

```
/* Einführung in die Statistische Software WS 2016/17 - SAS -  
Cornelia Oberhauser */
```

```
/******/  
/* Datensatz: Mietspiegel München 2003  
http://www.stat.uni-muenchen.de/service/datenarchiv/miete/miete03.html
```

Zahlreiche deutsche Städte erstellen sogenannte Mietspiegel, um Mietern, Vermietern, Mietberatungsstellen und Sachverständigen eine "objektive" Entscheidungshilfe in Mietfragen zur Verfügung zu stellen.

Die Mietspiegel werden dabei insbesondere zur Ermittlung der "ortsüblichen Vergleichsmiete" (Nettomiete in Abhängigkeit von Wohnungsgröße, -ausstattung, -alter, etc.) herangezogen.

Bei der Erstellung von Mietspiegeln wird aus der Gesamtheit aller in Frage kommenden Wohnungen eine repräsentative Zufallsstichprobe gezogen (im Fall der Stadt München durch Infratest), und die interessierenden Daten werden von Interviewern anhand von Fragebögen ermittelt.

Der vorliegende Datensatz stellt einen Ausschnitt aus dem Mietspiegel München des Jahres 2003 dar und enthält die Daten von 2053 Wohnungen.

Variablen:

nm	Nettomiete in EUR
nmqm	Nettomiete pro m ² in EUR
wfl	Wohnfläche in m ²
rooms	Anzahl der Zimmer in der Wohnung
bj	Baujahr der Wohnung
bez	Stadtbezirk
wohngut	Gute Wohnlage? (J=1,N=0)
wohnbest	Beste Wohnlage? (J=1,N=0)
ww0	Warmwasserversorgung vorhanden? (J=0,N=1)
zh0	Zentralheizung vorhanden? (J=0,N=1)
badkach0	Gekacheltes Badezimmer? (J=0,N=1)
badextra	Besondere Zusatzausstattung im Bad? (J=1,N=0)
kueche	Gehobene Küche? (J=1,N=0)

```
*/
```

```
/* **** */
/* Gliederung:
```

Zu SAS

- 1) Was ist SAS
- 2) Bestandteile von SAS
- 3) Hilfe zu SAS

Daten einlesen, speichern, exportieren

Maßzahlen

- 1) Übersicht über die Daten verschaffen
- 2) Häufigkeiten
- 3) Lage- und Streuungsmaße

Diagramme

- 1) Balkendiagramm
- 2) Kreisdiagramm
- 3) Histogramm
- 4) Balkendiagramm für Kontingenztafeln
- 5) Mosaicplot
- 6) Boxplot
- 7) Streudiagramm

Korrelation

Regression

*/

```
/* **** */
/*
```

Zu SAS

- 1) Was ist SAS
- 2) Bestandteile von SAS
- 3) Hilfe zu SAS

siehe Ausschnitt aus Skript (pdf-Datei)

zusätzliche Unterlagen sind auf der Homepage zur Statistischen Software (SAS-Kurs) verlinkt:

<http://www.stablab.stat.uni-muenchen.de/saskurs2016>

*/

```

/*****/

/* zum Arbeiten sollte man sich eine sinnvolle Ordnerstruktur anlegen
   hier: Projektordner "D:\Einf_Software WS2016\SAS"
   mit Unterordnern: Programme, Daten, Ergebnisse
*/

/* Zuweisung des Pfades zu einer Makro-Variablen - Projektordner hier angeben
*/
%LET pfad_projekt = D:\Einf_Software WS2016\SAS;
*%LET pfad_projekt = Z:\Einf_Software WS2016\SAS;

/* Ausgabe im log-Fenster zur Überprüfung (&NameMakroVariable.) */
%PUT &pfad_projekt.;

/*****/
/* Daten einlesen, speichern, exportieren */

/* Einlesen: Datei -> Daten importieren (miete.asc)
   Datei mit Trennzeichen - Trennzeichen: Tab */

PROC IMPORT OUT= WORK.miete
    /* "Name der Bibliothek" Punkt "Name des Datensatzes*/
    DATAFILE= "&pfad_projekt.\Daten\miete03.asc"
    /* Dateipfad \ Dateiname mit Endung */
    DBMS=DLM REPLACE;
    DELIMITER='09'x;
    GETNAMES=YES;           /* Variablenamen in erster Zeile? */
    DATAROW=2;              /* Nummer der ersten Zeile, in der Daten stehen */
RUN;

/* Speichern: Datei -> Daten exportieren */

PROC EXPORT DATA= WORK.MIETE
    /* "Name der Bibliothek" Punkt "Name des Datensatzes*/
    OUTFILE= "&pfad_projekt.\Daten\miete_export.csv"
    /* Dateipfad \ Dateiname mit Endung */
    DBMS=CSV REPLACE;
RUN;

/*****/
/* Output ab hier speichern - als PDF
   Befehl nachträglich per Hand einfügen
*/
ODS PDF FILE = "&pfad_projekt.\Ergebnisse\output_sas.pdf";

```

```

/*****/
/* Maßzahlen
1) Übersicht über die Daten verschaffen
2) Häufigkeiten
3) Lage- und Streuungsmaße
*/

/* 1) Übersicht über die Daten verschaffen */
proc contents data=miete;
run;
proc means data=miete;
run;
/*
proc freq data=miete;
run;
*/

/* 2) Häufigkeiten
- Tabelle von rooms */
proc freq data=miete;
tables ROOMS;
run;
/* - Kreuzabelle zwischen rooms und bez */
proc freq data=miete;
tables BEZ*ROOMS;
run;

/* 3) Lage- und Streuungsmaße
- von nm */
proc means data=miete;
var NM;
run;
proc univariate data=miete;
var NM;
run;
/* - von nm (getrennt nach rooms) */
proc means data=miete n mean std min max;
class ROOMS; /* Gruppierungsvariable */
var NM;
run;
proc univariate data=miete;
class ROOMS; /* Gruppierungsvariable */
var NM;
run;

/* der Output von PROC MEANS kann einfach angepasst werden */
proc means data=miete mean std min q1 median q3 max n ;
class ROOMS;
var NM;
run;
/* siehe auch

https://support.sas.com/documentation/cdl/en/proc/67916/HTML/default/viewer.htm#nlqnc9bddfvhzqn105kqitnf29cp.htm
insbesondere: statistic-keyword(s) */

```

```

/*****
/* Diagramme - meist mit PROC SGPLOT
http://support.sas.com/documentation/cdl/en/grstatproc/67909/HTML/default/viewer.htm#plt32i8511t1gfn17sw07yxtazad.htm

1) Balkendiagramm
2) Kreisdiagramm
3) Histogramm
4) Balkendiagramm für Kontingenztafeln
5) Mosaicplot
6) Boxplot
7) Streudiagramm
*/

/* 1) Balkendiagramm
   - von rooms */

/* absolute Häufigkeiten */
proc sgplot data=miete ;
    vbar ROOMS / stat=freq;
run;

/* relative Häufigkeiten */
proc sgplot data=miete ;
    vbar ROOMS / stat=percent;
run;

/* 2) Kreisdiagramm
   - rooms */
proc gchart data=miete ;
    pie ROOMS;
run;
quit;

/* 3) Histogramm
   - nm */
proc sgplot data=miete;
    histogram NM;
run;

/* 4) Balkendiagramm für Kontingenztafeln
   - z.B.: ww0 getrennt nach rooms */

/* Diagramme nach Gruppen: */
/* Optionen:
   GROUP=variable
   GROUPDISPLAY=STACK | CLUSTER
*/

/* Balken nebeneinander: GROUPDISPLAY=CLUSTER */
proc sgplot data= miete;
vbar rooms / group=ww0 groupdisplay=cluster;
run;

proc sgplot data= miete;
vbar rooms / group=ww0 groupdisplay=cluster stat=percent seglabel;
RUN;

```

```

/* Balken gestapelt: GROUPDISPLAY=STACK */
proc sgplot data= miete;
vbar rooms / group=ww0 groupdisplay=stack;
run;

/* zu 100% skalierte gestapelte Balken */
proc freq data=miete;
tables ww0*rooms / plots=freqplot(twoway=stacked scale=group_pct);
run;

/* 5) Mosaicplot */
proc freq data=miete;
tables ww0*rooms / plots=mosaicplot;
run;

/* 6) Boxplot
- von nm */
proc sgplot data=miete;
hbox nm;
run;
/* - von nm getrennt nach rooms */
proc sgplot data=miete;
hbox nm / group=rooms;
run;

/* 7) Streudiagramm
- von nm und wfl */
proc sgplot data=miete ;
scatter x=WFL y=NM;
run;
/* - von nm und wfl getrennt nach wohngut */
proc sgplot data=miete ;
scatter x=WFL y=NM / group=WOHNGUT;
run;

```

```

/*****
/* Korrelation
   - nach Pearson */
proc corr data=miete pearson;
    var NM NMQM WFL;
run;
/* - nach Spearman */
proc corr data=miete spearman;
    var NM NMQM WFL;
run;
/* - Kendall's tau-b */
proc corr data=miete kendall;
    var NM NMQM WFL;
run;

/* - alle zusammen */
proc corr data=miete pearson spearman kendall;
    var NM NMQM WFL;
run;

/*****
/* Regression */
proc reg data=miete;
    model NM=WFL;
run;
quit;

/*****
/* Output bis hier speichern
   Befehl nachträglich per Hand einfügen
*/
ods pdf close;

```

```

/*****
/*****
/* Kopieren von Graphiken
Rechtsklick auf Graphik -> Kopieren */

/* Speichern von Graphiken
Rechtsklick auf Graphik -> Bild speichern unter... */

/*****
/*****
/* elegante Lösungen zum Speichern der Ergebnisse (Tabellen und Graphiken in
einer Datei) */

/* Output als RTF-Datei speichern */
ODS RTF FILE = "&pfad_projekt.\Ergebnisse\dateiname.rtf";
/*
dann:
die ganze Syntax

dann:
*/
ODS RTF CLOSE;

/* z.B.: */

ODS RTF FILE = "&pfad_projekt.\Ergebnisse\output_sas_test.rtf";

proc means data=miete n mean std min max;
run;

proc freq data=miete;
    tables ROOMS;
run;

proc sgplot data=miete ;
    vbar ROOMS / stat=freq;
run;

ODS RTF CLOSE;

/* Alternative: Output als PDF-Datei speichern */
ODS PDF FILE = "&pfad_projekt.\Ergebnisse\output_sas_test.pdf";

proc means data=miete n mean std min max;
run;

proc freq data=miete;
    tables ROOMS;
run;

proc sgplot data=miete ;
    vbar ROOMS / stat=freq;
run;

ODS PDF CLOSE;

```