Data Collection
ENGR 1181
Class 7
Data Collection in the Real World

Data is collected all of the time, just think about it. When you are at the grocery store using your “loyalty card” or when the Google maps SUV is taking videos of your street.
Today's Learning Objectives

After today’s class, students will be able to...

• Understand the importance of collecting high-quality data.
• Recognize when a data set contains errors.
• Define accuracy, precision, and resolution.
• Define systematic variation and random variation.
Today's Learning Objectives

After today’s class, students will be able to...

- Identify how systematic variation and random variation might influence data collection.
- Choose an appropriate type of measurement when given a specific problem statement
- Choose an appropriate measurement tool for the type of measurement that will be collected
Today’s Plan

- Review of important terms from prep reading
- Data collection activity
- Analysis of data:
  - Accuracy, repeatability, resolution, outliers
  - Errors & variation: what types do we have?
  - Error propagation: how do errors affect subsequent calculations with our data?
Review of Important Terms

Accuracy

*The degree of closeness between a measured value and desired value.*
Review of Important Terms

Precision or Repeatability

*The degree of closeness between a set of measured values;*

*The ability of an instrument to reproduce a measurement*
Review of Important Terms

Systematic Variation

The variation of measured values that differ from the nominal value in a predictable pattern
Random Variation

*The unpredictable variation that occurs when an object is measured multiple times*
Data Collection Activity

The Scenario:

You are a product design engineer working with the RPAC to design a new vending machine that sells small bottles of shampoo.

You have designed most of the machine’s components already, but now you need to determine how many bottles of shampoo the machine can hold when full.
Data Collection Activity

Work in pairs to:

1. Choose one of the available measuring devices.

2. Take measurements of each type of bottle.

3. Write down your measurements.
Data Collection Activity

What did you measure to find the size of the bottle?

Why did you choose this aspect to measure?

Which measurement device did you use?

Did the device affect your results?
Data Collection Activity

Let’s measure with purpose.

We need to know how many bottles will fit in the chute that dispenses the product, which means we need to know the **diameter** of the bottle.

Now have each person measure the diameter of the cylindrical bottle using the calipers. Record the results.
Sample Data

- 8 sample measurements were collected
- The expected value is 1.020 inches
- Note the difference in resolution of the dial calipers and the scale
  - Dial calipers can be read to 3 decimals
  - Scale can be read to 1 decimal place

<table>
<thead>
<tr>
<th>Trial</th>
<th>Scale</th>
<th>Caliper</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.2</td>
<td>1.011</td>
</tr>
<tr>
<td>2</td>
<td>1.1</td>
<td>1.039</td>
</tr>
<tr>
<td>3</td>
<td>1.0</td>
<td>1.005</td>
</tr>
<tr>
<td>4</td>
<td>1.1</td>
<td>1.015</td>
</tr>
<tr>
<td>5</td>
<td>1.0</td>
<td>1.024</td>
</tr>
<tr>
<td>6</td>
<td>1.1</td>
<td>1.029</td>
</tr>
<tr>
<td>7</td>
<td>1.1</td>
<td>1.031</td>
</tr>
<tr>
<td>8</td>
<td>1.2</td>
<td>1.024</td>
</tr>
</tbody>
</table>
Sample Data

- Notice how the measurements fall above and below the specified value of 1.02”

- Which measurement tool provided more accurate values?

- How did the geometry of the bottle affect your ability to measure with the scale?
Accuracy & Repeatability

- Which measurement offer good accuracy?
- Which measurements offer good repeatability?
Systematic Variation

What are some possible sources of systematic variation in this data set?

- Human?
- Instrumentation?
Random Variation

Random variation is the unpredictable difference between measure and true values.

What are some possible sources of random variation in this data set??
Outliers

Outliers are values that differ greatly from the specified value.

Do we have any outliers here?

Outliers can occur from systematic or random variation, or they can be a misrepresented value.

The important thing is to figure out why the values are so different.
Propagation of Error

- The errors from inaccurate measurements can be magnified when the erroneous value is used in an equation.

- What seems like a small error in the initial measurement can propagate to a large overall error in the calculation of an equation.
Example: Propagation of Error

To calculate the moment of inertia of a cantilever beam, we must measure thickness and width of the beam.

Moment of inertia is calculated by:

\[ I = \frac{1}{12}wt^3 \]

- What if there is an error measuring \( w \)?
- What about an error measuring \( t \)?
Important Takeaways

- Knowing what to measure, and which tool to measure it with, is very important!

- Variation will occur in measured data, so it is important to know its possible sources.

- Small errors in measurement can magnify into larger errors of subsequent calculations.
Preview of Next Class

- Analyzing Measurement Data
  - Now that we know how to properly collect data, we will look how to analyze it.
  - We will use basic statistical methods and graphically represent our data.