Third Assignment

Introduction

The main goal of this third assignment is to gather web based data from at least two sources, merge the data sets, conduct basic descriptive and some inferential statistics on the data to address a relevant research question and briefly describe the results including with dynamically generated tables and figures.

We have gathered to different datasets from two different data sources. The first dataset is from the CIA World Fact book, and the second from the The Quality of Government Institute (QOG), which is part of the University of Gothenburg. Both datasets are open and can be found in their respective webpages.

In order to produce a tangible statistical analysis we have cleaned and merged the relevant data of both datasets. Then, we have generate tables and figures to relate the data with the aim of adressing, as well as possible, our researching question to begin to determine whether our hypothesis is correct or not.

Detailed process

The data introduced in the Markdown document was, on the one hand, imported from an excel document, the data of which had been previously cleaned and merged from the original QOG document. The second dataset has been gathered from the World Bank data and introduced from the WDI package, which is already included in 'R'.

Tables and Figures

Purpose: Gather, clean, and (begin to) analyse data Deadline: 11 November You will submit a GitHub repo that: - Gathers web-based data from at least two sources. Cleans and merges the data so that it is ready for statistical analyses. - Conducts basic descriptive statistics with the data to address a relevant research question. Could include inferential stats. - Briefly describes the results including with dynamically generated tables and figures. - Has a write up of 1,500 words maximum that describes the data gathering and analysis, It also will use literate programming

```
## Loading required package: rJava
```

```
## Loading required package: xlsxjars
```

Education and Internet Accessibility

Cleaning QOG data
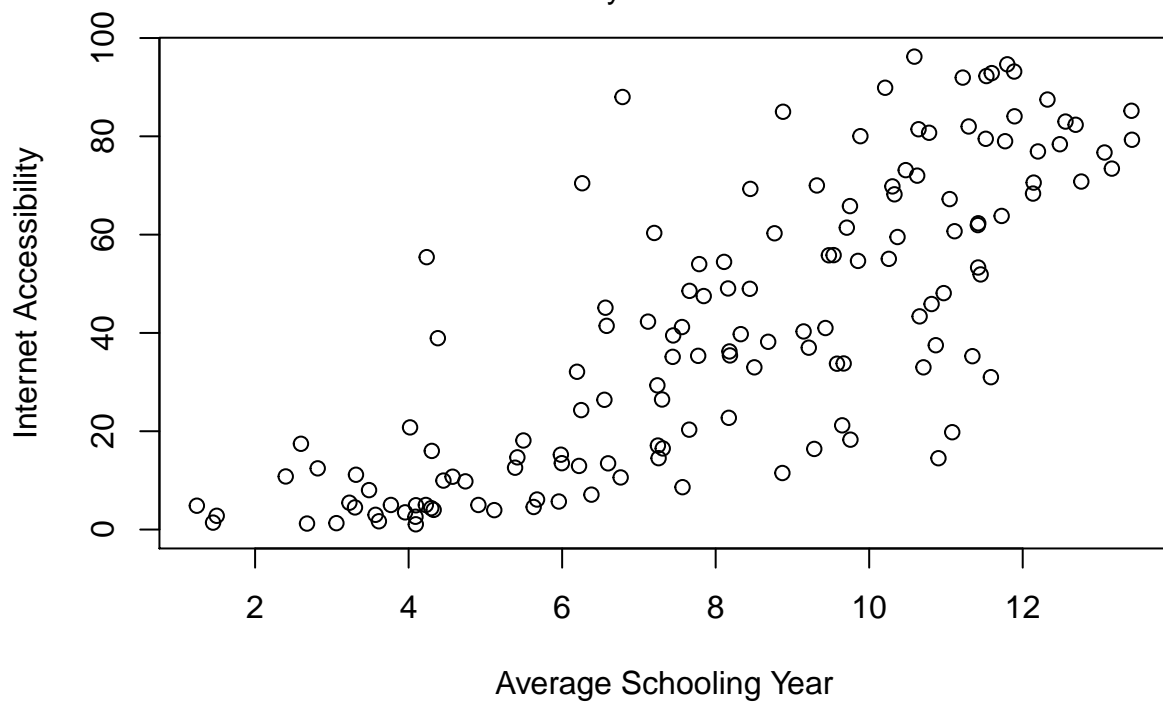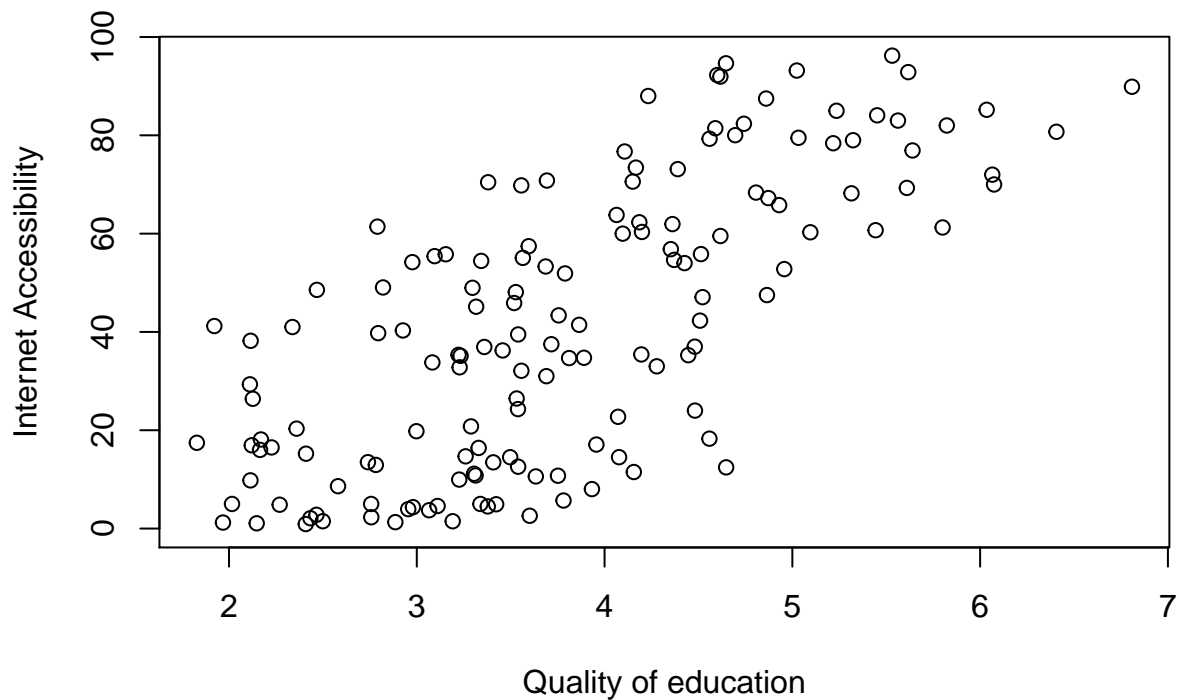
Conducting hypothesis tests

```
##
##  Pearson's product-moment correlation
```

```
##
## data:  Q and I
## t = 12.396, df = 146, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.6271975 0.7865483
## sample estimates:
##       cor
## 0.7160782

##
##  Pearson's product-moment correlation
##
## data:  Y and I
## t = 15.355, df = 139, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.7225957 0.8473676
## sample estimates:
##       cor
## 0.7931617
```
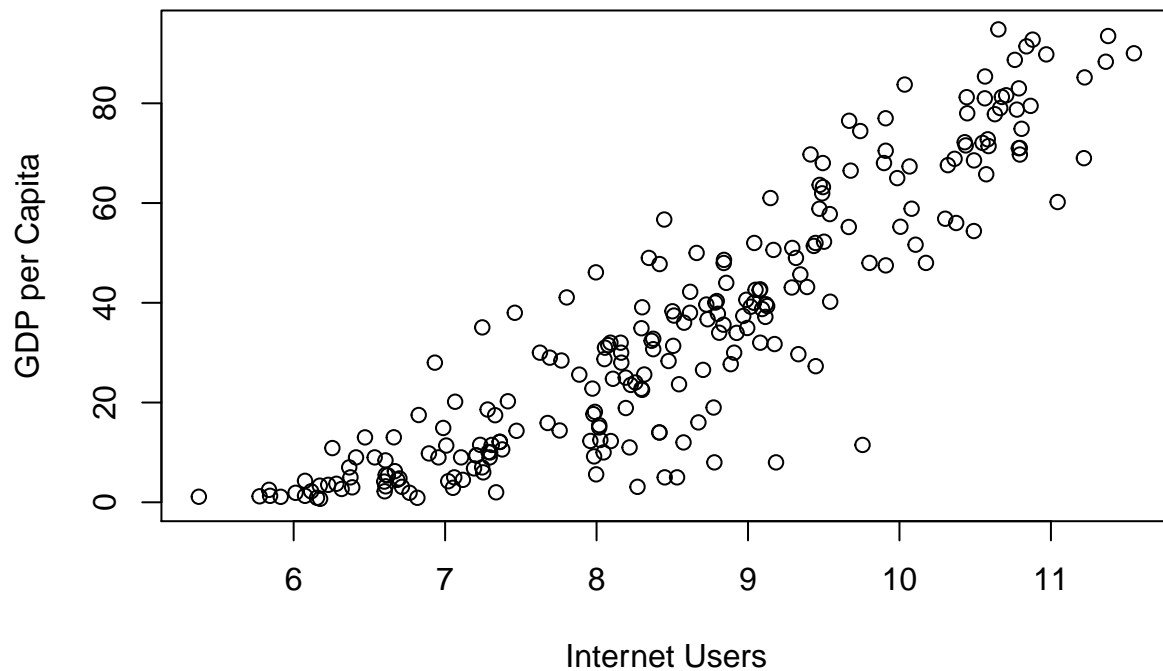
GDP and Internet Accessibility

Cleaning WDI data

```
## Loading required package: RJSONIO
```

Conducting basic descriptive statistics

conducting hypothesis tests

```
##
##  Pearson's product-moment correlation
##
## data:  G and I2
## t = 20.74, df = 232, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.7558780 0.8467173
## sample estimates:
##       cor
## 0.8059925
```