

# INTRODUCTION TO THE COURSE

---

Sergio I. Garcia-Rios

Government 3990: Statistics in the Social Science

## MAIN IDEAS

---

# 1. USE A SAMPLE TO MAKE INFERENCES ABOUT THE POPULATION

- Ultimate goal: make inferences about populations

# 1. USE A SAMPLE TO MAKE INFERENCES ABOUT THE POPULATION

- Ultimate goal: make inferences about populations
- Caveat: populations are difficult or impossible to access

# 1. USE A SAMPLE TO MAKE INFERENCES ABOUT THE POPULATION

- Ultimate goal: make inferences about populations
- Caveat: populations are difficult or impossible to access
- Solution: use a sample from that population, and use **statistics** from that sample to make inferences about the unknown population **parameters**

# 1. USE A SAMPLE TO MAKE INFERENCES ABOUT THE POPULATION

- Ultimate goal: make inferences about populations
- Caveat: populations are difficult or impossible to access
- Solution: use a sample from that population, and use **statistics** from that sample to make inferences about the unknown population **parameters**
- The better (more **representative**) sample we have, the more reliable our estimates and more accurate our inferences will be

# 1. USE A SAMPLE TO MAKE INFERENCES ABOUT THE POPULATION

- Ultimate goal: make inferences about populations
- Caveat: populations are difficult or impossible to access
- Solution: use a sample from that population, and use **statistics** from that sample to make inferences about the unknown population **parameters**
- The better (more **representative**) sample we have, the more reliable our estimates and more accurate our inferences will be

# 1. USE A SAMPLE TO MAKE INFERENCES ABOUT THE POPULATION

- Ultimate goal: make inferences about populations
- Caveat: populations are difficult or impossible to access
- Solution: use a sample from that population, and use **statistics** from that sample to make inferences about the unknown population **parameters**
- The better (more **representative**) sample we have, the more reliable our estimates and more accurate our inferences will be

## Your Turn

Suppose we want to know how many offspring female lemurs have, on average. It's not feasible to obtain offspring data from on all female lemurs, so we use data from the Duke Lemur Center. We use the sample mean from these data as an estimate for the unknown population mean. Can you see any limitations to using data from the Duke Lemur Center to make inferences about all lemurs?



# SAMPLING IS NATURAL



- When you taste a spoonful of soup and decide the spoonful you tasted isn't salty enough, that's **exploratory analysis**
- If you generalize and conclude that your entire soup needs salt, that's an **inference**
- For your inference to be valid, the spoonful you tasted (the sample) needs to be **representative** of the entire pot (the population)