



Statistik I

Übungsblatt 5

Abgabe: Donnerstag, 8.06.06, 12:00 Uhr, Briefkasten: Statistik I.

Dieses Blatt dient der Vorbereitung auf das Projekt. Insgesamt können 25 Punkte erreicht werden.
Eine erfolgreiche Bearbeitung ist **Voraussetzung** für eine Teilnahme am Projekt!

R

Mondrian

-
1. Laden Sie den Datensatz "EcoTestAuto.txt" von der Vorlesungs Webseite.
 2. Laden Sie die Daten in R (Load Dataset in JGR). Welcher R-Befehl wurde dafür benutzt?
 3. Beschreiben Sie kurz die Bedeutung aller Variablen und deren Variablen-Typen.
 4. Erstellen Sie möglichst kompakt und übersichtlich, Summaries für alle Variablen.
 5. Berechnen Sie Korrelationen zwischen den stetigen Variablen, und erstellen Sie 2-dim. Kreuztabellen für alle kategoriellen Daten.
 6. Untersuchen Sie die stetigen Variablen auf mögliche Zusammenhänge mittels Scatterplots, und beschreiben Sie die 2 wichtigsten Ergebnisse.
 7. Analysieren Sie die Variablen "Firma", "Klasse" und "Typ" mittels der R-Befehle `barplot` und `spine` (in Library `vcd`). Wie bewerten Sie ihre Ergebnisse?
 8. Plotten Sie alle stetigen Variablen in einem Parallelen Koordinaten Plot. Was fällt auf?
 9. Identifizieren Sie im Scatterplot für "CCM" gegen "kW" Fahrzeuge mit überdurchschnittlich viel oder wenig Leistung, gegeben "CCM". Welche Fahrzeuge sind dies? Wie begründen Sie Ihre Auswahl?
 10. Im Scatterplot für "FuelCons" gegen "CO2" sieht man 6-7 isolierte Gruppen:
Bestimmen Sie separierende Geraden, die die Gruppen trennen, und tragen Sie diese mittel `abline` in den Scatterplot ein. Verwenden Sie ggf. diese Geraden, um die Gruppen verschieden einzufärben und in anderen Plots weiter zu analysieren.

2. Laden Sie den Datensatz in Mondrian.

4. Erstellen Sie für alle Variablen passende univariate Graphiken. Wählen Sie sinnvolle Reihenfolgen in Barcharts und Parameter in Histogrammen.

5. Betrachten Sie die Zusammenhänge zwischen allen Variablen mittels Selektionen und dem korrespondierenden Highlighting. Notieren Sie die wichtigsten Ergebnisse.

7. Analysieren Sie die Variablen "Firma", "Klasse" und "Typ" mittels Mosaic Plots und deren Varianten.

Selektieren Sie die Gruppen (eine Selektion im "OR" Modus ist dabei hilfreich) und färben Sie sie ein. Kann man diese Gruppen in anderen Plots identifizieren?

Dokumentieren Sie Ihre Arbeit so weit wie möglich jedoch nur so weit wie nötig, um die Ergebnisse nach zu vollziehen. Ziel dieses Projektes ist auch eine erste Einübung in das Schreiben von Berichten.