

Name: _____
BSAD 210—Montgomery College

EXAM 1

Practice #2

- There are 110 possible points on this exam. The test is out of 100.
- You have one class period to complete this exam, but you should be able to complete it in less than that
- Please turn off all cell phones and other electronic equipment.
- Be sure to read all instructions and questions carefully.
- Remember to show all your work. You may print your formulas in Excel using the Show Formulas option in the Formulas tab. Printed versions of your work showing formulas *and* showing the results counts as showing your work. But you must include both with your test for “showing your work” to count this way. Write your name on both print outs.
- Try all questions! You get zero points for questions that are not attempted.
- Note the last sheet lists all the equations you will need for this exam.
- *Please print clearly and neatly.*

Part I: Matching. Write the letter from the column on the right which best matches each word or phrase in the column on the left. You will not use all the options on the right and you cannot use the same option more than once.

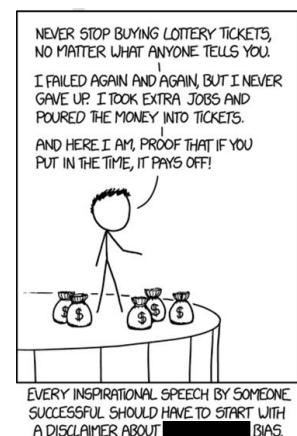
2 points each.

- | | |
|-----------------------------|---|
| 1. ___ Accurate sample | A. Example: Historic unemployment rates for different countries. |
| 2. ___ Cross-sectional data | B. Example: Historic unemployment rate for the United States. |
| 3. ___ Element | C. Example: Lisbeth thinks that she's more likely to have a power outage simply because she hasn't had one in a while. |
| 4. ___ Gambler's fallacy | D. Example: Muhammad believes he won't get rear-ended by another car because he hasn't been rear-ended in years. |
| 5. ___ Law of large numbers | E. Example: Unemployment rates in 2012 for different countries. |
| 6. ___ Hot-hand fallacy | F. Suggests a running average will tend to change less and less as each additional observation is included in the sample. |
| 7. ___ Time series data | G. This means the statistic does not systematically bias the parameter. |
| | H. This means the sample has a lot of observations. |
| | I. What's referenced to determine a variable's value. |

Part II: Multiple Choice. Circle the best answer to the following.

4 points each.

8. Imagine the average cost of living grew by 3% during 2015, 2% during 2016, and 4% during 2017. If the average cost of living was \$30,000 at the end of 2014, what is the average cost of living at the end of 2017?
- \$30,918.72
 - \$32,700.00
 - \$32,778.72
 - \$32,781.81
 - None of the above
9. Consider the following xkcd comic. I've blacked out the kind of bias mentioned in the caption. What kind of bias should go here?
- Self-selection
 - Accuracy
 - Under-coverage
 - Precision
 - None of the above



10. Suppose the average number of pounds of potatoes grown on five acres of land is exactly 50,000 per acre. If an additional acre of land is acquired and it produces exactly 50,000 pounds of potatoes, what happens to the average and standard deviation of potato production for this now six-acre farm?
- The average stays the same and the standard deviation decreases.
 - The average stays the same and the standard deviation stays the same.
 - The average stays the same and the standard deviation increases.
 - The average either increases or decreases and the standard deviation stays the same.
 - It is impossible to tell without knowing how many pounds each of the other acres produces.
11. In 2012, the average income in D.C. was about \$164,000, over three times the U.S. average. But because of how the statistic is calculated, this number is misleading compared to other states. How?
- Because a lot of jobs are high-paying government jobs.
 - Because a lot of people earn income in D.C. but don't live there.
 - Because the average is thrown off by a few very high income people.
 - Because there are many embassies and thus that income doesn't count as being earned "in" D.C.
 - None of the above
12. Use the Practice Exam 1 Data Set 2 for this question. It includes data from the Job Openings and Labor Turnover Survey (JOLTS) concerning the total number of hires and separations (divided into layoffs/discharges, quits, and other) per year and by sector (public or private). (See JOLTS Level tab.) Suppose you were curious which sector—public or private—had the most consistent number of people retiring (based on the coefficient of variation). While you can't directly measure it (retirement isn't an explicit variable), perhaps you might think you can get an idea of which is most consistent. What should you conclude?
- Private sector is more consistent, based on the quit rates.
 - Public sector is more consistent, based on the quit rates.
 - Private sector is more consistent, based on the other separation rates.
 - Public sector is more consistent, based on the other separation rates.
 - Since "retirements" is not a listed variable, there's no way of even estimating which was most consistent.
13. Use the Practice Exam 1 Data Set 2 for this question. It includes data from the Job Openings and Labor Turnover Survey (JOLTS) concerning the total number of hires and separations (divided into layoffs/discharges, quits, and other) per year and by sector (public or private). (See JOLTS Level tab.) You can determine the Total Separations by adding together Layoffs/Discharges, Quits, and Other Separations. Total Hires – Total Separations = Net Jobs. Which year had the highest number of Net Jobs for the public sector?
- 2011
 - 2014
 - 2015
 - 2016
 - None of the above

14. Use the Practice Exam 1 Data Set 2 for this question. It includes data from the Job Openings and Labor Turnover Survey (JOLTS) concerning the total number of hires and separations (divided into layoffs/discharges, quits, and other) per year and by sector (public or private). (See JOLTS Level tab.) What's the average number of total hires from 2008 to 2017?
- About 51.9 thousand
 - About 55.6 thousand
 - About 51.9 million
 - About 55.6 million
 - None of the above
15. A 100% stacked bar chart allows you to do what?
- Combine two different kinds of categorical data
 - Combine categorical data with quantitative data
 - Combine two different kinds of quantitative data
 - Combine cross-sectional data with quantitative data
 - None of the above
16. Which of the following is/are true?
- A weighted average is a normal average but with equal weights.
 - The sample's standard deviation can be greater than the sample's mean.
 - The geometric mean of a series of growth rates will overstate the true growth rate.
 - A & B
 - All of the above are true
17. According to the [Pew Research Center](#), adults who don't participate in elections (nonvoters) tend to be poorer, younger, more racial diverse, and less educated compared to adults in general. If we think of the people who voted as a sample of what the adult population wants, is there anything wrong with this sample? Why or why not?
- Yes; it is not precise.
 - Yes; it is not accurate.
 - Yes; it is neither precise nor accurate.
 - No; though it's over-represented in some areas, it's under-represented in others and the effects cancel out.
 - No; these factors don't matter.
18. "Freemium" software is free to use but you can pay to have an enhanced experience. Most users pay nothing and a few users pay a lot. Describe the median and the mean revenue per user.
- The median is zero and the mean is more than zero.
 - The mean is zero and the median is more than zero.
 - The mean is higher than the median; neither are zero.
 - The median is higher than the mean; neither are zero.
 - It is impossible to tell given the information provided.

19. The chance that any one driver will be in an accident is very hard to predict. But insurance companies can reasonably predict how many drivers they cover will be in an accident. What's this an example of?
- Central tendency
 - Range
 - Sampling bias
 - Sampling accuracy
 - None of the above

Part III: Short Answer. *Answer the following.*

16 points each.

20. Use the Practice Exam 1 Data Set 2 for this question. It includes data from the Job Openings and Labor Turnover Survey (JOLTS) concerning the number of hires and separations (divided into layoffs/discharges, quits, and other) from the previous year by sector (public or private). (See JOLTS Level tab). Create a histogram of the private sector quits. Make sure that:
- The bins are logically define;
 - The bins are clearly presented; and
 - That the histogram has appropriate labels.

Print out a copy of your histogram and attach it to the exam.

21. Use the Practice Exam 1 Data Set 2 for this question. It includes data from the Job Openings and Labor Turnover Survey (JOLTS) concerning the growth of hires and separations (divided into layoffs/discharges, quits, and other) from the previous year by sector (public or private). (See JOLTS Growth tab.) Determine the average growth rate for private sector layoffs/discharges. Express your answer as a percent to two decimal places (e.g. 45.62%).

Don't forget the last question on the back!

22. A metropolitan statistical area (MSA) is an economically integrated geographic definition with an urban core. The ten largest MSAs in the U.S. are indicated in the table below, along with the population density (people per square miles), average income, and population. (Because of the awkwardness of how the Census defines L.A.'s MSA, I've left out its population density.)

Use weighted average to determine the average density and average income for the top ten MSAs in the U.S.

| <i>Metro Statistical Area</i> | <i>Density (sq mi)</i> | <i>Average Income</i> | <i>2010 Census</i> |
|---------------------------------------|------------------------|-----------------------|--------------------|
| New York-Newark-Jersey City | 1,781.30 | 68,525 | 19,567,410 |
| Los Angeles-Long Beach-Anaheim | | 59,441 | 12,828,837 |
| Chicago-Naperville-Elgin | 1,318.00 | 56,423 | 9,461,105 |
| Dallas-Fort Worth-Arlington | 634.00 | 58,744 | 6,426,214 |
| Houston-The Woodlands-Sugar Land | 630.30 | 67,746 | 5,920,416 |
| Washington-Arlington-Alexandria | 1,084.00 | 76,712 | 5,636,232 |
| Miami-Fort Lauderdale-West Palm Beach | 1,004.00 | 45,041 | 5,564,635 |
| Philadelphia-Camden-Wilmington | 2,746.32 | 58,463 | 5,965,343 |
| Atlanta-Sandy Springs-Roswell | 631.18 | 51,993 | 5,286,728 |
| Boston-Cambridge-Newton | 1,375.00 | 72,494 | 4,552,402 |

Equation and Information Sheet

| <i>Function or Command</i> | <i>Result</i> |
|----------------------------|--|
| ABS | Returns the absolute value of an input |
| AVERAGE | Arithmetic mean of an array |
| CTRL + ` | Show formulas |
| CTRL + F | Find |
| CTRL + P | Print |
| CTRL + X | Cut highlighted area |
| CTRL + C | Copy highlighted area |
| CTRL + V | Past highlighted area |
| CTRL + Z | Undo |
| F4 | Makes cell reference absolute |
| GEOMEAN | Geometric mean of an array (adjustments must be added manually) |
| LARGE | Larger values of an array (k=1 is largest, k=2 is second largest, k=3 is third largest...) |
| MAX | Maximum value of an array |
| MEDIAN | Median of an array |
| MIN | Minimum value of an array |
| MODE | Mode of an array |
| QUARTILE | Returns the 0 th to 4 th quartile of an array |
| SQRT | Finds the square root of the value in question. |
| SMALL | Smaller values of an array (k=1 is smallest, k=2 is second smallest, k=3 is third smallest...) |
| STDEV.S | Standard deviation of a sample |
| SUMPRODUCT | The summed product of two or more arrays. |

Geometric Mean

$$\text{Geometric Mean} = \sqrt[n]{\prod_{i=1}^n (1 + x_i)} - 1$$

Standard deviation of a sample

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}$$

Weighted Average

$$\text{Weighted Average} = \frac{\sum_i^n (w_i x_i)}{\sum_i^n w_i}$$

Coefficient of Variation

$$CV_{\text{sample}} = \frac{s}{\bar{x}} (100)$$