Math 214 Statistics Fall 13 Sample Midterm 1

(1) The following comes from an online news source:

MONDAY, Feb. 18 (HealthDay News) – Young adults and teens who smoke, are obese and have high blood sugar levels may be more likely to die before they reach their 55th birthday, new government research suggests. The average life expectancy in the United States is 78.7, according to the CDC.

Dr. Saydah and her colleagues analyzed data on close to 9,250 people who took part in the third National Health and Nutrition Examination Survey.

Participants were aged 12 to 39 when the study was conducted. Of these, more than 15 percent were obese, and 30 percent were smokers. Overall, 298 of the participants died before they turned 55.

Those who smoked between the ages of 12 and 39 had an 86 percent greater risk of dying before 55, compared with those who did not, the data showed. Those who were obese when they were young had a 39 percent higher likelihood of dying before 55, compared with those not obese during these early years. In addition, the risk of dying before 55 tripled among those with high blood sugar levels between the ages of 12 to 39, the study showed. High blood pressure and high cholesterol levels, however, did not affect the risk of dying before age 55.

Answer the following questions using the text above.

(a) What is an experimental unit in the study?

(b) What is the sample size?

(c) What were the statistics collected?

(d) Which statistics were categorical and which were numerical?

(2) The following numbers are a sample of the monthly rental prices for apartments in Biloxi, Mississippi.

```
> rent
 [1] 290 180 320 260 160 980 470 150 120 850
```

(a) Arrange the data in a stem and leaf diagram. The stems should be in the hundreds place.

(b) Using the stem and leaf diagram, find the minimum, median and maximum.

(c) Using the stem and leaf diagram, do you expect the mean to be larger or smaller than the median? Explian why.

(3) A data set on the size, in millimeters, of a sample of 40 potato bugs produces the following histogram.

Histogram of bugs
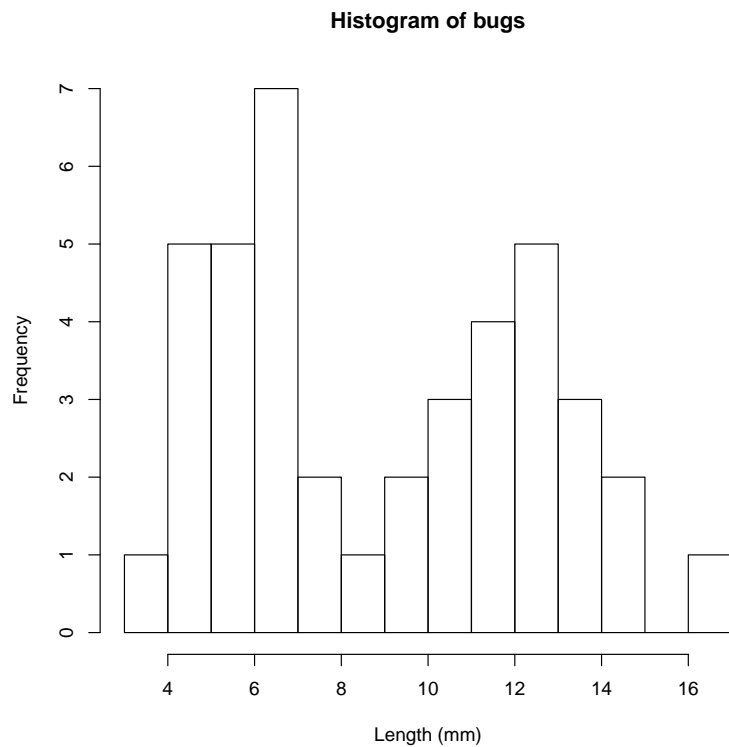
(a) What percentage of bugs have lengths less that 10mm? (None have length exactly 10mm.)

(b) Find the median, $Q1$, and $Q3$. Compute the interquartile range (IQR).
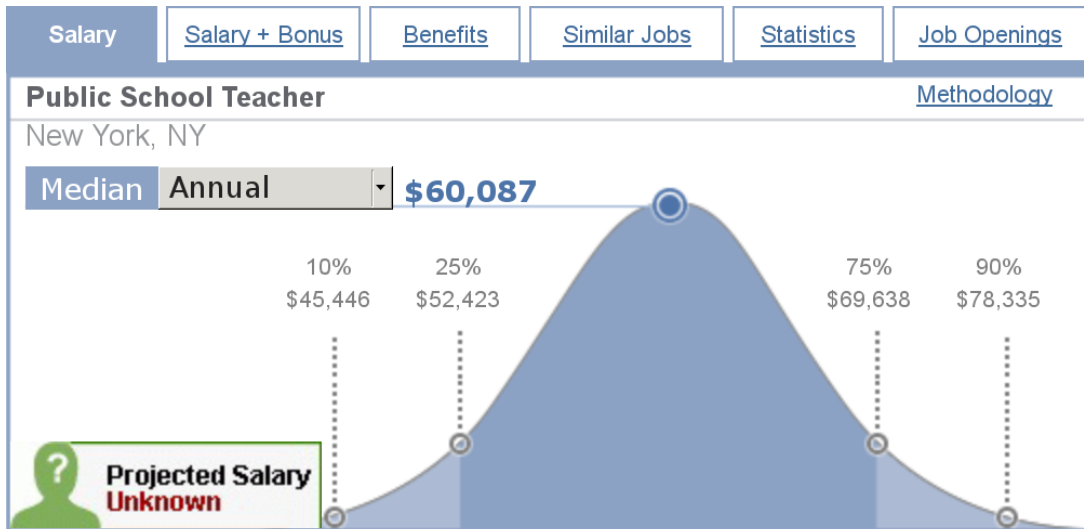
(c) Construct the corresponding box plot.

(4) The following data comes from a study of birtweights of newborn babies. The data set of over 1,000 newborns was broken up by the smoking status of the mother. For each group a boxplot was produced.

Birthweight of newborns by mother's smoking status
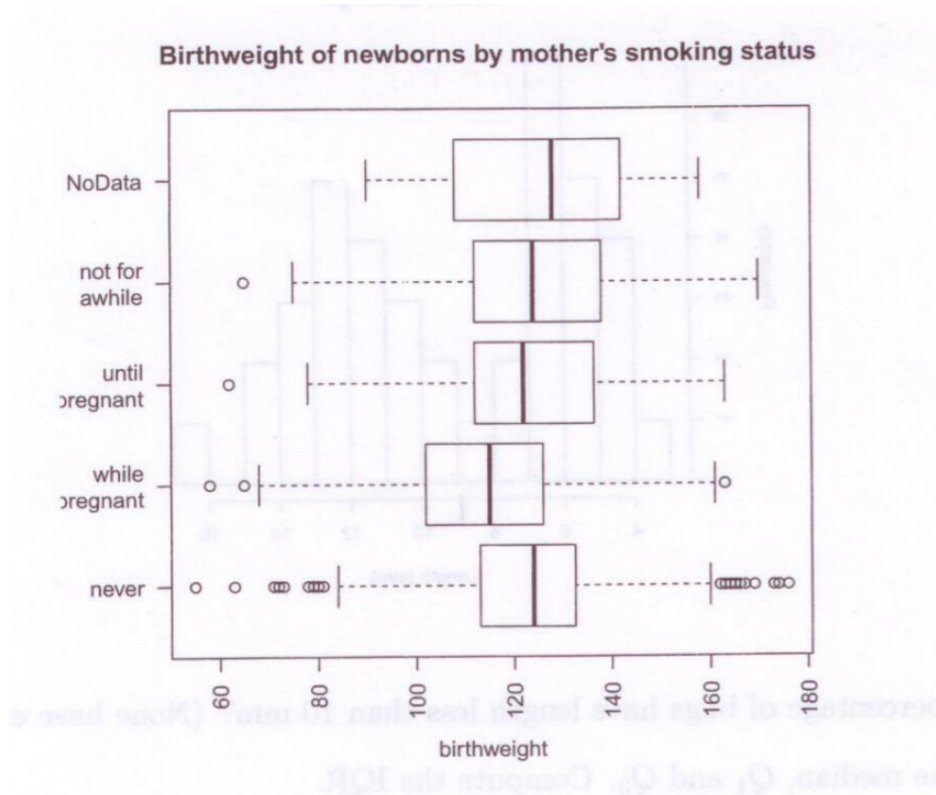
Based on the boxplots, answer the following questions.

(a) Which group has the smallest median birthweight?

(b) Which group contains the baby with the smallest birthweight?

(c) Approximately _____% of babies born to mothers who smoked during pregnancy were below the median birth weight of babies born to mothers who never smoked.

(d) True or False: The group with the largest range of birthweights has the smallest IQR of birthweights.

(e) Based on the boxplots, what would you advise a newly pregnant woman who smokes? Explain you reasoning.

(5) A game consits fo rolling two fair dice, a 4-sided tetrahedron, and a 6-sided cube, and then adding up their scores. The random variable X is the sum of the two numbers on the dice.

    (a) List the sample space of the game, i.e. list all possible outcomes of rolling the dice.

    (b) What is the probability that the sum (X) is equal to 3?

    (c) Find the probabilities $P(X < 5)$ and $P(x \neq 8)$.

    (d) Draw the probablity distribution of X.

(6) Let Z be a normally distributed random variable with mean $\mu = 0$ and standard deviation $\sigma = 1$.

    (a) Find $P(Z < -0.5)$.

    (b) Find $P(0 < Z < 1.5)$.

(7) The average time for an Empire penguin egg to hatch is 70 days, with a standard deviation of 5 days. Assume the distribution of times is normally distributed.

    (a) Find the probability that an egg hatches in 73 days or more.

    (b) Find the probability that an egg hatches between 65 and 75 days.

    (c) By what day have 90% of the eggs hatched?

(8) Below is a screenshot from salary.com showing the distribution of public school salaries in New York City.

| Salary | Salary + Bonus | Benefits | Similar Jobs | Statistics | Job Openings |

Public School Teacher   Methodology

New York, NY

Median Annual $60,087

10%   25%   75%   90%
$45,446 $52,423 $69,638 $78,335

Projected Salary Unknown

(a) What percentage of teachers of earn between $52,423$ and $78,335$?

(b) The distribution is actually skewed a little. Right or Left? What is bigger, mean or median?

(c) Assuming the distribution is approximately normal, what is the standard deviation?