

# Einfache statistische Auswertungen

Aus einer Inputquelle, Simulation etc. werden eindimensionale Daten geliefert (ich behandle hier nur eindimensionale Größen, also keine Korrelation etc.). Die beschreibende Statistik berechnet Kenngrößen, die Rückschlüsse auf die Verteilung dieser Werte zulassen. Für die Programmierung ist folgende Unterscheidung wichtig:

- 1) Die Kenngröße kann berechnet werden, ohne die Daten zu speichern.
- 2) Die Kenngröße kann nur berechnet werden, wenn alle Daten gespeichert wurden.

Beispiele für 1):

Minimum, Maximum, Mittelwert, Stichprobenvarianz

Beispiel für 2)

Median

## Minimum und Maximum

Zur Ermittlung des Minimums (das Maximum geht analog) braucht man nur eine Variable, etwa `min_val`. Jeden Inputwert `x` vergleicht man mit `min_val`. Ist `x > min_val`, macht man nichts, anderenfalls speichert man `x` in `min_val`:

```
if (x < min_val)
    min_val = x;
```

Man darf jedoch nicht vergessen, `min_val` am Beginn auf einen vernünftigen Anfangswert zu setzen. Dazu gibt es 2 sinnvolle Strategien:

- 1) Man initialisiert `min_val` mit dem ersten Inputwert
- 2) Man initialisiert `min_val` mit dem größtmöglichen Wert des Datentyps

ad 2) In C++ würde man `<limits>` inkludieren und man erhält den größtmöglichen Wert eines Datentyps (z.B. `double`) als `std::numeric_limits<double>::max()`