

In der Hausübung soll der Datensatz „miete.2015“ von der Homepage verwendet werden. Es handelt sich bei den Daten um einen Ausschnitt der Daten zum Münchner Mietspiegel 2015. Der Datensatz enthält die folgenden Variablen:

Nettomiete:	Nettomiete in Euro
Nettomiete_pro_Quadratmeter:	Nettomiete pro Quadratmeter in Euro
Wohnflaeche:	Wohnfläche in Quadratmetern
Anzahl_Zimmer:	Anzahl an Zimmern in der Wohnung
Baujahr:	Baujahr
Stadtbezirk:	Stadtbezirk
Wohnlage:	Wohnlage (1=normale, 2=gute, 3=beste Wohnlage)

Vorbereitung:

- Legen Sie auf dem Z-Laufwerk (im Cip-Pool) einen Ordner für die Hausübung an. Legen Sie alle Dateien, mit denen Sie arbeiten, in diesem Ordner ab.
- Auf der Homepage ist die Datei „Einf_Software_WS1516_Homepage.R“ verlinkt. Speichern Sie diese in Ihrem Ordner.
- Versuchen Sie, diese Syntax in R laufen zu lassen.
 - Weisen Sie dem Objekt `matrikelnummer` Ihre eigene Matrikelnummer zu.
 - Ziehen Sie wie in der Vorlage eine Stichprobe vom Umfang 2500 ohne Zurücklegen aus den Daten. Verwenden Sie dazu Ihre Matrikelnummer als Startwert.
 - Speichern Sie den resultierenden personalisierten Teil-Datensatz in den vier vorgegebenen Formaten. Verwenden Sie für die weiteren Aufgaben diesen Teildatensatz.
- Speichern Sie die Vorlage für die Abgabedatei „Abgabe_Einf_Software_Nachname_Vorname.docx“.
 - Ändern Sie den Dateinamen, indem Sie statt Nachname Ihren eigenen Nachnamen und statt Vorname Ihren eigenen Vornamen einsetzen, z.B. „Abgabe_Einf_Software_Oberhauser_Cornelia.docx“
 - Schreiben Sie in die ersten Zeilen des Dokumentes Ihren Nachnamen, Vornamen, Ihre Matrikelnummer und Ihre Email-Adresse.
 - Kopieren Sie unter Aufgabe 0) die R-Syntax, die Sie verwendet haben, um Ihren personalisierten Teil-Datensatz zu erzeugen.

Lösen Sie die folgenden Aufgaben in allen drei vorgestellten statistischen Softwarepaketen **R**, **SPSS** und **SAS**. Verwenden Sie für jedes Softwarepaket jeweils einen der gespeicherten, personalisierten Teil-Datensätze. (Verwenden Sie für jedes Softwarepaket eine Version des Datensatzes, die sich fehlerfrei einlesen lässt. Evtl. lassen sich alle Formate fehlerfrei in alle Softwarepakete einlesen, evtl. können bei einzelnen Formaten aber Probleme auftreten.)

Kopieren Sie für jedes Softwarepaket die Syntax und den Output (Tabellen, Graphiken) in Ihre Abgabedatei. Verwenden Sie dazu immer die Reihenfolge **R**, **SPSS** und **SAS**. Fangen Sie für jede Aufgabe eine neue Seite an. Falls irgendeine Form von Interpretation gefragt ist, schreiben Sie diese bitte ans Ende der jeweiligen Aufgabe.

Aufgabe 1:

Lesen Sie Ihren personalisierten Teil-Datensatz in das jeweilige Statistikprogramm ein. (Verwenden Sie dazu jeweils das Format des Teildatensatzes, welches Sie fehlerfrei in das jeweilige Programmpaket einlesen können.)

Aufgabe 2:

Erstellen Sie eine Häufigkeitstabelle von `Stadtbezirk`. Berechnen Sie sowohl die absoluten als auch die relativen Häufigkeiten.

Aus welchem Stadtbezirk sind die meisten Daten in der Stichprobe?

Aufgabe 3:

Stellen Sie die absoluten Häufigkeiten von `Stadtbezirk` in einem Balkendiagramm dar.

Aufgabe 4:

Erstellen Sie eine Kreuztabelle zwischen `Stadtbezirk` und `Wohnlage`. Erstellen Sie wahlweise eine oder zwei Kreuztabellen mit den absoluten und den auf die Variable `Stadtbezirk` bedingten relativen Häufigkeiten.

In welchem Stadtbezirk gibt es (absolut) die meisten Wohnungen mit bester Wohnlage?

Aufgabe 5:

Stellen Sie die Variablen `Stadtbezirk` und `Wohnlage` gemeinsam in einem Balkendiagramm dar. Verwenden Sie dazu ein auf 100% skaliertes Balkendiagramm der relativen bedingten Häufigkeiten.

In welchem Stadtbezirk ist der Anteil der Wohnungen in bester Wohnlage am größten?

Aufgabe 6:

Berechnen Sie für die `Wohnflaeche` und für die `Nettomiete_pro_Quadratmeter` jeweils den Mittelwert, die Varianz, die Standardabweichung und die 5-Punkte-Zusammenfassung.

Aufgabe 7:

Erstellen Sie ein Histogramm der `Wohnflaeche`.

Wie würden Sie die Verteilung charakterisieren?

Aufgabe 8:

Erstellen Sie einen Boxplot der `Wohnflaeche` getrennt nach `Wohnlage`.

In welcher Wohnlage ist die Wohnfläche am größten?

Aufgabe 9:

Erstellen Sie einen Boxplot der `Nettomiete_pro_Quadratmeter` getrennt nach `Wohnlage`.

In welcher Wohnlage ist die Nettomiete pro Quadratmeter am kleinsten?

Aufgabe 10:

Zeichnen Sie ein Streudiagramm zwischen `Nettomiete_pro_Quadratmeter` und `Wohnflaeche`. Färben Sie die Punkte nach der Variable `Wohnlage` ein.

Geben Sie die von Ihnen umbenannte Abgabedatei „Abgabe_Einf_Software_Nachname_Vorname.docx“ bis spätestens **22. Januar 2015** über Moodle ab. (Wahlweise kann auch eine pdf-Datei abgegeben werden.)