

FREC 408

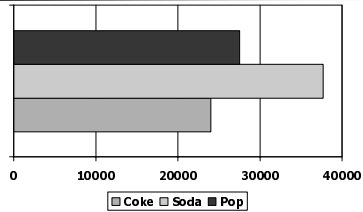
Group Exercise 2

What do you call a carbonated beverage?



Coke
Soda
Pop

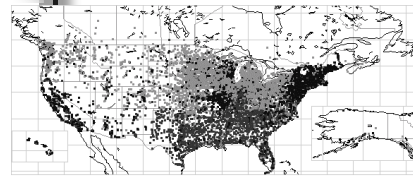
Based on an Internet Survey



Source: CNN.com

Coke, soda, or pop? Netizens drink up e-cola poll
September 12, 2002 Posted: 4:00 AM EDT (0800 GMT)

It is related to residence



Pop vs. Soda data as of September 13, 2002

Pop Soda Coke Other

Green areas order pop
Red areas call it Coke
Blue areas ask for soda

Source: **The Great Pop vs. Soda Controversy**
http://www.ugcs.caltech.edu/~almccoon/pop_soda/

GROUP EXERCISE

- Work together and share information
- Make sure everyone is involved and has a chance to
 - Contribute
 - Learn
- Seek me out if the group needs help
- Let me know if there are problems

THE GROUPS

GROUP 1	GROUP 2	GROUP 3	GROUP 4	GROUP 5
Albertus	Ash	Barkocy	Beaty	Bienstock
Cormier	DeWire	Dundas	Fausnaught	Frosch
Hubner	Lancos	Li	McClelland	Milligan
Sentoff	Smith	Spinner	Stehly	Thompson
Williamson	Yingst	Brown	Pisani	Ferber
GROUP 6	GROUP 7	GROUP 8	GROUP 9	
Calvanico	Campanicki	Ciappio	Colbert	
Goeller	Grezikowski	Hackler	Hansen	
O'Neill	Perza	Reilly	Riniker	
Tosatto	Tuttle	Veenema	Walker	
Blackwell				

GROUP EXERCISE 1 – Question 1

Mean	=	60.36
Median	=	64
Mode	=	84
Maximum	=	94
Minimum	=	11
Range	=	83
Variance	=	390.92
Std Dev	=	19.77

GROUP EXERCISE 1 – Question 1

- b. What is the *position* of the median value for this data?
 • $(115+1)/2 = 58^{\text{th}}$ case
- c. Does the mode make sense as a measure of Central Tendency for this data?
 • Not really

GROUP EXERCISE 1 – Question 1

- d. Based on what you know about getting into universities, why do you think this distribution is the way it is?

The spread is very large.

It might reflect differences between public and private institutions. Private institutions generally have lower acceptance rate. Public schools may have as part of their mission to have higher rates of acceptance to provide educational opportunities to citizens in the state.

We could think of this data as being two populations.

GROUP EXERCISE 1 – Question 2

- a. What is the primary disadvantage of using the range to compare the variability of data sets?
 • It only uses two values - the maximum and the minimum - to calculate the range. It can be very sensitive to outliers.
- b. Describe the sample standard deviation using words rather than a formula.
 • It is the average deviation around the center (in this case the center is the mean).

GROUP EXERCISE 1 – Question 2

- c. Can the variance of a data set ever be negative? Explain.
 • Since the variance is based on a squared measure, no, it cannot be negative
- d. Can the variance ever be smaller than the standard deviation? Explain.
 • In most cases the standard deviation is less than the variance since it is a square root of the variance. However, in the special case where the variance is between 0 and 1, the standard deviation will be more than the variance.
- For example, if
- $S^2 = .5$
 - $s = .71$

GROUP EXERCISE 1 – Question 3

Group Exercise 2 for STAT 200

Data	Squared	Data		
1	10	100		
2	11	121	Mean	14.55
3	12	144	Standard Error	0.62
4	13	169	Median	15.00
5	15	225	Mode	15.00
6	15	225	Standard Deviation	2.73
7	15	225	Sample Variance	7.47
8	16	256	Kurtosis	-1.02
9	17	289	Skewness	-0.35
10	18	324	Range	8
11	18	324	Minimum	10
			Maximum	18
Sum	160	2402	Sum	160
			Count	11

**GROUP EXERCISE 1 –
Question 3**

- a. Mean = $160/11 = 14.55$
- b. Median = 6th position of ordered data = 15.00
- c. Mode = most frequent value = 15
- d. Range = $18 - 10 = 8$
- e. Variance = $[2402 - (160)^2/11]/10 = 7.47$
- f. Std. Deviation = 2.73

Z-Scores

$$z_i = \frac{(x_i - \bar{x})}{s}$$

- This is a method of transforming the data to reflect relative standing of the value
- We subtract the mean and divide by the standard deviation
- The result represents the distance between a given measurement x and the mean, expressed in standard deviations

**GROUP EXERCISE 1 –
Question 3**

- g. z-score for a value of 17

$$Z = (17 - 14.55)/2.73 = .897$$