

ECON 0150 | Spring 2025 | Homework 05 + 06

Due: Friday, Feb 21, 5PM (late due: Tuesday, Feb 25, 5PM)

Homework is designed to both test your knowledge and challenge you to apply familiar concepts in new applications. Answer clearly and completely. You are welcomed and encouraged to work in groups so long as your work is your own. Use the provided datasets to answer the following questions. Then submit your figures and answers to Gradescope.

Q1. US Cities Analysis

Using the provided dataset of US cities that includes:

- City name
- Population
- Latitude/longitude coordinates
- Average temperature

a) Plot all city locations (using lat/lng) on a scatterplot (which represents a map in this case). In a separate color, on the same figure, plot all cities in the 'America/New_York' timezone. (*Interesting but not required: represent the population of the city with the size of the point.*)

b) Create a scatter plot showing the relationship between latitude and average temperature for cities in the 'America/New_York' timezone. How would you describe this relationship?

- Strongly Positive
- Weakly Positive
- Unclear
- Weakly Negative
- Strongly Negative

Q2. Normal Random Variable

Consider a model where customer wait times (in minutes) at a restaurant location follows a normal distribution with mean (μ) = 12 minutes and standard deviation (σ) = 2.5 minutes

a) What is the theoretical mean (μ) wait time (*warm up question*)?

b) What is the theoretical variance for the wait time (*warm up question*)?

c) What is the 77th percentile of the wait time distribution (*use python*)?

Q3. Sampling Variation For A Normal Random Variable

Consider a model where customer wait times (in minutes) at a restaurant location follows a normal distribution with mean (μ) = 12 minutes and standard deviation (σ) = 2.5 minutes

- a) Take samples of 10,000 from the distribution and create a histogram of this sample.
- b) What is the computed difference between the sample mean (\bar{x}) and the theoretical mean (μ) wait time?
- c) What is the computed difference between the sample standard deviation (S) and the theoretical standard deviation (σ) for the wait time?
- d) What is the computed difference between the 77th quartile for the sample and the random variable?

Q4. Sampling Variation From An Unknown Random Variable

Using the provided dataset of customer wait times (in minutes) from two different restaurant locations:

- a) Create a histogram of sample wait times for Location A.
- b) What is the sample mean (\bar{x}) wait time for Location A?
- c) What is the sample standard deviation (S) for the wait time for Location A?
- d) What is the sample size (n) for the sample wait times for Location A?
- e) What is the 77th quartile for the sample wait times for Location A?