

# EXAM REVIEW 1st Semester

## Unit 1: One Variable Statistics

Name: KEY

Hour:

### ocabulary:

- Numerical Statistical Question → collect data (numbers)
  - Categorical Data Statistical Question → collect data (words)  
*average / typical height*
  - Non-Statistical Question  
*percent of students with a pet*
- one correct answer → don't need to collect data to answer  
(how many students go to HC)

### Know how to make:

- Dot plot
- Histogram (Frequency Table)
- Find 5 number summary
- Box Plot (5 number summary)

### Hints:

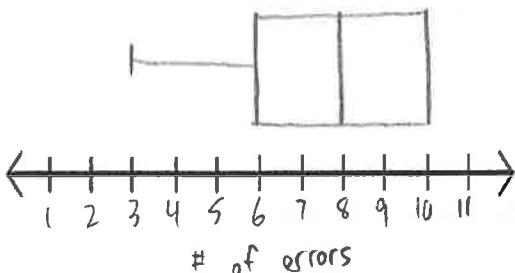
- Put data in numerical order if it isn't done already.
- Make a frequency table
- Include labels

### Problem Set:

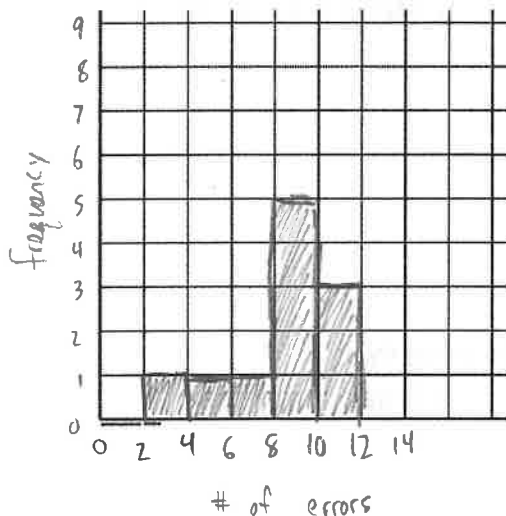
Use the data set to represent the number of errors on a typing test: ~~3, 5, 6, 8, 10, 8, 8, 9, 10, 10, 9~~  
3, 5, 6, 8, 8, 8, 9, 9, 10, 10, 10

1. Create a box plot:

Min: 3  
Q<sub>1</sub>: 6  
Med: 8  
Q<sub>3</sub>: 10  
Max: 10



2. Create a histogram:



# of errors	Frequency
0-2	0
2-4	1
4-6	1
6-8	1
8-10	5
10-12	3

3. Create a Dot Plot



**Vocabulary: Draw an example of**

Skewed to the right:



Skewed to the left:



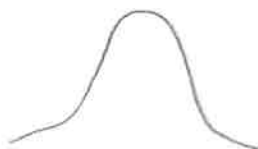
Bimodal:



Uniform:



Bell - Shaped:



Describe what a Symmetric graph would like look:

foldable using a vertical line



- When describing the center of a graph, when do you use the mean?

*symmetric*

- When describing the center of a graph, when do you use the median?

*skewed or outliers*

**Problem Set:**

4. Draw a dot plot with at least 8 data points such that:

a. mean = median

b. Mean > median

c. mean < median

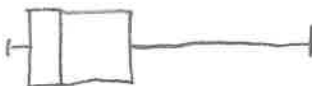
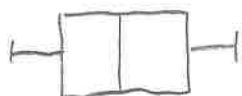


5. Draw a box and whiskers plot such that:

a. mean = median

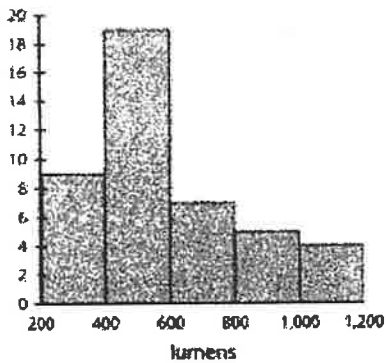
b. Mean > median

c. mean < median



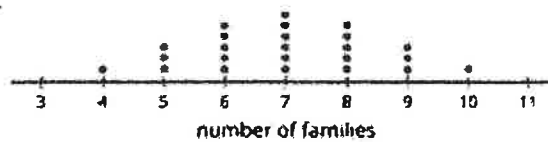
6. Describe the shape in as many ways as possible of the following graphs:

a.



skewed right

b.



symmetric, bell-shaped

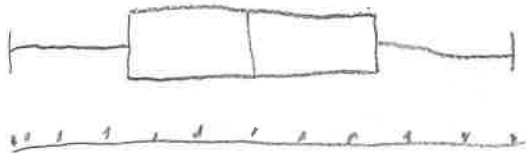
c.



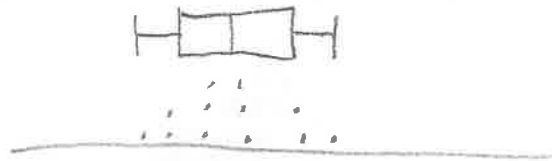
symmetric, bimodal

### COMPARING AND CONTRASTING DATA DISTRIBUTIONS (Draw box plots)

Draw a graph that has high variability



Draw a graph that has medium variability.



Draw a graph that has low variability



What does variability mean in a graph?

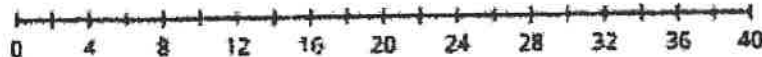
amount of spread from center data has

### Problem set:

7: Rank the box plots from the greatest variability (1) to the least variability (4):



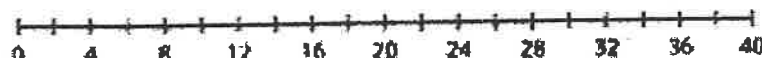
(3) tie



(3)



(1)



(2)



## STANDARD DEVIATION (Draw dot plots)

Draw a graph that has a SMALL standard deviation. (Explain why)



data is all close to center  
(not much spread)

Draw a graph with a standard deviation of 0. (Explain why)



data is all at the center (no variability)

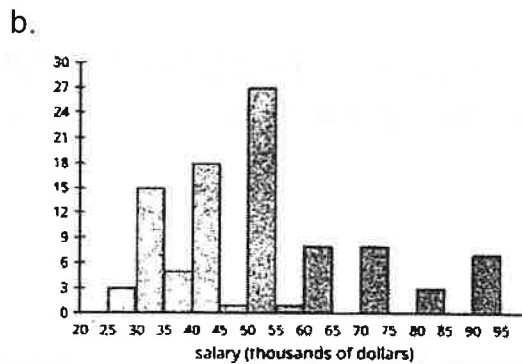
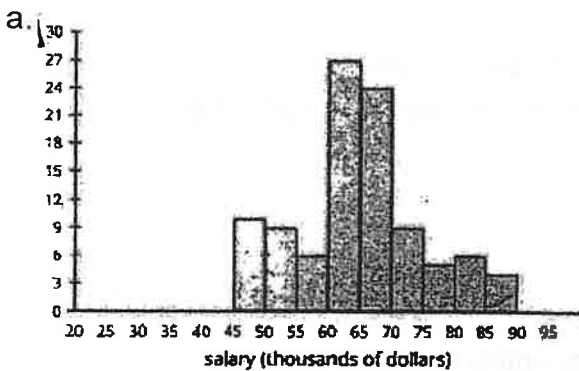
Draw a graph that has a BIG standard deviation. (Explain why)



data is spread out from center

## Problem Set:

8: Which graph has a higher standard deviation and why?



## OUTLIERS

How to calculate outliers given a set of data?

$$IQR = Q_3 - Q_1$$

Low Outlier:  $Q_1 - 1.5(IQR)$

High Outlier:  $Q_3 + 1.5(IQR)$

When do you include an outlier in a set of data? When do you take out an outlier from a set of data?

if it is actual/accurate data

when you know the outlier is an error (not accurate)

## Problem Set:

9: Calculate if there are any outliers in the following data set (SHOW ALL WORK!)

6      6      7       $Q_1$  8      8      8      9       $M$  10

10      12      13       $Q_3$  14      15      16      30

$$IQR = 14 - 8 = 6$$

Low boundary:  $Q_1 - 1.5(IQR)$

$$8 - 1.5(6) = 8 - 9 = -1$$

nothing lower than -1

high boundary:  $Q_3 + 1.5(IQR)$

$$14 + 1.5(6) = 14 + 9 = 23$$

$30 > 23$ , so 30 is an outlier