



### Key Words

- **Parameter** – measure of dependent variable for **population**
  - e.g., average crashes in *Mario Kart* level for everyone
  - Usually what we want to know, but can't get easily
- **Statistic** – measure of dependent variable in **sample**
  - e.g., average crashes in Mario Kart level for IMGD 2905 class
- **Statistics** - set of numerical methods for getting information about **population** based on data from **sample**, usually to get information about population **parameters**

We want to know about these

Population

Parameter  $\mu$   
(Population mean)

<https://psh.ecampus.utd.edu/main-simg-0587913631f0d9a93398331d1d1d1c>

We have these to work with

Sample

Statistic  $\bar{x}$   
(Sample mean)

"Statistics - a branch of mathematics dealing with the collection, **analysis**, **interpretation**, and presentation of masses of numerical data."  
-- Merriam-Webster dictionary

### Sources of Data

- **Published** – generally made available from those that collected it
  - e.g., Riot's *League of Legends* data
  - e.g., Metacritic reviews and ratings
  - e.g., Kaggle dataset on *FIFA 18*
- **Experiments** – multiple trials to collect data
  - Can be in laboratory or "real world" setting
  - e.g., play shooter, add lag and play again
- **Survey** – ask people to answer questions
  - e.g., self-rating as gamer, difficulty with level, ...
  - Ethical issues with stress and use of data
  - **Institute Review Board (IRB)** for approval with human subjects

CONTROL TEST EXPERIMENTAL

<https://i.ytimg.com/v/jqtLn8t6BAQ/maxresdefault.jpg>

Survey:  
Excellent  
Good  
Fair  
Poor

<http://www.mayesmemorial.com/pictures/content/122253.jpg>

### Sampling Concepts

- **Sampling** – process by which members of population are selected for sample
  - e.g., choose ½ class based on spacing, or choose ½ class based on alphabet
- **Probability sampling** – sampling considering likelihood of selection
  - e.g., survey for intended Champ, ask ½ class, but when tournament starts, result different. Why? → sample didn't consider League players! (e.g., often similar analogy for voter polls)
  - e.g., voluntary polls/surveys
  - Use probability sampling whenever possible, but sometimes it is not (cost) or not known
- **Sampling with replacement** – once sample, put back in pool
  - e.g., die roll to see which attack boss makes
- **Sampling without replacement** – once sample, won't sample again
  - e.g., user survey – don't allow to submit twice
  - E.g., deck of 52 cards for blackjack

### Using Sample Data

- Word "sample" comes from same root word as "example"
  - Similarly, one **sample** does not prove a theory, but rather is an **example**
- Basically, in general, definite statement **cannot** be made about characteristics of all systems
- Instead, make **probabilistic statement** about range of most systems

→ That's where statistics come in!