**Stats Homework**

**A few homework hints before you start:**

* Please write neatly, or type, and label all of your problems clearly.
* Provide sufficient space for your work; do not cram all the problems onto one page.
* Review your work and verify that you have answered ALL required problems.
* You should round all of your *final* answers to 2 decimals (1.75, not 1.749).
* When writing or explaining definitions of key terminology, rephrase the definitions in your own words; do not copy directly from the notes or provided answers because doing so will inhibit your understanding. Also, when asked to provide examples of certain concepts, avoid using those provided by me or by the author of your readings. Rather, generate your own examples so that you can deepen your understanding of the material; generating examples is linked with learning.
* *Optional Questions*: Some questions may ask you to “explain this concept to someone who has never taken a course in statistics”. These questions will challenge you to understand statistical principles from a conceptual perspective rather than from a computational one. Conceptual questions will appear on the exams, so I very much encourage you to attempt answering these questions by simply speaking your response aloud to yourself; you need not draft a written answer to such questions although writing the question and your well thought our answer on a flashcard should prove beneficial. If you cannot provide a response in a timely manner, then you know you may struggle on the exam.
* Show all your work/steps for each the problem so that you can see where you make any errors. Copying answers to problems without answering them yourself will be considered plagiarism; doing so will also disservice you in terms of learning and on exam performance. I encourage you to solve the problems on your own and then compare your answers to those in the answers provided in order to verify that you answered them correctly.
* Answer all questions unless they are ~~crossed out.~~ The answers to these problems are also provided if you wish to challenge yourself with additional concepts.

**Homework #1: Measurement, Frequency *&* Probability Distributions**

**Homework 1**: Questions 1 – 6

**Conceptual Questions:**

1. A client rates her satisfaction with her vocational counselor as a 3 on a 4-point scale. The scale ranges from 1 = *not at all satisfied* to 4 = *very satisfied*. Identify the (a) variable, (b) possible values, and (c) client’s score.
2. Name the *scale of measurement* for each of the following variables: (a) ethnic group to which a person belongs, (b) number of times a mouse makes a wrong turn in a laboratory maze, and (c) position one finishes a race.
3. Provided below are the scores on a test of sensitivity to smell taken by 25 chefs attending a national conference: 96, 83, 59, 64, 73, 74, 80, 68, 87, 67, 64, 92, 76, 71, 68, 50, 85, 75, 81, 70, 76, 91, 69, 83, 75
   1. The average/mean of the data is about 75 points. Rank order the data from lowest to highest score so that you can imagine what a frequency distribution/histogram with a mean of 75 might look like for these data. Is it shaped symmetrically, skewed, etc.?
   2. Describe the general shape of the grouped frequency distribution (histogram) below; this description should also provide information about the mode.



* 1. If the shape of the distribution was skewed negatively, what would it look like? Would most chefs’ sensitivities be less than or greater than the average sensitivity?

1. Draw an example of each of the following distributions
   1. Symmetrical
   2. Rectangular
   3. Skewed positively
2. Describe the four levels or *scales of measurement* in your own words. Based on your descriptions, identify the scale for each of the following variables:
   1. Number of downloads of different bands’ songs on iTunes
   2. Names of the bands that were downloaded
   3. Positions on the iTunes download chart
   4. The money earned by the bands from the downloads
3. Describe what a normal distribution, a positively skewed distribution, and a negatively skewed distribution look like.

**Homework #1: Answers**

1. **Client data**
   1. The *variable* is the client’s level so satisfaction.
   2. The possible *values* are 1,2,3, and 4
   3. The client’s *score* is 3.
2. **Measurement scales** 
   1. the scale is nominal
   2. the scale is ratio; an absolute 0 exists
   3. the scale is ordinal (rank-ordered)
3. **Sensitivity to Smell**
   1. Your imagined histogram should have the y-axis labeled frequency (*f*) and the x-axis labeled with the name of the measured variable or dependent variable. If you imagined a grouped histogram, each bar would labeled with the midpoint of the interval rather than an individual score. The height of the bars in the histogram should reflect the frequency of scores occurring within a particular interval on the graph. Depending on how you imagine your histogram, the shape may differ.
   2. This grouped (10-point bins) histogram is *unimodal* in shape. There is only *one* mode, or most frequently occurring, score. In this instance, you can think of the class interval with midpoint 75 as the mode. The distribution is not exactly symmetrical, but is fairly symmetrical, because the shape of the distribution to the right and the left of the middle bar almost mirror each other.



* 1. If the smell sensitivity data happened to produce a graph with a distribution that was skewed negatively, it would not look so symmetrical as it currently does. Rather, most of the scores would be on the high (right) end of the distribution and fewer scores would be on the low (left) side. Because the extreme low scores pull the mean of negative distributions to the left (negatively), most of the chef’s would have higher sensitivities than the average of all chefs’ sensitivities.

1. **Drawings for distributions:** 
   1. Symmetrical



* 1. rectangular (e.g., the outcomes associated with of rolling a fair die; or any other case for which the frequencies of events would be uniform or identical)



* 1. skewed positively (e.g., scores on a challenging physics exam)



1. **Scales of measurement**

*Nominal Scale* - Variables with more than two qualitatively different categories.

*Ordinal Scale* - Variables that have ranked order only; numbers are not interval.

*Interval Scale* - Variables that have ranked order and equal intervals between values. No true zero.

*Ratio Scale* - Variables represent ranked order, equal intervals between values, and a true zero point.

* 1. Number of downloads of different bands’ songs on iTunes – Ratio Discrete
  2. Names of the bands for the songs that were downloaded – Nominal/Categorical
  3. Positions on the iTunes download chart - Ordinal
  4. The money earned by the bands from the downloads – Ratio Continuous

1. **Shapes of Distributions** 
   1. *Symmetrical* and *Normal* Distributions have scores that are distributed symmetrically around the mean. The majority of the scores are around the center of the distribution, with fewer scores in the tails.
   2. *Positively Skewed* Distributions have scores that are clustered at the lower end of the scale, with tails pointing to more positive scores.
   3. *Negatively Skewed* Distributions have scores that are clustered at the upper end of the scale, with tails pointing to more negative scores.