Homework:
Introduction to Statistical Sampling

MATH 150

Due: Jan 29, 2024

Sepehr Akbari

Problem 1

Identify the sampling technique used in each of the following. If applicable, identify possible sources of sampling bias.

(a) Following an earthquake in San Francisco, a relief agency selects forty city blocks at random and then sends agents door-to-door to assess damage.

(b) A corn field is divided into 25 one-acre plots and a random sample is taken from each one to estimate crop yield.

(c) From an alphabetical list of customers, a car dealership sends satisfaction surveys to every tenth person.

(d) A student gathers data on satisfaction with a school's cafeteria by surveying all of their friends and carefully recording answers in a numerical spreadsheet.

(e) At a basketball game, ten seats are chosen at random and the occupants given prizes.

Answer

(a) Clustering. There is a potential for non-response bias as some people might not answer the door. Another problem could be convenience bias, if researchers don't want to travel to further block.

(b) Stratified Random Sampling. Just like the previous case non-response and convenience bias can take place.

(c) Systematic Sampling. In this case too, there is an obvious non-response bias, moreover, survey responses have other factors as well. For example unsatisfied costumers are highly more expected to respond, and that itself could create errors in the data analysis.

(d) Convenience Sample. Convenience Bias, as they might only survey their friends or people they know. That also introduces further problems if they all have similar opinions.

(e) Simple Random Sampling (SRS). There is a convenience bias, as they all might be chosen from a closer section of the stadium. There is also a sort of non-response bias, as some seats randomly chosen might be empty.

Problem 2

A university wants to determine what fraction of its undergraduate student body support a new $25 annual fee to improve the student union. For each proposed method below, indicate whether the method is reasonable or not.

(a) Survey a simple random sample of 500 students.

(b) Stratify students by their field of study, then sample 10% of students from each stratum.

(c) Cluster students by their ages (e.g. 18 years old in one cluster, 19 years old in one cluster, etc.), then randomly sample three clusters and survey all students in those clusters.

Answer

(a) Valid. As the 500 student sample is taken of just the undergraduate body at random.

(b) Valid. If the sample taken from each field is done randomly, then there is no issue in stratifying based on field of study. However, the population of each field is not that bold to be clustered. Moreover, for the goal of this research, this is not required.

(c) Not Valid. It will disregard / ignore some sage groups which is a bold bias.

Homework:
Describing Data Qualitatively

MATH 150

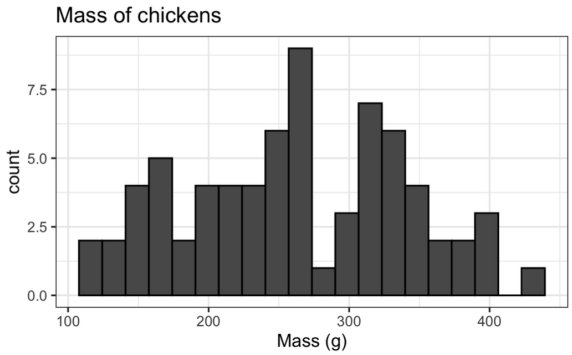
Due: Jan 29, 2024

Sepehr Akbari

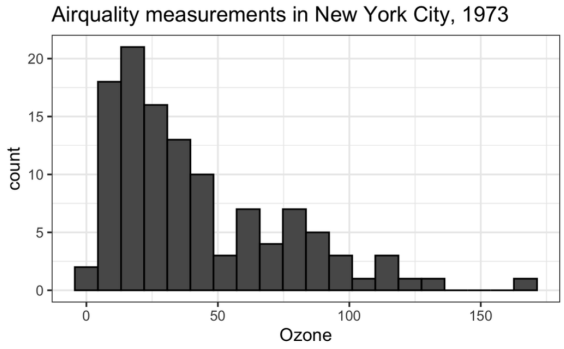
Problem 1

Qualitatively describe the following plots using the vocabulary developed in class.

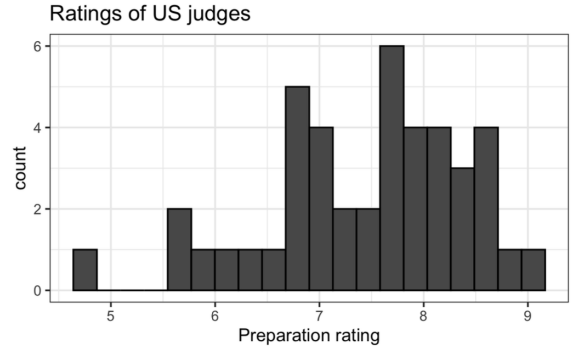
(a)



(b)



(c)



Answer

(a) Symmetric Histogram

(b) Right-Skewed with a potential outlier

(c) Left-Skewed with a potential outlier