

Übungsserie 2

Dateien, die wir Ihnen zur Verfügung stellen, finden Sie im Verzeichnis
Z:\users\s\sfs\datasets\Systat.

1. In untenstehender Tabelle sind die Farbintensitäten (**B**) von Äpfeln bei verschiedenen Lagerungstemperaturen (**TEMP**) und Lagerungsdauer (**TIME**) aufgeführt. Die Daten finden Sie in der Datei **apfelfar.syd**.

Lagerungsdauer	Lagerungstemperatur			
	1	10	18	27
0	40.98	41.14	41.08	40.56
7	40.84	41.17	42.02	40.29
14	41.14	42.60	43.82	40.56
21	40.70	44.41	45.22	41.11

- a) Welcher Intensitätswert (**B**) wird von ziemlich genau 25% aller Beobachtungen unterschritten? Wieviel Prozent der Beobachtungen weisen eine Intensität über 43 auf?

Im Histogrammenü die richtige Variable auf der richtigen Achse aktivieren, **Options / Histogram Options...** anwählen und Balkenbreite auf 0.2 setzen. Je nachdem, ob wir **Display cumulative frequency** aktivieren oder nicht, erhalten wir eine andere Grafik. Welche eignet sich besser, um obige Fragen zu beantworten?

(Am Besten markieren Sie die X-, Y- Achse mit Striche ausserhalb der Achse: **Axes / All Axes** anwählen und **Tick Mark Style** auf **Outside axes** setzen.)

- b) Erstellen Sie ein Stamm- und Blatttdiagramm der 16 gemessenen Farbintensitäten. Welche Werte werden als "Outside Values" angegeben?

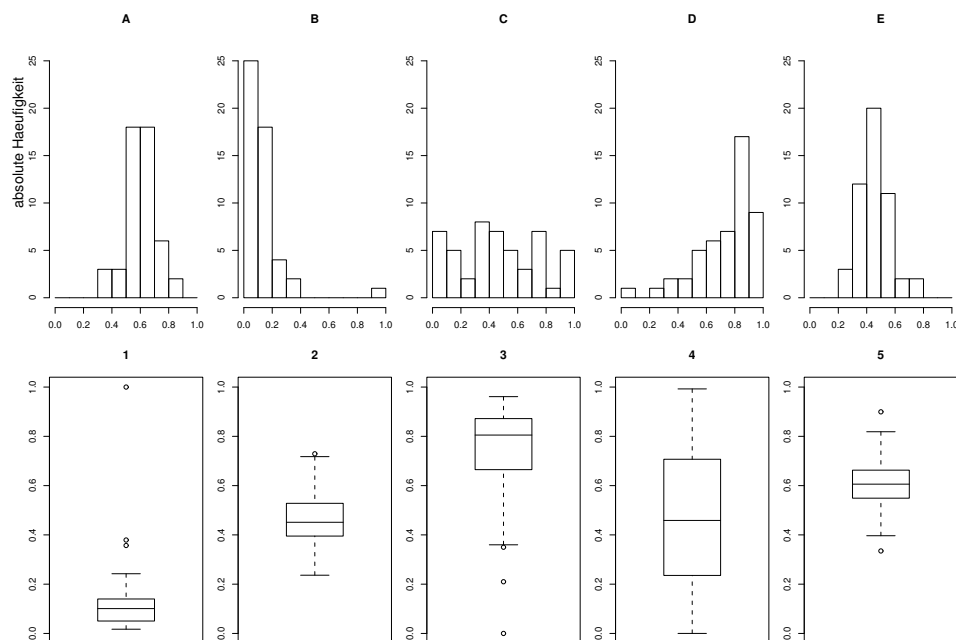
Im Hauptmenü **Statistics / Descriptive Statistics / Stem-and-Leaf ...** anwählen und die richtige Variable aktivieren.

- c) Zeichnen Sie die Farbintensität (**B**) gegen die Lagerungsdauer (**TIME**) für die vier verschiedenen Temperaturen auf. Wählen Sie für jede Temperaturstufe ein anderes Symbol und eine andere Farbe. Vergleichen Sie, wie sich bei den verschiedenen Temperaturstufen die Farbintensität mit der Lagerungsdauer verändert. Gibt es Unterschiede?

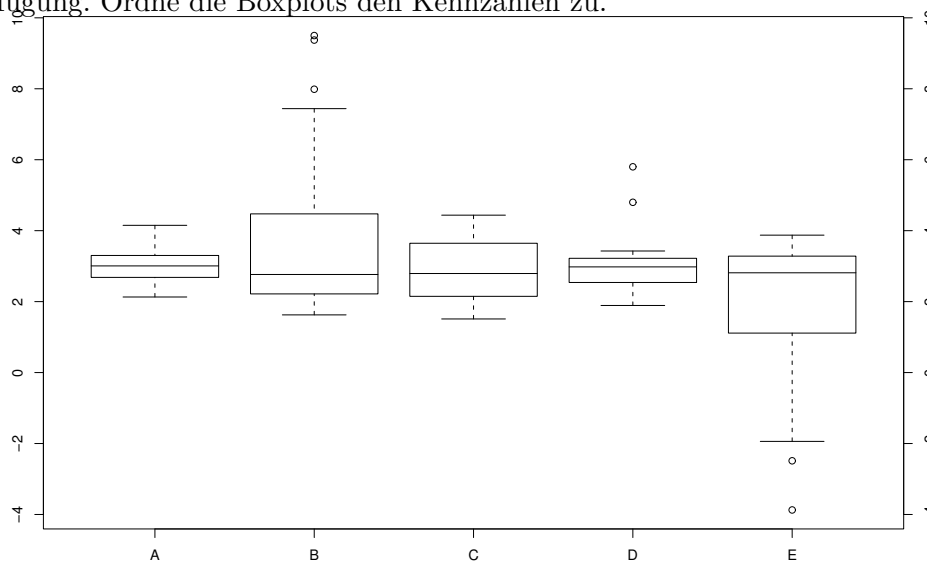
Im Scatterplotmenü die Variablen für die X- und Y-Achse aktivieren, dann in **Appearance / Color and Fill...** und **Appearance / Symbol and Label...** die Farbe und das Symbol temperaturabhängig machen, das heisst abhängig von **TEMP** oder **TEMP-CODE**. Um eine zu grosse Farbpalette zu vermeiden, kann vorgängig im Menü **Data / Category**, die Variable ausgewählt werden.

- d) Zeichnen Sie analog zu c) die Farbintensität gegen die Lagerungstemperatur (**TEMP**) für die vier verschiedenen Lagerungsdauern auf. Wählen Sie für jede Lagerungsdauer ein anderes Symbol und eine andere Farbe. Was stellen Sie hier fest?

2. a) Für fünf Stichproben vom Umfang $n = 50$ wurden je ein Histogramm und ein Boxplot gezeichnet. Ordne die fünf Boxplots den entsprechenden Histogrammen zu.



- b) Von fünf Stichproben vom Umfang $n = 30$ stehen je ein Boxplot sowie drei Kennzahlen zur Verfügung. Ordne die Boxplots den Kennzahlen zu.



Stichprobe	1	2	3	4	5
Mittelwert	1.93	3.8	3.04	2.92	3
Varianz	4.01	5.13	0.26	0.68	0.55
1. Quartil	1.17	2.24	2.69	2.16	2.55

3. Bei der Messung von Zelldurchmessern erhielt man folgende Werte:

1. Serie : 36, 68, 90, 49, 7, 20, 59, 85, 165, 29, 14, 43
2. Serie : 78, 51, 118, 147, 33, 203, 48, 63, 32, 40, 126, 21

a) Erzeuge aus obigen Werten ein Systat-Datenblatt.

Ein neues leeres Datenblatt kann mit **File / New / Data** geöffnet werden. Am bestens definiert man die 2 Variablen **Serie** und **Durchmesser** in 2 verschiedenen Spalten. Um einer Spalte einen Namen zu geben, genügt es in eine beliebige Zelle der Spalte zu klicken und in **Data / Variable Properties ...** den Namen zu geben.

b) Stelle die Daten der 1. und 2. Serie mit einem Box-Plot dar.

Der Box-Plot Befehl findet man unter **Graph / Density Displays** oder als "Quick Button". Als X-Variable **Serie** angeben, so dass die Box-Plots im gleichen Plot gezeichnet werden, und als Y-Variable ...

c) Fasse die Daten zusammen. Berechne den Median und das arithmetische Mittel.

Unter **Statistics / Descriptive Statistics / Basic Statistics** die richtigen Felder aktivieren.

4. Bei einer Untersuchung werden Wasserproben auf Verunreinigungen untersucht. Da nur 2 Prozent aller Proben verunreinigt sind, wird vorgeschlagen, von 10 Einzelproben jeweils die Hälfte der Probe zu einer Sammelprobe zusammenzufassen und zunächst nur die Sammelprobe zu untersuchen. Wird in der Sammelprobe keine Verunreinigung festgestellt, so ist die Untersuchung für die 10 Einzelproben beendet. Im anderen Fall werden alle 10 übriggebliebenen Hälften in 10 Einzeluntersuchungen geprüft.

a) Wie gross ist die Wahrscheinlichkeit, in der Sammelprobe keine Verunreinigung zu finden (unter der Annahme, dass die Einzelproben unabhängig voneinander sind)?

b) Sei die Zufallsvariable Y die Gesamtzahl benötigter Analysen. Welche Werte kann Y annehmen, und mit welchen Wahrscheinlichkeiten treten sie auf?

c) Wieviele Analysen werden "im Durchschnitt" für die gesamte Untersuchung benötigt (d.h. wie gross ist $\mathbf{E}[Y]$)? Wieviele Analysen werden durch die Bildung von Sammelproben "im Durchschnitt" eingespart?

Punkteverteilung:

Aufgabe 1: (1 Punkt in a), in b), in c) + d)) Antworten auf die gestellten Fragen.

Aufgabe 2: (2 Punkte) Zuordnung.

Aufgabe 3: (1 Punkt in b) und in c)) Antworten auf die gestellten Fragen.

Aufgabe 4: (1 Punkt je Unteraufgabe) Antworten auf die gestellten Fragen.

Abgabe: 1 Woche nach der Übungsstunde in der Vorlesung.

Präsenz: Montag, 12:00-13:00, LEO C12.1.