Datenanalyse

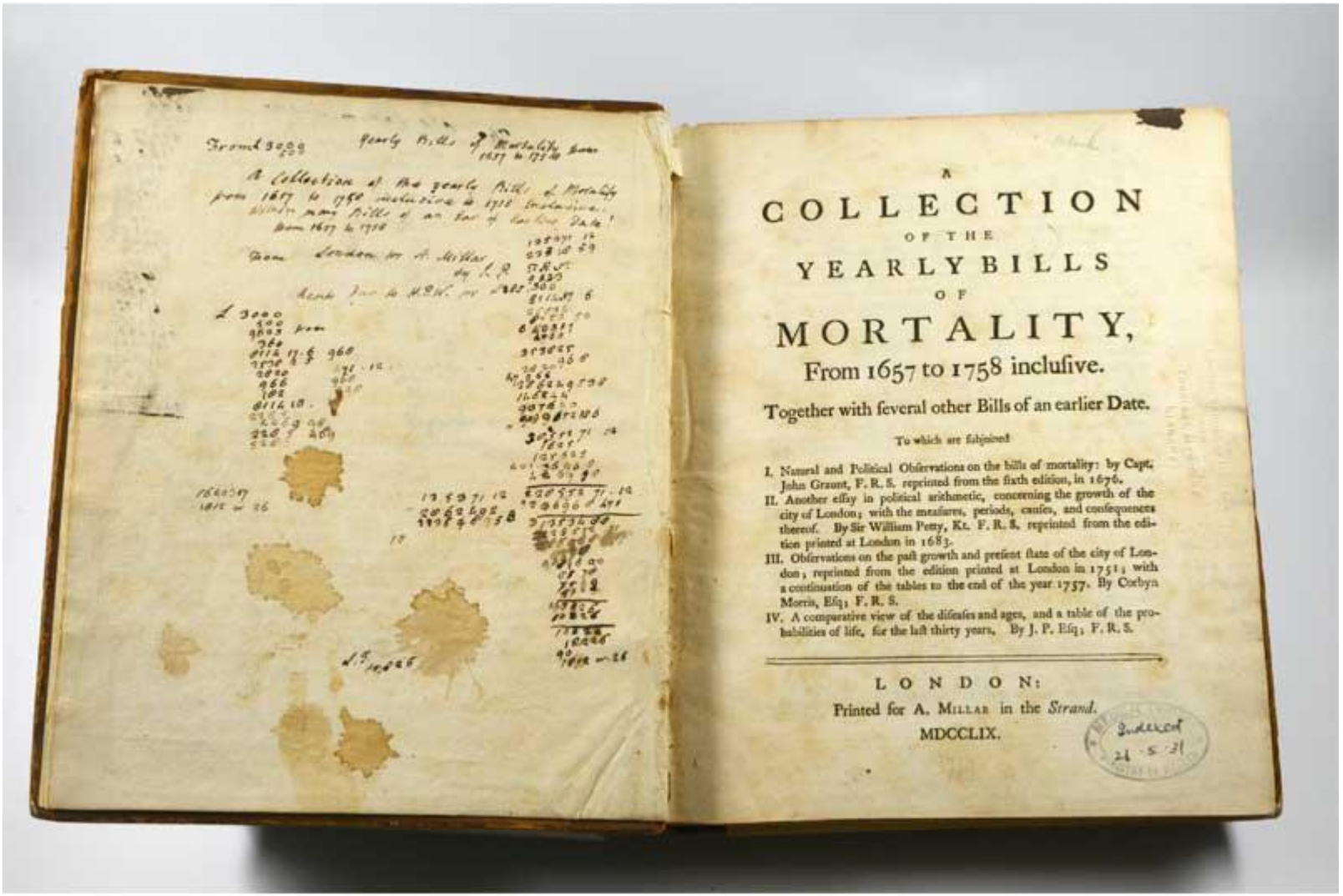
.consulting .solutions .partnership

Datenanalyse

- Google´s Chef-Ökonom Hal Varian
  - „The next sexy job“
  - „The ability to take data – to be able to understand it, to process it, to extract value from it, to communicate it – that´s going to be a hugely important skill.“
  - New York Times, 2009

- „Hot new gig in tech“ – Fortune

Bills of Mortality



A
COLLECTION
OF THE
YEARLY BILLS
OF
MORTALITY,
From 1657 to 1758 inclusive.

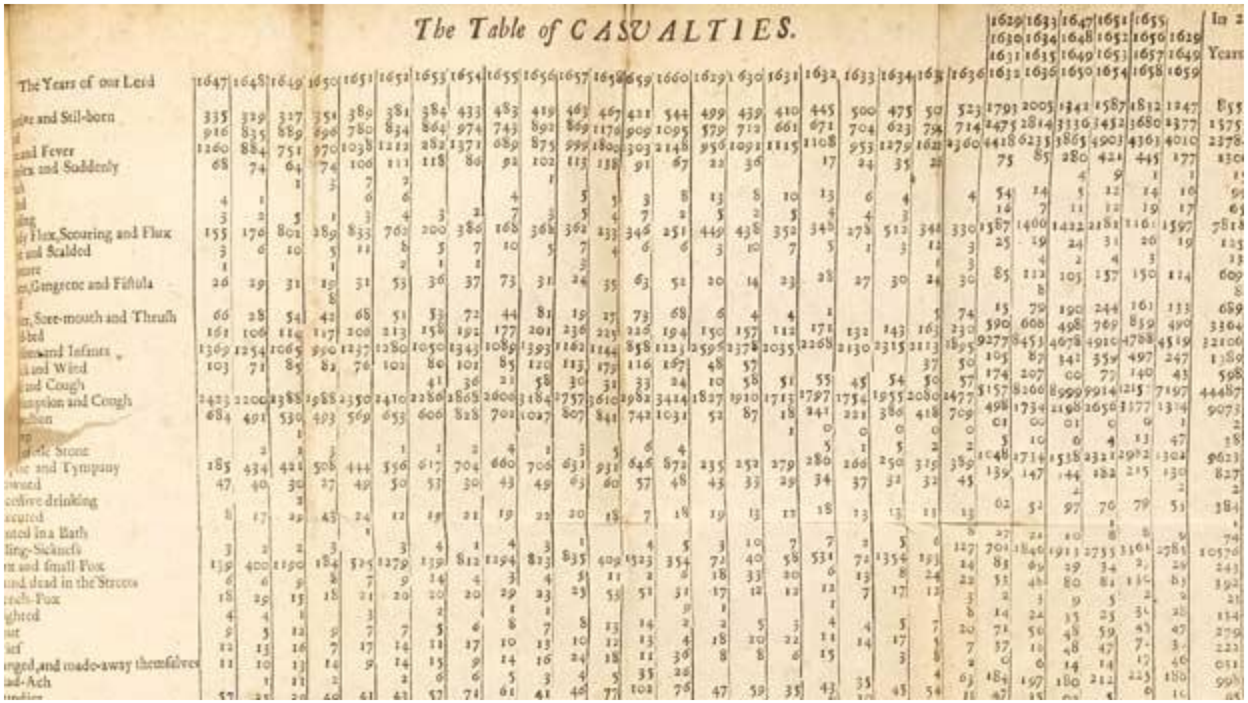
Together with several other Bills of an earlier Date.

To which are subjoined

I. Natural and Political Observations on the bills of mortality: by Capt. John Graunt, F. R. S. reprinted from the fixth edition, in 1676.
II. Another eſſay in political arithmetic, concerning the growth of the city of London; with the meaſures, periods, cauſes, and conſequences thereof. By Sir William Petty, Kt. F. R. S. reprinted from the edition printed at London in 1683.
III. Obſervations on the paſt growth and preſent ſtate of the city of London; reprinted from the edition printed at London in 1751; with a continuation of the tables to the end of the year 1757. By Corbyn Morris, Eſq; F. R. S.
IV. A comparative view of the diſeaſes and ages, and a table of the probabilities of life, for the laſt thirty years. By J. P. Eſq; F. R. S.

LONDON:
Printed for A. MILLAR in the *Strand*.
MDCCLIX.

Bills of Mortality



Wöchentliche Todesstatistiken

US Census Income Data

Explorative Datenanalyse

Daten

adult.data

```
1  39, State-gov, 77516, Bachelors, 13, Never-married, Adm-clerical, Not-in-family, White, Male, 2
2  50, Self-emp-not-inc, 83311, Bachelors, 13, Married-civ-spouse, Exec-managerial, Husband, White
3  38, Private, 215646, HS-grad, 9, Divorced, Handlers-cleaners, Not-in-family, White, Male, 0, 0,
4  53, Private, 234721, 11th, 7, Married-civ-spouse, Handlers-cleaners, Husband, Black, Male, 0, 0
5  28, Private, 338409, Bachelors, 13, Married-civ-spouse, Prof-specialty, Wife, Black, Female, 0,
6  37, Private, 284582, Masters, 14, Married-civ-spouse, Exec-managerial, Wife, White, Female, 0,
7  49, Private, 160187, 9th, 5, Married-spouse-absent, Other-service, Not-in-family, Black, Female
8  52, Self-emp-not-inc, 209642, HS-grad, 9, Married-civ-spouse, Exec-managerial, Husband, White,
9  31, Private, 45781, Masters, 14, Never-married, Prof-specialty, Not-in-family, White, Female, 1
10 42, Private, 159449, Bachelors, 13, Married-civ-spouse, Exec-managerial, Husband, White, Male,
11 37, Private, 280464, Some-college, 10, Married-civ-spouse, Exec-managerial, Husband, Black, Mal
12 30, State-gov, 141297, Bachelors, 13, Married-civ-spouse, Prof-specialty, Husband, Asian-Pac-Is
```

- Erste 12 Instanzen mit
- 15 Variablen

Daten

CensusData.R × rawData ×

32,561 observations of 15 variables

	id	employerKind	fnlwgt	degree	yearsOfEd	maritalStatus	occupation	relationshipRole
1	39	State-gov	77516	Bachelors	13	Never-married	Adm-clerical	Not-in-family
2	50	Self-emp-not-inc	83311	Bachelors	13	Married-civ-spouse	Exec-managerial	Husband
3	38	Private	215646	HS-grad	9	Divorced	Handlers-cleaners	Not-in-family
4	53	Private	234721	11th	7	Married-civ-spouse	Handlers-cleaners	Husband
5	28	Private	338409	Bachelors	13	Married-civ-spouse	Prof-specialty	Wife
6	37	Private	284582	Masters	14	Married-civ-spouse	Exec-managerial	Wife
7	49	Private	160187	9th	5	Married-spouse-absent	Other-service	Not-in-family
8	52	Self-emp-not-inc	209642	HS-grad	9	Married-civ-spouse	Exec-managerial	Husband
9	31	Private	45781	Masters	14	Never-married	Prof-specialty	Not-in-family
10	42	Private	159449	Bachelors	13	Married-civ-spouse	Exec-managerial	Husband
11	37	Private	280464	Some-college	10	Married-civ-spouse	Exec-managerial	Husband
12	30	State-gov	141297	Bachelors	13	Married-civ-spouse	Prof-specialty	Husband

- Erste 12 Instanzen als Data-Frame mit
- 8 von 15 Variablen

Daten

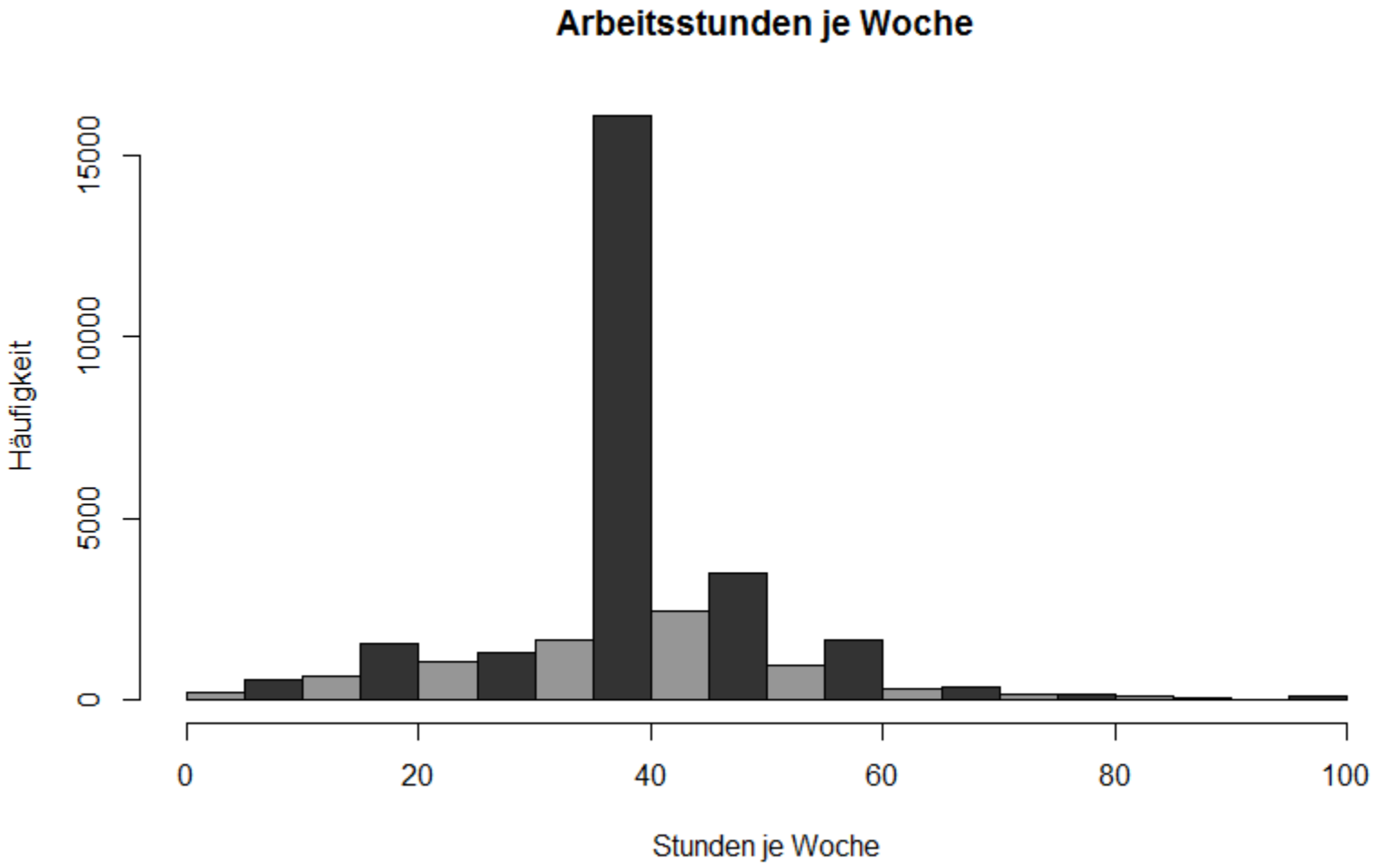
CensusData.R × | rawData ×

32,561 observations of 20 variables

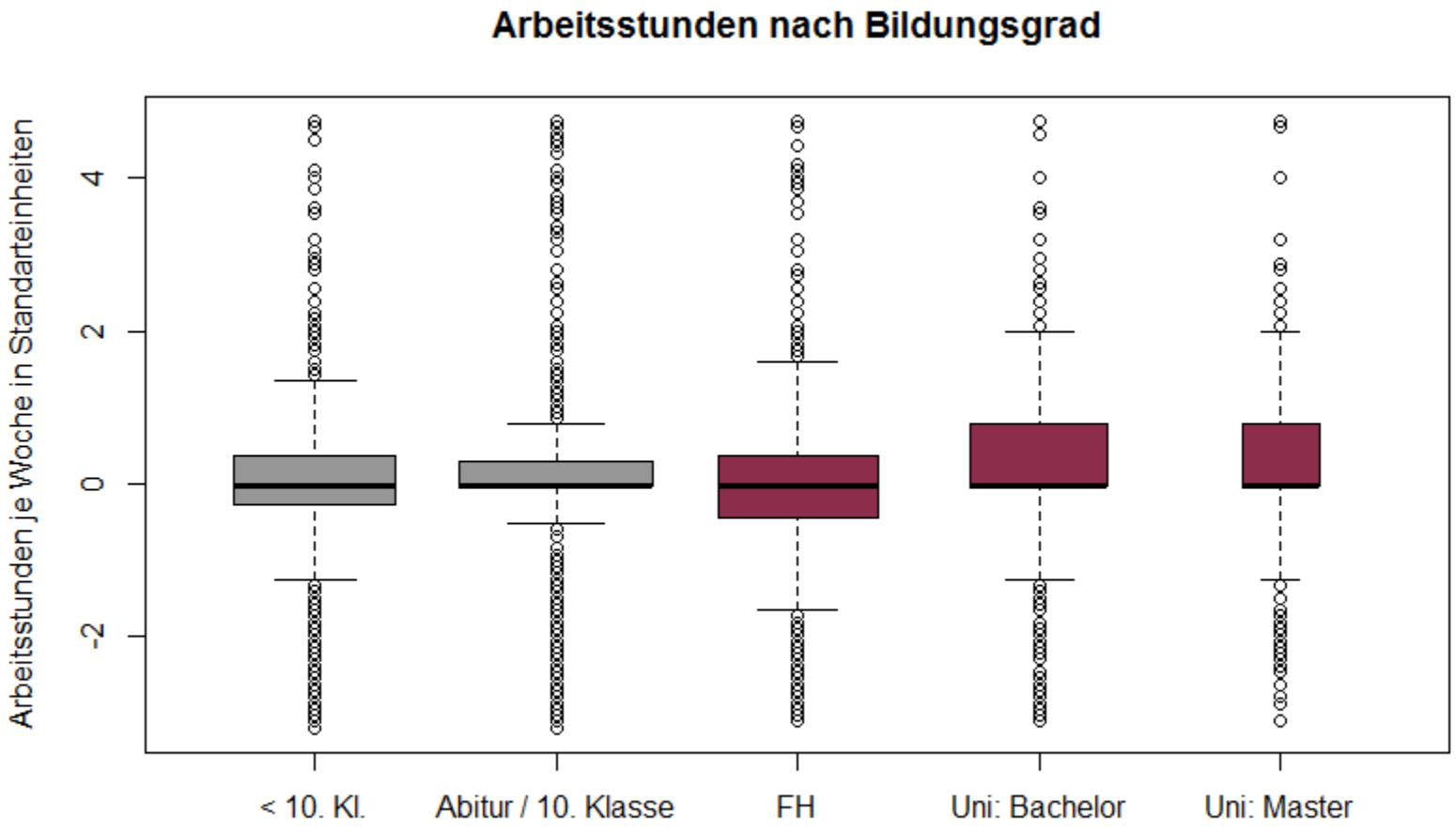
incomeGroup	academicLvl	incomeMoreThan50K	capitalDeviation	yearsOfEdStdUnits	workingHoursAWeekStdUnits
<=50K	Bachelor	TRUE	2.7562412	1.13472134	-0.03542890
<=50K	Bachelor	TRUE	-0.3071748	1.13472134	-2.22211900
<=50K	Highschool	TRUE	-0.3071748	-0.42005317	-0.03542890
<=50K	1	TRUE	-0.3071748	-1.19744043	-0.03542890
<=50K	Bachelor	TRUE	-0.3071748	1.13472134	-0.03542890
<=50K	Master	TRUE	-0.3071748	1.52341497	-0.03542890
<=50K	1	TRUE	-0.3071748	-1.97482769	-1.97915343
>50K	Highschool	FALSE	-0.3071748	-0.42005317	0.36951371
>50K	Master	FALSE	3.5009221	1.52341497	0.77445632
>50K	Bachelor	FALSE	3.1020895	1.13472134	-0.03542890
>50K	College	FALSE	-0.3071748	-0.03135955	3.20411198
>50K	Bachelor	FALSE	-0.3071748	1.13472134	-0.03542890

5 Sekundärvariablen

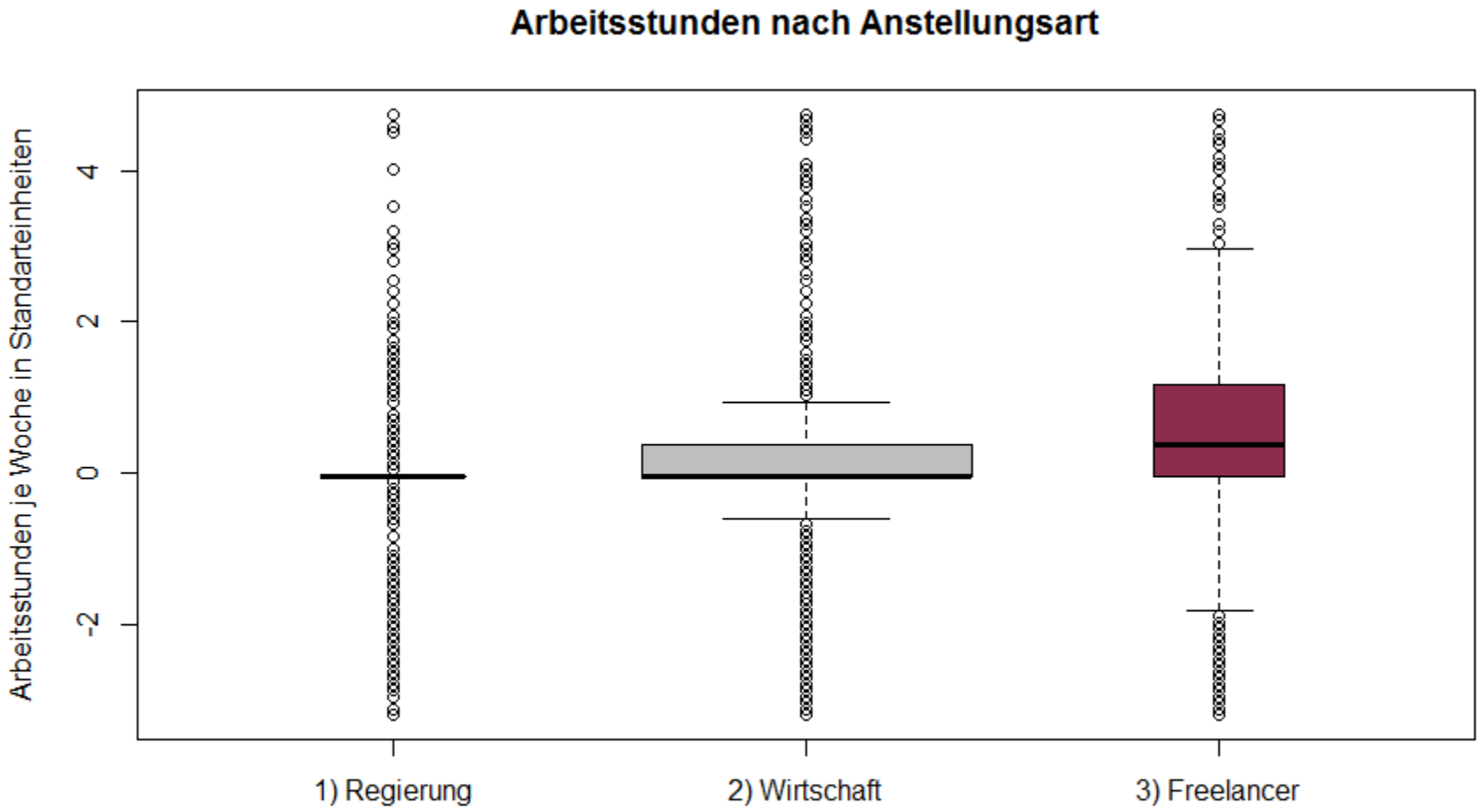
Visualisierungen



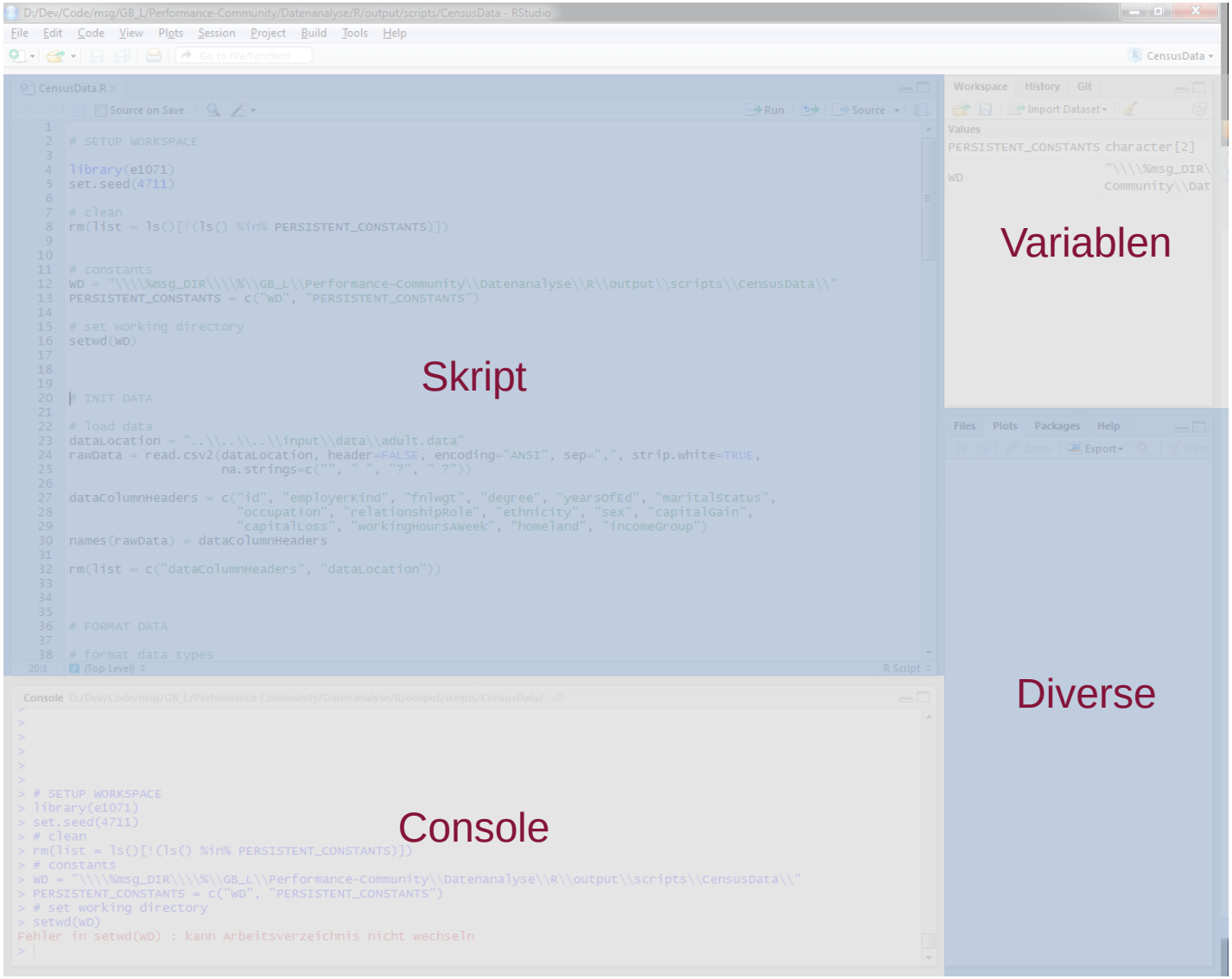
Visualisierungen



Visualisierungen



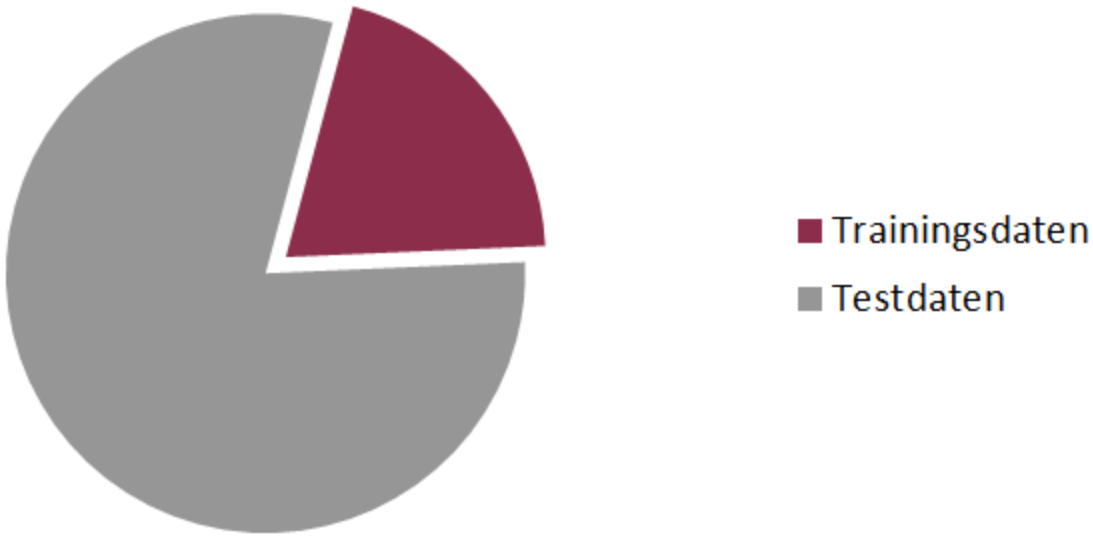
RStudio



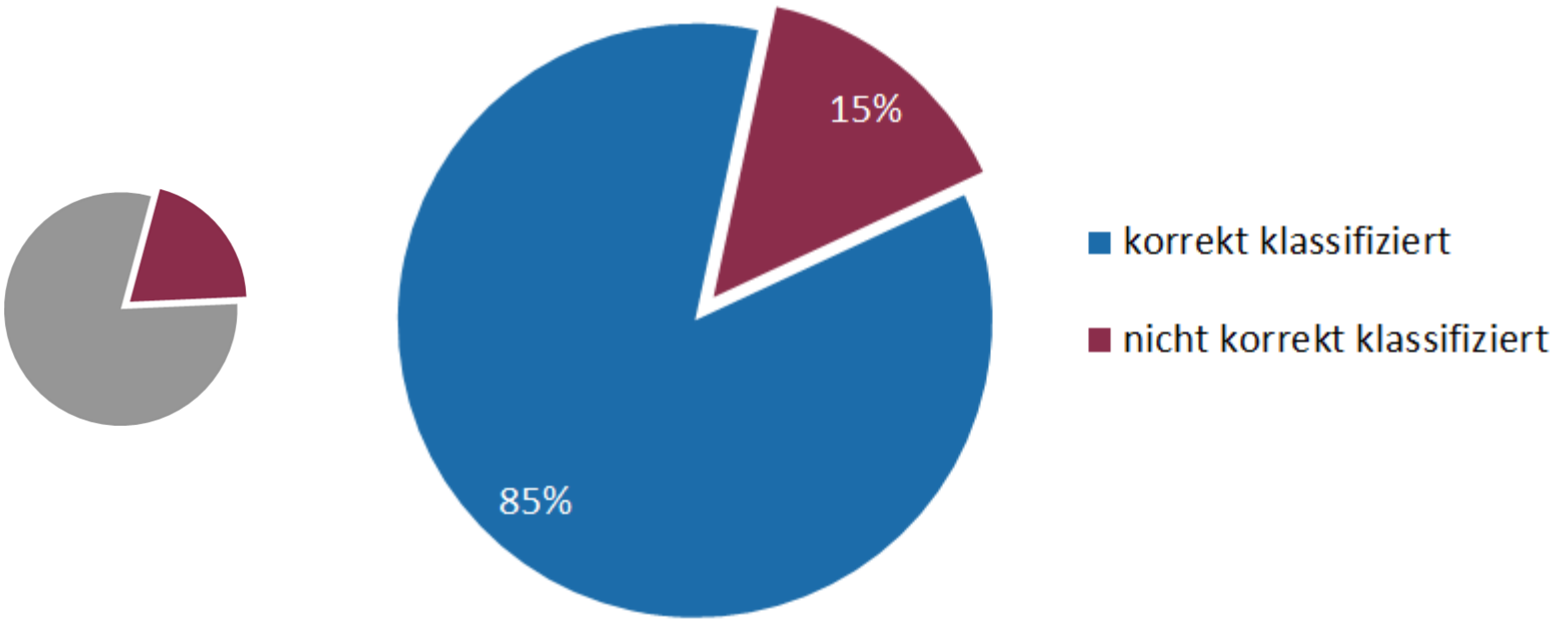
US Census Income Data

Klassifikation

Klassifizierung



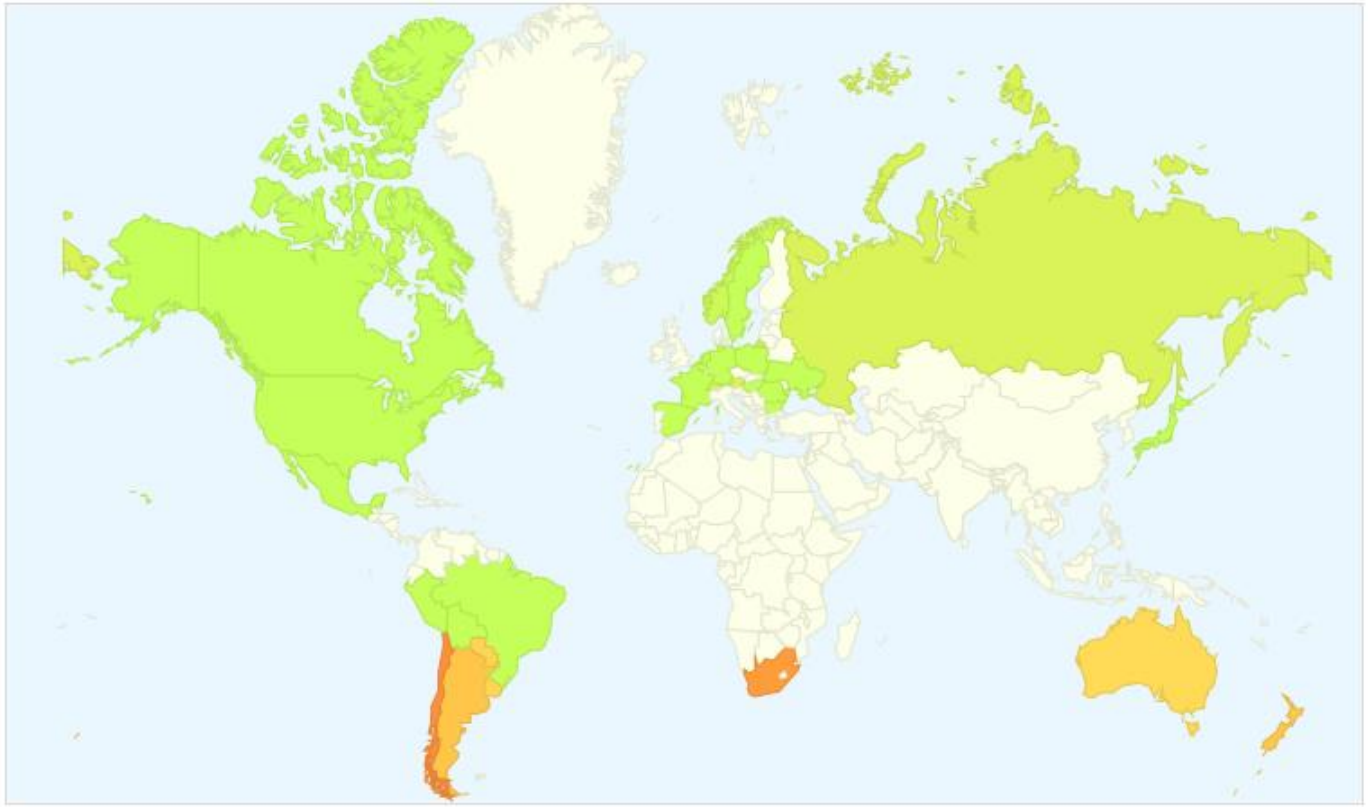
Klassifizierung



Anwendung

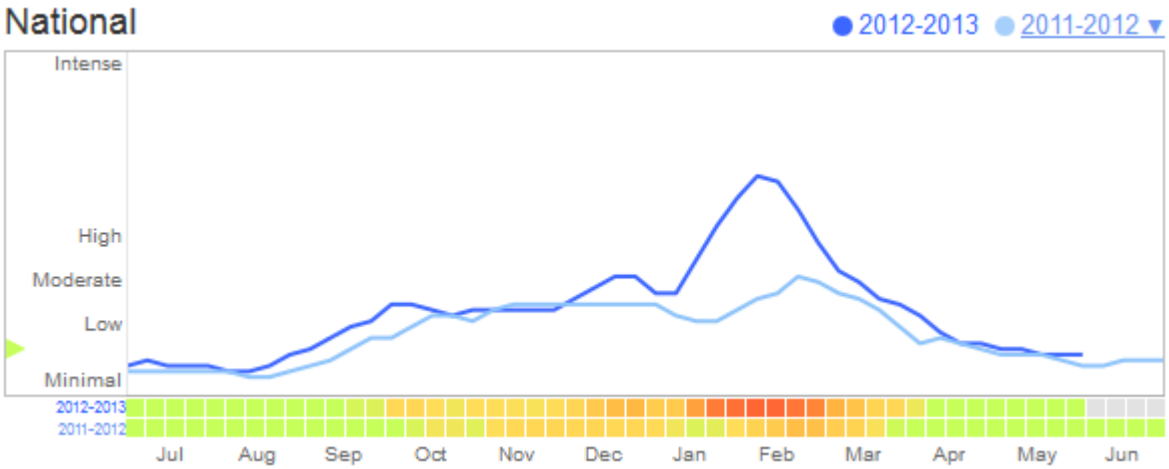
Anwendung

Google



Anwendung

Google



Datenquellen

UCI Machine Learning Repository
archive.ics.uci.edu/ml

Datenquellen

kaggle.com

Ressourcen

github
SOCIAL CODING

LearningR

Training Dataset for R Beginners

Last updated 3 minutes ago

AddHealth-Data-Analysis

The analysis of biases and influencer in attendance of religious services

Last updated a month ago

github.com/danielschulz/LearningR

Vielen Dank für Ihre Aufmerksamkeit

.consulting .solutions .partnership