Data and slope and y-intercepts, Oh My!

Linear Regression in the Common Core

Jared Derksen
mrmathman.com/talks
jared@mrmathman.com

Outline

I. Data collection-Cheerios

II. History lesson

III. More data collection-wire and red vines

IV. Math Chat

V. More examples and resources

VI. Technology resources

Thanks to: http://purpleprontopups.wordpress.com/ and http://statteacher.blogspot.com for sharing the correlation station ideas with me via Twitter and their blogs.

Common Core Standards

8.SP Investigate patterns of association in bivariate data.

1. Construct and interpret scatter plots for bivariate measurement data to investigate patterns of association between two quantities. Describe patterns such as clustering, outliers, positive or negative association, linear association, and nonlinear association.

2. Know that straight lines are widely used to model relationships between two quantitative variables. For scatter plots that suggest a linear association, informally fit a straight line, and informally assess the model fit by judging the closeness of the data points to the line.

3. Use the equation of a linear model to solve problems in the context of bivariate measurement data, interpreting the slope and intercept. For example, in a linear model for a biology experiment, interpret a slope of 1.5 cm/hr as meaning that an additional hour of sunlight each day is associated with an additional 1.5 cm in mature plant height.

S.ID Summarize, represent, and interpret data on two quantitative variables

6. Represent data on two quantitative variables on a scatter plot, and describe how the variables are related.

a. Fit a function to the data; use functions fitted to data to solve problems in the context of the data. Use given functions or choose a function suggested by the context. Emphasize linear, quadratic, and exponential models.

b. Informally assess the fit of a function by plotting and analyzing residuals.

c. Fit a linear function for a scatter plot that suggests a linear association.

S.ID Interpret linear models

7. Interpret the slope (rate of change) and the intercept (constant term) of a linear model in the context of the data.

8. Compute (using technology) and interpret the correlation coefficient of a linear fit.

9. Distinguish between correlation and causation.

Station One

On Station One's poster, with x values 0-31 and y values 0-31, x equals the day of the month of your birthday and y represents the number of days left after your birthday in your birth month.

Number of Days in each Month:

January 31 days	February - 29 days
March 31 days	April 30 days
May 31 days	June 30 days
July 31 days	August 31 days
September 30 days	October 31 days
November 30 days	December 31 days

For example, Mr. Derksen's birthday is on January 28th (FYI, I love doughnuts) so his point would be at:

X = L1 = Birth day = 28

Y = L2 = Number of days left in month = 3

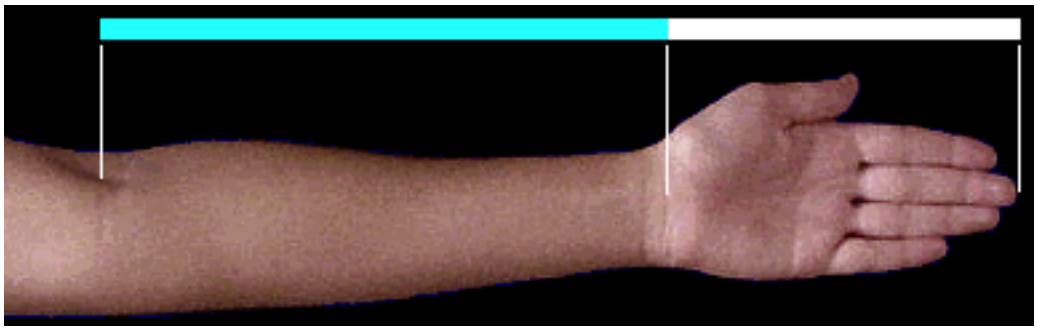
Station Two

Measure your forearm in millimeters.

Measure your forearm + your hand in millimeters.

X = L1 = forearm

Y = L2 = forearm + hand

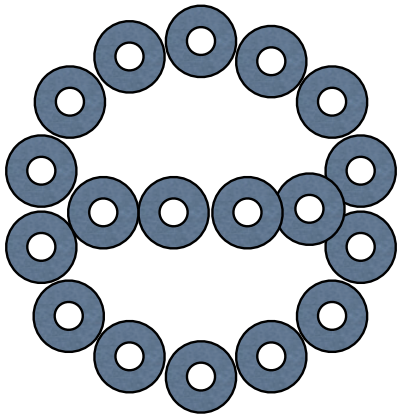


Station Three

Make a circle of cheerios. Count the number it takes to make the circle. Count the number it takes to make the diameter.

X = L1 = diameter

Y = L2 = # around

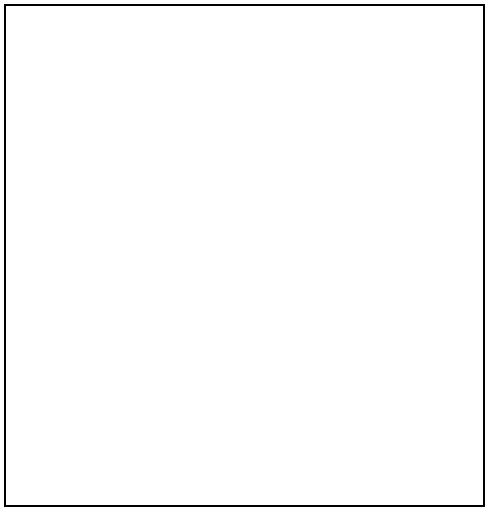


Station Four

Make a circle FULL of cheerios. Count the number it takes to fill entire the circle. Count the number it takes to make the radius.

X = L1 = radius

Y = L2 = total # to make the circle



Station Five

Roll the die. This is the number of knots you will tie in the wire.
After tying the knots, measure the length of the wire in inches.
Don't forget to untie the knots before leaving this station!!
X = L1 = # of knots
Y = L2 = length of wire

Station Six

Take a bite of a red vine.
Measure the remaining piece in cm.
Weigh it.
Repeat (two data points per team)
X = L1 = length (cm)
Y = L2 = weight (g)

Station Seven

X = L1 = # of letters in your first name
Y = L2 = length of the longest hair on your head (inches)

Station Eight

X = L1 = shoe size
Y = L2 = length of the longest hair on your head (inches)

Other data sets

car	hp	mpg
Audi A4	211	30
BMW 3 series	230	28
Buick LaCrosse	182	30
Chevy Cobalt	155	37
Chevy Suburban 1500	320	21
Ford Expedition	310	20
GMC Yukon	320	21
Honda Civic	140	34
Honda Accord	177	31
Hyundai Elantra	138	35
Lexus IS 350	306	25
Lincoln Navigator	310	20
Mazda Tribute	171	28
Toyota Camry	169	33
Volkswagen Beetle	150	28

Team	total runs	attendance
New York Yankees	859	46491
Boston Red Sox	818	37610
Tampa Bay Rays	802	22758
Texas Rangers	787	30928
Minnesota Twins	781	39798
Toronto Blue Jays	755	20068
Chicago White Sox	752	27091
Detroit Tigers	751	30385
Los Angeles Angels	681	40133
Kansas City Royals	676	20191
Oakland Athletics	663	17511
Cleveland Indians	646	17435
Baltimore Orioles	613	21662
Seattle Mariners	513	25746

	interest	amount
1985	11.93	65.3
1986	10.09	75.5
1987	9.52	81.9
1988	10.04	82.8
1989	10.21	96.7
1990	10.06	97.4
1991	9.38	101.0
1992	8.21	104.4
1993	7.27	101.9
1994	7.98	96.1
1995	8.01	99.4
1996	7.81	107.2
1997	7.73	118.7
1998	7.05	124.9
1999	7.32	126.4
2000	8.14	129.2
2001	7.03	143.9
2002	6.62	148.9
2003	5.83	155.4
2004	5.95	158.2
2005	6.00	183.0
2006	6.60	200.3
2007	6.44	210.7
2008	6.09	213.1
2009	5.06	215.8
2010	4.84	212.6

Stats, Modeling the World
Bock, Velleman, DeVeaux

Technology Resources

Excel
slope(y, x)
intercept(y, x)
correl (x, y)

Statcrunch
www.statcrunch.com

StatKey
http://lock5stat.com/statkey/index.html

Casio AP Stats Guide
http://www.casioeducation.com/educators/activities

Texas Instruments
Make sure to go to Catalog-Diagnostic On

NCTM
http://www.nctm.org/resources/content.aspx?id=32706

Useful Applets
http://media.pearsoncmg.com/aw/aw_bock_statsmodel_3/cw/smw3e_references.html

Data Sources
http://lib.stat.cmu.edu/DASL/DataArchive.html

Smarter Balanced Calculator (includes a regression tool!)
http://sbac.portal.airast.org/practice-test/calculators/

Guess My Correlation
http://istics.net/stat/correlations/