Statistics Workflow

Understand our sample data

Descriptive Statistics

Mean, Median, Mode

What we have in sample data

sample → Descriptive stats → uncertainty

sample data is not all data

purpose: to see if sample fits any distribution

then we can use it as model

Probability Distributions

if sample data fits any existing prob. dist., use it as model for population

sample data ≠ prob. dist

CLT - Central limit Theorem

Confidence Intervals

Try to make estimate about population on sample

- This is the actual stats
- Closely related to CI
- leverage CLT to make conclusions

Hypothesis Tests

- use additional variables to increase our accuracy

Regression Analysis

- helps us to make predication based on the relationships