

Notes Day 1: 1-Vars Stat, Dot Plots and Histograms, Shape, Center, Spread, Outlier, Descriptive Analysis

Warm Up:

1. At a certain intersection, the light for eastbound traffic is red for 15 seconds, yellow for 5 seconds, and green for 30 seconds. Find the probability that out of the next eight eastbound cars that arrive randomly at the light, exactly three will be stopped by a red light.

$$\text{binompdf}(8, .3, 3) = 0.254 \quad 25.4\%$$

2. A company makes electronic gadgets. One out of every 50 gadgets is faulty, but the company doesn't know which ones are faulty until a buyer complains. Suppose the company makes a \$3 profit on the sale of any working gadget, but suffers a loss of \$80 for every faulty gadget because they have to repair the unit. Check whether the company can expect a profit in the long term.

$$\begin{array}{r|l} x & 3 \quad -80 \\ \hline P(x) & \frac{49}{50} \quad \frac{1}{50} \end{array} = 1.34$$

3. Jones is the Chairman of a committee. In how many ways can a committee of 5 be chosen from 10 people given that Jones must be one of them?

$${}^9C_4 = 126$$

Notes: In Statistics, we work with two types of variables:

Variable: Any characteristic whose value may change from one variable to another

- categorical

Qualitative \Rightarrow Variable that places an individual into a category.

- quantitative

Numerical \Rightarrow Variables that take on numerical values

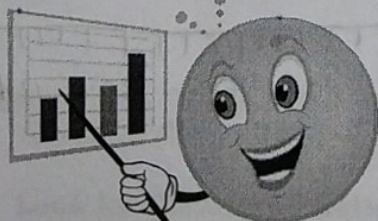
Identify the Following:

- gender C
- age Q
- hair color C
- smoker C
- systolic blood pressure Q
- number of girls in class Q

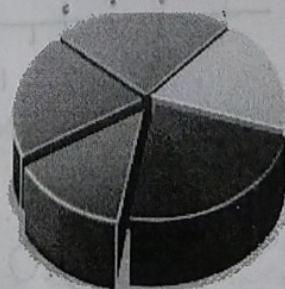
How do we display these two types of data?

Categorical Data:

Bar Graphs



Pie Charts



Refresher on Bar Graphs:

Take the data from your class survey and make a bar graph from each person's favorite music.

fav. music

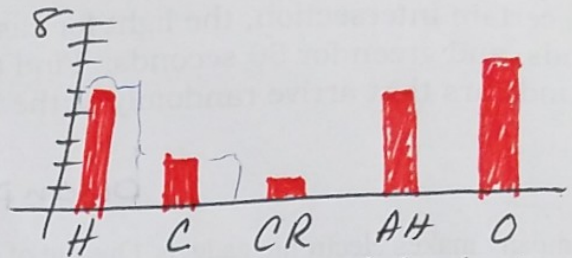
Hip hop 5

Country 2

Classic Rock 1

AH. 5

Other. 6

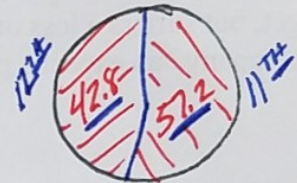


Refresher on Pie Chart:

Take the data from your class survey and make a pie chart from the grade level category in percents!

$$\frac{11}{21} = 52.4\% \quad \frac{9}{21} = 42.9\%$$

$$\frac{12}{21} = 57.1\% \quad 42.8\% \quad 57.2\%$$



String Theory Dot Plots

To describe these two types of graphical representation, let's take a look at the class survey!

Class Survey

Is there a way to take the quantitative data from of class survey and represent it graphically? Yes!

String Theory -- Dot Plots

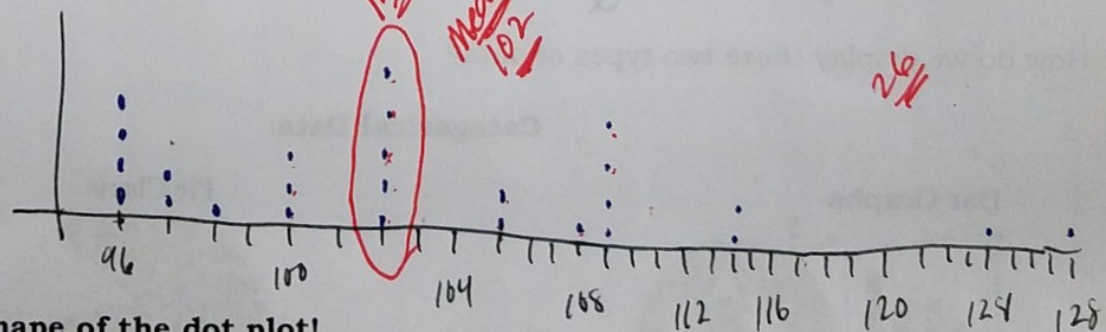
Step 1: How should we organize the data?

~~96, 97, 124, 98, 114, 102, 100, 96, 100~~
~~108, 102, 105, 102, 102, 108, 97, 114, 102~~
~~102, 96, 96, 108, 108, 105, 128, 100~~

Step 2: What can help us with organizing the data?

~~96, 96, 96, 96, 97, 97, 98, 100, 100, 100~~
~~102, 102, 102, 102, 102, 105, 105, 107, 108, 108~~
~~108, 108, 114, 114, 124, 128~~

Create Your Dot Plot!



Comment on the shape of the dot plot!

Shape
outliers
Center
Spread

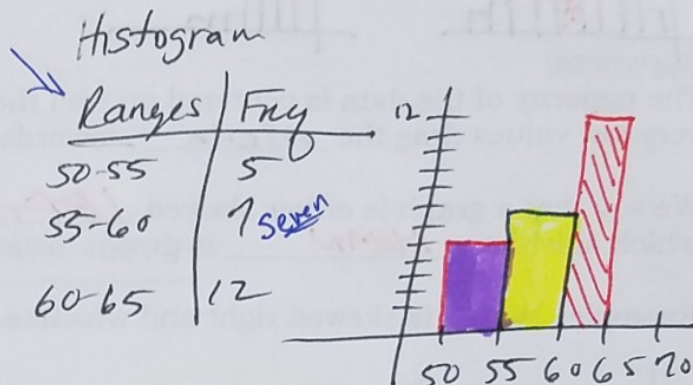
S.O.C.S



Histograms are another type of graph we can use to represent quantitative data.

- Using the information from the Getting to Know You Survey, create a histogram showing the distribution of the height in inches for members of this class.
- For histograms, you have equal-sized categories (width of the bars).
- Remember to label your graph. Specify the classes
- You should have at least 5 bars that touch

Comment on the shape of the histogram!



Let's try it with the calculator!

You Try! Make a dot plot and histogram from another category from your class survey and compare each graph. Use the same category for both the dot plot and histogram.

Dot Plot

Histogram

Compare each graph and discuss the pros and cons of each one. Is one better than the other?

When describing a distribution, remember your **S.O.C.S!!!**

S HAPE
O UT L I N E S
C ENTER
S PREAD



who wears these!?

Examining a Distribution

- Describe the pattern of a distribution by its Shape, center, spread.
- Look for the overall pattern and for any Deviation from that pattern.

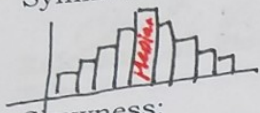
↑ outliers
Standard Deviation
Mean Deviation

Displaying quantitative variables

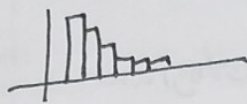
- Dot plots, stem plots, → Histograms
- To describe the overall pattern don't forget your SOCs!

Shape

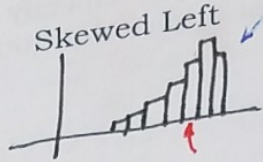
Symmetric



Skewed Right



Skewed Left



Uniform

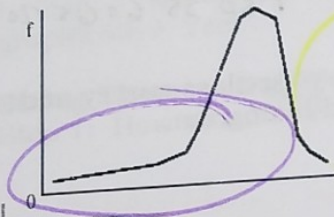


Skewness:

The majority of the data is centered around the Median while a few very high or very low values drag the Mean towards that side.

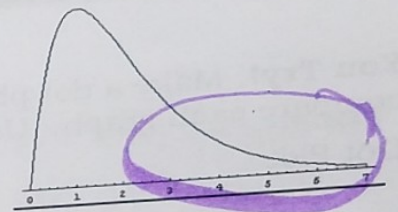
We say that a graph is either skewed LEFT or skewed RIGHT telling us which side the MEAN is drawn towards.

Example: Which is skewed right and which is skewed left?



Skewed left

Skewed Right



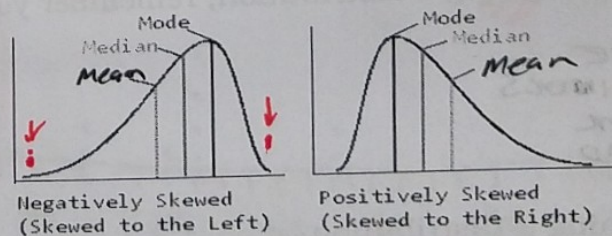
You Try! You look at real estate ads in Naples, FL. There are many houses ranging from \$200,000 to \$500,000 in price. The few houses on the water, however, have prices up to \$15 million. The distribution of these houses will be

- (a) skewed to the right
- (b) roughly symmetric
- (c) skewed to the left
- (d) too high

Outliers - is an individual DEVIATION that falls outside the overall pattern of the graph.

Center

MEAN = AVERAGE
MEDIAN = MIDDLE NUMBER
MODE = MOST REPETITIONS



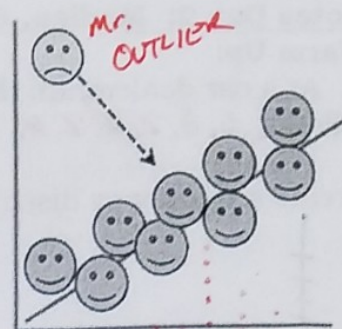
Spread

- Just use the RANGE (maximum - minimum)

Big - small = RANGE

Deviations from the Pattern

- outliers
- clusters
- gaps
- Anything else unusual



MEAN → Average

$$\bar{X} = \mu = \frac{\text{ADD ALL NUMBERS}}{\# \text{ of NUMBERS}}$$

MEDIAN → MIDDLE

EX: $\overset{3}{\text{5, 7, 11, 19, 21, 25, 35}}$
 $\underset{1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7}{\quad}$

92, 93, 94, 94.5, 95, 100, 100

$$\frac{94 + 95}{2}$$

Range

5, 7, 11, 19, 21, 25, 35

$$35 - 5 = 30$$

MODE → MOST

$$\text{MODE} = \underline{16}$$

16, 18, 17, 17, 16, 16, 16, 17, 17, 16, 16, 16, 16, 16, 16, 16, 16,
 16, 16, 17, 16, 16, 16, 17, 16, 17, 36