some texts.