

2 Plan design & data: Assignment

Instructions

Submit your answers to the eight first exercises through Aula Global. The remainder of your assignments concern self-studies.

Describing a study

Quick reminder from class: Kanwal et al. (2017) taught subjects an artificial language. This language only has three (made up) words: *zopudon*, *zopekil* and *zop*. The first two words referred to distinct objects (think: *zopudon* means apples and *zopekil* means bananas). The short form *zop* could mean either (think: fruit). That is, *zop* is ambiguous between the two meanings, and thus can lead to misunderstandings. They had subjects communicate about the two meanings (apples/bananas) in pairs, alternating who was the speaker and who was the receiver. The core motivation for this study was to see if people would use the ambiguous word, *zop*, even though it is more risky than the unequivocal but longer alternatives *zopudon* and *zopekil* under certain experimental manipulations.

Here's a glimpse of the data of the speakers:

```
df <- read.csv('https://tinyurl.com/2s3p9s2z')
head(df)
```

##	pairnum	IP	trial	display	label
## 1	1	67.85.42.18	1	0	zop
## 2	1	67.85.42.18	2	3	zopudon
## 3	1	67.85.42.18	3	0	zop
## 4	1	67.85.42.18	4	0	zopekil
## 5	1	67.85.42.18	5	2	zopudon
## 6	1	67.85.42.18	6	1	zopekil

The column **pairnum** identifies each pair of subjects (alternating speaker and hearer roles); the **IP** column identifies each subject; **trial**s keep track of the order in which the turns happened (trial 1 is the first time a subject is a speaker, trial 2 is the second, and so on); **display** codes whether one type of object (0/1) or another type of object (2/3) was displayed to the speaker; and **label** shows what the speaker actually said to communicate this object.

1. What kind of study is this? Observational, experimental, or simulation?
2. What kind of variable is **pairnum**?
3. What kind of variable is **trial**?
4. What kind of variable is **label**?

Inspect the sender data from Kanwal et al. (2017) yourself.

5. How many times did the sender with IP 67.85.42.18 say *zop*?
6. How many unique pairs participated in the experiment?
7. How many sender trials did the experiment have for each subject?
8. Is this data *tidy*, *untidy* or *almost tidy*?

Self-study

1. Work through Chapter 4 and 6 of Introduction to Data Analysis (Franke 2021). If you need a refresher on descriptive statistics, also look at Chapter 5;
2. Apply the terminology from this session to your analysis question of interest. Change your research question if you have come across another that you find more interesting

References

- Franke, Michael. 2021. *An Introduction to Data Analysis*.
- Kanwal, Jasmeen, Kenny Smith, Jennifer Culbertson, and Simon Kirby. 2017. “Zipf’s Law of Abbreviation and the Principle of Least Effort: Language Users Optimise a Miniature Lexicon for Efficient Communication.” *Cognition* 165: 45–52. <https://doi.org/10.1016/j.cognition.2017.05.001>.