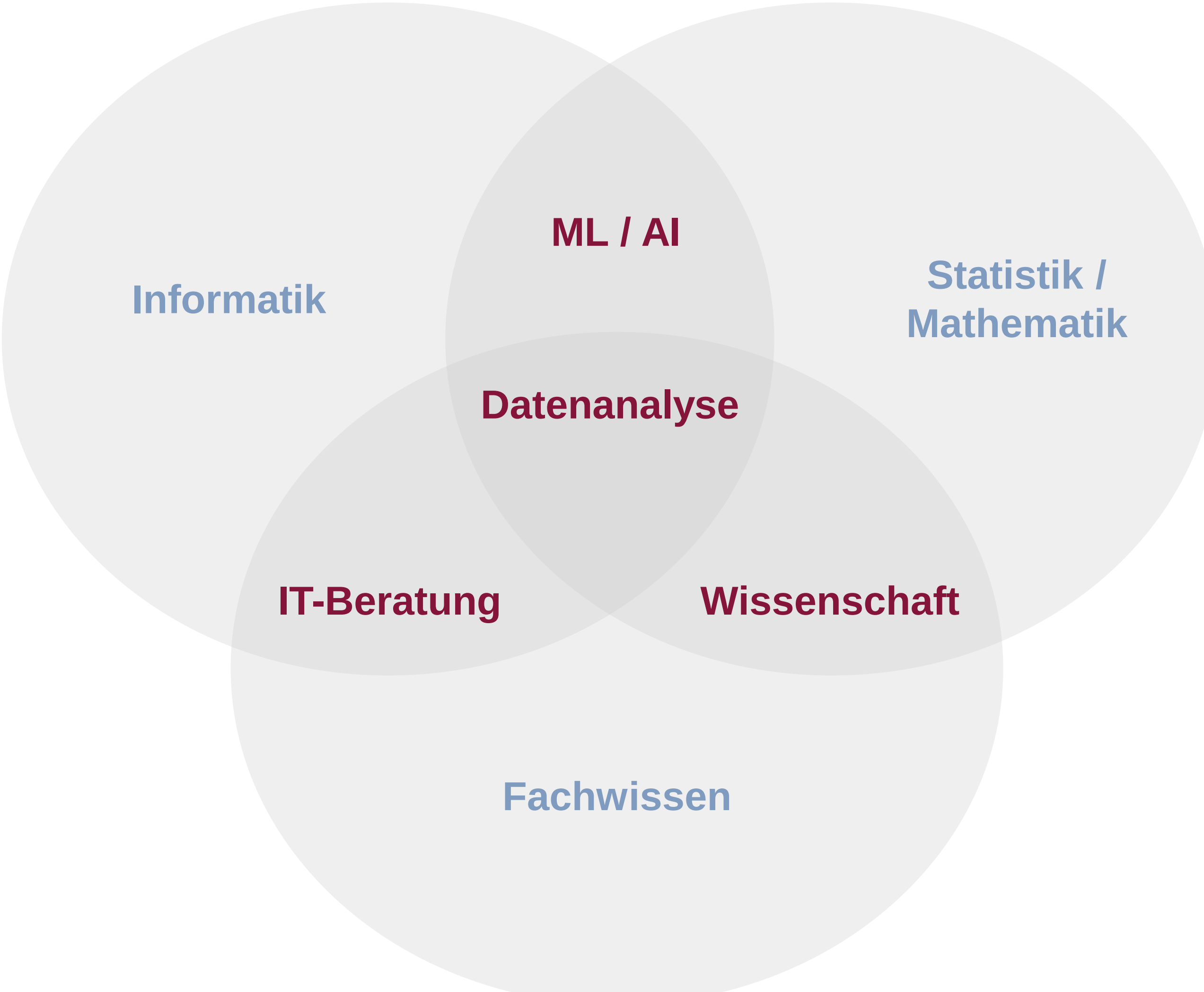
Datenanalyse

Automotive Communications Financial Services Government Insurance Life Science & Healthcare Travel & Logistics Utilities Automotive Communications Financial Services Government
Insurance Life Science & Healthcare Travel & Logistics Utilities Automotive Communications Financial Services Government Insurance Life Science & Healthcare Travel & Logistics
Utilities Automotive Communications Financial Services Government Insurance Life Science & Healthcare Travel & Logistics Utilities Automotive Communications Financial Services
Government Insurance Life Science & Healthcare Travel & Logistics Utilities Automotive Communications Financial Services Government Insurance Life Science & Healthcare Travel
& Logistics Utilities Automotive Communications Financial Services Government Insurance Life Science & Healthcare Travel & Logistics Utilities Automotive Communications
Financial Services Government Insurance Life Science & Healthcare Travel & Logistics Utilities Automotive Communications Financial Services Government Insurance Life
Science & Healthcare Travel & Logistics Utilities Automotive Communications Financial Services Government Insurance Life Science & Healthcare Travel & Logistics Utilities
Automotive Communications Financial Services Government Insurance Life Science & Healthcare Travel & Logistics Utilities Automotive Communications Financial Services
Government Insurance Life Science & Healthcare **Travel & Logistics** Utilities Automotive Communications Financial Services Insurance Life Science & Healthcare
Travel & Logistics Utilities Automotive Communications Financial Services Government Insurance Life Science & Healthcare Travel & Logistics Utilities Automotive
Communications Financial Services Government Insurance Life Science & Healthcare Travel & Logistics Utilities Automotive Communications Financial Services Government
Insurance Life Science & Healthcare Travel & Logistics Utilities Automotive Communications Financial Services Government Insurance Life Science & Healthcare Travel & Logistics
Utilities Automotive Communications Financial Services Government Insurance Life Science & Healthcare Travel & Logistics Utilities Automotive Communications Financial
Services Government Insurance Life Science & Healthcare Travel & Logistics Utilities Automotive Communications Financial Services Government Insurance Life Science &
Healthcare Travel & Logistics Utilities Automotive Communications Financial Services Government Insurance Life Science & Healthcare Travel & Logistics Utilities Automotive
Communications Financial Services Government Insurance Life Science & Healthcare Travel & Logistics Utilities Automotive Communications Financial Services Government Insurance
Life Science & Healthcare Travel & Logistics Utilities Automotive Communications Financial Services Government Insurance Life Science & Healthcare Travel & Logistics Automotive Utilities

.consulting .solutions .partnership

Datenanalyse nach Drew Conway



US Census Income Data

Explorative Datenanalyse

Daten

adult.data

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

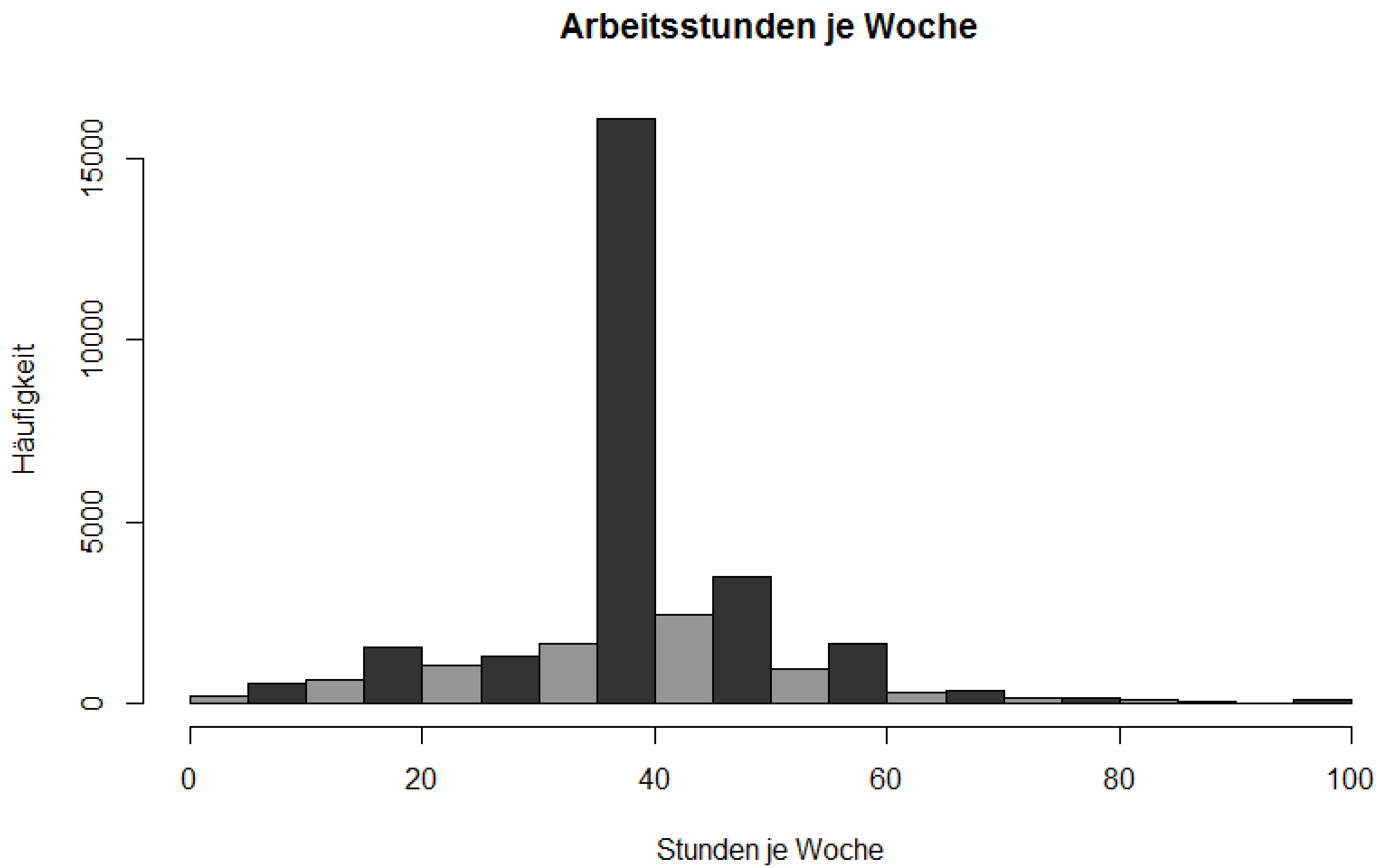
CensusData.R | rawData

32,561 observations of 20 variables

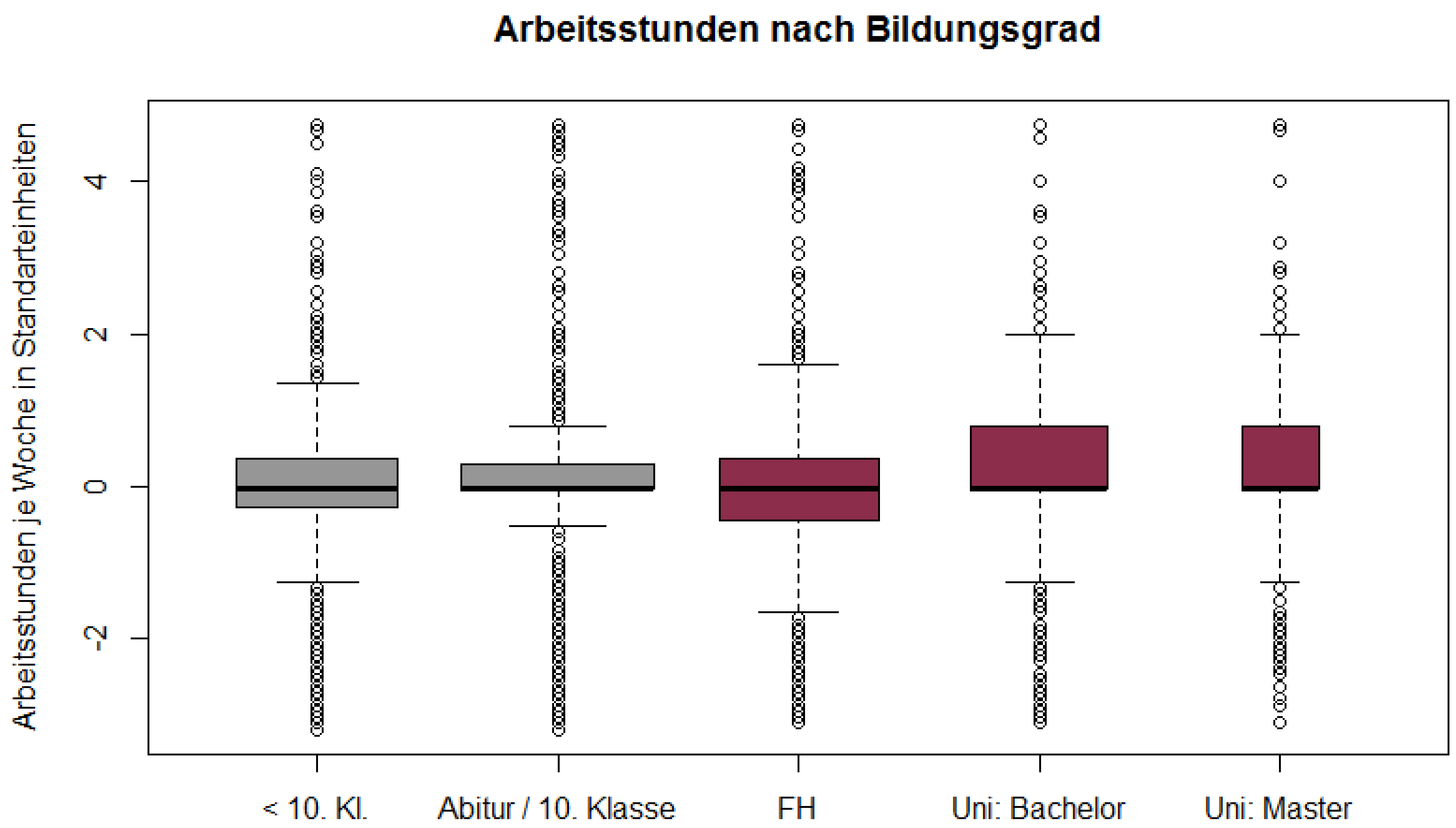
incomeGroup	academicLvl	incomeMoreThan50K	capitalDeviation	yearsOfEdStdUnits	workingHoursAWeekStdUnits
<=50K	Bachelor	TRUE	2.7562412	1.13472134	-0.03542890
<=50K	Bachelor	TRUE	-0.3071748	1.13472134	-2.22211900
<=50K	Highschool	TRUE	-0.3071748	-0.42005317	-0.03542890
<=50K	1	TRUE	-0.3071748	-1.19744043	-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

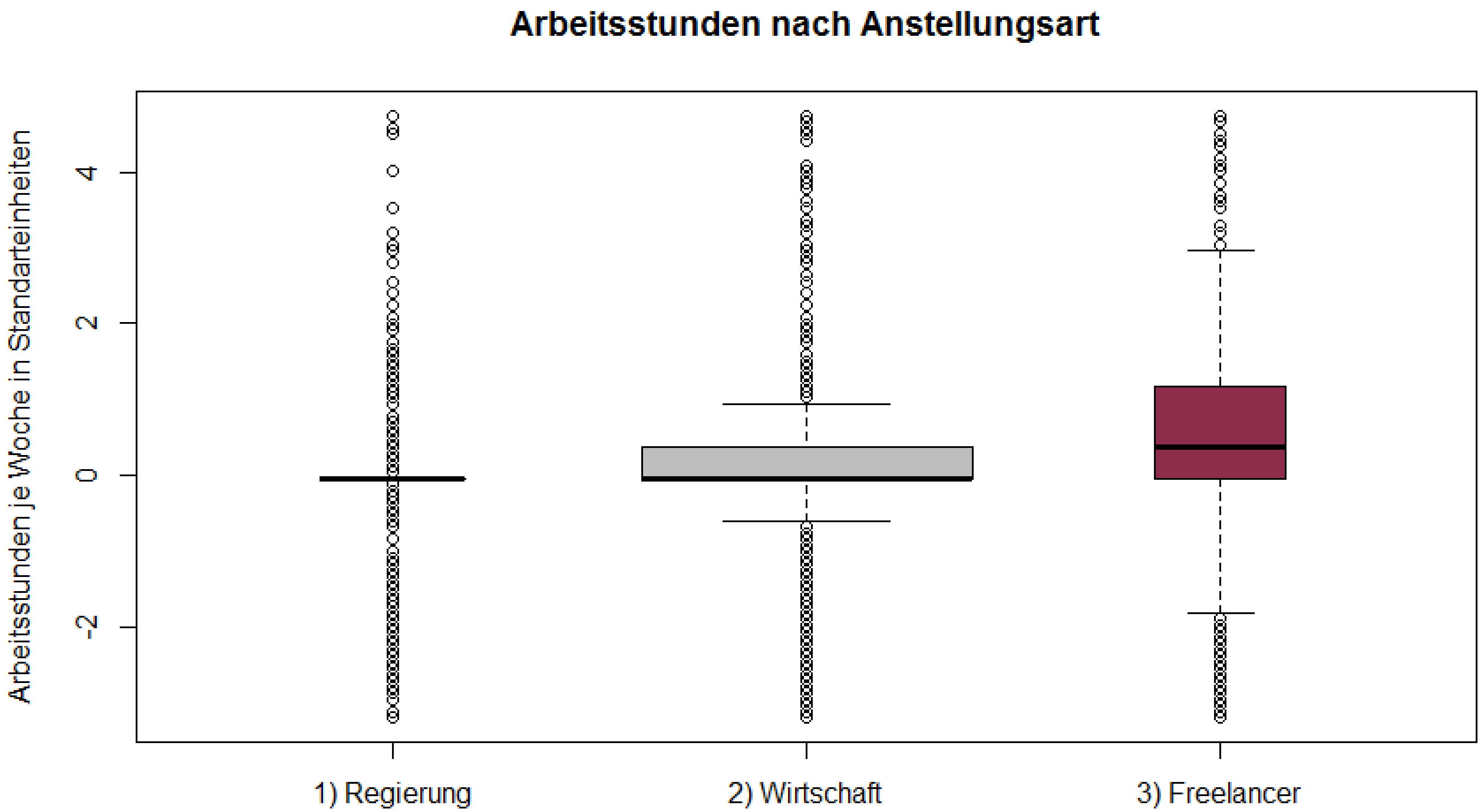
Visualisierungen



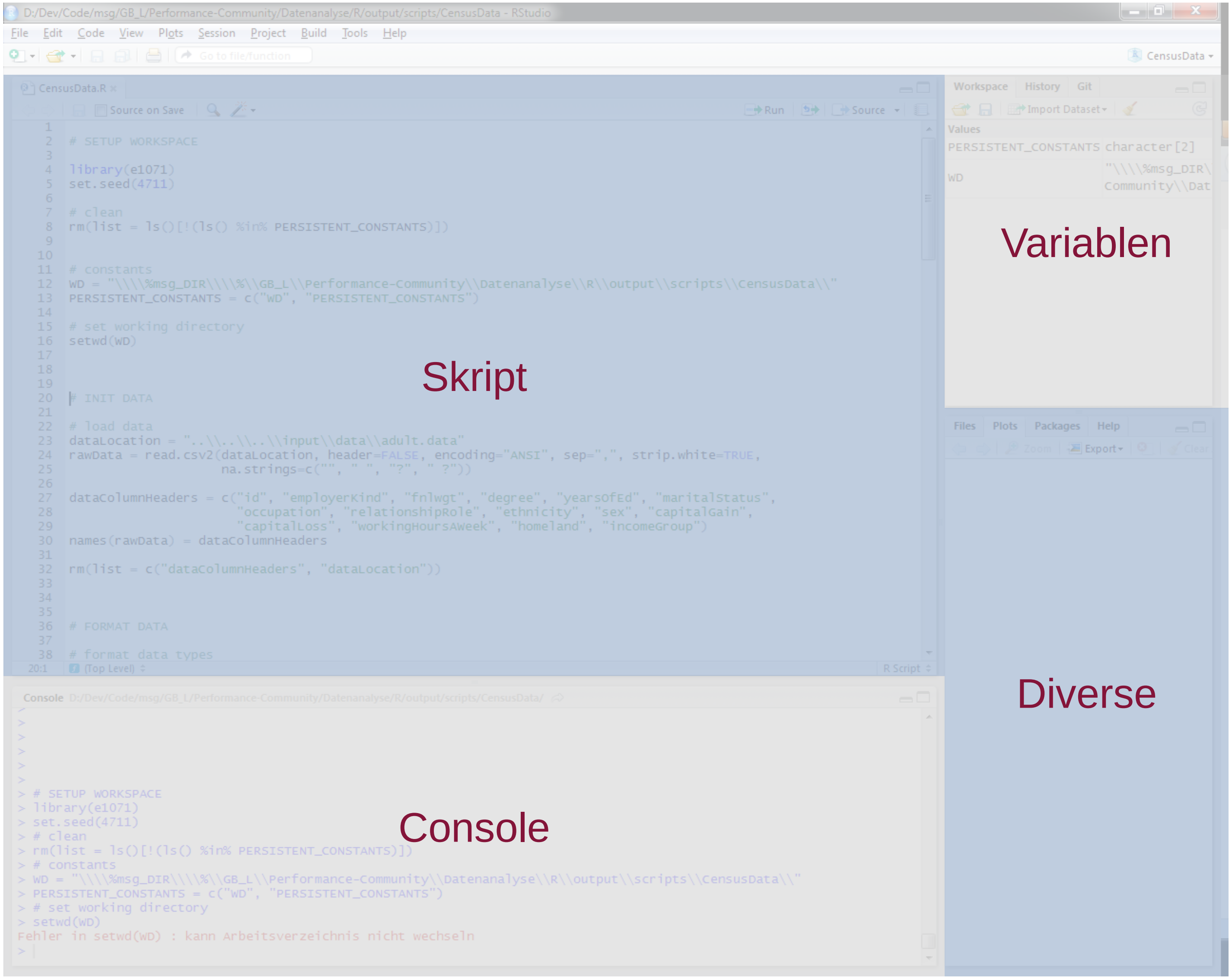
Visualisierungen



Visualisierungen



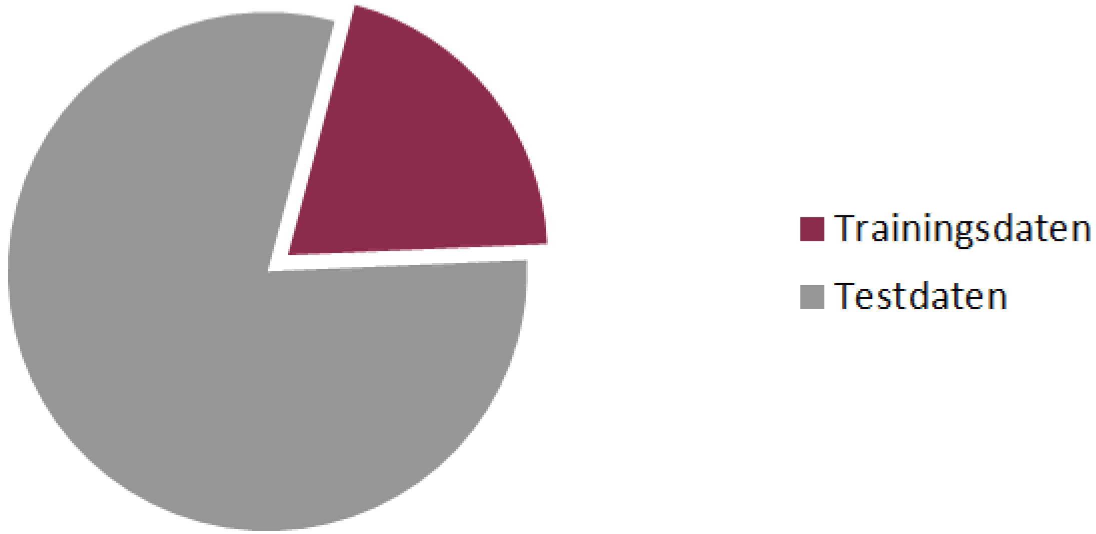
RStudio



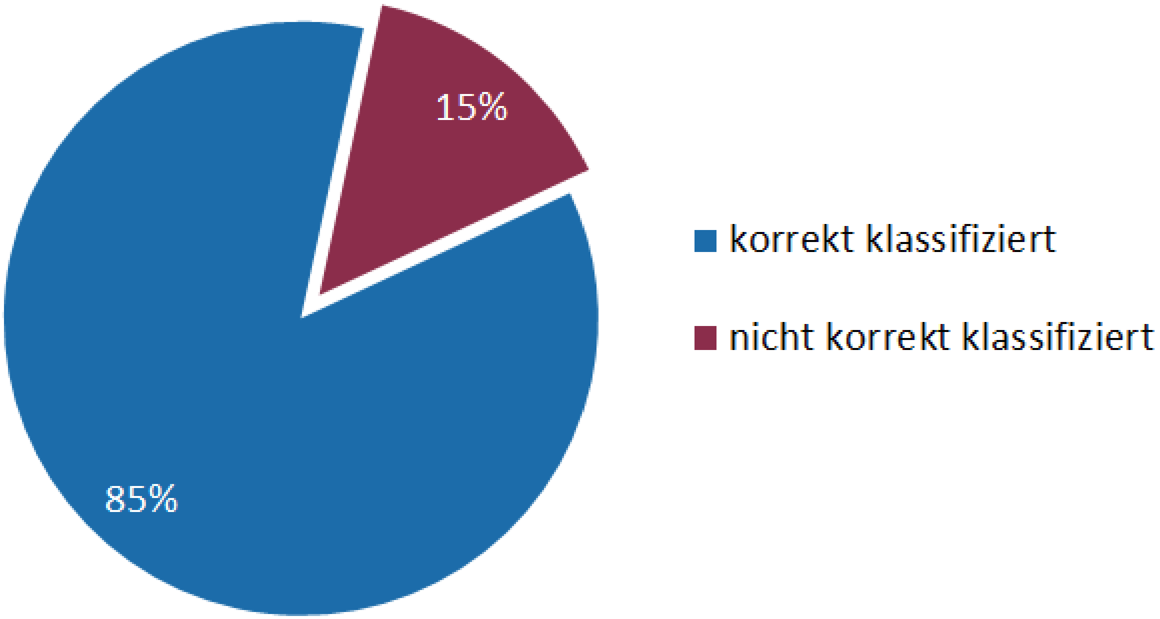
US Census Income Data

Klassifikation

Klassifizierung



Klassifizierung



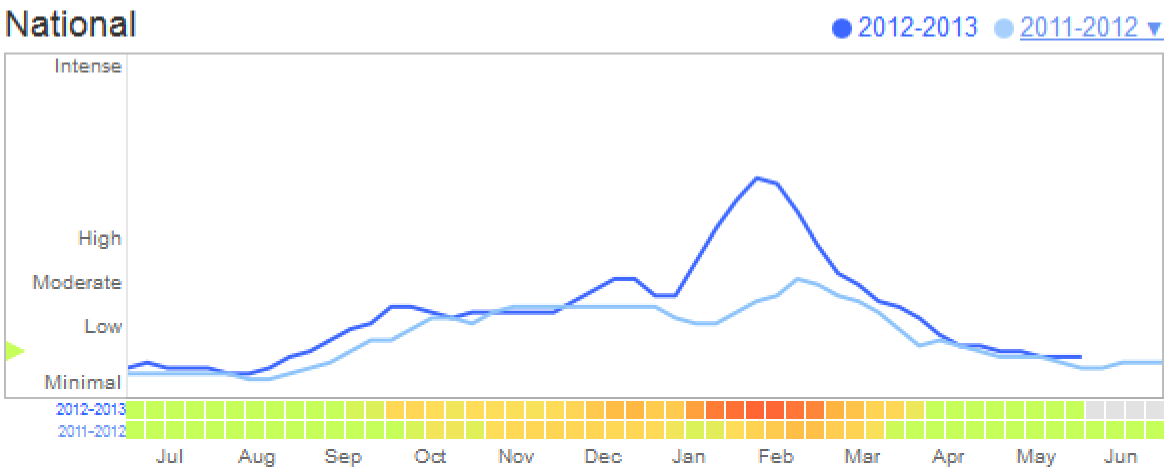
Anwendungsbeispiele

Anwendungsbeispiele

Google

Anwendungsbeispiele

Google



Resourcen

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

Datenquellen

UNIVERSITY of CALIFORNIA | IRVINE

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

Datenquellen

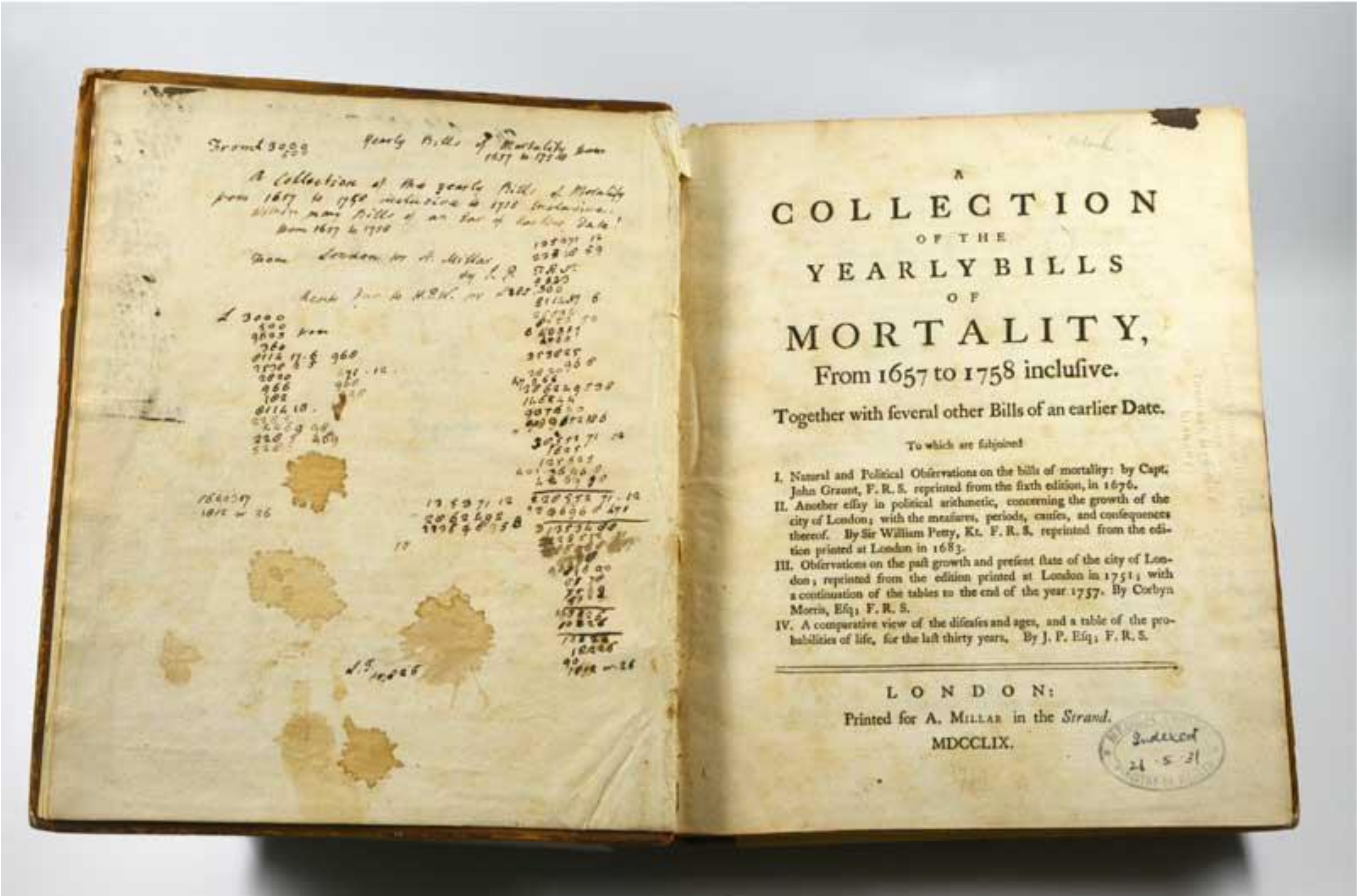
kaggle

kaggle.com

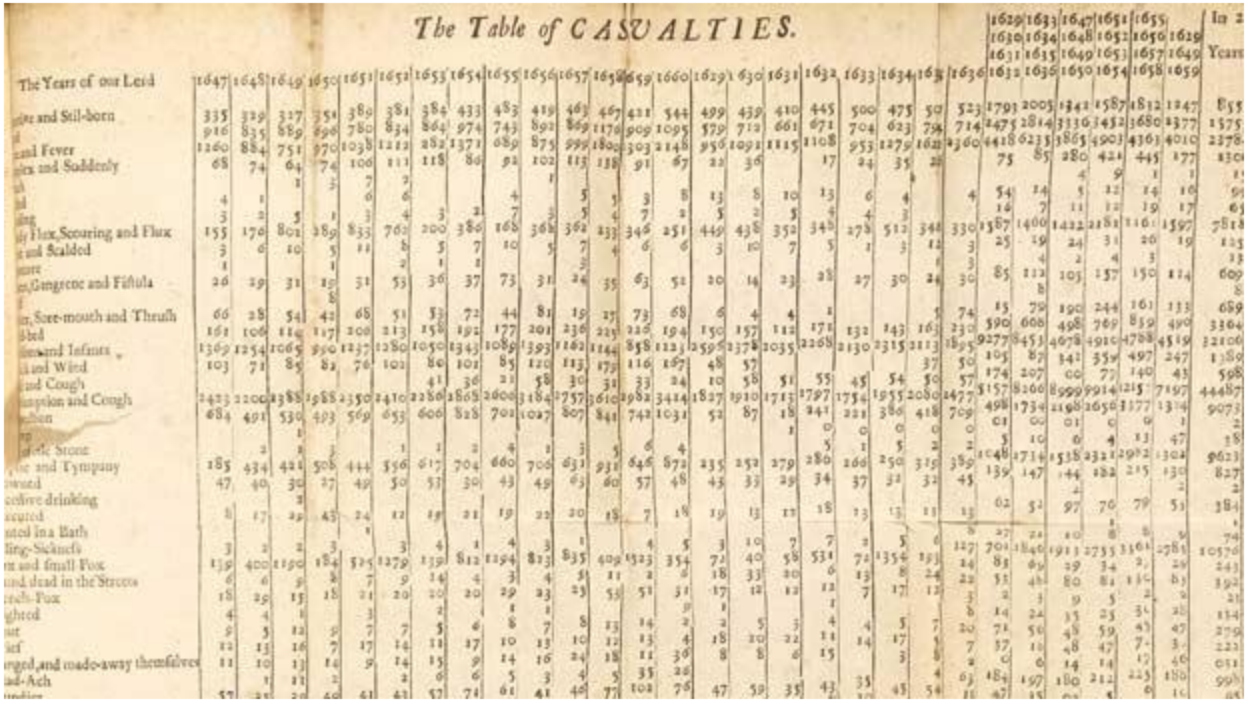
Datenquellen

data.gov

Bills of Mortality



Bills of Mortality



Wöchentliche Todesstatistiken

Resumée

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

Vielen Dank für Ihre Aufmerksamkeit

.consulting .solutions .partnership

.msg systems