

Date: _____

Quiz 7 – B

1. Explain what a sampling frame problem is and how it is different than an under-/overcounting problem.
2.
 - A. Define weak randomness.
 - B. Define strong randomness.
 - C. Define cluster sampling.
 - D. Define stratified sampling.
 - E. Explain where cluster sampling is on the weak to strong random scale.
 - F. Explain where stratified sampling is on the weak to strong random scale.
3. When is cluster sampling appropriate? Explain.

4. Explain why oversampling is different than overcounting.

5. Non-response doesn't necessarily create bias, although it often does. Explain what non-response is, when it can create bias, and when it doesn't create bias.

6. From the table below, calculate the overall sample result \hat{p} for a survey of Portlanders for support for new bikeways in north and east Portland. Show your work.

Race	n	\hat{p}	Demographic weight
White	528	0.32	74%
Hispanic/Latinx	392	0.67	10%
Asian American	159	0.82	9%
African American	167	0.73	6%
Native American	5	0.80	1%

Quiz 7 – B – answer key

1. A sampling frame problem is when the list of people who could be sampled isn't one-to-one with the population of interest. An undercounting problem is when too few people on the list from a certain group are selected to be representative of the sampling frame (opposite for overcounting).
2.
 - A. Every person has an equal chance to be chosen in the sample.
 - B. Every combination of people has an equal chance to be chosen in the sample.
 - C. A few (representative) geographic locations are chosen, and people are randomly sampled from each location.
 - D. Quotas are set for each demographic group, and the pollsters continue to randomly sample until every quota is met.
 - E. Cluster sampling is weak random at best. Assuming every geographic location could have been chosen equally, then every person should have an equal chance of being chosen. However, because certain geographic locations are not chosen, some combinations of individuals are impossible once the locations have been picked.
 - F. Stratified sampling is weak moving toward strong. Every person has an equal likelihood of being chosen. Strictly speaking, it's not strong random because certain combinations are impossible due to violating the quotas. However, that means only the most unrepresentative samples are ruled out—most combinations are still allowed.
3. Cluster sampling is appropriate, for example, when interviews need to be conducted in-person and transportation is difficult. It works for rural areas without connectivity, in war zones where moving around is dangerous, and for surveying students (pick a few schools to be the clusters).
4. Oversampling occurs when pollsters intentionally survey more people than necessary in a subpopulation; the reason is to have a smaller margin of error for that subpopulation. Overcounting is an unknown accident when the randomization method fails or the contact method causes too many people from a subpopulation to be sampled.
5. Non-response is when people are reluctant to participate in a survey. It can create bias if the people who don't respond are consistently different than the people who do respond (e.g., they speak a different language). Non-response doesn't create bias if there's no systematic pattern to who responds (e.g., if it's just luck who doesn't have five minutes to respond to the survey).
6. 42.94%