Statistics Worksheet 6 Answers

1. D
2. A
3. A
4. C
5. A
6. A
7. C
8. B
9. B
10. Both histograms and box plots are used to explore and present the data in an easy and understandable manner. Histograms are preferred to determine the underlying probability distribution of a data. Boz plots on the other hand are more useful when comparing between several data sets. They are less detailed than histograms and take up less space.
11. – Think about goals to determine which metrics we need to track, we have to first think about what goals you want to achieve from the data.
    - Choose metrics that are simple and accurately measurable.
12. Researchers use a measurement known as the p-value to determine statistical significance: if the p-value falls below the significance level, then the result is statistically significant. The p-value is a function of the means and standard deviations of the data samples.
13. A. Allocation of wealth among individuals.
    b. Values of oil reserves among oil fields (many small ones, a small number of large ones).

14. Let's say you run a customer satisfaction survey with a sample of 9 and rate their overall satisfaction scores on a scale of 1 to 10. You get an average of 5.22. You know that in general, you tend to retain customers with a score over 3, so you're satisfied, because this indicates that you're still above where you want to be. But then, suddenly, you lose 6 of those 9 customers. You go back to look at your data, and you find these scores:
    1, 3, 3, 3, 3, 5, 9, 10, 10

    The median of this group is a 3, indicating that at least half of your customers or more were unhappy. The scores became lopsided because of the unexpected 10's, and you missed out on an important part of your data – the midpoint that indicated that as many as half of your customers or more were dissatisfied with your company.
15. The likelihood function represents the probability of random variable realizations conditional on particular values of the statistical parameters. Thus, when evaluated on a given sample, the likelihood function indicates which parameter values are more likely than others, in the sense that they would have made the observed data more probable.