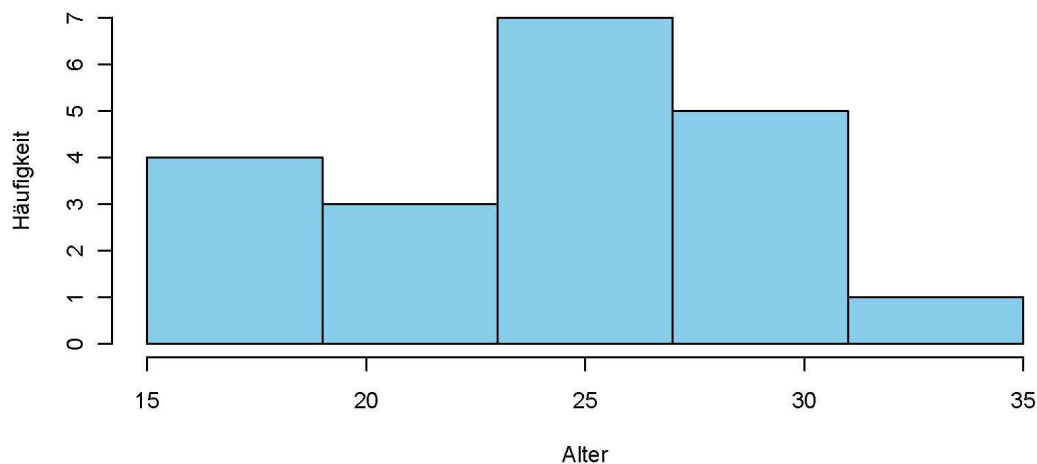


4 Datenvisualisierung

Aufgabe 6

Neben Stab-, Balken- und Kreisdiagrammen haben Sie auch das **Stamm-Blatt-Diagramm** sowie das **Histogramm** kennen gelernt.

- (a) Wie lässt sich zwischen Stamm-Blatt-Diagramm und Histogramm bzw. Dichtekurve eine Beziehung herstellen?
- (b) Welchen Aussagen bezüglich des Histogramms können Sie zustimmen?
- (b1) Bei einem Histogramm müssen immer die relativen Häufigkeiten angegeben werden.
 - (b2) Sind die Klassen des Histogramms gleich breit, so ist es äquivalent zum Balkendiagramm.
 - (b3) Die Klassengrenzen müssen so gewählt werden, dass keine leere Klasse entsteht.
 - (b4) Das Histogramm folgt dem Prinzip der Flächentreue.
 - (b5) Fällt ein Wert genau auf die Klassengrenze, zählt er je halb zur unteren und zur oberen Klasse.
 - (b6) Bei der Entscheidung über die Klassenbreiten entscheidet im Endeffekt der subjektive Eindruck.
- (c) Im unten stehenden Histogramm ist das Merkmal 'Alter' für 20 zufällig ausgewählte Besucher eines Pop-Konzerts dargestellt. Welchen der folgenden Aussagen können Sie *nicht* widersprechen? [Anmerkung: Die Klassen sind alle genau gleich breit gewählt.]
- (c1) Der Altersunterschied zwischen der ältesten und der jüngsten Person beträgt genau 20 Jahre.
 - (c2) Der Median liegt bei 27.
 - (c3) Mehr als die Hälfte der Personen ist über 20.
 - (c4) Die Personen sind zwischen 15 und 33 Jahren alt.
 - (c5) Der Altersdurchschnitt beträgt rund 22 Jahre.
 - (c6) Es gibt nur eine Person, die älter als 28 ist.
 - (c7) Der Modus liegt bei 19.
 - (c8) Der Modus liegt bei 31.
 - (c9) Die Gruppe der 15- bis 24-Jährigen ist doppelt so stark vertreten wie die der 25- bis 34-Jährigen.



- (d) Sehen Sie (nun) auch Vorteile des Stamm-Blatt-Diagramms gegenüber dem Histogramm?

Aufgabe 7

Ende September wurde in Bayern ein neuer Landtag gewählt. Die folgende Tabelle gibt eine Übersicht über die Stimmenanteile der verschiedenen Parteien - wie aber lässt sich dieses Wahlergebnis geeignet visualisieren? [Anmerkung: Parteien, die weniger als 5% der Stimmen erhielten, wurden unter 'Sonstige' zusammengefasst.]

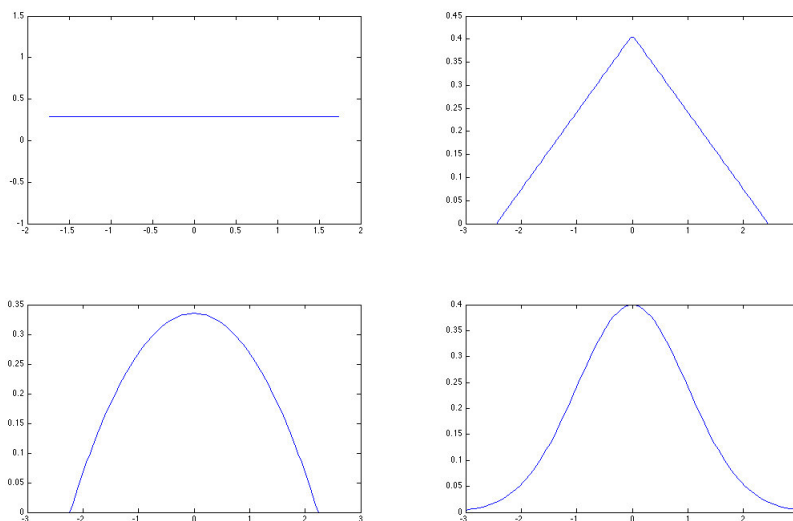
Partei	Stimmenanteil
CSU	43.4%
SPD	18.6%
FW	10.2%
Grüne	9.4%
FDP	8.0%
Sonstige	10.4%

5 Kerndichteschätzung

Aufgabe 8

Während bei Histogrammen die Klassenbreiten und besonders die Klassengrenzen über deren Form entscheiden, wird bei gleitenden Histogrammen mit einer variablen Klasseneinteilung gearbeitet. Eine Verallgemeinerung der gleitenden Histogramme wiederum stellen Kerndichteschätzer dar.

- (a) Erklären Sie das Prinzip der Kerndichteschätzung in eigenen Worten!
(b) Um welche Kerne handelt es sich bei den hier abgebildeten? Kennen Sie weitere?



- (c) Welche drei Bedingungen müssen alle Funktionen $K(u)$ erfüllen, damit sie der Definition eines Kerns genügen?
(d) Weisen Sie - ausgehend von Teilaufgabe (c) - nach, dass der Epanechnikow-Kern der Definition nach tatsächlich ein Kern ist!

Epanechnikow-Kern:

$$K(u) = \frac{3}{4}(1 - u^2) \text{ für } -1 \leq u \leq 1, 0 \text{ sonst.}$$

- (e) Gegeben seien die Daten $x = (x_1, x_2, x_3, x_4, x_5) = (1, 2, 4, 4, 6)$. Entscheiden Sie sich für eine Bandbreite von $h = 1$ und skizzieren Sie den zugehörigen Epanechnikow-Kerndichteschätzer!

Kerndichteschätzer für $f(x)$:

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right), x \in \mathbb{R}$$