

## Introduction to Univariate Data

Definitions & Examples:

- ❖ Variable: Any characteristic whose value may change from one individual to another
- ❖ Data: Results from making observations on the variable
- ❖ Univariate Data Set: Observations on a single variable
  - Categorical: non-numerical, values can't be interpreted as numbers
    - Example: hair color, car model
  - Numerical: Data that is counted or measured
    - Example: # of siblings, # of hours, money
    - Discrete: Data that is counted. Possible values are isolated points
      - # of shoes a person owns, # of speeding tickets
    - Continuous: Data that is measured. All possible values from an entire interval of observations
      - Height, temp, Time... Things where in between values make sense
- ❖ Bivariate & Multivariate Data Sets
  - Bivariate - observations on two variables
    - ex: Height and weight of football players
  - Multivariate: observations on two or more variables
    - ex: Heart, weight, heart rate of soccer players

determine whether they are discrete or continuous.

CAT or Numeric. For Numeric, discrete or continuous

1. Brand of vehicle purchased by a customer **C**
2. Price of a CD **N, D (can't go between pennies)**
3. Number of students in a class of 30 who prefer Peanut M&Ms over Plain M&Ms. **N, D**
4. Phone number of all the students enrolled in school **Cat**
5. The length of a 1-year-old bird **N, Continuous**

II. For the following numerical attributes, state whether each is discrete or continuous.

1. The number of suitcases lost by an airline **N, Discrete**
2. The height of corn plants **Continuous**
3. The number of ears of corn produced **Discrete**
4. The number of red Skittles in a bag **Discrete**
5. The time it takes for a car battery to die **Continuous**
6. The production of tomatoes by weight **Continuous**

III. For each of the following situations, give some possible data values that might arise from making the observations described.

1. The brand of cookies purchased by each of 20 customers.

**Crisps Ahoy  
oreo**

2. The shoe brand for each of the next 10 runners to cross the finish line is observed.

**Nike, Adidas, New Balance**

3. Fifteen television stations are monitored during a 2-hour period, and the amount of time devoted to commercials is determined for each.

**8, 10, 12.5**

**16, 20, 24, 40 mins (esp w/ movie near end!)**

IV. For your statistics project, you must collect numerical data. Which, if any, of the situations observed in Part III above could you use for this project?

**Commercial times**