Varianzzerlegung / Streuungszerlegung:
Varianz bei geschichteten Daten.

Zur Erinnerung: Daten liegen oft in Schichten vor (v.a. bei Sekundär- und Tertiärerhebungen). Beispiel: Daten über Einkommensverteilung geschichtet nach Bundesland. Bei der Berechnung von $\bar{x}$ waren die einzelnen Besetzungszahlen sehr wichtig.

Schicht	$1, \ldots, l, \ldots, z$	
Besetzungszahlen	$n^{(1)}, \ldots, n^{(l)}, \ldots, n^{(z)};$	$\sum_{l=1}^{z} n^{(l)} = n$
Mittelwerte	$\bar{x}^{(1)}, \ldots, \bar{x}^{(l)}, \ldots, \bar{x}^{(z)}$	
Varianzen	$\tilde{s}^{2^{(1)}}, \ldots, \tilde{s}^{2^{(l)}}, \ldots, \tilde{s}^{2^{(z)}}$	

Für das arithmetische Mittel gilt

$$\bar{x} = \frac{1}{n} \sum_{l=1}^{z} n^{(l)} \bar{x}^{(l)}.$$

Seien nun

$$\tilde{s}^2_{innerhalb} := \frac{1}{n} \sum_{l=1}^{z} n^{(l)} \tilde{s}^{2^{(l)}}$$

sowie

$$\tilde{s}^2_{zwischen} := \frac{1}{n} \sum_{l=1}^{z} n^{(l)} (\bar{x}^{(l)} - \bar{x})^2$$

Varianzzerlegung Es gilt

$$\begin{array}{ccccc} \text{Gesamtvarianz} & = & \text{Varianz in. d. Schichten} & + & \text{Varianz zw. d. Schichten} \\ \tilde{s}^2 & = & \tilde{s}^2_{innerhalb} & + & \tilde{s}^2_{zwischen}. \end{array}$$