1. The median is often a better representative of the central value of a data set when the data set:
2. Source
3. Is bimodal.
4. Has a high standard deviation.
5. Is highly skewed.
6. Has no outliers.
7. The data in the Excel spreadsheet linked below provide information on the nutritional content (in grams per serving) of some leading breakfast cereals.

For which nutrients is the mean nutrient content per serving greater than the median nutrient content per serving?

1. Proteins only.
2. Complex carbohydrates only.
3. Both nutrients.
4. Neither nutrient.
5. The histogram below plots the carbon monoxide (CO) emissions (in pounds/minute) of 40 different airplane models at take-off.

 The distribution is best described as is:



1. Uniform.
2. Heteroskedastic.
3. Normal.
4. Skewed right.
5. The histogram below plots the carbon monoxide (CO) emissions (in pounds/minute) of 40 different airplane models at take-off.

Which of the following statements is the best inference that can be drawn from this histogram?



1. The mean amount of carbon monoxide emissions is greater than the median amount of carbon monoxide emissions.

B. The mean amount of carbon monoxide emissions is less than the median amount of carbon monoxide emissions.

C. The mean and median amounts of carbon monoxide emissions are about equal.

D. The relative sizes of the mean and median amounts of carbon monoxide emissions cannot be inferred from the histogram.

1. In the data set shown below, the correlation coefficient of the two variables is:



A. -1.0.

B. -0.5

C. 0.0

D. None of the above

1. The data in the Excel spreadsheet linked below give the ages and salaries of the chief executive officers of 59 companies with sales between $5 million and $350 million.

The correlation between age and salary can be characterized as:

1. Strong and positive.
2. Strong and negative.
3. Weak and positive.
4. Weak and negative.
5. For a given set of data, the standard deviation measures:
6. The difference between the mean and the data point farthest from the mean.
7. The difference between the mean and the data point nearest to the mean.
8. The difference between the mean and the median.
9. None of the above.
10. The data in the Excel spreadsheet linked below give the average exchange rates of four currencies to the US dollar during October 2002.

Over this period, which currency was most volatile relative to the US dollar, as measured by the coefficient of variation?

1. The Brazilian Real.
2. The Euro.
3. The Japanese Yen.
4. The South Korean Won
5. A national business magazine intends to survey its subscribers to determine who they think is the "CEO of the Century." Subscribers are invited to complete an online survey.
6. They will be biased because the magazine's subscribers are unqualified to determine who the CEO of the century is.
7. They will accurately reflect the opinion of the nation as a whole about who the CEO of the century is.
8. They will be unbiased because the respondents will be self-selected.
9. They will be biased because subscribers who have access to the online survey are not representative of the population of subscribers as a whole.
10. A political consultant conducts a survey to determine what position the mayoral candidate she works for should take on a proposed smoking ban in restaurants.

Which of the following survey questions will deliver an unbiased response?

1. Should the city ban smoking in restaurants to protect our children from second-hand smoke?
2. Should tobacco smoke, a known cause of lung cancer, be banned from public spaces such as restaurants?
3. Does the city have the right to restrict recreational activities, such as moderate consumption of alcohol or tobacco, on the premises of privately-owned businesses?
4. None of the above.
5. In a recently administered IQ test, the scores were distributed normally, with mean 100 and standard deviation 15.
6. About 68%.
7. About 84%.
8. About 95%.
9. About 99.5%.
10. Which of the following is not true about the Normal Distribution?
11. It is symmetric.
12. Its mean and median are equal.
13. It is completely described by its mean and its standard deviation.
14. It is bimodal.
15. The histogram below shows the underlying distribution pattern of the results of a rolled die. Suppose the die is rolled 50 times and the results are averaged. Suppose this process — rolling 50 dice and averaging their results — is repeated 100 times.

Which of the following best describes the distribution of these 100 averages?



1. Skewed right.
2. Skewed left.
3. Normal.
4. Uniform.
5. A nutrition researcher wants to determine the mean fat content of hen's eggs. She collects a sample of 40 eggs. She calculates a mean fat content of 23 grams, with a sample standard deviation of 8 grams. From these statistics she calculates a 90% confidence interval of [20.9 grams, 25.1 grams].
6. Increase the confidence level.
7. Decrease the confidence level.
8. Decrease the sample size
9. None of the above.
10. In a random sample of 321 senior citizens, 61 were found to own a home computer.

Based on this sample, the 95% confidence interval for the proportion of computer-owners among senior citizens is:

1. [2.6%; 7.4%].
2. [13.4%; 24.6%].
3. [14.7%; 23.3%].
4. The answer cannot be determined from the information given.
5. Preliminary estimates suggest that about 58% of students at a state university favor implementing an honor code.

To obtain a 95% confidence interval for the proportion of all students at the university favoring the honor code, what is the minimum sample size needed if the total width of the confidence interval must be less than 5 percentage points (i.e., the confidence interval should extend at most 2.5 percentage points above and below the sample proportion)?

1. 375.
2. 264.
3. 1,498.
4. The answer cannot be determined from the information given.
5. In a survey of twelve Harbor Business School graduates, the mean starting salary was $93,000, with a standard deviation of $17,000.

The 95% confidence interval for the average starting salary among all Harbor graduates is:

1. [$83,382; $102,618].
2. [$82,727; $103,327].
3. [$82,199; $103,801].
4. [$59,000; $127,000].
5. In a survey of 53 randomly selected patrons of a shopping mall, the mean amount of currency carried is $42, with a standard deviation of $78.

What is the 95% confidence interval for the mean amount of currency carried by mall patrons?

1. [$39.1; $44.9].
2. [$24.4; $59.6].
3. [$21.0; $63.0].
4. [$14.4; $69.6].
5. A filling machine in a brewery is designed to fill bottles with 355 ml of hard cider. In practice, however, volumes vary slightly from bottle to bottle. In a sample of 49 bottles, the mean volume of cider is found to be 354 ml, with a standard deviation of 3.5 ml.

At a significance level of 0.01, which conclusion can the brewer draw?

1. The true mean volume of all bottles filled is 354 ml.
2. The machine is not filling bottles to an average volume of 355 ml.
3. There is not enough evidence to indicate that the machine is not filling bottles to an average volume of 355 ml.
4. The machine is filling bottles to an average volume of 355 ml.
5. To conduct a one-sided hypothesis test of the claim that houses located on corner lots (corner-lot houses) have higher average selling prices than those located on non-corner lots, the following alternative hypothesis should be used:
6. The average selling price of a corner-lot house is higher than it is commonly believed to be.
7. The average selling price of a corner-lot house is higher than the average selling price of all houses.
8. The average selling price of a corner-lot house is the same as the average selling price of a house not located on a corner lot.
9. The average selling price of a corner-lot house is higher than the average selling price of a house not located on a corner lot.
10. The data in the Excel spreadsheet linked below indicate the selling prices of houses located on corner lots ("corner-lot houses") and of houses not located on corner lots.

Conduct a one-sided hypothesis test of the claim that corner-lot houses have higher average selling prices than those located on non-corner lots. Using a 99% confidence level, which of the following statements do the data support?

1. Upscale, expensive neighborhoods have more street corners.
2. The average selling price of a corner-lot house is higher than that of the average house not located on a corner lot.
3. The average selling price of a corner-lot house is no more than that of the average house not located on a corner lot.
4. There is not enough evidence to support the claim that the average selling price of a corner-lot house is higher than that of the average house not located on a corner lot.
5. Two semiconductor factories are being compared to see if there is a difference in the average defect rates of the chips they produce. In the first factory, 250 chips are sampled. In the second factory, 350 chips are sampled. The proportions of defective chips are 4.0% and 6.0%, respectively.

Using a confidence level of 95%, which of the following statements is supported by the data?

1. There is not sufficient evidence to show a significant difference in the average defect rates of the two factories.
2. There is a significant difference in the average defect rates of the two factories.
3. The first factory's average defect rate is lower than the second factory's on 95 out of 100 days of operation.
4. None of the above.
5. the regression analysis below relates average annual per capita beef consumption (in pounds) and the independent variable "average annual beef price" (in dollars per pound).

The coefficient on beef price tells us that:



1. For every price increase of $1, average beef consumption decreases by 9.31 pounds.
2. For every price increase of $1, average beef consumption increases by 9.31 pounds.
3. For every price increase $9.31, average beef consumption decreases by 1 pound.
4. For price increase of $9.31, average beef consumption increases by 1 pound.
5. SEE CHART ABOVE: The regression analysis below relates average annual per capita beef consumption (in pounds) and the independent variable "average annual beef price" (in dollars per pound).

In a year in which the average price of beef is at $3.51 per pound, we can expect average annual per capita beef consumption to be approximately:

A. 55.2 pounds

B. 52.6 pounds

C. 53.6 pounds

D. 117.9 pounds

1. The regression analysis below relates average annual per capita beef consumption (in pounds) and the independent variable "average annual per capita pork consumption" (in pounds).

At what level is the coefficient of the independent variable pork consumption significant?



A. 0.10.

B. 0.05.

C. 0.01.

D. None of the above.

1. SEE CHART ABOVE: The regression analysis below relates average annual per capita beef consumption (in pounds) and the independent variable "average annual per capita pork consumption" (in pounds).

Which of the following statements is true?

1. Beef consumption can never be less than 65.09 pounds.
2. Beef consumption can never be greater than 65.09 pounds.
3. The y-intercept of the regression line is 65.09 pounds.
4. The x-intercept of the regression line is 65.09 pounds.
5. The regression analysis at the bottom relates average annual per capita beef consumption (in pounds) and the independent variables "average annual per capita pork consumption" (in pounds) and "average annual beef price" (in dollars per pound).

Which of the independent variables is significant at the 0.01 level?



1. Beef price only.
2. Pork consumption only.
3. Both independent variables.
4. Neither independent variable.
5. SEE CHART ABOVE: The regression analysis at the bottom relates average annual per capita beef consumption (in pounds) and the independent variables "annual per capita pork consumption" (in pounds) and "average annual beef price" (in dollars per pound).

The coefficient for beef price, -12, tells us that:

1. For every $1 increase in beef price, average beef consumption decreases by 12 lbs, not controlling for pork consumption.
2. For every $12 drop in beef price, average beef consumption decreases by 1 lbs, not controlling for pork consumption.
3. For every $1 increase in beef price, average beef consumption decreases by 12 lbs, controlling for pork consumption, i.e. holding pork consumption constant.
4. For every $12 decrease in beef price, average beef consumption decreases by 1 lbs, controlling for pork consumption, i.e. holding pork consumption constant.
5. The data in the Excel spreadsheet linked below give the seasonally adjusted value of total new car sales (in millions of dollars) in the United States, total national wage and salary disbursements (referred to here as "compensation") (in billions of dollars), and the employment level in the non-agricultural sector (in thousands) for 44 consecutive quarters. An auto industry executive wants to know how well she can predict new car sales two quarters in advance using the current quarter's compensation data.

How many data points can she use in a regression analysis using the data provided?

1. 41.
2. 42.
3. 43.
4. 44.
5. The Excel spreadsheet linked below contains the simple regressions of total new car sales (in millions of dollars) on each of two independent variables: "compensation" (in billions of dollars) and "employment level in the non-agricultural sector" (in thousands) .

Which of the following independent variables explains more than 90 percent of the observed variation in new car sales?



|  |  |  |
| --- | --- | --- |
| **Seasonally Adjusted Quarterly Car Sales** |  |  |
| **vs. Compensation** |  |  |  |  |
|  |  |  |  |  |  |
| *Regression Statistics* |  |  |  |  |
| Multiple R | 0.9557 |  |  |  |  |
| R2 | 0.9133 |  |  |  |  |
| Adjusted R2 | 0.9112 |  | *F test results* |  |  |
| Standard Error | 6078.1809 |  | *F* | *Signif. F* |  |
| Observations | 44 |  | 442.49 | 0.0000 |  |
|  |  |  |  |  |  |
|  | *Coefficients* | *Std Error* | *t Stat* | *P-value* |  |
| Intercept | -18,381.02 | 4,118.12 | -4.46 | 0.0001 |  |
| Compensation | 45.51 | 2.16 | 21.04 | 0.0000 |  |
|  |  |  |  |  |  |

1. Compensation only.
2. Employment level only.
3. Both independent variables.
4. Neither independent variable.
5. The regression analysis below relates the value of new car sales (in millions of dollars) to compensation (in billions of dollars) and the employment level in the non-agricultural sector (in thousands) for 44 consecutive quarters.

Which of the two independent variables is statistically significant at the 0.05 level?



|  |  |
| --- | --- |
| **Seasonally Adjusted Quarterly Car Sales** |  |
| **vs. Compensation and Employment Level** |  |
|  |  |  |  |  |
| *Regression Statistics* |  |  |  |
| Multiple R | 0.9559 |  |  |  |
| R2 | 0.9137 |  |  |  |
| Adjusted R2 | 0.9094 |  | *F test results* |
| Standard Error | 6139.6071 |  | *F* | *Signif. F* |
| Observations | 44 |  | 216.92 | 0.0000 |
|  |  |  |  |  |
|  | *Coefficients* | *Std Error* | *t Stat* | *P-value* |
| Intercept | -32,562.59 | 35,287.07 | -0.92 | 0.3615 |
| Compensation  | 42.33 | 8.15 | 5.19 | 0.0000 |
| Employment Level | 0.21 | 0.51 | 0.40 | 0.6878 |

1. Compensation only.
2. Employment level only.
3. Both independent variables.
4. Neither independent variable.
5. The regression analysis below relates the value of new car sales (in millions of dollars) and the independent variables "compensation" (in billions of dollars) and "employment level in the non-agricultural sector" (in thousands) for 44 consecutive quarters. Compare this multiple regression to the simple regressions with compensation and employment level as the respective independent variables.

Which of the following is the likely culprit of the dramatic increase in the p-value for employment level in the multiple regression?



|  |  |
| --- | --- |
| **Seasonally Adjusted Quarterly Car Sales** |  |
| **vs. Compensation** |  |  |  |
|  |  |  |  |  |
| *Regression Statistics* |  |  |  |
| Multiple R | 0.9557 |  |  |  |
| R2 | 0.9133 |  |  |  |
| Adjusted R2 | 0.9112 |  | *F test results* |  |
| Standard Error | 6078.1809 |  | *F* | *Signif. F* |
| Observations | 44 |  | 442.49 | 0.0000 |
|  |  |  |  |  |
|  | *Coefficients* | *Std Error* | *t Stat* | *P-value* |
| Intercept | -18,381.02 | 4,118.12 | -4.46 | 0.0001 |
| Compensation | 45.51 | 2.16 | 21.04 | 0.0000 |

1. Multicollinearity.
2. Heteroskedasticity.
3. Nonlinearity.
4. None of the above
5. The regression analysis below relates the value of new car sales (in millions of dollars) and the independent variables "compensation" (in billions of dollars) and "employment level in the non-agricultural sector" (in thousands) for 44 consecutive quarters.

The coefficient for employment level, 0.21, describes:



|  |  |
| --- | --- |
| **Seasonally Adjusted Quarterly Car Sales** |  |
| **vs. Compensation and Employment Level** |  |
|  |  |  |  |  |
| *Regression Statistics* |  |  |  |
| Multiple R | 0.9559 |  |  |  |
| R2 | 0.9137 |  |  |  |
| Adjusted R2 | 0.9094 |  | *F test results* |
| Standard Error | 6139.6071 |  | *F* | *Signif. F* |
| Observations | 44 |  | 216.92 | 0.0000 |
|  |  |  |  |  |
|  | *Coefficients* | *Std Error* | *t Stat* | *P-value* |
| Intercept | -32,562.59 | 35,287.07 | -0.92 | 0.3615 |
| Compensation  | 42.33 | 8.15 | 5.19 | 0.0000 |
| Employment Level | 0.21 | 0.51 | 0.40 | 0.6878 |
|  |  |  |  |  |

1. The behavior of car sales as the employments level changes, controlling for compensation.
2. The behavior of the employment level as car sales change, not controlling for compensation.
3. The behavior of car sales as the employment level changes, not controlling for compensation.
4. The behavior of the employment level as car sales change, controlling for compensation.
5. A new blood pressure treatment is being tested. The regression analysis below describes the relationship between the 41 test subjects' diastolic blood pressure and the dummy variable "medication." When a test subject is taking the new drug, the value of medication is 1, when not, the value of medication is 0.

Which of the following can be inferred from the regression analysis?



1. The medication has no statistically significant effect (at a 0.01 significance level).
2. The use of medication accounts for around 42% of the variation in diastolic blood pressure.
3. On average, test subjects taking the medication report a diastolic blood pressure level about 5 points lower than those not taking the medication.
4. None of the above.
5. In a regression analysis, a residual plot is:
6. A scatter diagram that plots the values of the residuals against the values of the dependent variable.
7. A scatter diagram that plots the values of the residuals against the values of an independent variable.
8. A histogram that plots the frequency of certain value ranges of the residuals.
9. None of the above.
10. In a regression analysis, if a new independent variable is added and R-squared increases and adjusted R-squared decreases precipitously, what can be concluded?
11. The new independent variable improves the predictive power of the regression model.
12. The new independent variable does not improve the predictive power of the regression model.
13. The regression was performed incorrectly. It is impossible for R-squared to increase and adjusted R-squared decrease simultaneously.
14. The new independent variable's coefficient is not significant at the 0.01 level.
15. The table below displays data the First Bank of Silverhaven (FBS) has collected on the personal savings accounts of its job-holding customers. The table includes data on the distribution of the number of accounts held by Homeowners vs. Non-Homeowners, and by whether the customer is Self-Employed or is Employed by a firm in which he or she does not have an ownership stake.

What is the probability that a given account-holder is self-employed?



1. 15%
2. 12%
3. 3%
4. None of the above.
5. The table below displays data the First Bank of Silverhaven (FBS) has collected on the personal savings accounts of its job-holding customers. The table includes data on the distribution of the number of accounts held by Homeowners vs. Non-Homeowners, and by whether the customer is Self-Employed or is Employed by a firm in which he or she does not have an ownership stake.

What is the conditional probability that an account-holder is employed by a firm in which he or she does not have an ownership stake, given that the account-holder is a homeowner?



A. 82.9%

B. 68.2%

C. 59.5%

D. None of the above

1. The table below displays data the First Bank of Silverhaven (FBS) has collected on the personal savings accounts of its job-holding customers. The table includes data on the distribution of the number of accounts held by Homeowners vs. Non-Homeowners, and by whether the customer is Self-Employed or is Employed by a firm in which he or she does not have an ownership stake.

Which of the following statements is true?



1. Home ownership and employment status are statistically dependent.
2. The fact that a given personal account is held by a homeowner tells us nothing about the account holder's employment status.
3. The fact that a given personal account is held by a self-employed person tells us nothing about the account-holder's home ownership status.
4. None of the above.
5. Electronics manufacturer SE must decide whether or not to invest in the development of a new type of battery. If the development succeeds, the market for the battery may be large or small. If it doesn't succeed, the development efforts may or may not generate minor innovations that would offset some of the battery's development costs. The tree below summarizes the decision.

What is the expected monetary value of developing the new battery?



A. $0.3 million

B. -$0.5 million

C. $0.9 million

D. $1.5 million

1. Electronics manufacturer SE must decide whether or not to invest in the development of a new type of battery. If the development succeeds, the market for the battery may be large or small. If it doesn't succeed, the development efforts may or may not generate minor innovations that would offset some of the battery's development costs. The tree below summarizes the decision.

The EMV of developing the new battery is $300,000. Based on EMV, SE should develop the battery. If the manager chooses not to develop the battery, which of the following best describes the manager's attitude towards this decision?



1. Risk averse.
2. Risk neutral.
3. Risk seeking.
4. Cowardly
5. Electronics manufacturer SE must decide whether or not to invest in the development of a new type of battery. If the development succeeds, the market for the battery may be large or small. If it doesn't succeed, the development efforts may or may not generate minor innovations that would offset some of the battery's development costs. The tree below summarizes the decision.

The EMV of developing the new battery is $300,000. Based on EMV, SE should develop the battery. For what values of p = Prob[success] does developing the battery have a lower EMV than not developing the battery?



1. p < 25%

1. p > 25%

1. p < 75%
2. None of the above
3. Electronics manufacturer SE must decide whether or not to invest in the development of a new type of battery. If the development succeeds, the market for the battery may be large or small. If it doesn't succeed, the development efforts may or may not generate minor innovations that would offset some of the battery's development costs. The tree below summarizes the decision.

The EMV of developing the new battery is $300,000. Based on EMV, SE should develop the battery. Given that development is successful, for what endpoint values for a large market does developing the battery have a lower EMV than not developing the battery?



1. Less than $1.2 million.
2. Greater than $1.2 million.
3. Greater than $0.75 million.
4. None of the above.
5. The manager of the Eris Shoe Company must decide whether or not to contract a controversial sports celebrity as its spokesperson. The new spokesperson's value to Eris depends heavily on consumers' perception of him. An initial decision analysis based on available data reveals that the expected monetary value of contracting the new spokesperson is $260,000. For $50,000 Eris can engage a market research firm that will help Eris learn more about how consumers might react to the celebrity. The EMV of buying this sample information (assuming it is free) for this decision is $300,000.

The tree below summarizes Eris's decision. Based on EMV analysis, Eris's manager should:



1. Hire the research firm.
2. Not hire the research firm, but contract the new spokesperson.
3. Not hire the research firm and not contract the new spokesperson.
4. The answer cannot be determined from the information provided.