

# Statistische Software (R)

Paul Fink, M.Sc.

Institut für Statistik

Ludwig-Maximilians-Universität München

*Grafiken*



## **Sinnvolle, leicht zu verstehende Grafiken**

Leitfragen:

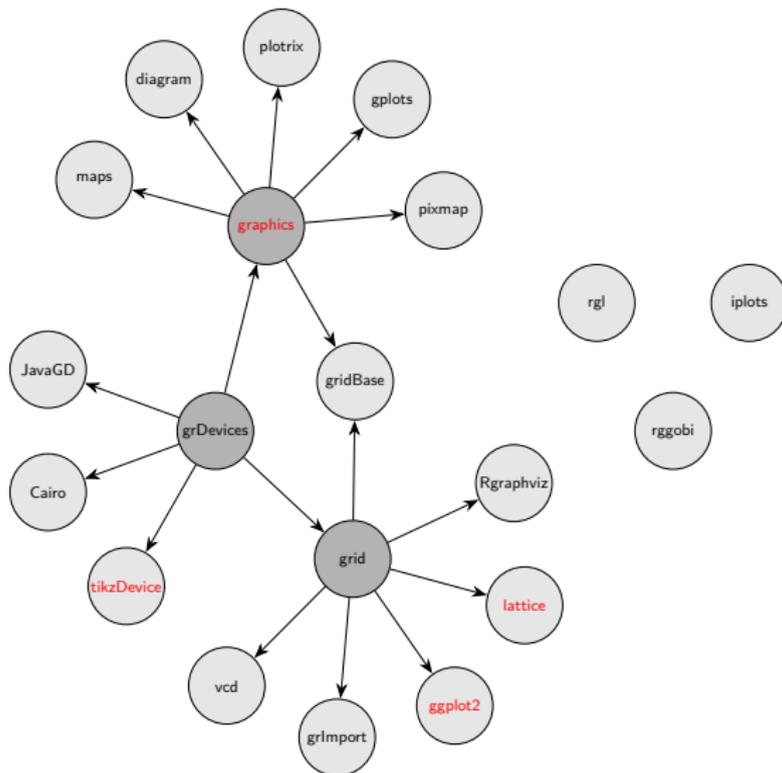
1. Welchen Grafiktyp verwenden?
2. Ist der Inhalt der Grafik klar dargestellt?
3. Einsatz von (einheitlichen) Farben?
4. Ist der Text lesbar?
5. Legende?

Paul Murrel (2011): R Graphics Second Edition, CRC Press

Link zu den im Buch dargestellten Grafiken und dem zugehörigem R-Code:

<http://www.stat.auckland.ac.nz/~paul/RG2e/>

# Übersicht ausgewählter Grafikpakete (nach Murrel 2011)



- Die Grafikausgabe erfolgt in ein sogenanntes Gerät (Device)
- Die öffnende Funktion bestimmt das Gerät
- Standardmäßig zur Verfügung stehen u.a.: `bmp()`, `jpeg()`, `pdf()`, `png()`, `postscript()`, `x11()` für Bildschirmfenster
- Zusatzpaket `tikzDevice` stellt Funktion `tikz()` zur Verfügung.
- Tatsächliche Ausgabe bei Datei-Geräten erst nach Schließen

Neues Grafikfenster auf dem Bildschirm:

```
> x11()
```

Zeichnen der Dichte der Standard-Normalverteilung mit  
Grafikfunktion `curve()`:

```
> curve(from = -3, to = 3, dnorm(x), main = "Dichte N(0,1)-Vtlg.")
```

Schließen des Geräts (Fenster):

```
> dev.off()
```

# Grafikausgabe – PDF-Datei

---

Anlegen der PDF-Datei *dichteN01.pdf* im aktuellen Arbeitsverzeichnis:

```
> pdf(file = "dichteN01.pdf")
```

Zeichnen der Dichte der Standard-Normalverteilung mit Grafikfunktion `curve()`:

```
> curve(from = -3, to = 3, dnorm(x), main = "Dichte N(0,1)-Vtlg.")
```

Schließen des Geräts (Tatsächliches Erstellen des Datei-Inhalts):

```
> dev.off()
```

# Traditionelle Grafiken (`graphics` Paket)

---

Zwei Klassen von Grafik-Funktionen

- *High-level* Grafik-Funktionen als vorgefertigte Grafiken: Boxplots, Histogramme, Streu-, Balkendiagramme, ...
- *Low-level* Grafik-Funktionen als Basis aller Grafiken: Punkte, Linien, Rechtecke, Segmente, Beschriftung, Koordinaten-Achsen, ..., ...

⇒ Grafiken nach Baukasten-Prinzip selbst erstellen oder High-level Grafiken erweitern

# Die `plot()` Funktion

---

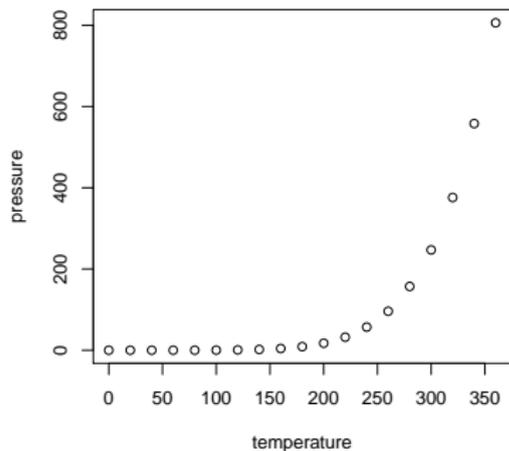
`plot()` ist wichtigste traditionelle *High-level* Funktion

Oft einfachste Variante um Basis-Grafiken zu erstellen

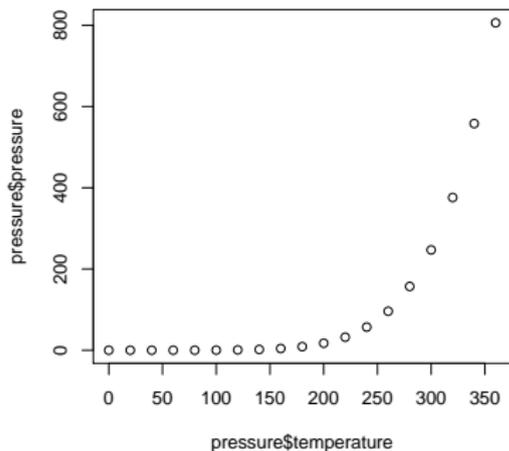
Je nach Datentyp der übergebenen Objekte liefert die Funktion eine andere Grafik:

*Die `plot()` Funktion ist generisch.*

# Beispiele – Grafiken

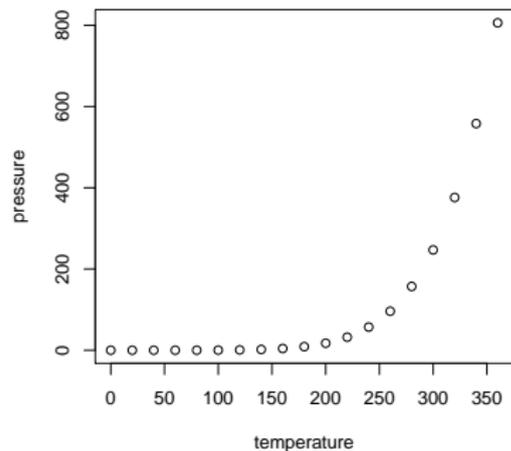


```
> plot(pressure)
```

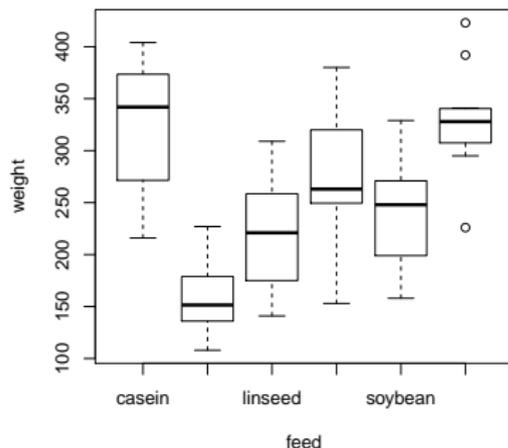


```
> plot(pressure$temperature,  
+      pressure$pressure)
```

# Beispiele – Grafiken



```
> plot(pressure ~ temperature,  
+      data = pressure)
```



```
> plot(weight ~ feed,  
+      data = chickwts)
```

# Wichtige High-Level Grafikfunktionen

<b>Funktion</b>	<b>Datentyp(en)</b>	<b>Beschreibung</b>
<code>plot()</code>	numeric (, numeric)	Scatterplot
<code>plot(), pairs()</code>	data.frame	Scatterplot Matrix
<code>sunflowerplot()</code>	numeric, numeric	Scatterplot (diskret)
<code>plot()</code>	factor oder 1-dim. table	Barplot
<code>barplot()</code>	numeric (Höhe der Balken)	Barplot
<code>barplot()</code>	matrix	Barplot
<code>hist()</code>	numeric	Histogramm
<code>boxplot()</code>	(list of) numeric	(bedingte) Boxplot
<code>plot()</code>	factor, numeric	bedingte Boxplots
<code>plot()</code>	factor, factor	Spineplot
<code>plot()</code>	2-dim. table	Mosaic plot
<code>mosaicplot()</code>	n-dim. table	Mosaic plot

# Argumente für Grafikfunktionen

Anpassung von Aussehen über Argumente der Grafikfunktion, z.B. Titel, Achsenbeschriftung, Farbe, ...

Argument	Beschreibung
<code>main</code>	Haupttitel der Grafik
<code>xlab, ylab</code>	Titel der X-Achse bzw Y-Achse
<code>xlim, ylim</code>	Vektor mit Minimum/Maximum für Werte in der Plot-Region in X bzw. Y Richtung
<code>cex</code>	Generelle Vergrößerung
<code>cex.main, cex.axis, cex.lab</code>	Vergrößerung von Titel, Achsenbeschriftung und -titeln relativ zu <code>cex</code>
<code>col</code>	Farbe der Objekte in der Plot-Region
<code>axes</code>	Bei <code>FALSE</code> werden keine Achsen eingezeichnet
<code>xaxt="n", yaxt="n"</code>	Kein Einzeichnen von X bzw. Y-Achse
<code>lty, lwd</code>	Linientyp, Linienbreite
<code>pch</code>	Zeichen für Punkte

# Aufgaben

---

1. Laden Sie den Datensatz `iris`.
2. Welche Variablen sind in dem Datensatz enthalten und von welchem Typ sind sie?
3. Erstellen Sie ein Histogramm für die Variable `Sepal.Width`.
4. Erstellen Sie ein Streudiagramm von `Petal.Length` und `Petal.Width` auf zwei verschiedene Arten.
5. Erstellen Sie ein Streudiagramm von `Petal.Length` und `Petal.Width`, in welchem die verschiedenen Gruppen der Variable `Species` in unterschiedlichen Farben dargestellt werden.