Name: Eddie S. Jackson

1. In the following situation identify the implied population.

**A recent report on the weekly news presented the findings of a study on the effectiveness of Onglyza, along with diet and exercise, for treating diabetes.**my answer: The implied population is all the people who are diabetics. Diabetics were used to test the effectiveness of Onglyza, diet and exercise.

2. In the following scenario identify the type of statistical study that was conducted.

**A Gallop poll surveyed 1,018 adults by telephone, and 22% of them reported that they smoked cigarettes within the past week.**my answer: This is an observational study because the sample population that was studied was *not* influenced by the researcher themselves.

3. In the following scenario what is the statistic and the parameter (population as a whole) it would estimate.

**A recent study of 460 drivers age 70 and over by the National Highway Traffic Safety Administration reported that 75% of those drivers had uncorrected vision problems.**my answer: 75% (345 people) was the reported statistic from the 460 drivers of 70 or older drivers.  
  
my answer: The assumed parameter would be the population of drivers that had uncorrected vision that were 70 years old or older.

4. What type of sampling procedure was used to collect the data in the MM207 Student Data Set?   
my answer: I would say convience sampling because this was a population that was readily available; college students. You have the ability to pull many data variables from college students like gender, gpa, and hair color as examples.

5. From the MM207 Student Data Set identify one variable that is discrete and one variable that is continuous. Explain your reasoning. How do you know?

1. Discrete:   
   my answer: Gender would be discrete because it’s basically one or the other, male or female.   
    From our studies it states discrete consists of indivisible categories, just like male or   
    female.
2. Continuous:  
   my answer: Height, because height is not a single number…it has many variances in inches. From our studies it says a continuous variable is divisible into an infinite number of fractional parts. For example when I used StatCrunch, the returned mean for height was 65.46471.

6. From the MM207 Student Data Set identify the following:

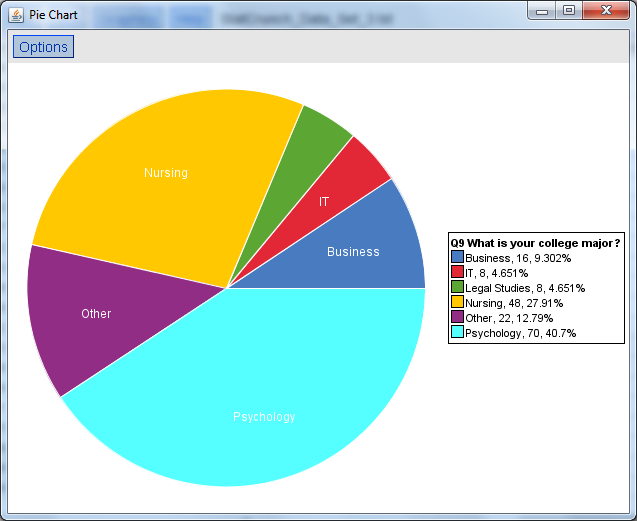
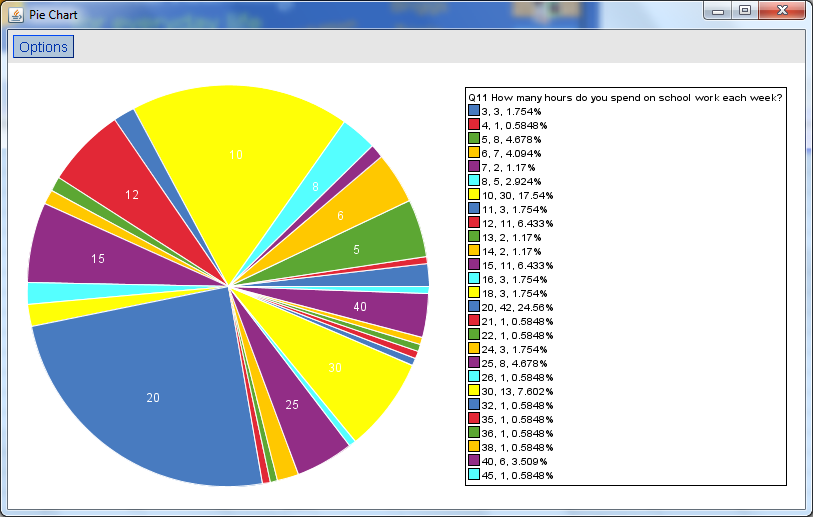
1. A variable measured at the nominal level of measurement.

my answer: gender, this has been assigned a name.

1. A variable measured at the ratio level of measurement.  
   my answer: height would be a good ratio measurement because it is an interval scale and has an absolute zero.
2. A variable measured at the ordinal level of measurement.  
   my answer: political philosophy because they can be ordered liberal, moderate, conservative, or other levels of order.

7. What is the approximate percentage of students represented in the data set who are between the ages of 29 and 45 inclusive?  
my math is:   
175 people between ages 21 to 65 = 100% of students represented  
86 people between ages 29 to 45 = 49.14% of students represented  
  
I came to my answer by using StatCrunch to first return summary states on “How Old” which gave me the total represented 172.  
  
Then I used StatCrunch to create a frequency table and just counted the total number of students represented from 29-45 which is 86. 86/172 = 0.4914285714285714 \* 1 = 49.14%  
my answer: is **49.14%**

8. Assume that you are preparing a report on MM207 statistics students at Kaplan University to present to the Kaplan Board of Trustees. Prepare an appropriate graphical representation for each of the following variables.

1. **The different majors of students taking MM207.**  
   my answer:  
   
2. **The number of hours spent on school work by students in MM207.**  
   my answer: I tried multiple charts and graphs, none of them seemed good enough. I just chose this final one. For example, represented in darker - blue 42 students spent 20 hours a week in studies – which was 24.56% of the students.  
   

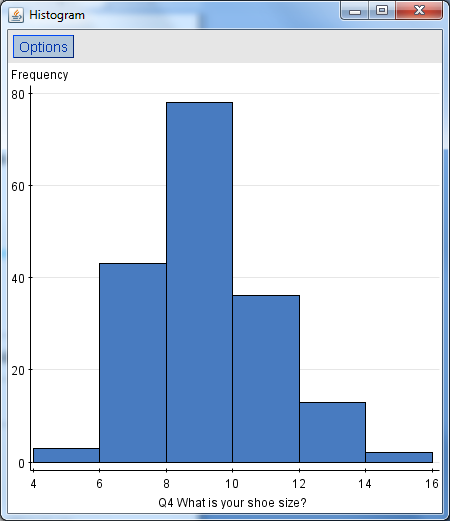
9. Using the range rule of thumb estimate the standard deviation for the number of credit hours students in this sample are taking and the shoe sizes of the females in the class. Then using StatCrunch compute the actual standard deviation. Compare the results.

**Does the range rule of thumb overestimate or underestimate the standard deviation for number of credit hours?**  
range rule of thumb: (18-3)/4 = 3.75  
standard deviation: 2.5635269  
my answer: the range rule of thumb underestimates the standard deviation for the number of credit hours.  
  
Summary statistics for Q6 How many credit hours are you taking this term?:   
Group by: Gender

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Gender | n | Mean | Variance | Std. Dev. | Std. Err. | Median | Range | Min | Max | Q1 | Q3 |
| Female | 138 | 10.797101 | 6.5716705 | 2.5635269 | 0.2182219 | 10 | 15 | 3 | 18 | 10 | 12 |
|  |  |  |  |  |  |  |  |  |  |  |  |

**Does the range rule of thumb overestimate or underestimate the standard deviation for shoe sizes of females?**   
  
range rule of thumb: (12-5)/4 = 1.75  
standard deviation: 1.294457  
my answer: the range rule of thumb overestimates the standard deviation for the shoe size.  
  
Summary statistics for Q4 What is your shoe size?:   
Group by: Gender

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Gender | n | Mean | Variance | Std. Dev. | Std. Err. | Median | Range | Min | Max | Q1 | Q3 |
| Female | 138 | 8.271739 | 1.6756189 | 1.294457 | 0.110191494 | 8.5 | 7 | 5 | 12 | 7 | 9 |

**How does the shape of the distribution impact what you conclude?**my answer: That the majority of females have a shoe size between the sizes of 6-12 and very few females have shoe sizes below size 6 or above size 12. The chart shows 1 mode, skewed left, wide variation.  
****

10. Using measures of center and measures of variability compare the number of hours on school work (Q11) and the number of hours watching television (Q14).

1. **Which variable has greater variability? How do you know?**

Q11 has a 82.21 hour variance and Q14 has a 35.31 hour variance according to StatCrunch column statistics.  
my answer: would be Q11

**What number of hours on school work would be at the 10th percentile?**   
my answer: is 6 hours  
  
- Summary statistics:

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Column | n | Mean | Variance | Std. Dev. | Std. Err. | Median | Range | Min | Max | Q1 | Q3 | **10th Per.** |
| Q11 How many hours do you spend on school work each week? | 171 | 17.146198 | 82.20791 | 9.066858 | 0.69336 | 16 | 42 | 3 | 45 | 10 | 20 | 6 |

**What number of hours watching television would be at the 90th percentile?**my answer: is 16 hours of television watching.  
Summary statistics:

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Column | n | Mean | Variance | Std. Dev. | Std. Err. | Median | Range | Min | Max | Q1 | Q3 | 90th Per. |
| Q14 How many hours of television did you watch last week? | 173 | 7.5433526 | 35.3077 | 5.9420285 | 0.45176405 | 6 | 25 | 0 | 25 | 3 | 10 | 16 |