Data Output/Access

Data Analysis
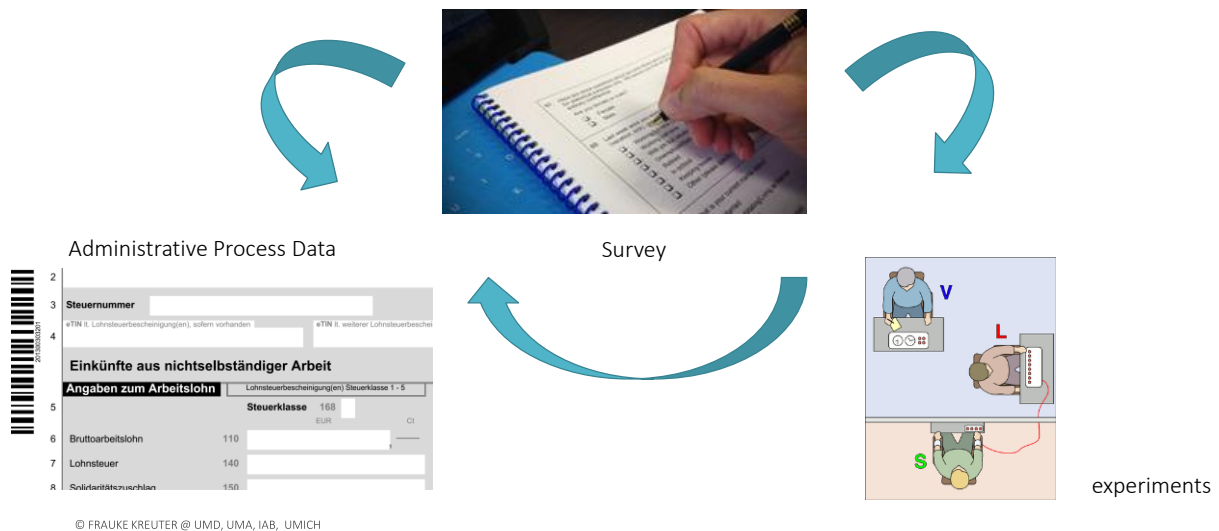
Data Curation/Storage

Data Generating Process

Research Questions

Beyond Stats 101

- Means, percentages, totals
- Differences in means
- Regressions
- Clustering
- .....

and a few challenges

How to combine/link different data sources?

Administrative Process Data

Survey

experiments
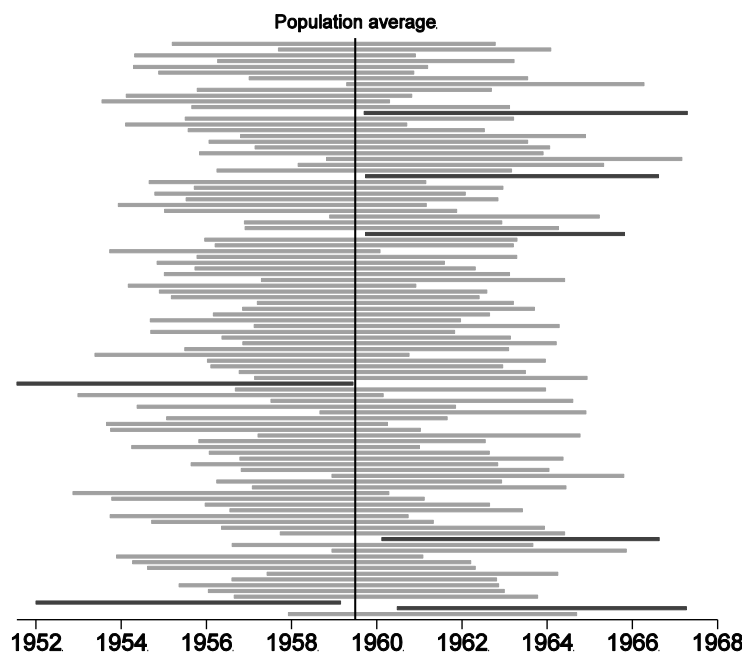

How to properly capture variability?
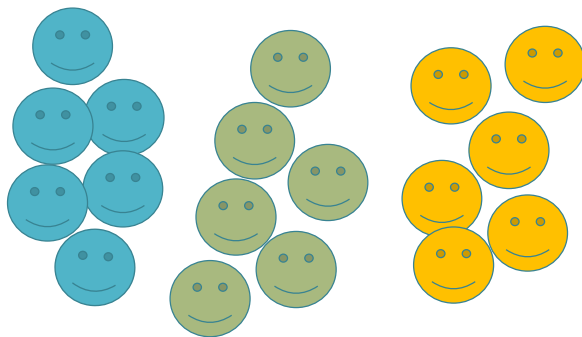
Or check this out https://economicsnetwork.ac.uk/statistics/sample_means_households.htm

Many Samples

Each sample is presented here with a confidence interval around the mean. For now you can think of this as the mean of the sample with an error band. How that is estimated does not matter for the purpose of this demonstration.
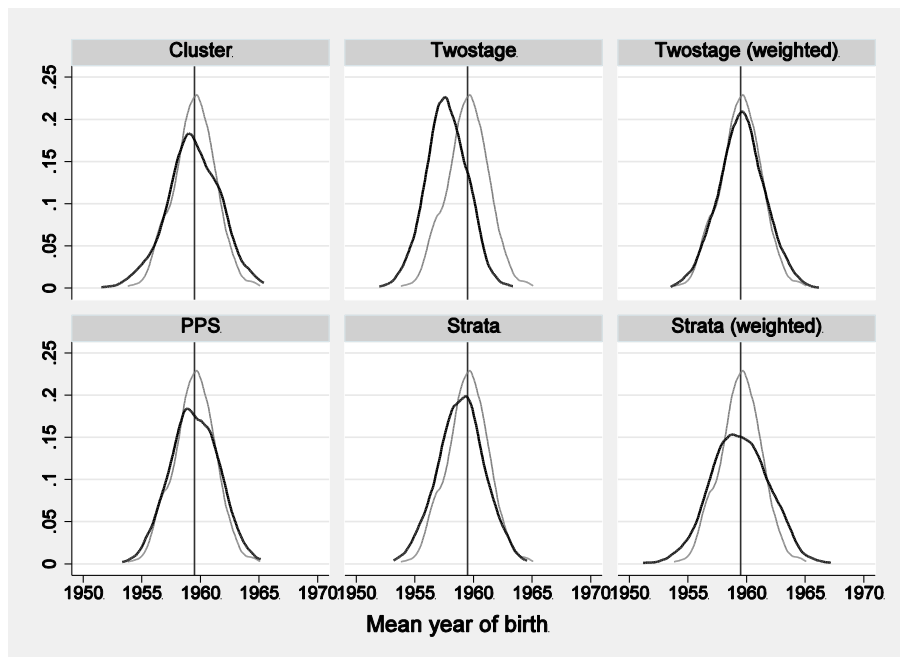
Source:
Kohler & Kreuter 2012

Population average

| | | | | | | | | |
|1952|1954|1956|1958|1960|1962|1964|1966|1968|

Distribution of Means

Source:
Kohler & Kreuter 2012

Cluster   Twostage   Twostage (weighted)

PPS   Strata   Strata (weighted)

Mean year of birth

Resources
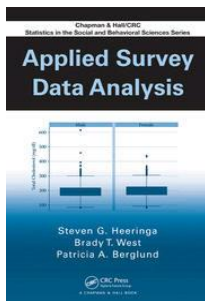
Our course:
Combining and Analyzing Complex Data