Assignment 3 Solutions

1 **Choice C is correct.** The mean and median values of a data set are equal when there is a symmetrical distribution. For example, a normal distribution is symmetrical. If the mean and the median values are not equal, then the distribution is not symmetrical. Outliers are a small group of values that are significantly smaller or larger than the other values in the data. When there are outliers in the data, the mean will be pulled in their direction (either smaller or larger) while the median remains the same. The example in the question has a mean that is larger than the median, and so an appropriate conjecture is that large outliers are present in the data; that is, that there are a few homes that are valued much more than the rest.

Choice A is incorrect because a set of home values that are close to each other will have median and mean values that are also close to each other. Choice B is incorrect because outliers with small values will tend to make the mean lower than the median. Choice D is incorrect because a set of data where many homes are valued between $125,000 and $165,000 will likely have both a mean and a median between $125,000 and $165,000.

2 **Choice B is correct.** The graph of $y = ax^b$, where a is positive and b is negative, would show a trend that is decreasing, but with a rate of decrease that slows as x increases. Of the scatterplots shown, only the one in choice B would be appropriately modeled by such a function.

Choice A is incorrect, as this scatterplot is appropriately modeled by a linear function. Choice C is incorrect, as this scatterplot is appropriately modeled by an increasing function. Choice D is incorrect, as this scatterplot shows no clear relationship between x and y.

3 **Choice B is correct.** The standard deviation is a measure of how far the data set values are from the mean. In the data set for City A, the large majority of the data are in three of the five possible values, which are the three values closest to the mean. In the data set for City B, the data are more spread out, with many values at the minimum and maximum values. Therefore, by observation, the data for City B have a larger standard deviation.

Alternatively, one can calculate the mean and visually inspect the difference between the data values and the mean. For City A the mean is $\frac{1,655}{21} \approx 78.8$, and for City B the mean is $\frac{1,637}{21} \approx 78.0$. The data for City A are closely clustered near 79, which indicates a small standard deviation. The data for City B are spread out away from 78, which indicates a larger standard deviation.

Choices A, C, and D are incorrect and may be the result of misconceptions about the standard deviation.

4 **Choice D is correct.** To interpret what the number 61 in the equation of the line of best fit represents, one must first understand what the data in the scatterplot represent. Each of the points in the scatterplot represents a large US city, graphed according to its population density (along the horizontal axis) and its relative housing cost (along the vertical axis). The line of best fit for this data represents the expected relative housing cost for a certain population density, based on the data points in the graph. Thus, one might say, on average, a city of population density x is expected to have a relative housing cost of y%, where $y = 0.0125x + 61$. The number 61 in the equation represents the y-intercept of the line of best fit, in that when the population density, x, is 0, there is an expected relative housing cost of 61%. This might not have meaning within the context of the problem, in that when the population density is 0, the population is 0, so there probably wouldn't be any housing costs. However, it could be interpreted that for cities with low population densities, housing costs were likely around or above 61% (since below 61% would be for cities with negative population densities, which is impossible).

Choice A is incorrect because it interprets the values of the vertical axis as dollars and not percentages. Choice B is incorrect because the lowest housing cost is about 61% of the national average, not 61% of the highest housing cost. Choice C is incorrect because one cannot absolutely assert that no city with a low population density had housing costs below 61% of the national average, as the model shows that it is unlikely, but not impossible.

5 **Choice B is correct.** Because Nick surveyed a random sample of the freshman class, his sample was representative of the entire freshman class. Thus, the percent of students in the entire freshman class expected to prefer the Fall Festival in October is appropriately estimated by the percent of students who preferred it in the sample, 25.6%. Thus, of the 225 students in the freshman class, approximately $225 \times 0.256 = 57.6$ or about 60 students would be expected to prefer having the Fall Festival in October.

Choices A, C, and D are incorrect. These choices may be the result of misapplying the concept of percent or of calculation errors.

6 **Choice B is correct.** The median of a data set is the middle value when the data points are sorted in either ascending or descending order. There are a total of 600 data points provided, so the median will be the average of the 300th and 301st data points. When the data points are sorted in order:

- Values 1 through 260 will be 0.
- Values 261 through 450 will be 1.
- Values 451 through 540 will be 2.
- Values 541 through 580 will be 3.
- Values 581 through 600 will be 4.

Therefore, both the 300th and 301st values are 1, and hence the median is 1.

Choices A, C, and D are incorrect and may result from either a calculation error or a conceptual error.

7 **Choice C is correct.** When survey participants are selected at random from a larger population, the sample statistics calculated from the survey can be generalized to the larger population. Since 10 of 300 students surveyed at Lincoln School have 4 siblings, one can estimate that this same ratio holds for all 2,400 students at Lincoln School. Also, since 10 of 300 students surveyed at Washington School have 4 siblings, one can estimate that this same ratio holds for all 3,300 students at Washington School. Therefore, approximately $\frac{10}{300} \times 2{,}400 = 80$ students at Lincoln School and $\frac{10}{300} \times 3{,}300 = 110$ students at Washington School are expected to have 4 siblings. Thus, the total number of students with 4 siblings at Washington School is expected to be $110 - 80 = 30$ more than the total number of students with 4 siblings at Lincoln School.

Choices A, B, and D are incorrect and may result from either conceptual or calculation errors. For example, choice A is incorrect; even though there is the same ratio of survey participants from Lincoln School and Washington School with 4 siblings, the two schools have a different total number of students, and thus, a different expected total number of students with 4 siblings.

8 **Choice B is correct.** According to the table, there are $18 + 7 = 25$ graduates who passed the bar exam, and 7 of them did not take the review course. Therefore, if one of the surveyed graduates who passed the bar exam is chosen at random, the probability that the person chosen did not take the review course is $\frac{7}{25}$.

Choices A, C, and D are incorrect. Each of these choices represents a different probability from the conditional probability that the question asks for. Choice A represents the following probability. If one of the surveyed graduates who passed the bar exam is chosen at random, the probability that the person chosen did take the review course is $\frac{18}{25}$. Choice C represents the following probability. If one of the surveyed graduates is chosen at random, the probability that the person chosen passed the bar exam is $\frac{25}{200}$. Choice D represents the following probability. If one of the surveyed graduates is chosen at random, the probability that the person chosen passed the exam and took the review course is $\frac{7}{200}$.

9 **Choice C is correct.** Because there are 16 ounces in 1 pound, a 3-pound pizza weighs $3 \times 16 = 48$ ounces. One half of the pizza weighs $\frac{1}{2} \times 48 = 24$ ounces, and one-third of the half weighs $\frac{1}{3} \times 24 = 8$ ounces.

Alternatively, since $\frac{1}{2} \times \frac{1}{3} = \frac{1}{6}$, cutting the pizza into halves and then into thirds results in a pizza that is cut into sixths. Therefore, each slice of the 48-ounce pizza weighs $\frac{1}{6} \times 48 = 8$ ounces.

Choice A is incorrect and is the result of cutting each half into sixths rather than thirds. Choice B is incorrect and is the result of cutting each half into fourths rather than thirds. Choice D is incorrect and is the result of cutting the whole pizza into thirds.

10 **Choice B is correct.** According to the table, in 2012 there was a total of 14,766 + 47,896 = 62,662 registered voters between 18 and 44 years old, and 3,453 + 11,237 = 14,690 of them were from the Midwest region. Therefore, the probability that a randomly chosen registered voter who was between 18 and 44 years old in 2012 was from the Midwest region is $\frac{14,690}{62,662} \approx 0.234$. Of the given choices, 0.25 is closest to this value.

Choice A is incorrect; this is the probability of selecting at random a registered voter from the Midwest who is 18 to 24 years old. Choice C is incorrect; this is the probability of selecting at random a registered voter from the Midwest who is 18 to 44 years old. Choice D is incorrect and may be the result of errors made when choosing the correct proportion or in calculating the probability.

11 **Choice C is correct.** The mosquito population starts at 100 in week 0 and then is multiplied by a factor of 10 every 5 weeks. Thus, if $P(t)$ is the mosquito population after t weeks, then based on the table, $P(t) = 100\ (10)^{\frac{t}{5}}$, which indicates an exponential growth relationship.

Choice A is incorrect. Increasing linearly means that the estimated population grows by the same amount every 5 weeks. According to the table, from week 0 to week 5, the estimated population grows by 900 mosquitoes, and from week 5 to week 10, it grows by 9,900 mosquitoes. Therefore, the estimated population is not increasing linearly. Choices B and C are incorrect because according to the table, the estimated population is increasing, not decreasing.

12 **Choice D is correct.** According to the given formula, the amount of money generated for a year at 5% interest, compounded monthly, is $1,000\left(1 + \frac{5}{1,200}\right)^{12}$, whereas the amount of money generated at 3% interest, compounded monthly, is $1,000\left(1 + \frac{3}{1,200}\right)^{12}$. Therefore, the difference between these two amounts, $1,000\left(1 + \frac{5}{1,200}\right)^{12} - 1,000\left(1 + \frac{3}{1,200}\right)^{12}$, shows how much additional money is generated at an interest rate of 5% than at an interest rate of 3%.

Choices A, B, and C are incorrect and may be the result of misinterpreting the given formula. For example, the expression in choice C gives how many times as much money, not how much additional money, is generated at an interest rate of 5% than at an interest rate of 3%.

13 **Choice B is correct.** The quality control manager selects 7 lightbulbs at random for inspection out of every 400 lightbulbs produced. A quantity of 20,000 lightbulbs is equal to $\frac{20,000}{400} = 50$ batches of 400 lightbulbs. Therefore, at the rate of 7 lightbulbs per 400 lightbulbs produced, the quality control manager will inspect a total of $50 \times 7 = 350$ lightbulbs.

Choices A, C, and D are incorrect and may result from calculation errors or misunderstanding of the proportional relationship.

14 **Choice C is correct.** For a data point to be above the line $y = x$, the value of y must be greater than the value of x. That is, the consumption in 2010 must be greater than the consumption in 2000. This occurs for 3 types of energy sources shown in the bar graph: biofuels, geothermal, and wind.

Choices A, B, and D are incorrect and may be the result of a conceptual error in presenting the data shown in a scatterplot. For example, choice B is incorrect because there are 2 data points in the scatterplot that lie below the line $y = x$.

15 **Choice B is correct.** Reading the graph, the amount of wood power used in 2000 was 2.25 quadrillion BTUs and the amount used in 2010 was 2.00 quadrillion BTUs. To find the percent decrease, find the positive difference between the two amounts, divide by the earlier amount (from 2000), and then multiply by 100: $\frac{2.25 - 2.00}{2.25} \times 100 = \frac{0.25}{2.25} \times 100 \approx 11.1$ percent. Of the choices given, 11% is closest to the percent decrease in the consumption of wood power from 2000 to 2010.

Choices A, C, and D are incorrect and may be the result of errors in reading the bar graph or in calculating the percent decrease.

16 **Choice C is correct.** To find the atomic weight of an unknown element that is 20% less than the atomic weight of calcium, multiply the atomic weight, in amu, of calcium by (1 − 0.20). This gives $40)(1 - 0.20) = (40)(0.8) = 32$.

Choice A is incorrect. This value is 20% of the atomic weight of calcium, not an atomic weight 20% less than that atomic weight of calcium. Choice B is incorrect. This value is 20 amu less, not 20% less, than the atomic weight of calcium. Choice D is incorrect. This value is 20% more, not 20% less, than the atomic weight of calcium.

17 **Choice A is correct.** The number of miles Earth travels in its one-year orbit of the Sun is 580,000,000. Because there are about 365 days per year, the number of miles Earth travels per day is $\frac{580{,}000{,}000}{365} \approx 1{,}589{,}041$. There are 24 hours in one day, so Earth travels at $\frac{1{,}589{,}041}{24} \approx 66{,}210$ miles per hour. Therefore, of the choices given, 66,000 miles per hour is closest to the average speed of Earth as it orbits the Sun.

Choices B, C, and D are incorrect and may result from calculation errors.

18 **Choice C is correct.** Since the biomass of the lake doubles each year, the biomass starts at a positive value and then increases exponentially over time. Of the graphs shown, only the graph in choice C is of an increasing exponential function.

Choice A is incorrect because the biomass of the lake must start at a positive value, not zero. Furthermore, this graph shows linear growth, not exponential growth. Choice B is incorrect because the biomass of the lake must start at a positive value, not zero. Furthermore, this graph has vertical segments and is not a function. Choice D is incorrect because the biomass of the lake does not remain the same over time.

19 **Choice B is correct.** Tony reads 250 words per minute, and he plans to read for 3 hours, which is 180 minutes, each day. Thus, Tony is planning to read $250 \times 180 = 45{,}000$ words of the novel per day. Since the novel has 349,168 words, it will take Tony $\frac{349{,}168}{45{,}000} \approx 7.76$ days of reading to finish the novel. That is, it will take Tony 7 full days of reading and most of an 8th day of reading to finish the novel. Therefore, it will take Tony 8 days to finish the novel.

Choice A is incorrect and may result from an incorrect calculation or incorrectly using the numbers provided in the table. Choice C is incorrect and may result from taking the total number of words in the novel divided by the rate Tony reads per hour. Choice D is incorrect and may result from taking the total number of words in the novel divided by the number of pages in the novel.

20 **Choice D is correct.** According to the graph, in the interval from 0 to 10 minutes, the non-insulated sample decreased in temperature by about 18°C, while the insulated sample decreased by about 8°C; in the interval from 10 to 20 minutes, the non-insulated sample decreased in temperature by about 9°C, while the insulated sample decreased by about 5°C; in the interval from 40 to 50 minutes, the non-insulated sample decreased in temperature by about 1°C, while the insulated sample decreased by about 3°C; and in the interval from 50 to 60 minutes, the non-insulated sample decreased in temperature by about 1°C, while the insulated sample decreased by about 2°C. The description in choice D accurately summarizes these rates of temperature change over the given intervals. (Note that since the two samples of water have equal mass and so must lose the same amount of heat to cool from 60°C to 25°C, the faster cooling of the non-insulated sample at the start of the cooling process must be balanced out by faster cooling of the insulated sample at the end of the cooling process.)

Choices A, B, and C are incorrect. None of these descriptions accurately compares the rates of temperature change shown in the graph for the 10-minute intervals.

21 **Choice C is correct.** According to the graph, the horizontal line that represents 550 billion miles traveled intersects the line of best fit at a point whose horizontal coordinate is between 2000 and 2005, and slightly closer to 2005 than to 2000. Therefore, of the choices given, 2003 best approximates the year in which the number of miles traveled by air passengers in Country X was estimated to be 550 billion.

Choice A is incorrect. According to the line of best fit, in 1997 the estimated number of miles traveled by air passengers in Country X was about 450 billion, not 550 billion. Choice B is incorrect. According to the line of best fit, in 2000 the estimated number of miles traveled by air passengers in Country X was about 500 billion, not 550 billion. Choice D is incorrect. According to the line of best fit, in 2008 the estimated number of miles traveled by air passengers in Country X was about 600 billion, not 550 billion.

22 **Choice D is correct.** Survey research is an efficient way to estimate the preferences of a large population. In order to reliably generalize the results of survey research to a larger population, the participants should be randomly selected from all people in that population. Since this survey was conducted with a population that was not randomly selected, the results are not reliably representative of all people in the town. Therefore, of the given factors, where the survey was given makes it least likely that a reliable conclusion can be drawn about the sports-watching preferences of all people in the town.

Choice A is incorrect. In general, larger sample sizes are preferred over smaller sample sizes. However, a sample size of 117 people would have allowed a reliable conclusion about the population if the participants had been selected at random. Choice B is incorrect. Whether the population is large or small, a large enough sample taken from the population is reliably generalizable if the participants are selected at random from that population. Thus, a reliable conclusion could have been drawn about the population if the 117 survey participants had been selected at random. Choice C is incorrect. When giving a survey, participants are not forced to respond. Even though some people refused to respond, a reliable conclusion could have been drawn about the population if the participants had been selected at random.

23 **Choice A is correct.** According to the graph, the animal with the longest gestation period (60 days) has a life expectancy of 3 years.

Choices B, C, and D are incorrect. All the animals that have a life expectancy of 4, 8, or 10 years have a gestation period that is shorter than 60 days, which is the longest gestation period.

24 **Choice A is correct.** The ratio of life expectancy to gestation period for the animal represented by point A is approximately $\frac{7 \text{ years}}{23 \text{ days}}$, or about 0.3 years/day, which is greater than the ratio for the animals represented by the other labeled points (the ratios for points B, C, and D, in units of years of life expectancy per day of gestation, are approximately $\frac{8}{44}$, $\frac{8}{51}$, and $\frac{10}{51}$ respectively, each of which is less than 0.2 years/day).

Choices B, C, and D are incorrect and may be the result of errors in calculating the ratio or in reading the graph.

25 **Choice A is correct.** According to the table, of the 50 movies with the greatest ticket sales in 2012, 4 are comedy movies with a PG-13 rating. Therefore, the proportion of the 50 movies with the greatest ticket sales in 2012 that are comedy movies with a PG-13 rating is $\frac{4}{50}$, or equivalently, $\frac{2}{25}$.

Choice B is incorrect; $\frac{9}{50}$ is the proportion of the 50 movies with the greatest ticket sales in 2012 that are comedy movies, regardless of rating. Choice C is incorrect; $\frac{2}{11} = \frac{4}{22}$ is the proportion of movies with a PG-13 rating that are comedy movies. Choice D is incorrect; $\frac{11}{25} = \frac{22}{50}$ is the proportion of the 50 movies with the greatest ticket sales in 2012 that have a rating of PG-13.