

IPRES Tutorial 9: Descriptive Statistics II



Milan Babic

m.babic@uva.nl

Department of Political Science - Political Economy and Transnational Governance

26 April 2019



Agenda

- 1.) Check-in and recap
- 2.) Spreads and Standardization
- 3.) Exercise
- 4.) Wrap-up and questions

Check-in (also Google doc)



Where we are

Part 2:

- Prerequisites of quantitative research: Experiments and comparative case studies (**Tutorial 7**)
- Descriptive Statistics I: LoM and data hands-on (**Tutorials 8**)
- Descriptive Statistics II: Spreads (**Tutorial 9**)
- Samples and Sampling (**Tutorial 10**)
- Inferential Statistics (**Tutorial 11**)
- Limits of numbers and ethics/assignment prep. (**Tutorial 12**)

Recap: Levels of measurement

| Categorical | | Quantitative | | | |
|-------------|--|--------------|------------|----------|------------|
| | | | | | |
| | | discrete | continuous | discrete | continuous |

Recap: Levels of measurement

| Categorical | | Quantitative | | | |
|-------------|---------|--------------|------------|----------|------------|
| Nominal | Ordinal | Interval | | Ratio | |
| | | discrete | continuous | discrete | continuous |

Spreads



Spreads

Three ways of making sense of our data:

- Central tendencies (Tutorial 8)
- Spreads: Range, IQR and SD
- Symmetries

Spreads: Number of legs

| Sample 1: Humans | Sample 2: Aliens |
|------------------|------------------|
| 2 | 0 |
| 2 | 2 |
| 2 | 1 |
| 2 | 5 |
| $\sigma = 2$ | $\sigma = 2$ |

Spreads: Number of legs

| Sample 1: Humans | Sample 2: Aliens |
|------------------|------------------|
| 2 | 0 |
| 2 | 2 |
| 2 | 1 |
| 2 | 5 |
| $\sigma = 2$ | $\sigma = 2$ |

Range, IQR and
SD?

Spreads: How to calculate SD

$$\sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2}{N-1}}$$

σ = SD

x_i = observations (all!)

\bar{x} = mean

N = number of observations

Spreads: Exercise

- Money (in Euro) that students have at their disposal (monthly):

840, 680, 490, 890, 540, 620, 770, 590, 830, 600, 900, 650

- Calculate:
 - What is the mean and median?
 - What is the range, IQR and variance/SD?
 - What type of distribution (normal, uniform, skewed?)

Spreads: Problems of SD

- SD increases with variability (sq. root of the *average* distance from the mean)
- Hence: Outliers affect SD
- How to account for this?
 - E.g.: Visualization through boxplots: <https://www.youtube.com/watch?v=b2C9I8HuCe4>

Spreads: Problems of SD

- SD increases with variability (sq. root of the *average* distance from the mean)
- Hence: Outliers affect SD
- How to account for this?
 - E.g.: Visualization through boxplots: <https://www.youtube.com/watch?v=b2C9I8HuCe4>
 - Missing in the video: skewness of distribution and outliers

Transforming data: standardization

$$Z_x = \frac{X_i - \overline{X}}{S_x}$$

Z_x = z-score

X_i = observation (one!)

\overline{X} = mean (of population)

S_x = SD (of pop.)

Exercise: see handout

- **First** round: everyone on their own
- **Second** round: in pairs
- **Third** round: discussion in class

Sources (all pictures CC license)

Image Wave Emoji: https://upload.wikimedia.org/wikipedia/commons/thumb/4/48/Emojione_1F44B.svg/2000px-Emojione_1F44B.svg.png

Image Pindakaas: https://farm1.staticflickr.com/33/36543980_b0f2d8081f_b.jpg