

# Finding out that 9 out of 10 Dentists Do Agree

## Data Collection

Charlie Chiccarine

PaTTAN Computer Science Praxis Prep

# Data Collection

## **What is it?**

Collecting data for a specific purpose

## **Why would we do it?**

Research, get information or feedback on a product

# Publicly Available Data Sets

- Saves time in data collection in the fact that you don't actually have to collect the data
- Have to check that the data is credible
- Good for use in the classroom to practice manipulating data
- [Kaggle Datasets](#)
- [Data.gov](#)

# Types of Data

- Many different types of data and information
- Each is applicable to its own situation
- Numbers vs words vs true/false

# Types of Data - Sensors

- In engineering and electronics, sensors provide lots of data to us and to the circuits they're used in
- Digital Signals
  - On or off signals, 0 or 1
  - Buttons
  - Switches
- Analog Signals
  - Provides a range of signals, 0-x
  - Light level sensors
  - Microphones

# Types of Data - Surveys

- In statistics, we use survey data to get information from a group of people
- Numerical Responses
  - "How old are you?"
  - Can perform mathematical operations on this data
  - Mean, Median, Range
- Multiple Choice Responses
  - "Choose, from this list, where you live"
  - Easy to tally up responses and perform operations on them
- Open Ended Responses
  - "What is your opinion on x"
  - Can be answered with anything
  - Much harder to automate responses analysis

# Crowdsourcing

- Collecting information from a group on the internet
- Companies use this to survey how a product is doing
- Waze is an app that uses crowdsourced information to share travel times
- StackOverflow is a place you can ask a question and get an answer from a crowd
- Requires crowd participation to be successful

# Applicability

- It's important to have credible, appropriate, and accurate data
- If data is not applicable, then it becomes useless
- When surveying, make sure the format of the questions as well as the material is beneficial
- When asking for numerical data, don't use and open ended response



# Recap

- Data is useful in many situations
- Data is only useful if collected and used properly