Descriptive statistics and diagrams

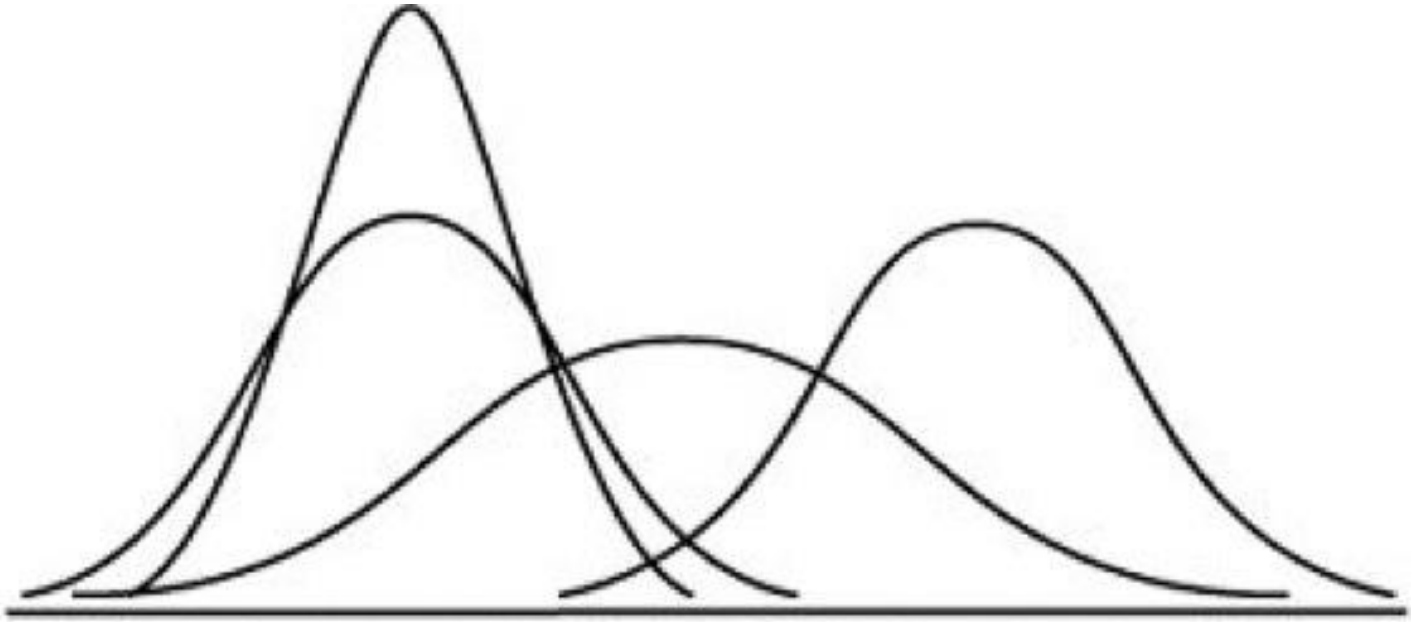
Lara Kataja

Room 1.13

l.kataja@edutech.uni-saarland.de

Mean and measures of dispersion

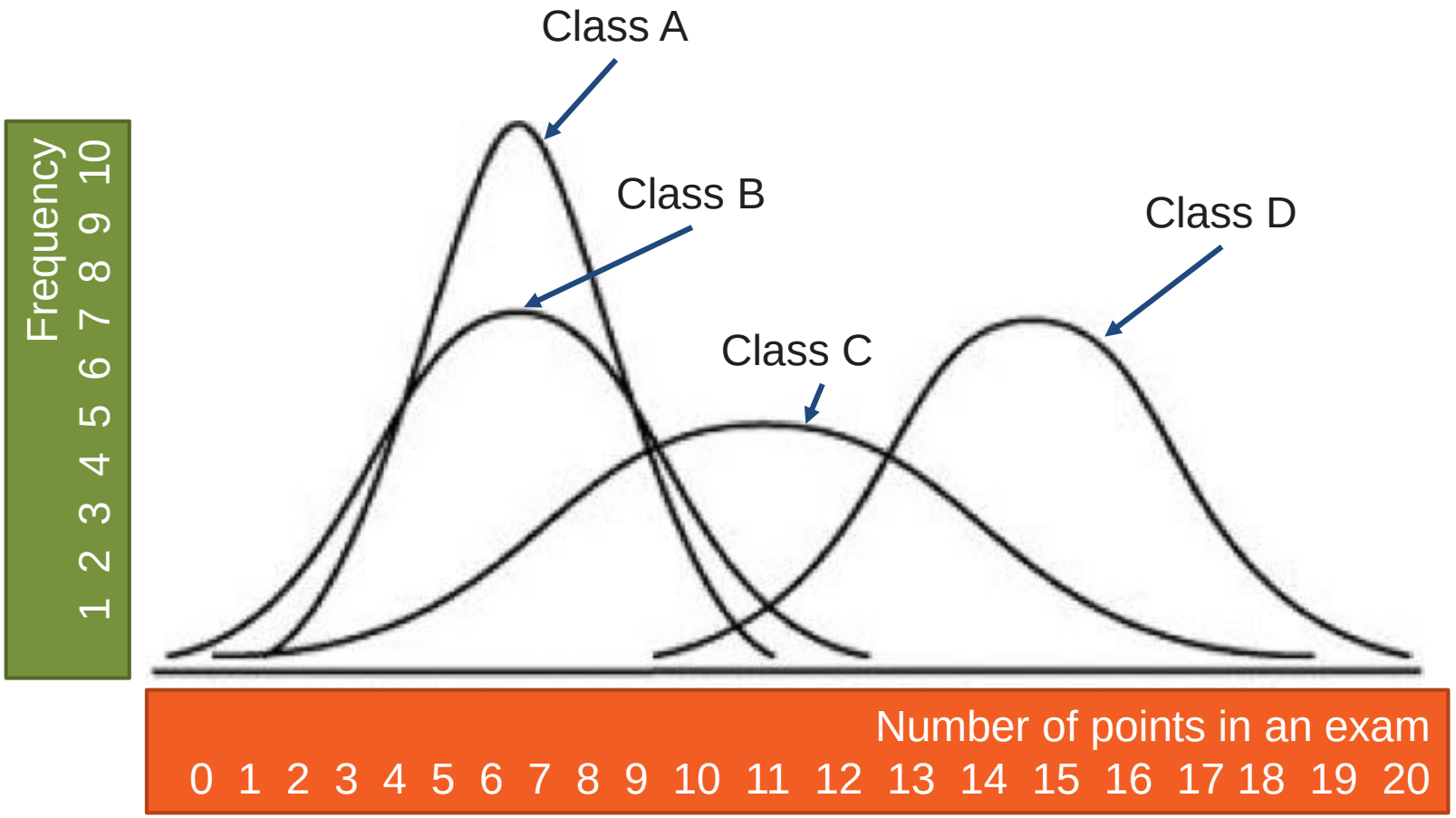
Why do you need measures of dispersion, e.g., standard deviation (σ), additionally to the mean (μ)?



(Normal distributions with different μ und σ)

http://www.spektrum.de/lexika/images/psycho/f2f209_w.jpg

Mean and measures of dispersion



http://www.spektrum.de/lexika/images/psycho/f2f209_w.jpg

Measures of central tendency

Data: 5, 4, 5, 2

Mean

$\frac{5+4+5+2}{4} = \frac{16}{4} = 4$

Mode

5, 4, 5, 2

→ most frequent: 5

Median

Order the values: 5, 4, 5, 2 → 2 4 5 5

→ median = 4.5

Measures of central tendency

Data: 5, 4, 5, 2, 20

→Mean = $\frac{5+4+5+2+20}{5}$ = 7.4

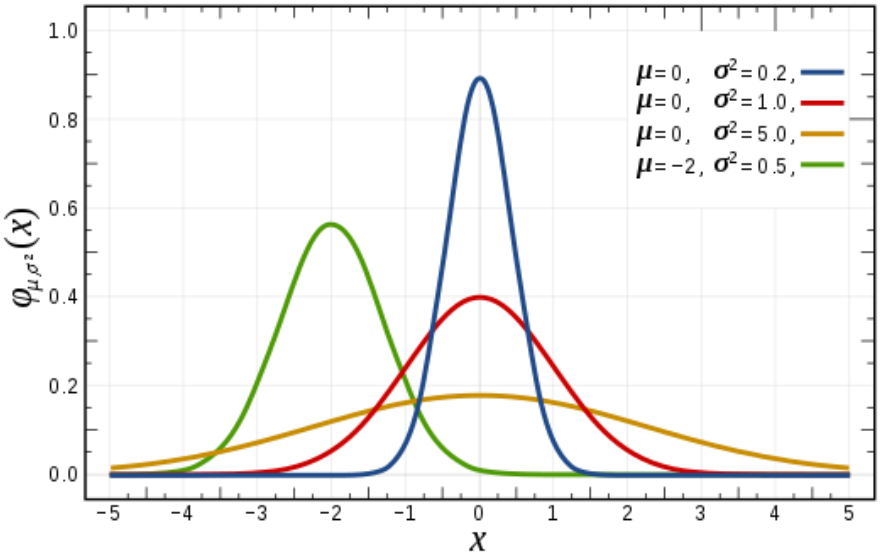
→Mode: 5, 4, 5, 2, 20 → 5

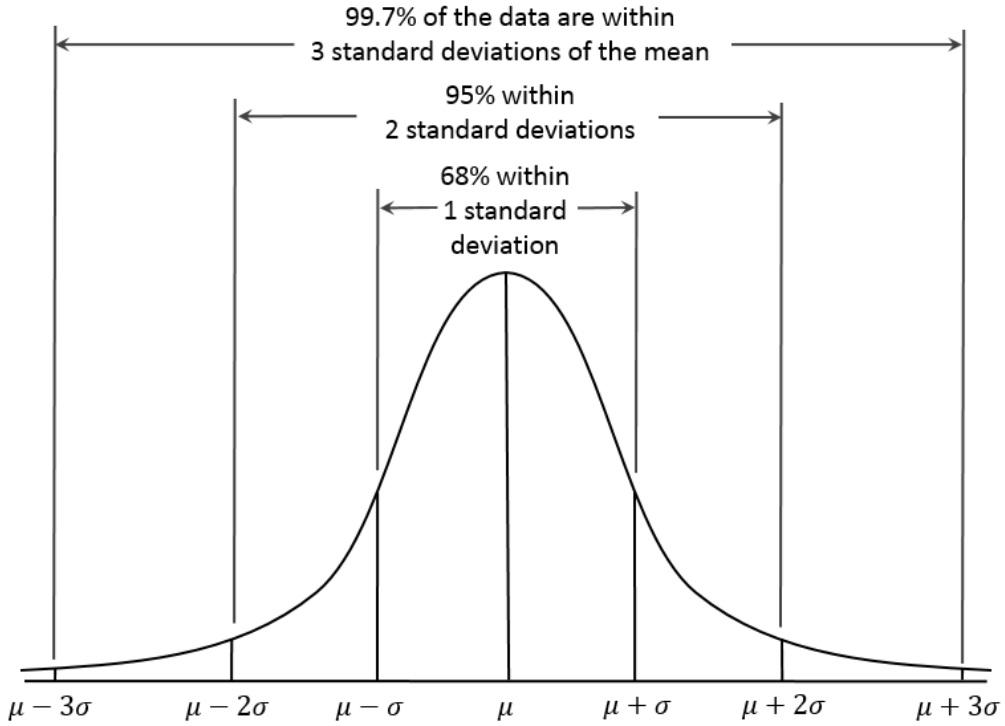
→Median: 2 4 5 5 20 → 5

→Median and mode are less sensitive to extreme values (compared to the mean)

Normal distribution

- A probability distribution that underlies many variables in the population (e.g. intelligence)
- Symmetric, unimodal, around 68% of the values are within the first SD, around 95% within 2 SD, and around 99% within 3 SD
- Normality is a basic assumption for many statistical tests





(**Normal distribution.** In *Wikipedia*. Retrieved December 09, 2019, from https://en.wikipedia.org/wiki/Normal_distribution)

Variance and standard deviation (SD)

- Variance (s^2) = sum of the squared deviations of the mean, divided by n

- Standard deviation (SD or s) = square root of the variance

Variance: Example

$$s^2 = \frac{\sum (X - \bar{X})^2}{N - 1}$$

Data: 5, 4, 5, 2

→ Mean = $\frac{5+4+5+2}{4} = 4$

$$s^2 = \frac{(5-4)^2+(4-4)^2+(5-4)^2+(2-4)^2}{4-1} = \frac{1+0+1+4}{3} = \frac{6}{3} = 2$$

→ Variance $s^2 = 2$

→ Standard deviation $s = \sqrt{2} = 1{,}4$

Example: Reporting Mean + SD

they posted. The control group ($M = 36.6$, $SD = 9.91$) and individual preparation-no script ($M = 30.86$, $SD = 9.26$) posted more, followed by argumentation scripts ($M = 19.19$, $SD = 6.84$) and the combination condition ($M = 15.26$, $SD = 6.11$). To take these differences

Tsovaltzi, D., Judele, R., Puhl, T., & Weinberger, A. (2017). Leveraging social networking sites for knowledge co-construction: Positive effects of argumentation structure, but premature knowledge consolidation after individual preparation. *Learning and Instruction*, *52*, 161-179.

Table 1. Formal argumentative dimension by experimental group: Mean percentages and standard deviations of grounds and counterarguments.

	Grounds		*Counterarguments*	
Experimental group	*M*	*SD*	*M*	*SD*
Control group	12.08 %	11.48	2.46 %	3.67
Script for the construction of single arguments	33.80 %	11.19	5.36 %	8.07
Script for the construction of argumentation sequences	16.36 %	17.78	5.99 %	3.95
Combined condition	30.64 %	6.10	13.00 %	6.59

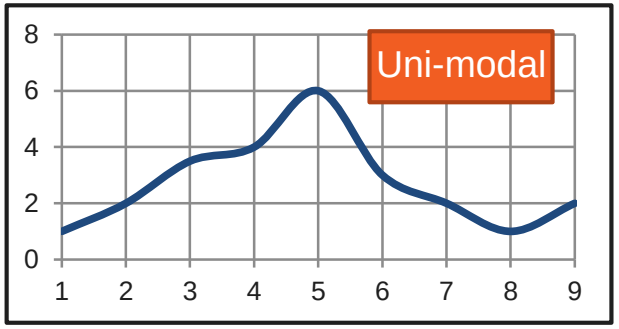
Weinberger, A., Stegmann, K., & Fischer, F. (2005). Computer-Supported Collaborative Learning in Higher Education : Scripts for Argumentative Knowledge Construction in Distributed Groups. In *The Next 10 Years! Proceedings of the 2005 Conference on Computer Support for Collaborative Learning, CSCL '05* (pp. 717–726).

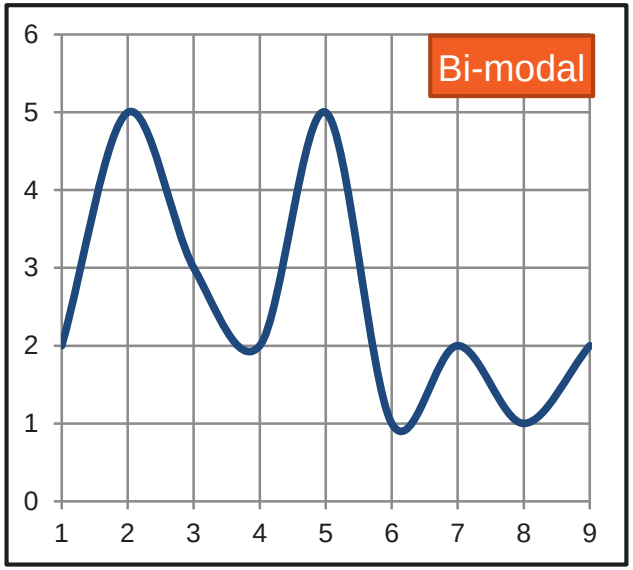
Mode

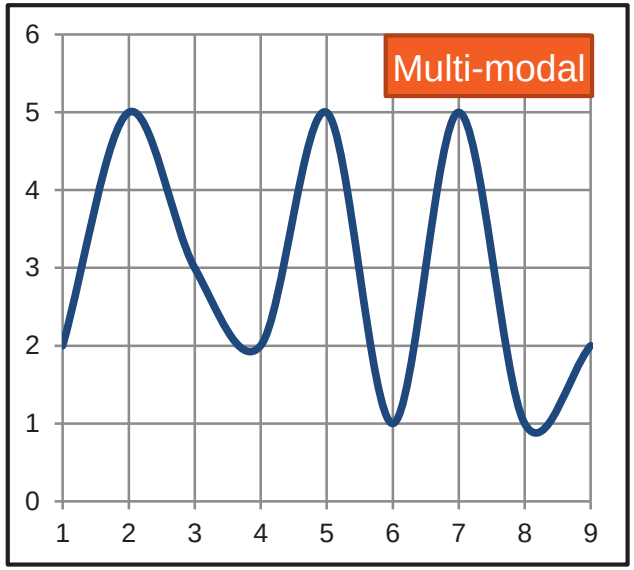
What if there are several values with the same high frequency? Which one is the mode?

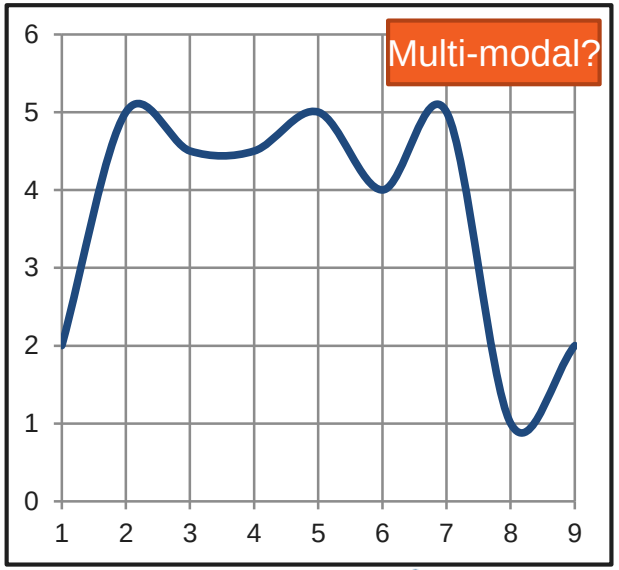
<u>Two approaches:</u>

- “bi-modal“, “multi-modal“ vs.
- “cannot be determined “



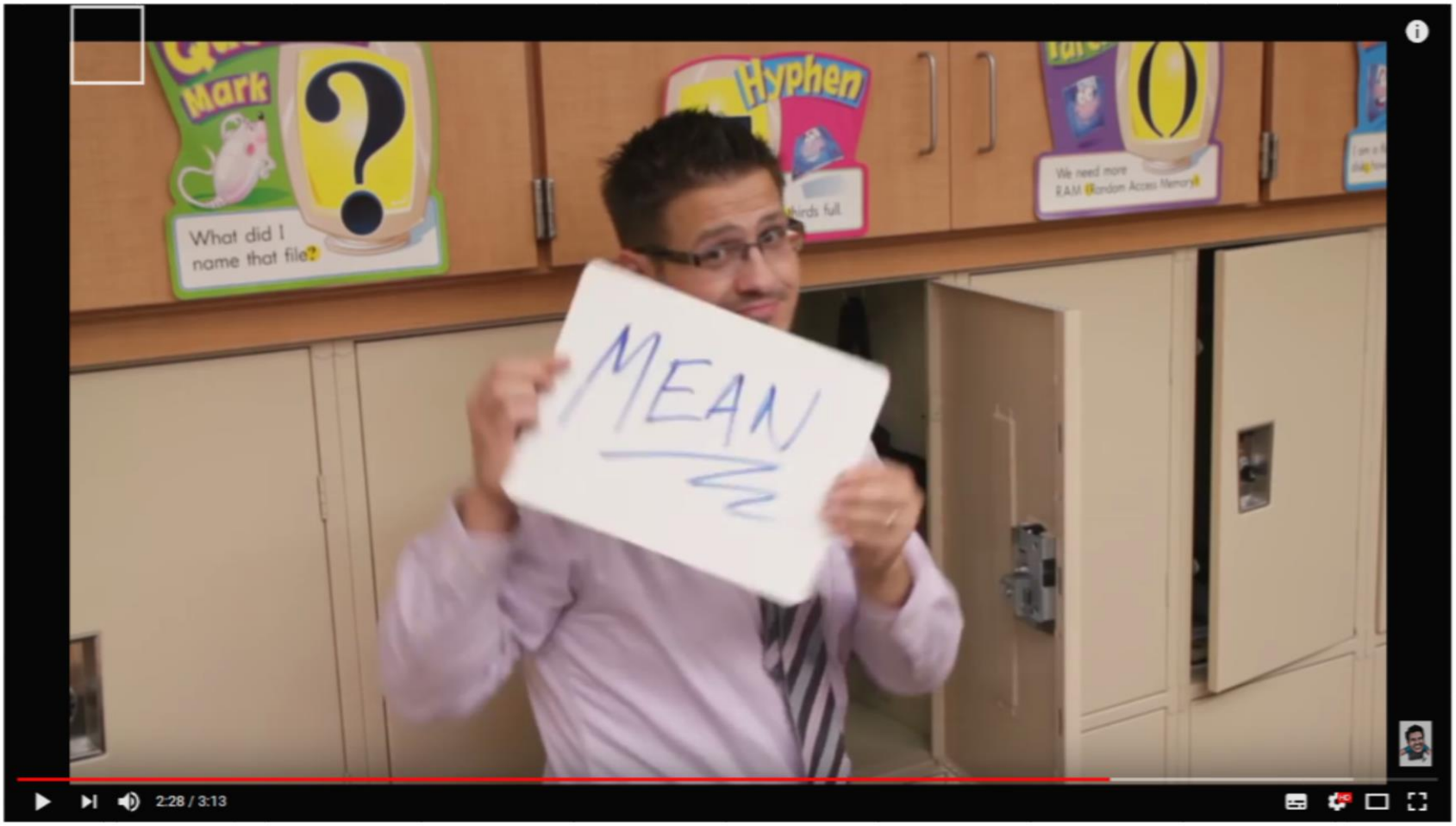






Mean, Median & Mode

https://www.youtube.com/watch?v=A7MxGyEaN64



Exercises (SPSS)

- Open Beispieldatensatz
- Calculate: Mean, median, and mode for “positiv“

  → Analyze > Descriptive Statistics > Frequencies

Statistics

positiv

N	Valid	150
	Missing	0
Mean		3,44
Median		3,00
Mode		2

edutech

Exercises (SPSS)

- Open Beispieldatensatz
- Calculate: Mean and standard deviation for “positiv“

  → Analyze > Descriptive Statistics > Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
positiv	150	0	10	3,44	2,071
Valid N (listwise)	150				

Exercises (SPSS)

- Open Beispieldatensatz
- Build a bar chart with ‘bed’ as independent variable (x axis) and ‘positive’ as dependent variable (y axis)

  →Graphs > Chart Builder