

1. (Modifikation von Aufgabe 3 auf Blatt 5.) Zur Arbeitsweise von `read.table`:
  - a) Lesen Sie die kopierte Datei **SMSA** aus Ihrem Arbeitsverzeichnis nun mit Hilfe der Funktion `read.table` in einen Data Frame namens **SMSA.df** erneut ein, und zwar so, dass die Spalte **ID** der SMSA-Identifikationscodes zur Zeilenbenennung und nicht als Datenspalte verwendet wird!
  - b) Benennen Sie die Spalten (= Variablen) des Data Frames **SMSA.df** (analog zu Aufgabe 3e auf Blatt 5) mit kurzen, aber sinnvollen und ihrer Bedeutung entsprechenden Namen!
  - c) Führen Sie für **SMSA.df** die analogen Schritte der Aufgaben 3f bis 3h auf Blatt 5 durch!
2. (Fortsetzung von Aufgabe 1.) Univariate, grafische explorative Analyse (EDA) der SMSA-Daten: Überlegen Sie sich eine (oder mehrere) grafische Darstellungsmethode(n), mit der (oder denen) die Verteilungen der
  - a) Anzahlen der praktizierenden Ärzte und Ärztinnen,
  - b) Prozentsätze an High-School-AbsolventInnen und
  - c) privaten Gesamteinkommennach geografischen Regionen aufgeteilt geeignet und sinnvoll (!) veranschaulicht werden können! Erstellen Sie die Grafiken mithilfe der Formelvariante(n) der von Ihnen ggf. verwendeten Grafikfunktion(en) (oder zur Not alternativ mithilfe der Funktion `split`)!
3. Kopieren Sie sich die Datei **MietenFrankfurt** aus Stud.IP in Ihr eigenes, übliches **R**-Arbeitsverzeichnis und lesen Sie sie geeignet in **R** ein!
  - a) Fertigen Sie mit `stem` ein “stem-and-leaf”-Diagramm für diese Daten an! Probieren Sie die Wirkung des Arguments `scale` aus, indem Sie es mit (wenigen) Werten kleiner und größer als 1 versehen!
  - b) Führen Sie Analoges mit flächennormierten (!) Histogrammen für diese Daten durch, indem Sie das Argument `breaks` von `hist` zunächst nicht verwenden, dann nur die eine oder andere natürliche Zahl zwischen 4 und 40 an `breaks` übergeben (und dadurch die ungefähre Anzahl der äquidistanten Intervalle, in die die „Mietskala“, sprich *x*-Achse eingeteilt werden soll, steuern) sowie schließlich verschiedene, auch irreguläre Gitter an `breaks` in Form eines Vektors zuweisen!
  - c) Experimentieren Sie entsprechend mit der Implementation des Kern-Dichteschätzers durch `density` (in Kombination mit `plot` wie in §4.2.2), indem Sie zunächst kein weiteres Argument von `density` verwenden und dann für ihr Argument `bw` (wie “bandwidth”) wenige verschiedene Werte zwischen 25 und 400 ausprobieren!  
Wie hängt das qualitative Aussehen, konkret die „Glattheit“ des Graphen des Kern-Dichteschätzers von der “bandwidth” tendenziell ab?