

Academic writing and scientific data collection

Dr. Hwayeong Jeong

Serhat Demirtaş

Prof. Dr. Jamie Paik

Reconfigurable Robotics Laboratory

EPFL, Switzerland

Evaluation

70% Project (20 % Final Report, 20 % Final Presentation, and 30 % Working Prototype)

Scientific data collection

Data

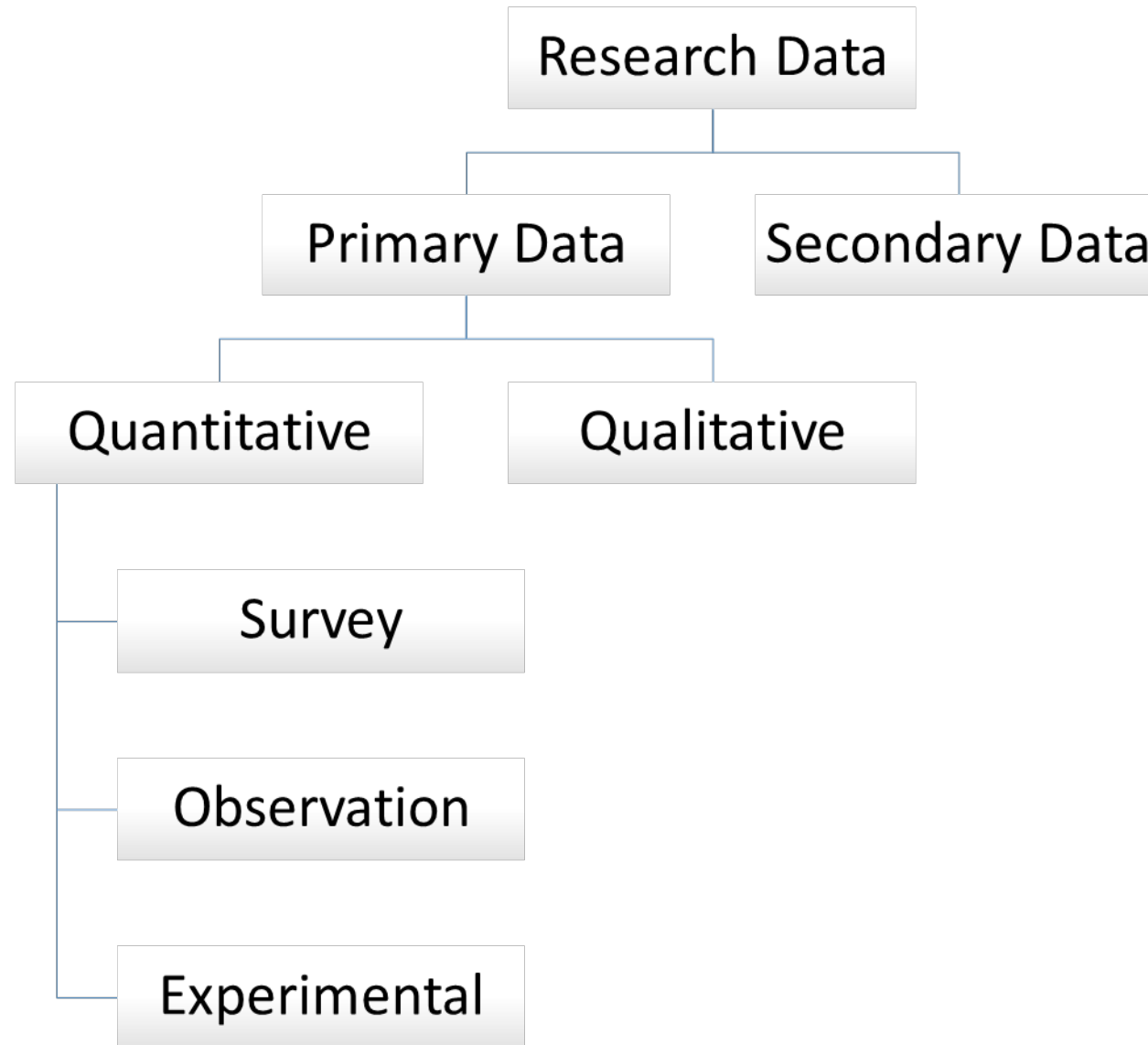


- **Data** is a collection of **discrete or continuous values** that convey **information**, describing the **quantity, quality, fact, statistics**, other basic units of meaning, or simply sequences of symbols that may be further interpreted formally.



- Factual information (such as **measurements** or **statistics**) used as a basis for **reasoning, discussion, and calculation**.

Types of data





Quantitative or **numerical data** are the values of numerical characteristics associated with items in a sample.



These are typically either counts of the number of occurrences of a phenomenon of interest or measurements of some physical property of the items.

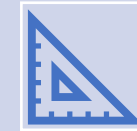
Data collection



What are you measuring? Why?



Which data to collect?



How to collect?



Who to collect?



How much to collect?



When to collect?



Measurement

Data collection - measurement



Make **operational definitions**. What and why?



Choose right equipment.



«A **measurement or measuring method** is called **valid** if it **usefully or appropriately represents** the **feature** of an object or system that is of engineering importance.»

Data collection

Measurement

Sampling

Data collection - sampling

“How much data do I need?”

“It depends.”

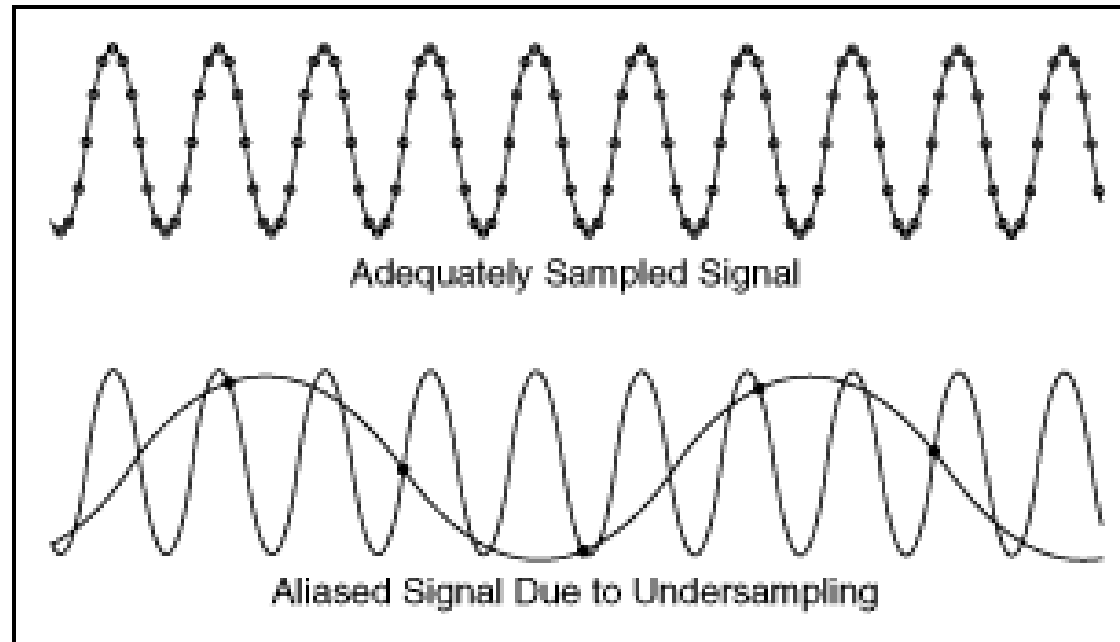
“How many repetitions do I need?”

“It depends.”

“On what?”

Data collection - sampling

- Sampling frequency refers to how often data is collected over time.
- **Too low:** You may miss important changes or underestimate variability (aliasing).
- **Too high:** May lead to redundant data, larger storage needs, and unnecessary complexity.



Data collection

Measurement

Sampling

Recording

Data collection - recording

- The format of data recording significantly affects its usability.
- Search for common tools.

Data collection - recording

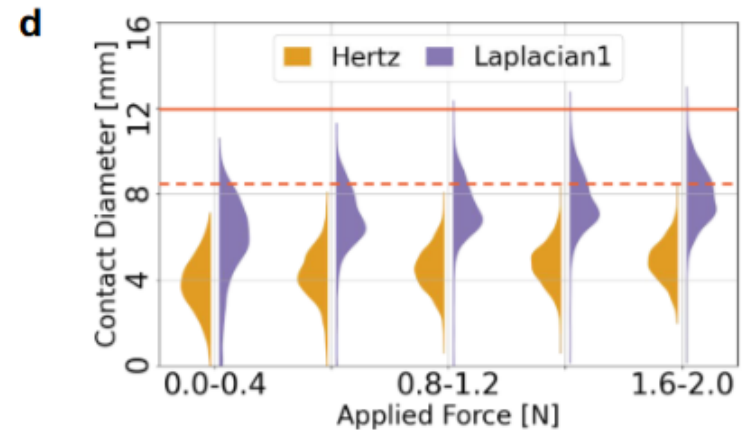
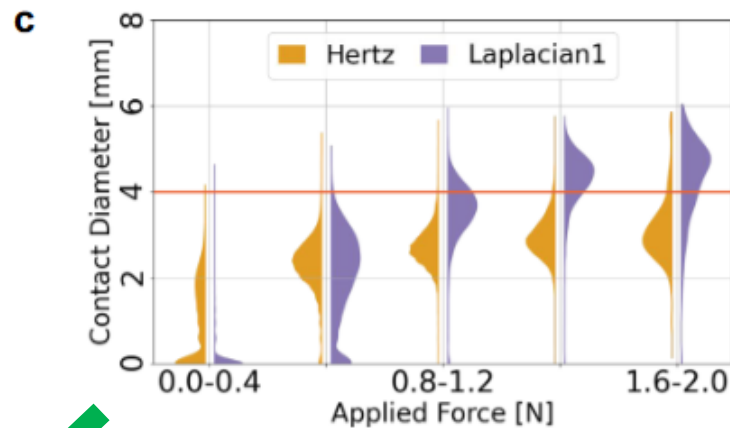
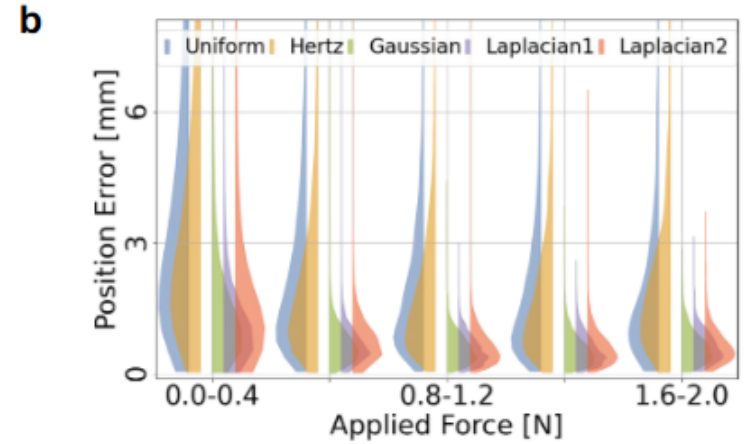
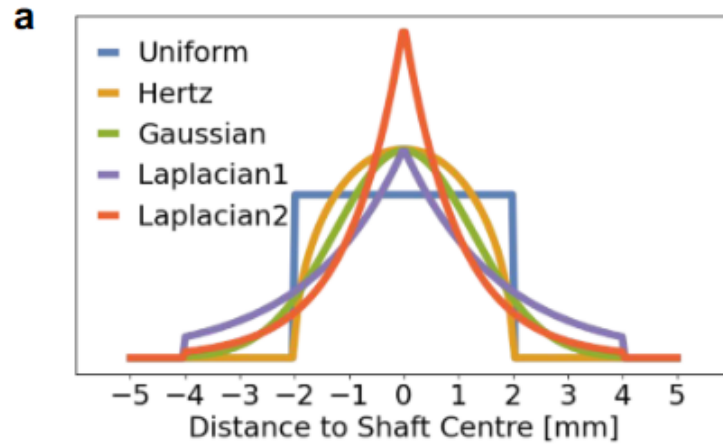
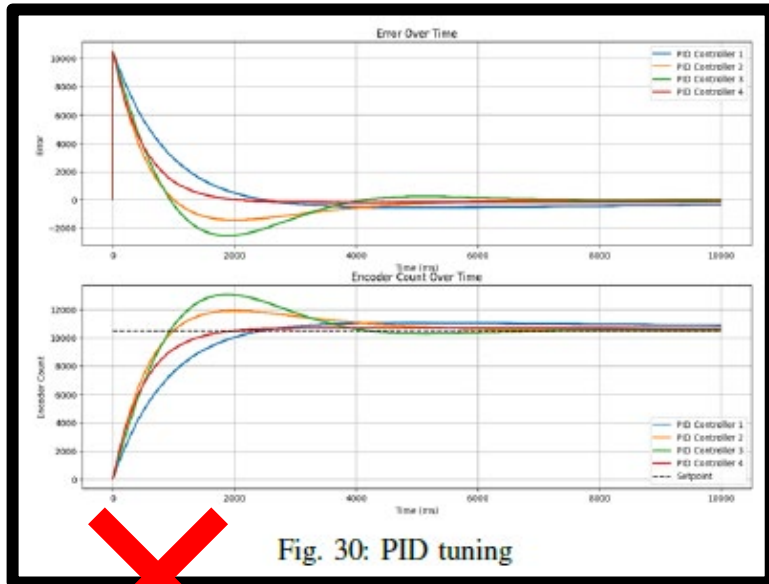
- Be very careful about the **visualization**
- Choose the right plot for the data type

Distribution of values	Histogram, Boxplot
Relationship between variables	Scatter Plot
Change over time	Line Chart
Comparison between groups	Bar Chart
Confidence intervals	Error Bars

Data collection - recording

- Use consistent axis scales, meaningful labels, and units
- Color responsibly: avoid using red-green combinations for accessibility
- Font size in figures should exceed (match at the least) the smallest font in your main text
- Don't hesitate to combine multiple related figures with subfigures

Data collection - recording



Under data Fig. 2 | Force distribution approximation and evaluation. **a**, The five approximate force distribution methods that we tested for the 4 mm hemi-spherical indenter. The illustration is in one dimension and was revolved to distribute the measured contact force vector across the local surface the force map around the contact point. The uniform (blue) and Hertz theory^[4] (orange) distribution curves are strongly localized with a radius of 2 mm. The green curve follows a truncated Gaussian distribution, and the two other curves (purple and red) follow Laplacian distributions that also stop at a radius of 4 mm. **b**, The position inference performance using the different methods. **c** and **d**, The diameters of the contact area prediction for two indenter sizes (4 mm and 12 mm) using two different approximation methods (Hertz and Laplacian1). The dashed line indicates a reference diameter from complementary Fig. 5 a) based on the fact that the large indenter penetrates into the sensor only partially.

- Take pictures
- Take notes
- Label your data
- Label your samples
- Avoid common mistakes
 - Just bringing the video and say “yes it is working”
 - Deleting the raw data, never do that!
 - Poor operational definitions – ‘better’, ‘efficient’
- CONTACT THE TAs

Reference

- Vardeman, S., & Jobe, J. M. (2001). Basic engineering data collection and analysis.



serhat.demirtas@epfl.ch