Name: *Caleb McWhorter — Solutions* **MAT 121**
Summer 2018 **Homework 1**

Problem 1: Determine the problem with the following: a magazine called *People Today* conducted a survey of 523 people about their satisfaction with the government. It was reported that exactly 63% of respondents were unsatisfied with their government's performance.

Solution. *We know 63% of 523 is...*

$$0.63 \cdot 523 = 329.49.$$

However, 329.49 people cannot respond with 'unsatisfied' on a survey. It must be a whole number. Either the 523 is wrong, the 63% is wrong, or the data was fabricated.

Problem 2: Determine whether the following variables are quantitative or categorical:

(a) $Quantitative$: Revenue of a company

(b) $Categorical$: Brand of computer

(c) $Categorical$: Win or loss (coded into a spreadsheet as $-1 =$ loss, $1 =$ win)

(d) $Categorical$: Zip code

(e) $Quantitative$: Years of schooling completed

(f) $Categorical$: Highest level of schooling completed.

Problem 3: Determine whether the following data is discrete or continuous:

(a) $Discrete$: Number of stars in the universe

(b) $Discrete$: Number of arrests in the U.S. each day

(c) $Continuous$: Time a computer takes to complete a task

(d) $Continuous$: Distance between objects

Problem 4: Determine whether the following measurements are nominal, ordinal, interval, or ratio:

(a) _____*Nominal*_____ : Gender

(b) _____*Interval*_____ : Analog clock time

(c) _____*Ratio*_____ : Income

(d) _____*Ordinal*_____ : "Spicyness" of food

(e) _____*Nominal*_____ : Hair color

(f) _____*Interval*_____ : Calendar year

(g) _____*Ratio*_____ : Weight

(h) _____*Ordinal*_____ : Satisfaction level

Problem 5: Determine whether the following are a random sample, a simple random sample, or both.

(a) A group of people are broken into 10 groups of six. The numbers one through ten are written on identical slips of paper and drawn from a hat. The number grabbed is the group chosen.

This is a random sample but is not *a simple random sample.*

(b) In a group of 60 people, the 5 people with the highest yearly income are chosen.

This is neither a random sample nor a simple random sample.

(c) The names of 60 people are written on slips of paper, shuffled, and a group of 6 is chosen from the first six cards in the shuffled pile.

This is both *a random sample and a simple random sample.*

Problem 6: Determine whether the following samples are random, systematic, convenience, stratified, or cluster:

(a) $Stratified$: A study of elementary school students conducted in a large city by sampling thirty students from three different elementary schools in a local school district.

(b) $Random$: A phone survey conducted by calling individuals from a list generated by a computer algorithm.

(c) $Cluster$: A survey of dining hall satisfaction at a university by surveying everyone in 3 of the 7 dorms.

(d) $Systematic$: A survey of algae levels in a lake by taking a sample of every 1,000 square foot of water.

(e) $Convience$: A large phone company sends a text message survey to its 15 million users (receiving over 3 million responses) about whether the user had heard of Peter Scholze before.

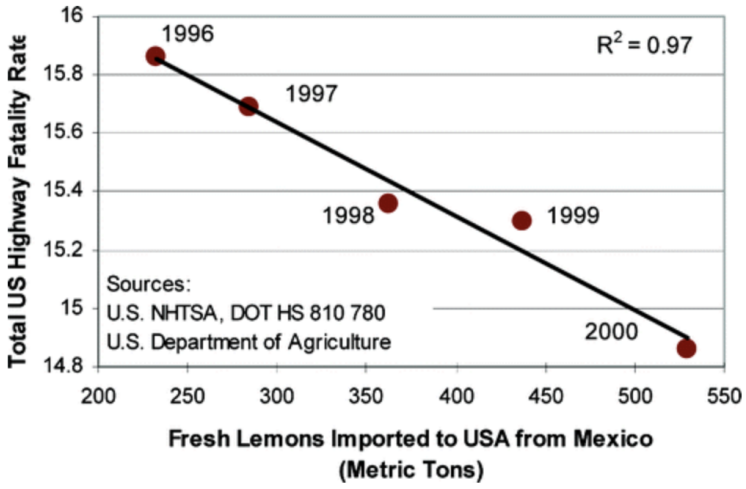
Problem 7: Define a *matched pairs design* experimental design.

Solution. *A matched pairs experimental design is an experiment in which subjects are paired into groups that share some characteristic. For example, one may do a 'twin study' or a 'before/after study'.*

Problem 8: In a study about weight loss, 17% of participants, or 16 individuals, stated they did not lose weight on the drug. How many people participated on the study? If the company redid the study using 50 individuals and wanted only 8% at most to fail to lose weight on the drug, at most how many people could not lose any weight on the drug?

Solution. *We know* $\dfrac{17}{100} = \dfrac{16}{N}$ *so that* $17N = 1600$*. But then* $N = \frac{1600}{17} \approx 94.12$*. Therefore, 94 people participated in the study. We know 8% of 50 is* $0.08 \cdot 50 = 4$*. Therefore, at most 4 people cannot lose weight on the drug.*

Problem 9: The following figure is an example that *correlation does not imply causation* .



Problem 10: Convert the following proportions to a percentage or percentages to a proportion:

(a) $94.3\% = 0.943$

(b) $0.821 = 82.1\%$

(c) $16.3\% = 0.163$

(d) $0.05 = 5\%$

(e) $1.23 = 123\%$

S. Johnson. *The Trouble with QSAR (or How I Learned To Stop Worrying and Embrace Fallacy)*. Journal of Chemical Information and Modeling, **2008**, 28 (1), 25–26.

Problem 11: Calculate the following to two decimal places:

(a) $$\frac{15.3 - 12.1}{1.3} = \frac{3.2}{1.3} = 2.46$$

(b) $$\frac{4212 - 5927}{845} = \frac{-1715}{845} = -2.03$$

(c) $$\frac{108.7 - 104.4}{\frac{13.6}{\sqrt{11}}} = \frac{4.3}{\frac{13.6}{3.3166}} = \frac{4.3}{4.1} = 1.05$$

Problem 12: A NYT article included the image below in an article suggesting that the number of guns in the U.S. may be the reason for the large number of mass shootings in the U.S.. Regardless of the validity of this statement, list some hidden variables which may/may not explain this unusual plot.

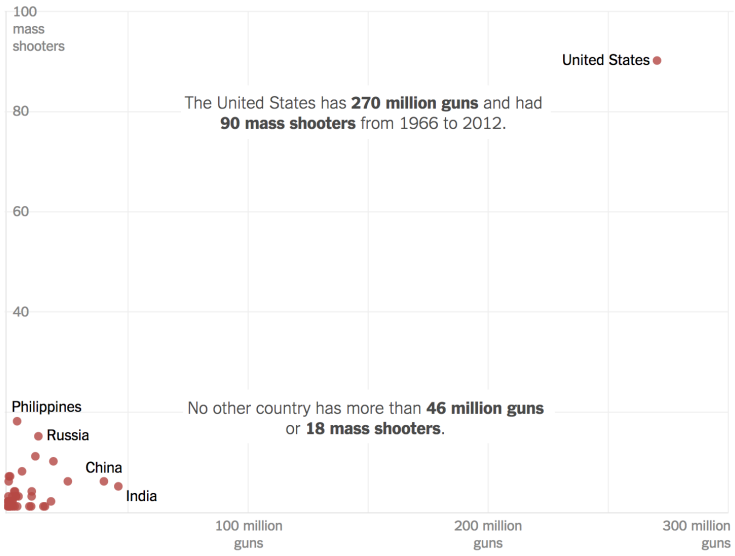


Figure 1: "Source: Adam Lankford, The University of Alabama (shooters); Small Arms Survey (guns). Note: Includes countries with more than 10 million people and at least one mass public shooting with four or more victims."

Solution. *Possible hidden variables are 'aggression' of Americans, mental health factors, violence in media, race/diversity factors, gun laws, etc..*

M. Fisher and J. Keller. "What Explains U.S. Mass Shootings: International Comparisons Suggest an Answer". *NY Times*, 11/07/2017, https://www.nytimes.com/2017/11/07/world/americas/mass-shootings-us-international.html.