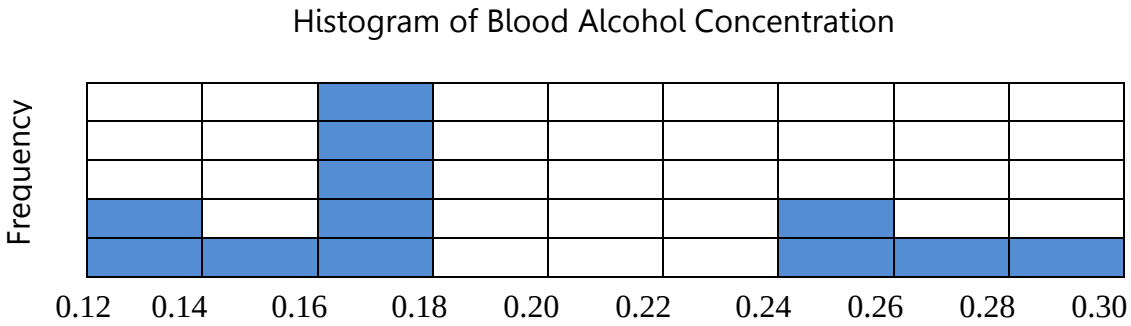
Sections 3.3-3.4, 4.1 Group Activity - SOLUTIONS

Group Activity

1. The following data are blood alcohol concentrations of 12 drivers involved in fatal crashes (data from the U.S. department of justice). We will analyze the shape, center and spread of this data, and whether there are any outliers.

a. On the first grid, make a histogram of the data using a bin-width of 0.02. Label your axes.

Blood alcohol concentration of drivers in fatal crashes
0.27
0.17
0.17
0.16
0.13
0.24
0.29
0.24
0.14
0.16
0.12
0.16



b. What is the shape of the histogram? If you are not sure yet, compare the mean and the median first.
The shape is bimodal and skewed to the right because the mean is greater than the median.

c. Find the mean, median and mode, include units.

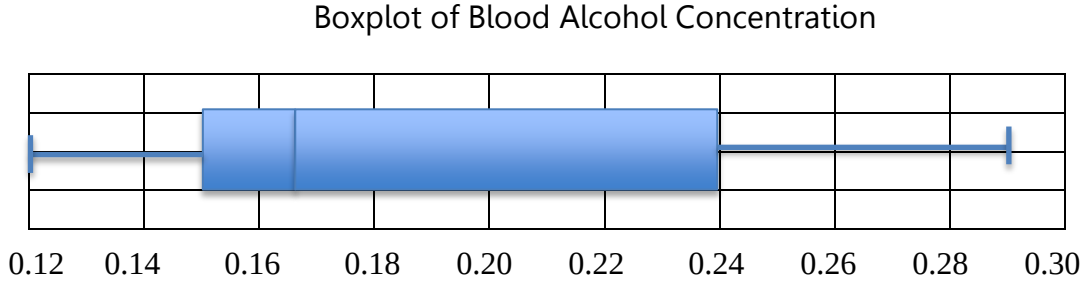
Put in order: 0.12, 0.13, 0.14, 0.16, 0.16, 0.16, 0.17, 0.17, 0.24, 0.24, 0.27, 0.29

The mean is 0.1875 blood alcohol concentration
The median is 0.165 blood alcohol concentration
The mode is the center of the highest peaks at 0.17 blood alcohol concentration

d. Find the 5-number summary, IQR and range, including units.

5-number summary: 0.12, 0.15, 0.165, 0.24, 0.29 blood alcohol concentration
IQR = Q3-Q1 = 0.24 – 0.15 = 0.09 blood alcohol concentration
Range = 0.29 – 0.12 = 0.17 blood alcohol concentration

e. Use the 5-number summary to draw a boxplot on the second grid below. Notice the horizontal scales match on both graph so we can see how they compare.



f. Do you think there are any outliers? Why or why not? There do not seem to be any outliers but there is an unusual gap.

Calculating Standard Deviation, *s*

g. Using your mean rounded to two decimal place, find the standard deviation, including units. The variable n refers to the number of data values.

Mean = ___**0.19 blood alcohol concentration**____, n = ___**12 drivers**_____

Test Score (points)	Deviation from the mean	Squared deviation
0.27 **– 0.19**	**0.08**	**0.0064**
0.17 **– 0.19**	**– 0.02**	**0.0004**
0.17 **– 0.19**	**– 0.02**	**0.0004**
0.16 **– 0.19**	**– 0.03**	**0.0009**
0.13 **– 0.19**	**– 0.06**	**0.0036**
0.24 **– 0.19**	**0.05**	**0.0025**
0.29 **– 0.19**	**0.10**	**0.0100**
0.24 **– 0.19**	**0.05**	**0.0025**
0.14 **– 0.19**	**– 0.05**	**0.0025**
0.16 **– 0.19**	**– 0.03**	**0.0009**
0.12 **– 0.19**	**– 0.07**	**0.0049**
0.16 **– 0.19**	**– 0.03**	**0.0009**
Sum of the squared deviations (numerator)		**0.0359**

$$s = \sqrt{\frac{\sum (x - mean)^2}{n-1}}$$

$$= \sqrt{\frac{0.0359}{12-1}}$$

$$\approx 0.0571 \text{ blood alcohol concentration}$$

4.1: Contingency Tables and Probability - SOLUTIONS

The survey data below is from four classes of Math 105 students. Their gender identities and modes of transportation to PCC are summarized in a contingency table.

	Bike	Bus	Drive Self	Ride with Another	Walk	Total
Female	0	12	25	5	3	45
Non-binary or Genderqueer	0	3	2	0	0	5
Male	1	4	16	3	1	25
Total	1	19	43	8	4	75

2. Find the following marginal, "and", and "or" probabilities.

If we were to randomly select a student who took the survey, what is the probability they:

a. identify as female?
P(female) = 45/75 = .6

b. identify as non-binary or genderqueer?
P(non-binary or genderqueer) = 5/75 = .0667

c. walk to PCC?
P(walk) = 4/75 = .0533

d. bus to PCC?
P(bus) = 19/75 = .2533

e. walk and identify as male?
P(walk and male) = 1/75 = .0133

f. identify as non-binary or genderqueer and drives them self to PCC?
P(non-binary or genderqueer and drives self) = 2/75 = .0267

g. identify as female or ride with another?
P(female or ride with another) = (12+25+5+3+3)/75 = 48/75 = .64

h. identify as male or walk to PCC?
P(male or walk) = (1+4+16+3+1+3)/75 = 28/75 = .3733

3. Calculate these conditional probabilities:

a. Given that a student from the survey identifies as female, what is the probability they take the bus to PCC?

P(bus given female) = 12/45 = .2667

b. If a student drives them self to campus, what is the probability they identify as non-binary or genderqueer?

P(non-binary or genderqueer given drives self) = 2/43 = .0465

c. What is the probability that a student walks, given they identify as male?

P(walk given male) = 1/25 = .04

d. Of those students who identify as female, what is the probability they ride to campus with another?

P(ride with another given female) = 5/45 = .1111

More Practice

1. Class Data. Below are the number of credit hours taken by students for two Math 105 Courses this term.

Spring T/Th Class:

15, 4, 12, 12, 12, 12, 12, 16, 11, 12, 18, 13, 12, 8, 15, 11, 13, 12, 14, 14, 11, 11, 15

Put in Order:
4, 8, 11, 11, 11, 11, 12, 12, 12, 12, 12, 12, 12, 12, 13, 13, 14, 14, 15, 15, 15, 16, 18

Fall M/W Class:
4, 12, 12, 11, 12, 8, 15, 12, 12, 12, 12, 11, 12, 12, 11, 8, 11, 13, 12, 13, 12, 16, 12, 12, 8

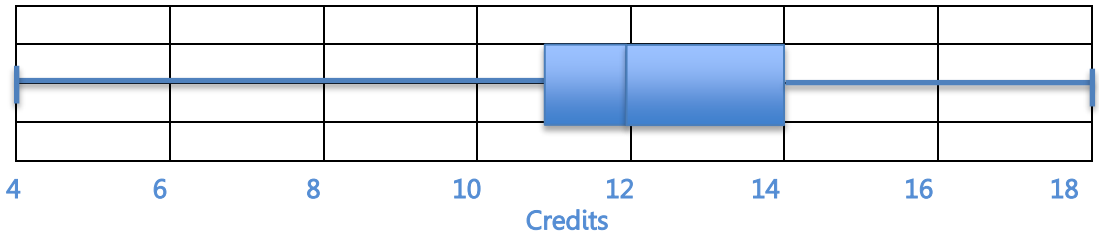
Put in Order:
4, 8, 8, 8, 11, 11, 11, 11, 12, 12, 12, 12, 12, 12, 12, 12, 12, 12, 12, 12, 12, 13, 13, 15, 16

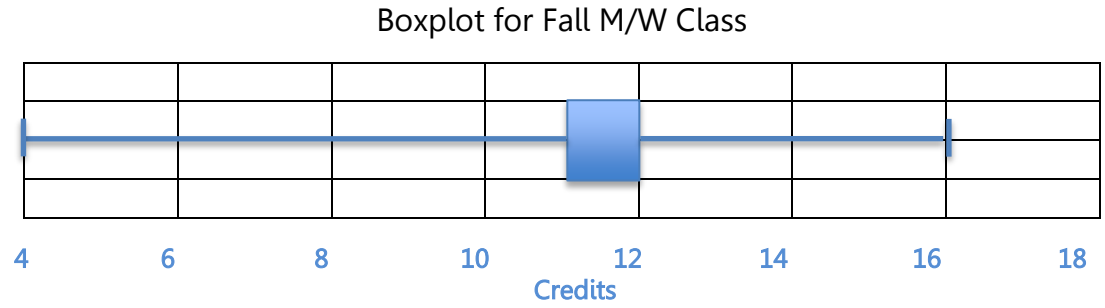
a. Find the mean, mode, 5-number summary, IQR and range for each class.

Spring T/Th Class	Fall M/W Class:
5-number summary: 4, 11, 12, 14, 18 credits	4, 11, 12, 12, 16 credits
Mean: 12.4 credits	11.4 credits
Mode: 12 credits	12 credits
IQR: 14-11 = 3 credits	12-11 = 1 credit
Range: 18-4=14 credits	16-4 = 12 credits

b. Draw the boxplot for each class using the same scale.

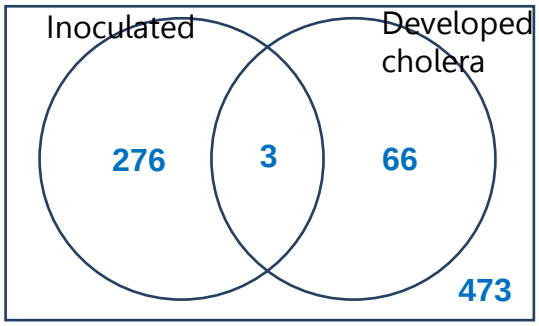
Boxplot for Spring T/Th Class





c. What is the shape of the data for each class? How would you describe the two classes?
Both classes are fairly symmetric but the Spring T/Th class is slightly skewed to the right and the Fall M/W class is slightly skewed to the left.

Cholera Inoculation Study, 1894-96:
A group of 818 people who were exposed to cholera in Calcutta, India in 1894-6 were studied. Of this group, 279 were inoculated with Haffkine's anti-cholera vaccine, while the remaining 539 had not been inoculated. Overall, 69 people developed cholera. Three of those who were vaccinated developed cholera. Use this information to complete the Venn diagram and the contingency table. (Source: https://mysite.du.edu/~jcalvert/econ/twobytwo.htm)



	Developed cholera	Did not develop cholera	Total
Inoculated	**3**	**276**	**279**
Not inoculated	**66**	**473**	**539**
Total	**69**	**749**	**818**

2. Find the marginal, "and," and "or" probabilities. If a randomly selected person from the study was chosen, what is the probability they:

MARGINAL

a. were inoculated?
P(inoculated) = 279/818 = .3411

b. developed cholera?
P(cholera) = 69/818 = .0844

AND

c. were inoculated and developed cholera?
P(inoculated and cholera) = 3/818 = .0037

d. were not inoculated and did not develop cholera?
P(not inoculated and no cholera) = 473/818 = .5782

OR

e. were inoculated or did not develop cholera?

P(inoculated or no cholera) = (3+276+473)/818 = .9193

f. were not inoculated or developed cholera?

P(not inoculated or cholera) = (3+66+473)/818 = .6626

CONDITIONAL

g. Given that a person was inoculated, what is the probability they developed cholera?

P(cholera given inoculated) = 3/279 = .0108

h. If a person was not inoculated, what is the probability they developed cholera?

P(cholera given not inoculated) = 66/539 = .1224

i. Do you think the vaccine was effective? Why or why not?

The chance of getting cholera if vaccinated is about 1% versus 12% if not vaccinated. It seems like the vaccine was effective.