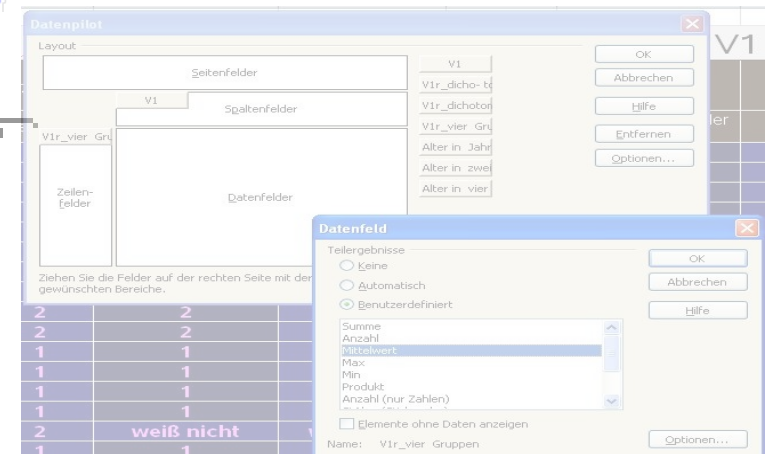
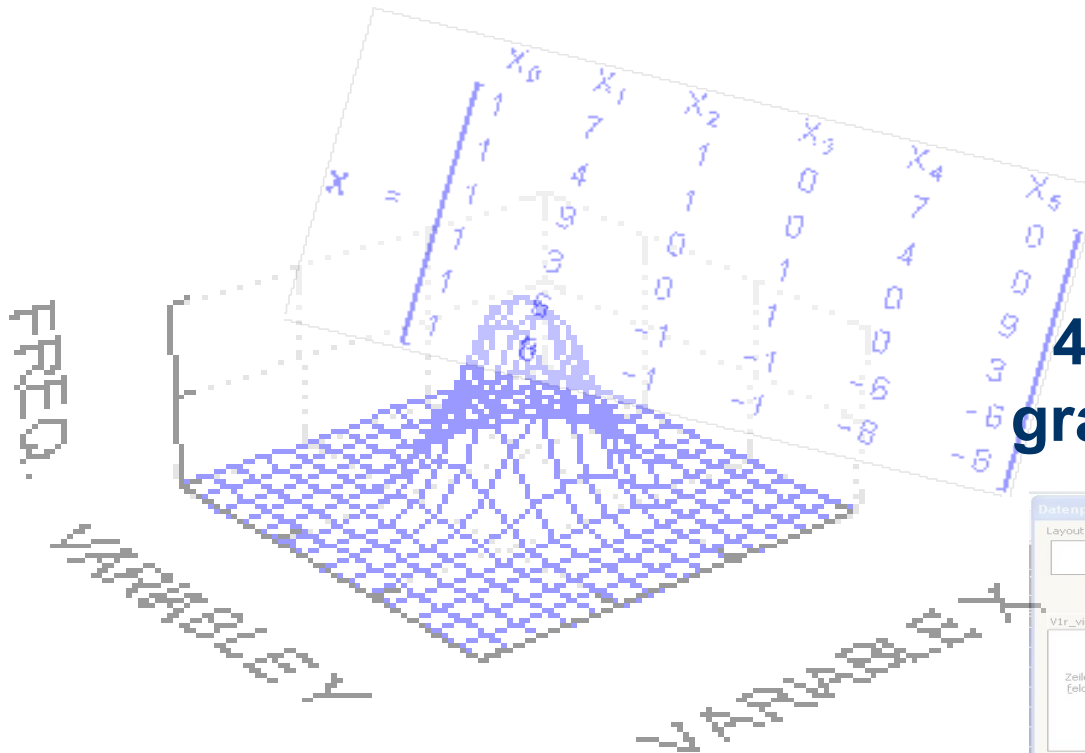


Statistik IV

4. Auswertungen und graphische Darstellung



Terminplanung

<i>Nr.</i>	<i>Termin</i>	<i>Inhalt</i>
1	14.04.09	Einführung • Organisatorisches und Scheinvoraussetzungen • Statistik mit Softwareunterstützung?
2	28.04.09	Grundlagen Stata • Aufbau des Programms • Umgang mit Daten, Dokumentation Datenaufbereitung für statistische Auswertungen I
3	12.05.09	Datenaufbereitung für statistische Auswertungen II
4	26.05.09	Deskriptive Statistik und graphische Darstellung
5	09.06.09	Zusammenhangsanalyse I
6	23.06.09	Zusammenhangsanalyse II
7	07.07.09	Zusammenhangsanalyse III Probeklausur Besprechung der Probeklausur
8	14.07.09	Klausur (90 Minuten)

Übersicht

- **Hausaufgabe 2**
- **Auswertungen und Gewichtung**
- **graphische Darstellung**

Hausaufgabe 2

Datenanalyse und Gewichte

Disproportionale Stichproben und Designgewichtung

Im ALLBUS werden überrepräsentativ viele Fälle aus Ostdeutschland erhoben (oversampling). Dies geschieht, um zu gewährleisten, dass auch im Hinblick auf seltene Ausprägungen in den neuen Bundesländern komplexe Aussagen (insb. Zusammenhänge) möglich sind.

Um einen solchen "Stichprobenfehler", der bereits vor der Stichprobenziehung festgelegt wurde, auszugleichen, werden Gewichtungsvariablen verwendet.

Vergleich

```
. tab v3
```

ERHEBUNGSGEBIET <WOHNGBIET>: WEST - OST	Freq.	Percent	Cum.
ALTE BUNDESLAENDER	2,392	68.95	68.95
NEUE BUNDESLAENDER	1,077	31.05	100.00
Total	3,469	100.00	

```
. tab v3 [iweight=v792]
```

ERHEBUNGSGEBIET <WOHNGBIET>: WEST - OST	Freq.	Percent	Cum.
ALTE BUNDESLAENDER	2,830.3288	81.59	81.59
NEUE BUNDESLAENDER	638.671188	18.41	100.00
Total	3,469	100.00	

Gewichte in STATA

- **[pweight=VARIABLE]**
 - „Probabilityweight“: Die GewichtungsvARIABLE enthält die Auswahlwahrscheinlichkeit der Fälle. Im Fall des ALLBUS liegt eine solche GewichtungsvARIABLE (v792) im Datensatz vor (im **tab** Befehl erfüllt **iweight** die Funktion von **pweight**)
 - **mean v3 [pweight=v792]**
 - **tab v3 [iweight=v792]**

Alternative Surveydatensatz

- Eine weitere Möglichkeit besteht darin, den Datensatz für Stata als Surveydatensatz zu definieren, dabei eine Gewichtungvariable anzugeben und Kommandos speziell für Surveydatensätze zu verwenden, indem jeweils dem Befehl ein **Präfix** vorangestellt wird.
 - `svyset [pweight=v792]`
 - **svy:** `tab v3`

Änderung der Darstellung mit **svy:**

```
svy: tab v3
```

```
(running tabulate on estimation sample)
```

```
Number of strata   =           1
Number of PSUs     =        3469
```

```
Number of obs      =        3469
Population size     =        3469
Design df          =        3468
```

```
ERHEBUNGS
GEBIET
<WOHN GEBI
ET>: WEST
- OST
```

```
proportions
```

```
ALTE BUN      .8159
NEUE BUN      .1841
```

```
Total        1
```

```
Key: proportions = cell proportions
```

Aufgabe 8

- Schreiben Sie den Do-File der Hausaufgabe so um, dass alle Ergebnisse korrekt mit v792 gewichtet werden.

Graphiken

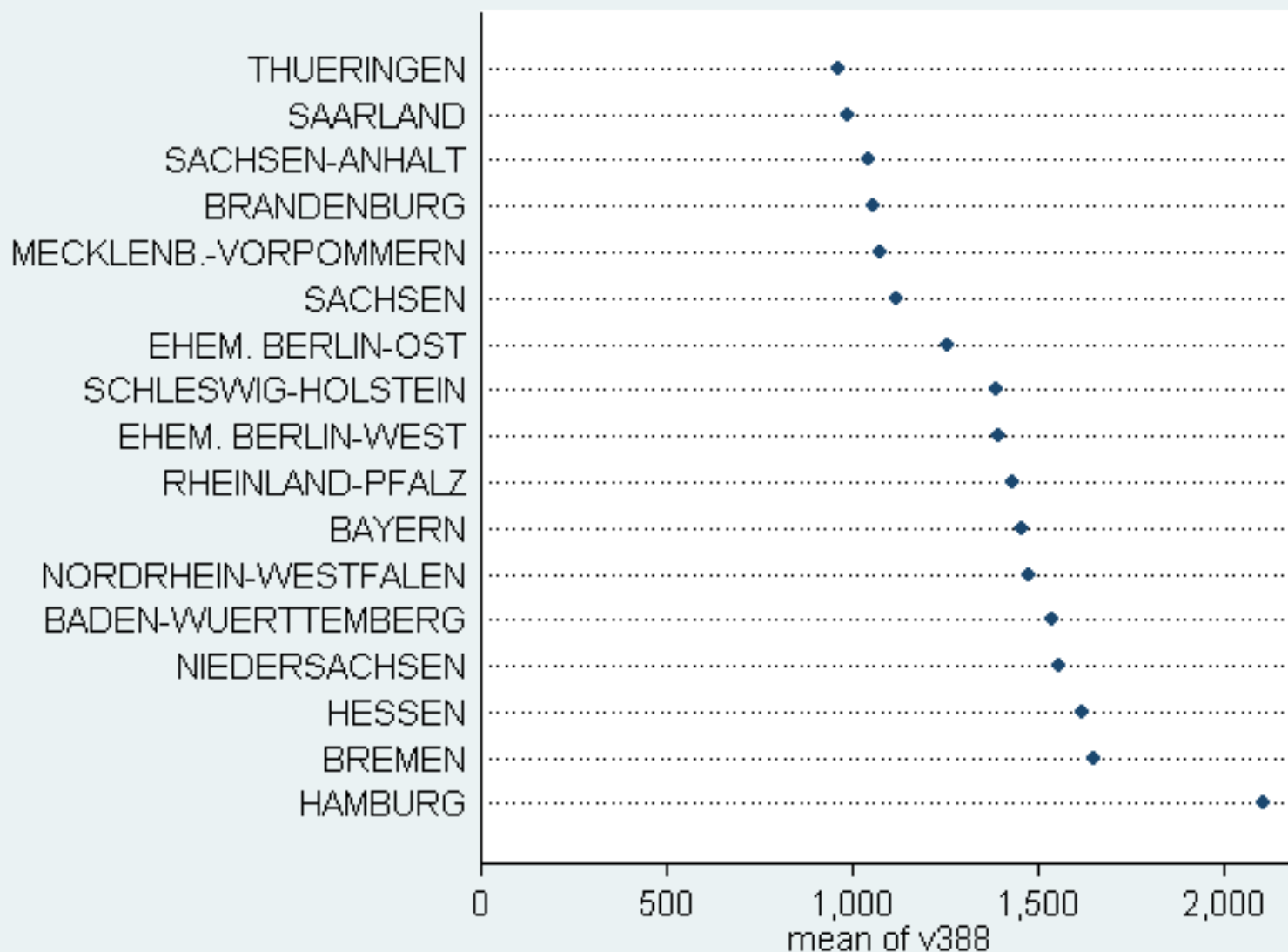
Graphiken und Skalenniveau

Skalenniveau	Ausgewählte Grafiken	Befehl
Kategorial	Balkendiagramm (Bar chart) Kreisdiagramm (Pie chart)	<i>histogram [variable], discrete</i> <i>graph pie, over(variable)</i>
Metrisch	Histogramm (Histogram) Kern-Dichte-Schätzer (kernel density estimation), Box-Plot (box and whisker plot)	<i>histogram [variable]</i> <i>kdensity [variable]</i> <i>graph box [variable]</i>

Dot und Unterbefehle

In Stata ist es über Unterbefehle und Optionen möglich fast alles an einer Grafik den eigenen Ansprüchen anzupassen, d.h. auch die Hilfedateien zu den Grafik-Befehlen sind recht umfangreich.

- `graph dot (mean) v388 [pweight=v792], over (v3)`
- `graph dot (mean) v388 [pweight=v792], over (v798)`
- `graph dot (mean) v388 [pweight=v792], over (v798, sort ((mean) v388))`
- `seperate v388, by (v151)`
 - `graph dot (mean) v3881 v3882 [pweight=v792], over (v798, sort ((mean) v388))`



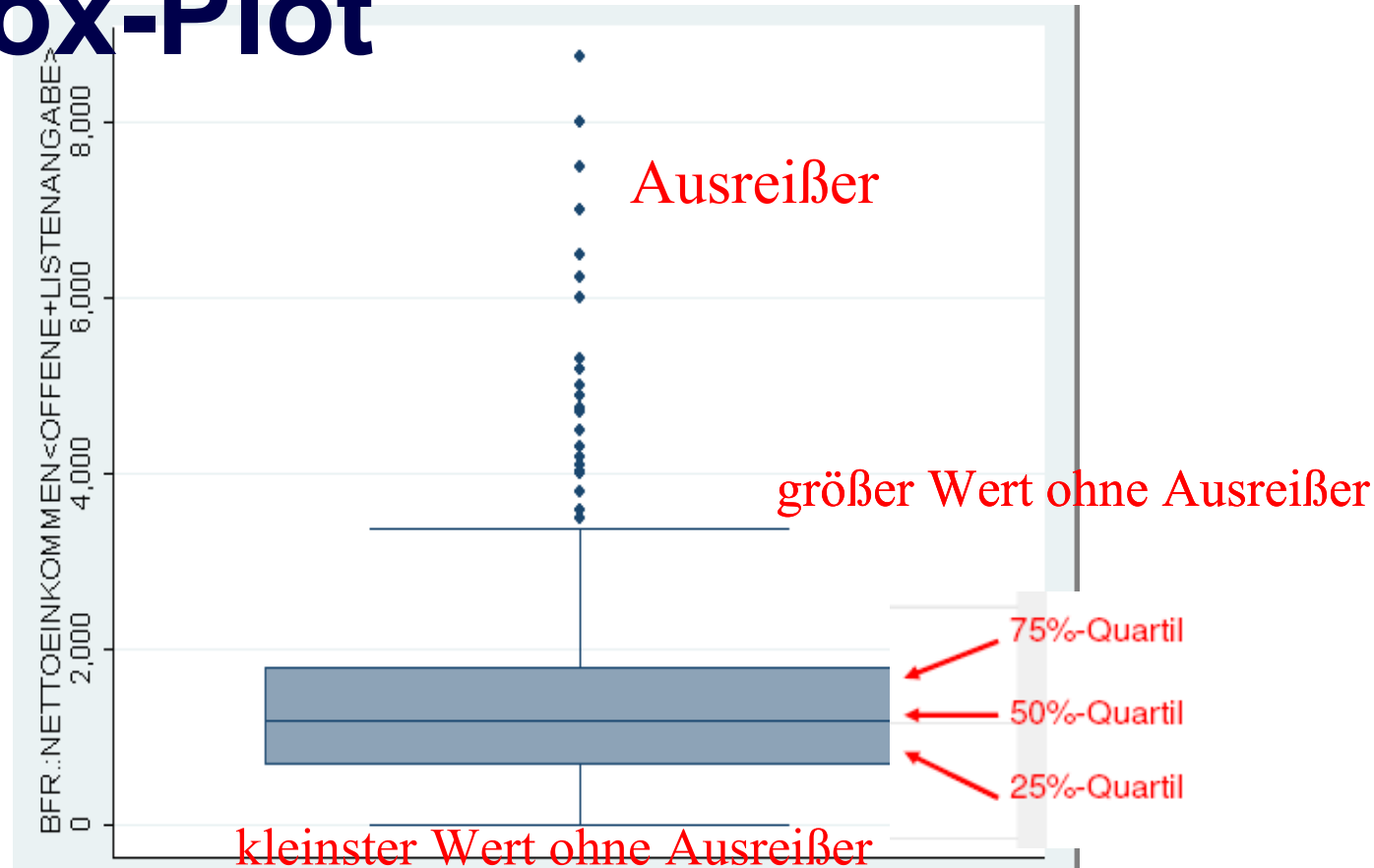
Säulendiagramme

- `graph bar v388 [pweight=v792], over(v3)`

Für Anteile:

- `tab v144, gen(nazi)`
- `graph bar nazi*, per`

Box-Plot

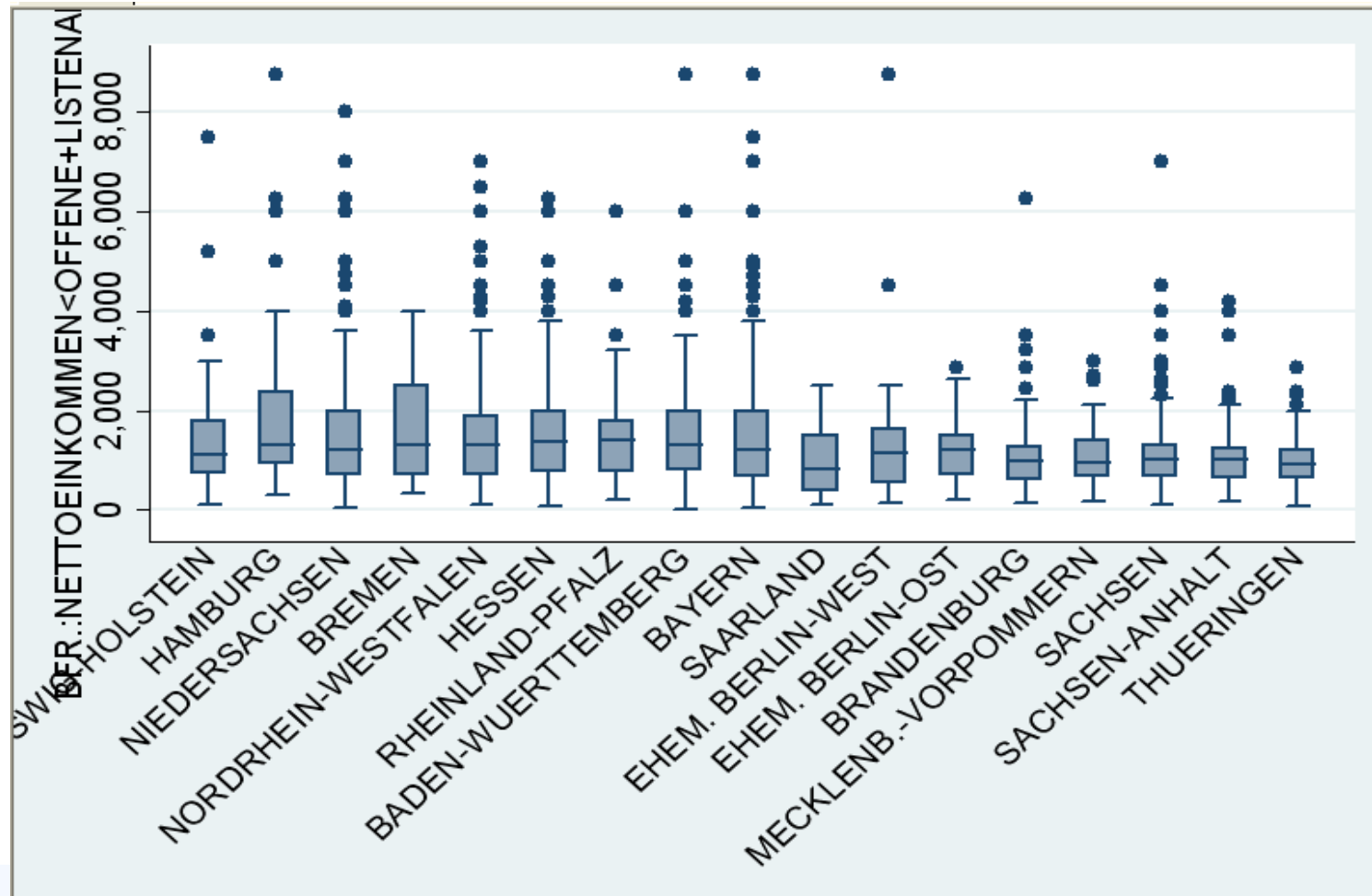


Maximaler Abstand der Striche für größten und kleinsten Wert von der Box: $1,5 \times \text{Boxlänge}$, sonst: als Ausreißer dargestellt.

Box-Plot

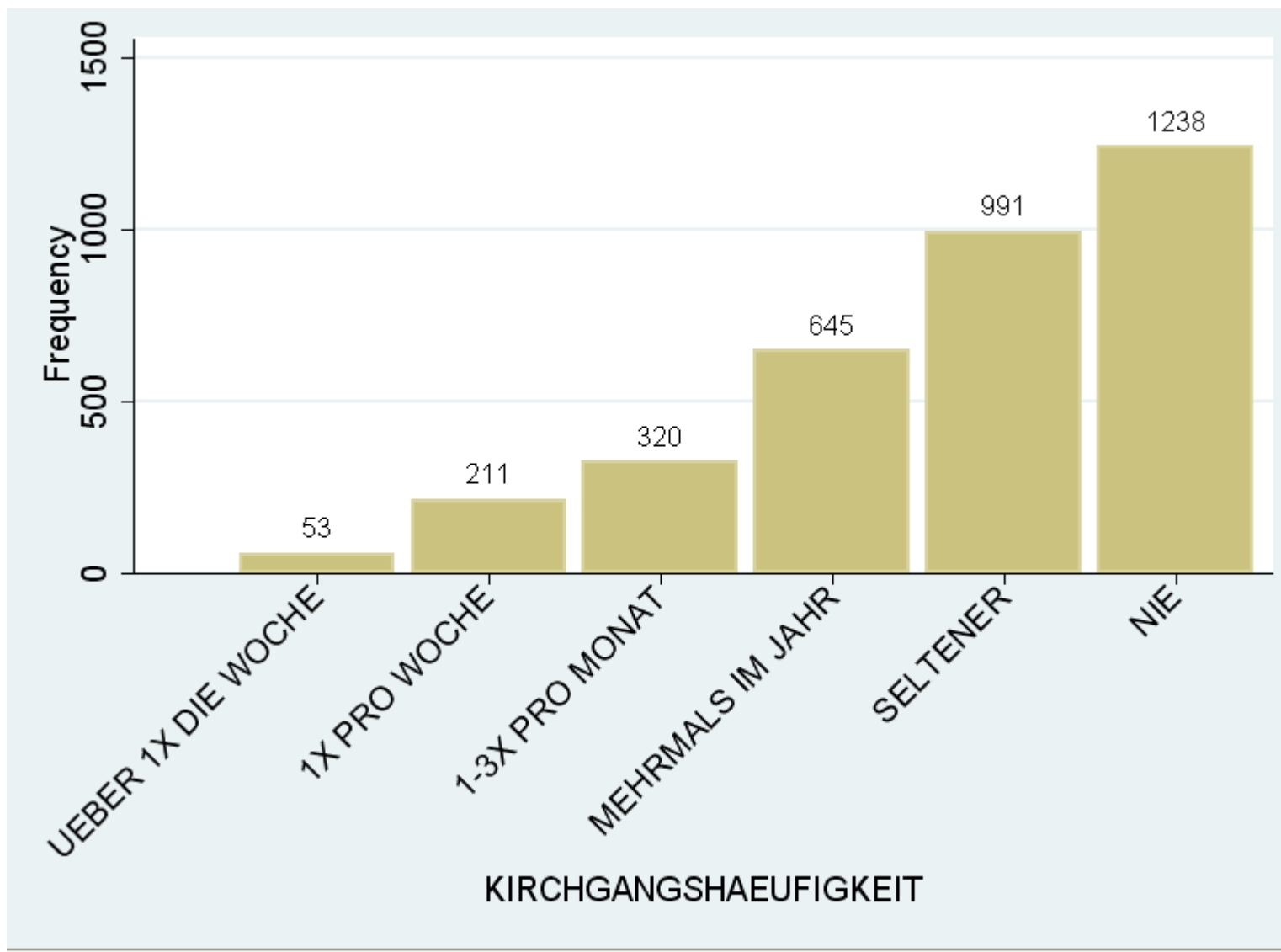
- `graph box v388 [pweight=v792]`
- `graph box v388 [pweight=v792], over(v3)`
- `graph box v388 [pweight=v792], over(v798)`
- `graph box v388 [pweight=v792], over(v798,
label(angle(45)))`

Box-Plots



Histogramme

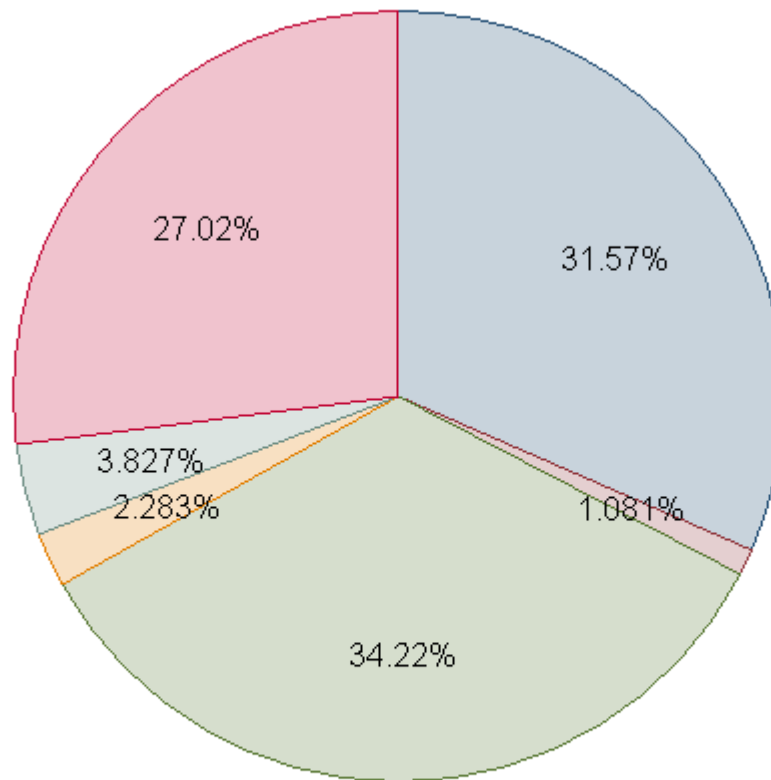
- Histogramme lassen keine Gewichtung zu!
 - histogram v526, freq
 - histogram v526, discrete freq
 - histogram v526, discrete percent
 - histogram v526, discrete percent barwidth(0.9)
 - histogram v526, discrete percent barwidth(0.9) addlabels
 - histogram v526, discrete freq barwidth(0.9) addlabels xlabel (1/6, valuelabel angle(45))



Kreisdiagramme

- `graph pie [pweight=v792], over(v524)`
- `graph pie [pweight=v792], over(v524) title(Konfession)`
- `graph pie [pweight=v792], over(v524) title(Konfession)
plabel(_all percent)`
- `graph pie [pweight=v792], over(v524) title(Konfession)
plabel(_all percent) intensity(*0.3)`

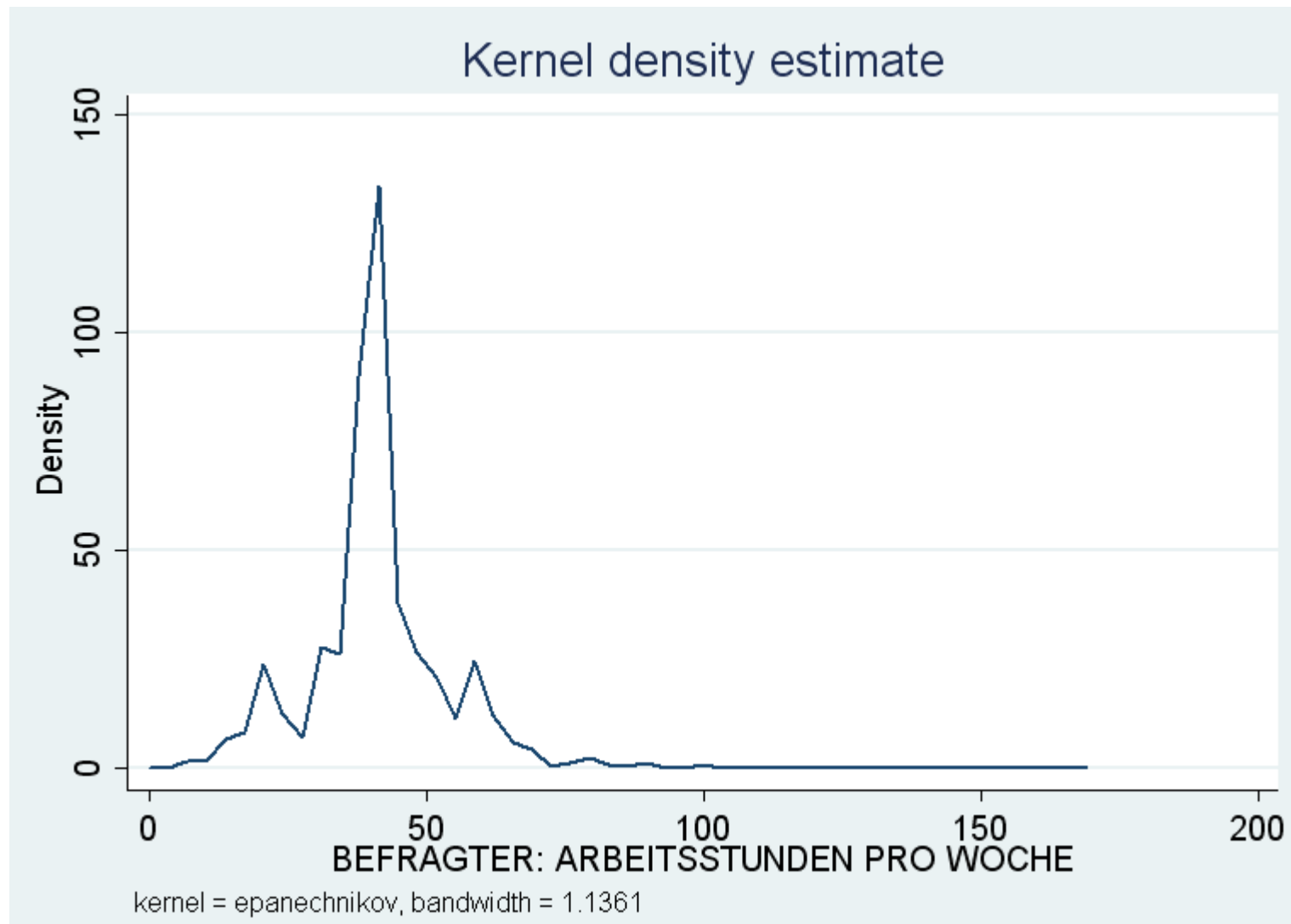
Konfession



EVANG.OHNE FREIKIRCH	EVANG.FREIKIRCHE
ROEMISCH-KATHOLISCH	AND.CHRISTL.RELIGION
AND.NICHT-CHRISTLICH	KEINER RELIGIONSGEM.

Histogramm und Kern-Dichte-Schätzer

- `mvdecode v207, mv(999.9=.a \ 0=.b) /* Missings setzen*/`
- `histogram v207`
- `histogram v207, width(20)`
- `histogram v207, kdensity`
- `kdensity v207 [iweight=v792]`



Aufgabe 9

- Erstellen Sie eine Graphik zur Wahlabsicht.
- Erstellen Sie Graphiken zum Alter der Befragten getrennt nach Geschlecht.
- Erstellen Sie Graphiken zum Nachrichtenkonsum der Befragten in den neuen und alten Bundesländern.
- Erstellen Sie Graphiken, die die Unterschiede von rechten und linken Befragten im Hinblick auf die Einstellung zum Nazionalsozialismus wiedergeben.
- Erstellen Sie getrennt nach Geschlechtern Graphiken, die über esoterische Ansichten Auskunft geben.

Hausaufgabe 3

- Versehen Sie Ihren Do-File mit anschaulichen Graphiken, die auch dem jeweiligen Skalenniveau entsprechen.