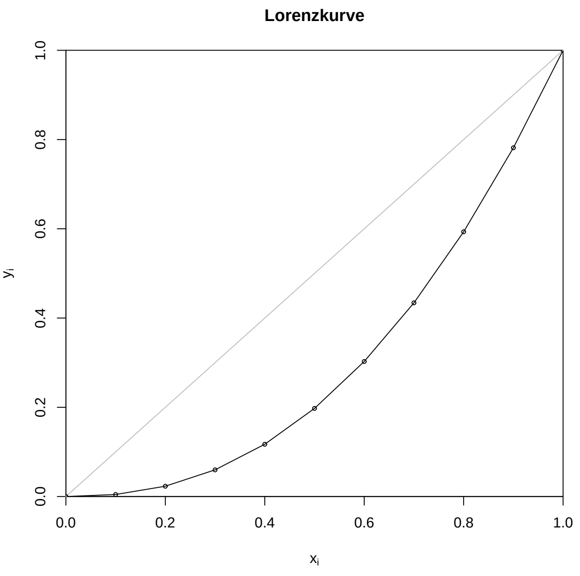
Lösungsnotizen Aufgabe 1

1. Modus (Modalwert)
   - mindestens Nominalskala
   - Beispiel: Haarfarbe
2. Median
   - mindestens Ordinalskala
   - Beispiel: Platzierung beim 100-m-Lauf
3. arithmetisches Mittel
   - mindestens Intervallskala
   - Beispiel: Temperatur in Grad Celsius
4. Harmonisches Mittel / geometrisches Mittel
   - mindestens Verhältnisskala
   - Beispiel: Temperatur in Grad Kelvin

Lösungsnotizen Aufgabe 2

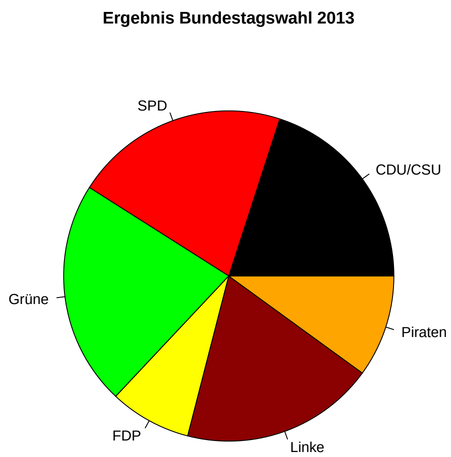


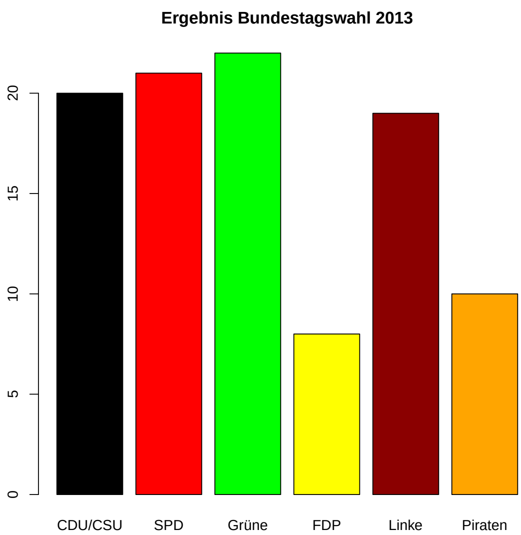
Ein Punkt (x_i, y_i) gibt an, dass die $x_i \cdot 100\%$ der Ärmsten über $y_j \cdot 100\%$ des Gesamteinkommens verfügen. Bei einer Gleichverteilung des gesammten Einkommens über alle Individuen würde die Winkelhalbierende

$y_i = x_i$ entstehen. Jede Lorenzkurve ist konvex und liegt unterhalb der Winkelhalbierenden. Der Gini-Koeffizient ist das doppelte der Fläche zwischen Winkelhalbierender und der Lorenzkurve und ist ein Maß für die Konzentration der Einkommensverteilung. Ein geringerer Gini-koeffizient steht dabei für eine gerechtere Verteilung und ein Koeffizient von 0 entspricht der Gleichverteilung.

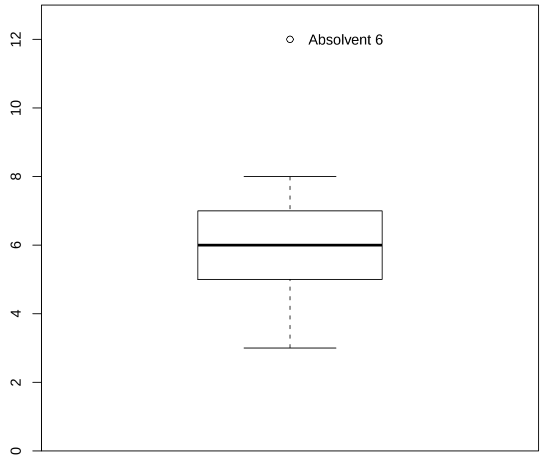
Lösungsnotizen Aufgabe 3

```
> Parteien = c("CDU/CSU", "SPD", "Grüne", "FDP", "Linke", "Piraten")
> Anteile = c(20, 21, 22, 8, 19, 10)
> pie(main = "Ergebnis Bundestagswahl 2013", x = Anteile, Parteien,
  col = c("black", "red", "green", "yellow", "darkred", "orange"))
> barplot(main = "Ergebnis Bundestagswahl 2013",
  height = Anteile, names.arg = Parteien,
  col = c("black", "red", "green", "yellow", "darkred", "orange"))
```





Lösungsnotizen Aufgabe 4



Interquartilsabstand $IQR = x_{0.75} - x_{0.25} = 7 - 5 = 2$
Außreißerkoordinaten

- oben: $x_o := x_{0.75} + 1.5 \cdot IQR = 7 + 3 = 10$
- unten: $x_u := x_{0.25} - 1.5 \cdot IQR = 5 - 3 = 2$

kleinste Beobachtung im Bereich $[x_u, x_o] : x = 3$
größte Beobachtung: $x = 8$
Einziger Ausreißer ist Absolvent 6 mit 12 Monaten bis zum ersten Job.

Der Boxplot dient allgemein dazu, eine Verteilung graphisch darzustellen und insbesondere einen schnellen Überblick über Lage- und Streumaße (Median, Interquartilsabstand) zu geben und weiterhin eine eventuell vorhandene Schiefe oder einen Exzess bzw. extreme Werte („Ausreißer“) sichtbar zu machen.

Lösungsnotizen Aufgabe 5

$$
\begin{aligned}
\bar{x}_m &= \frac{51+65+67+75}{4} = 64.5 \\
\bar{x}_w &= \frac{55+70+80+81}{4} = 71.5 \\
\bar{x} &= \frac{51+65+67+75+55+70+80+81}{8} = \frac{\bar{x}_m+\bar{x}_w}{2} = 68 \\
\tilde{s}^2_{zw} &= \frac{1}{n}\sum_{l=1}^{2} n_l(\bar{x}_l-\bar{x})^2 = \frac{1}{8}\cdot\left(4\cdot(64.5-68)^2+4\cdot(71.5-68)^2\right) = 12.25 \\
\tilde{s}^2 &= \frac{1}{n}\sum_{i=1}^{n} x_i^2-\bar{x}^2 = \frac{51^2+65^2+\ldots+81^2}{8}-68^2 = 104.25 \\
\text{Anteil } w &= \frac{\tilde{s}^2_{zw}}{\tilde{s}^2} = \frac{12.25}{104.25} \approx 11.75\%
\end{aligned}
$$

Lösungsnotizen Aufgabe 6

Gesamtstromproduktion 2007: $G = 54.6+106.9+164.7+77.3 = 403.5$

$$
\begin{aligned}
H &= \sum_{i=1}^{4} p_i^2 \\
&= \left(\frac{54.6}{403.5}\right)^2+\left(\frac{106.9}{403.5}\right)^2+\left(\frac{164.7}{403.5}\right)^2+\left(\frac{77.3}{403.5}\right)^2 \approx 0.2918
\end{aligned}
$$

Gesamtstromproduktion 2008: $G = 49.2+102.5+167.7+71.0 = 390.4$

$$
\begin{aligned}
H &= \sum_{i=1}^{4} p_i^2 \\
&= \left(\frac{49.2}{390.4}\right)^2+\left(\frac{102.5}{390.4}\right)^2+\left(\frac{167.7}{390.4}\right)^2+\left(\frac{71.0}{390.4}\right)^2 \approx 0.3024
\end{aligned}
$$

Im Jahr 2008 war die Konzentration stärker.