



Einführung in die Statistik

Praktische Übung – Jürgen Hermes – IDH – SoSe 2025

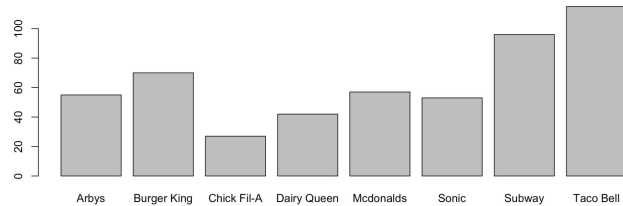
R – Die Grundlagen

- R und RStudio-Setup
- Einfache Datentypen in R
- Komplexe Datentypen in R
- Indizierung
- Funktionen
- Prädikate
- Zusatzpakete
- Daten einlesen
- Daten speichern
- **Daten visualisieren**

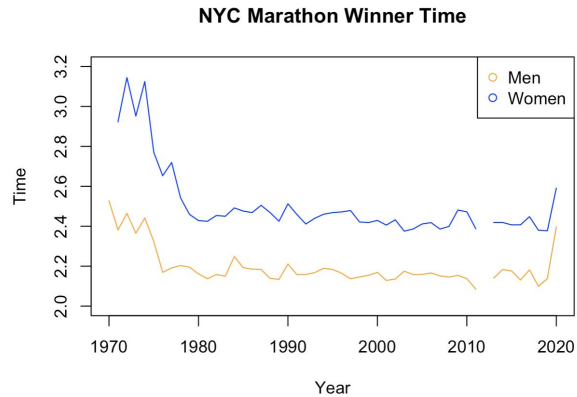


Visualisierungen (Exkurs)

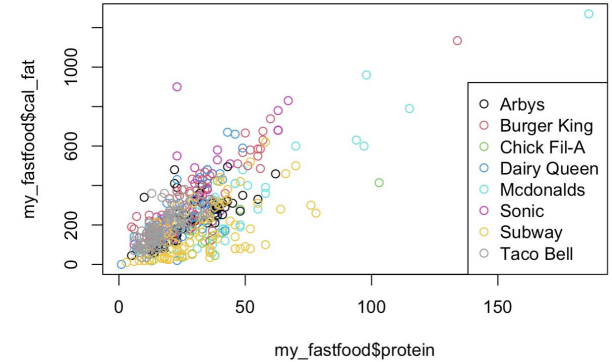
Barplots



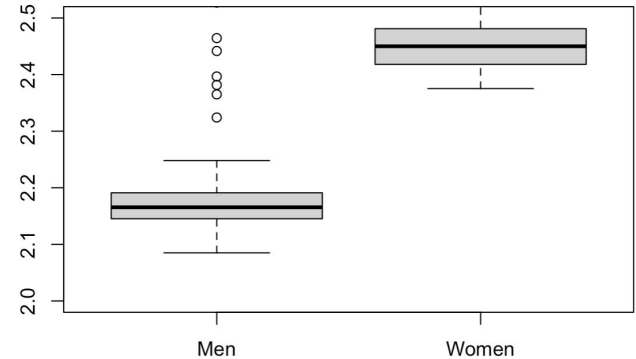
Linienplots



Punktplots

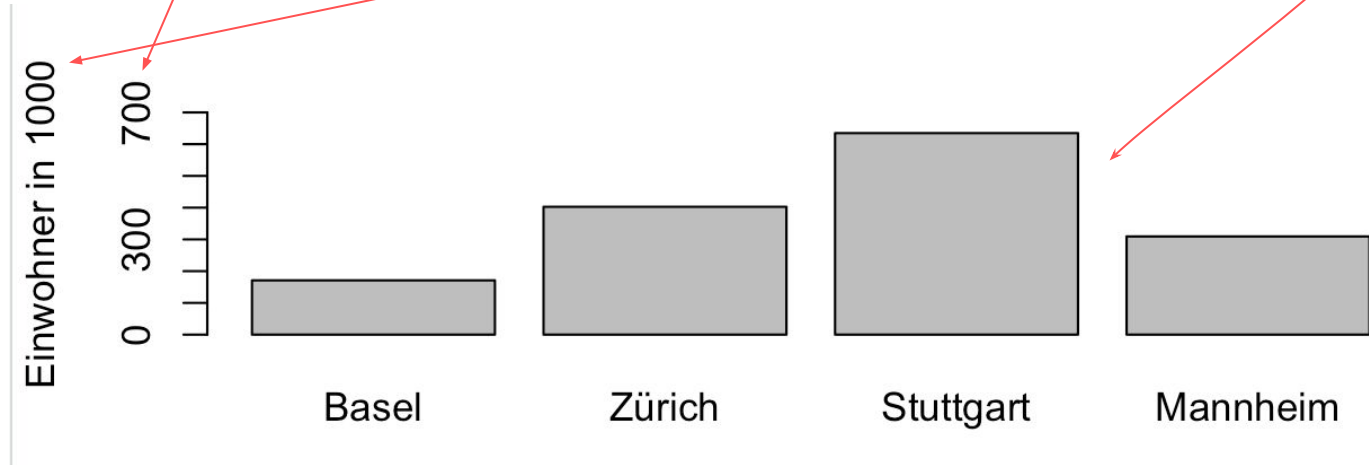


Boxplots



Visualisierung: Barplots

- `barplot(height = cit$einw/1000, names.arg = cit.name, ylim = c(0, 700), ylab= "Einwohner in 1000")`

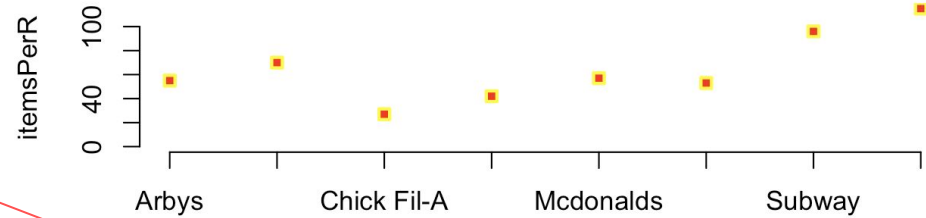


Visualisierung: Punktdiagramme

```
plot(itemsPerR, pch = 22,  
      bg = "red", color = "yellow")
```

plot (x, y, pch = _)

0	1	2	3	4	5	6	
□	○	△	+	×	◇	▽	
7	8	9	10	11	12	13	14
⊠	✱	⬡	⊕	⊗	⊞	⊗	⊞
15	16	17	18	19	20		
■	●	▲	◆	●	●		
21	22	23	24	25	21:25 mit bg füllbar		
●	■	◆	▲	▼			



Farben Übersicht (mehr in bF Anhang)



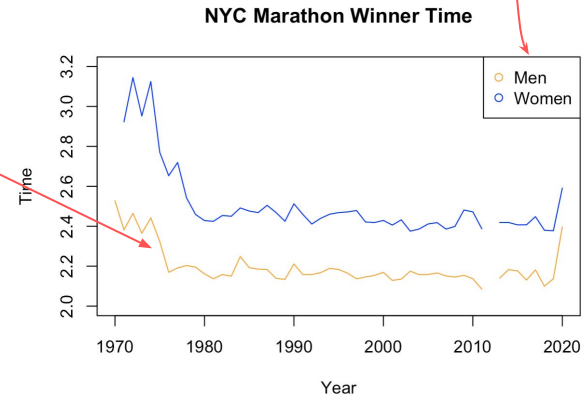
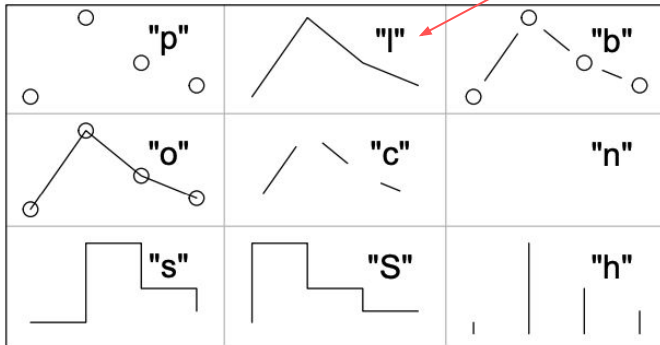
Visualisierung: Liniendiagramme

```
plot(y=nycm_women$time_hrs, x=nycm_women$year, col="blue", type = "l", ylim=c(2,3.2),  
     xlab = "Year", ylab="Time", main = "NYC Marathon Winner Time")
```

```
lines(y=nycm_men$time_hrs, x=nycm_men$year, col="orange")
```

```
legend("topright", levels(as.factor(nyc_marathon$division)),  
      col =c("orange","blue"), pch = 1)
```

plot (x, y, type = _)



Visualisierung: Boxplots

```
boxplot(nyc_marathon$time_hrs ~ nyc_marathon$division, ylim=c(2,2.5))
```

