# Global Warming or Global Cooling?



Mr. Chen

AP Statistics

# California Standards

- **8.0** Students determine the mean and the standard deviation of a normally distributed random variable.

- **12.0** Students find the line of best fit to a given distribution of data by using least squares regression.

- **13.0** Students know what the *correlation coefficient of two variables* means and are familiar with the coefficient's properties.

- **14.0** Students organize and describe distributions of data by using a number of different methods, including frequency tables, histograms, standard line graphs and bar graphs, stem-and-leaf displays, scatterplots, and box-and-whisker plots.

- **17.0** Students determine confidence intervals for a simple random sample from a normal distribution of data and determine the sample size required for a desired margin of error.

- **18.0** Students determine the *P*-value for a statistic for a simple random sample from a normal distribution.

# Data: Temperature (*Fahrenheit)*
# New York City, December, 1995-2009

- 1995    35.8
- 1996    48.7
- 1997    38.1
- 1998    39.0
- 1999    40.6
- 2000    28.1
- 2001    29.1
- 2002    36.8
- 2003    47.3
- 2004    42.2
- 2005    44.2
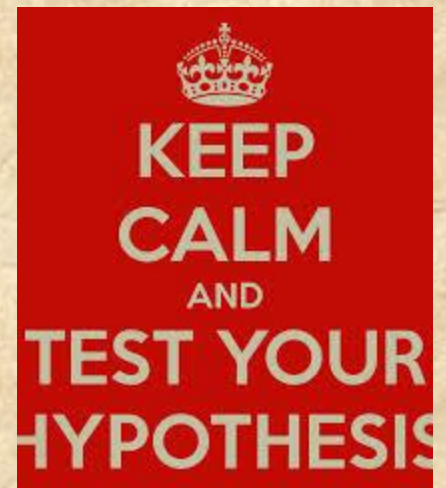- 2006    41.1
- 2007    41.7
- 2008    42.1
- 2009    23.6



University of Dayton, Temperature Archive
http://academic.udayton.edu/kissock/http/Weather/default.htm

# Hypothesis

- We predict the average temperature for New York City in December is $30^oF$, the standard deviation is $5^oF$.

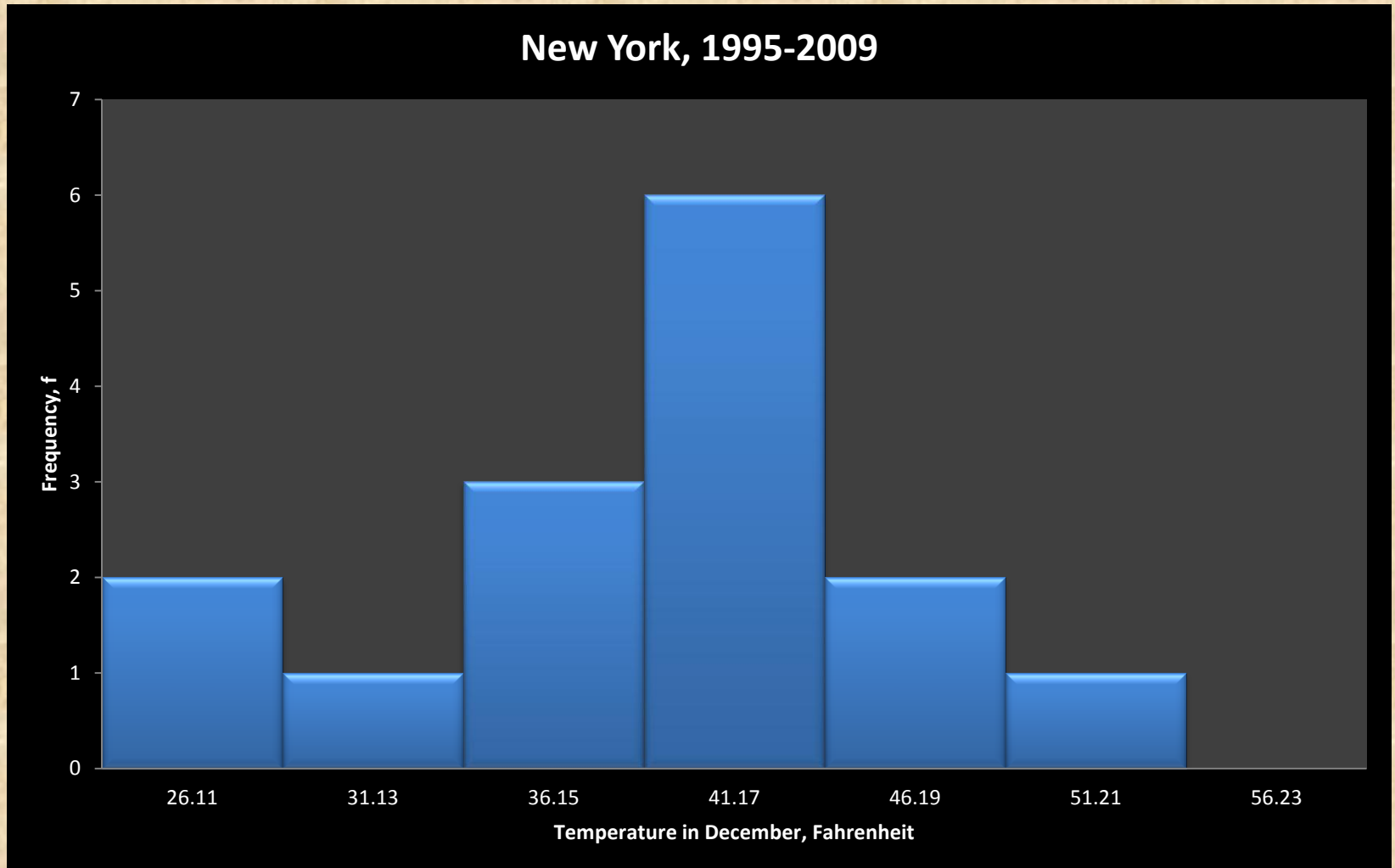- There might be a very weak or no correlation between time and temperature.

# Statistics

- Sample mean, $=38.56^o$F

- Sample Standard Deviation, $=7.0260^o$F

# Histogram

# Confidence Interval
# for Population Mean:

We are 95% confident the population mean temperature of New York in December is between 34.76 to 42.51 degrees Fahrenheit with a marginal error of 3.8 degrees

# Confidence Interval
# for Population Standard Deviation:

We are 95% confident the population standard deviation is between 34.76 to 42.51 degrees Fahrenheit with a marginal error of 3.8 degrees

$$x_R^2 = \frac{1-0.95}{2} = 0.025$$

$$x_L^2 = \frac{1+0.95}{2} = 0.975$$

n = 15, d.f. =14

$$x_R^2 = 26.119$$

$$x_L^2 = 5.629$$

$$\frac{(n-1)s^2}{x_R^2} < \sigma^2 < \frac{(n-1)s^2}{x_L^2}$$

$$\frac{(15-1)7.0260^2}{26.119} < \sigma^2 < \frac{(15-1)7.0260^2}{5.629}$$

$$26.46 < \sigma^2 < 122.78$$

$$5.14 < \sigma 11.08$$

# Hypothesis Test for Mean

Hypothesis:

$H_0 : \mu = 30 (claim)$

$H_a : \mu \neq 30$

Test Statistics:

**T-Test**

$p - value = 0.00033 < \alpha = 0.05$

Conclusion:

At 5% level of significance, we reject $H_0$. There is not enough evidence to support the claim the mean temperature in New York during the month of December is $30^0 F$

# Hypothesis Test
# for Standard Deviation

Hypothesis:

$H_0 : \sigma = 5 (claim)$

$H_a : \sigma \neq 5$

Test Statistics:

d.f.=14

since it is **two tailed**

$x_R^2 = \dfrac{1-0.95}{2} = 0.025$

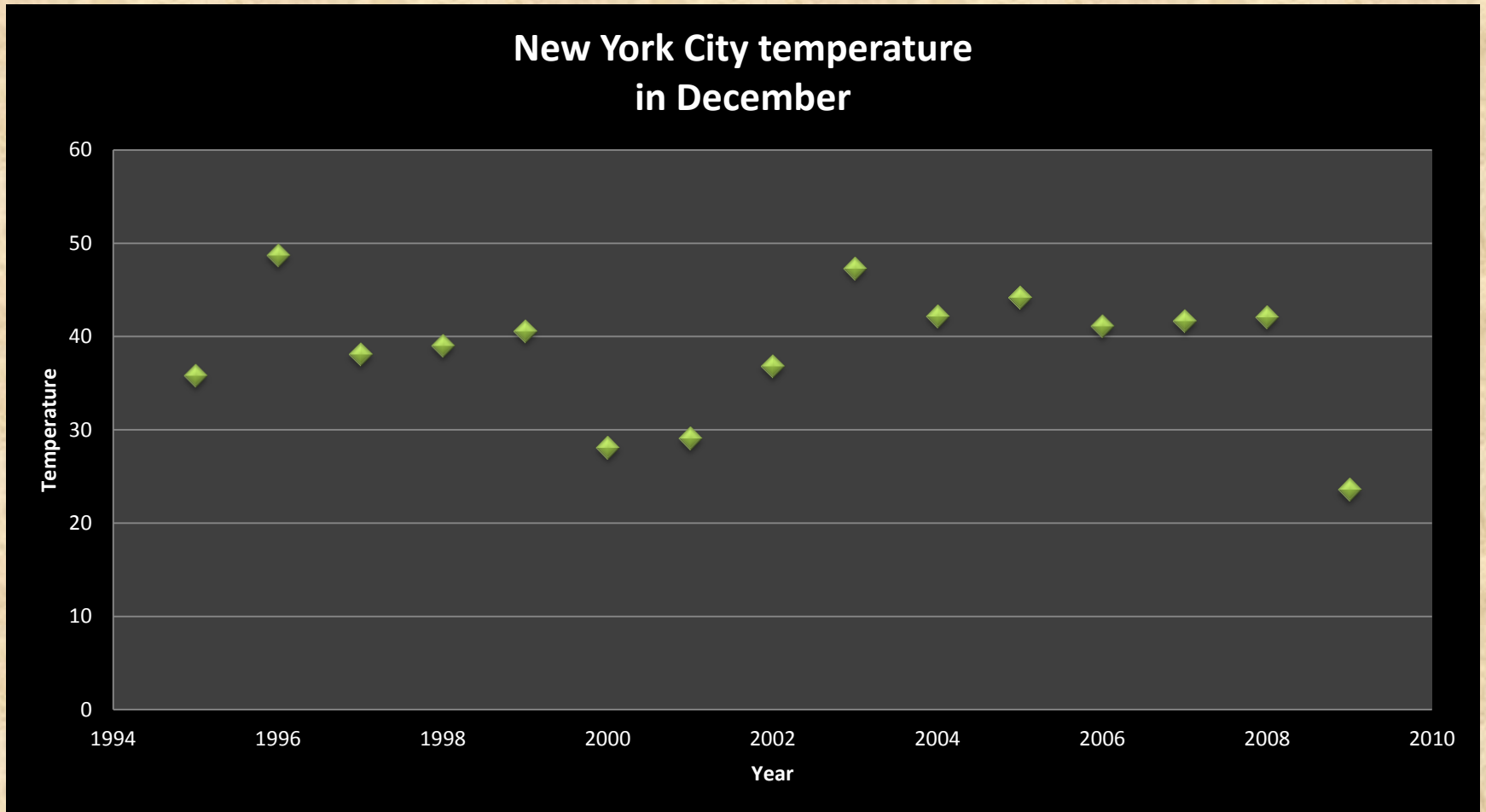$x_L^2 = \dfrac{1+0.95}{2} = 0.975$

n = 15, d.f. =14

$x_R^2 = 26.119$

$x_L^2 = 5.629$

$x^2 = \dfrac{(n-1)s^2}{\sigma^2} = \dfrac{(15-1)7.0260^2}{5^2} = 27.644$

Conclusion:

At 5% level of significance, we reject $H_0$. There is not enough evidence to support the claim the standard deviation of the temperature in New York during the month of December is $5^0 F$

# Scatter Plot



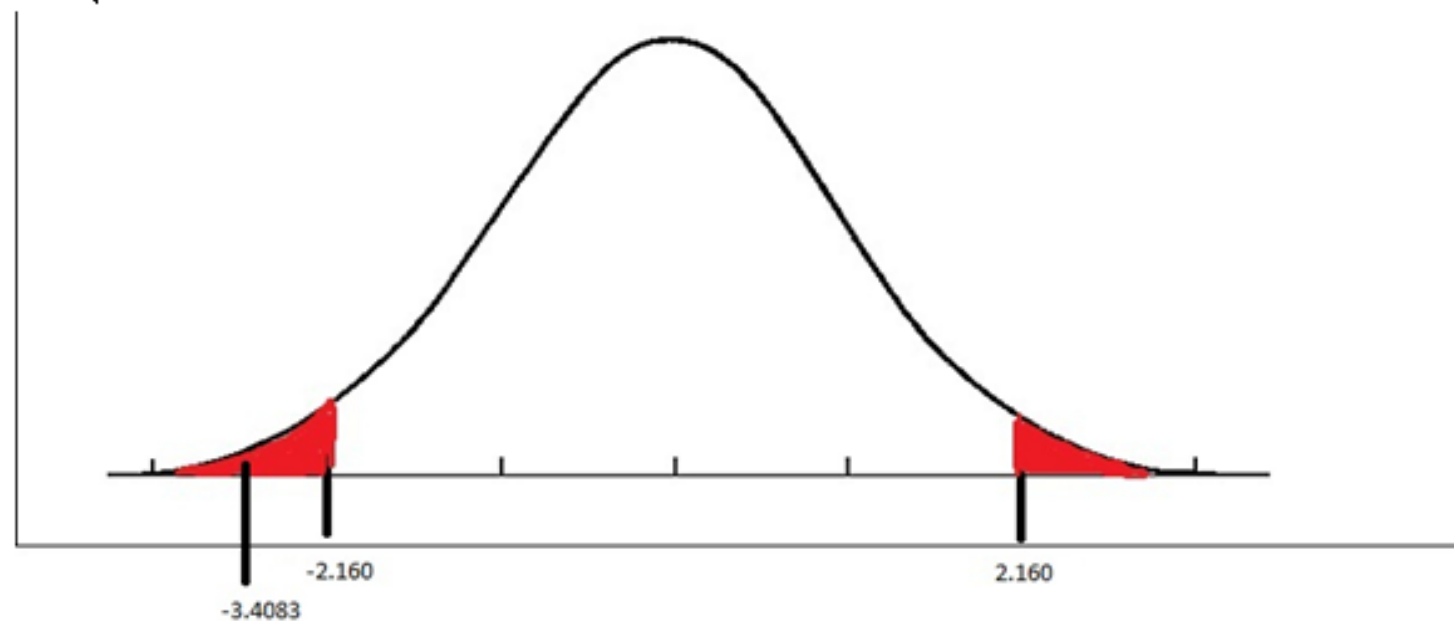New York City temperature in December

Hypothesis:

$H_0 : \rho = 0 (claim)$

$H_a : \rho \neq 0$

Test Statistics:

$\frac{1}{2}\alpha = 0.025$

**d.f.=n-2=15-2=13**

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} = -3.4083$$



Conclusion:

At 5% level of significance, we reject $H_0$. There is not enough evidence to support the claim; there is a correlation between time and temperature.

# Linear Model:
# A linear regression approach, y=ax+b

**Regression Line Equation**

**y = -0.1479x + 334.57**

**For every one year increase in time, there is a drop of temperature of -0.1479 degrees Fahrenheit for the past 15 years**

Correlation coefficient, r = -0.09411
and the coefficient of determination, is $r^2$ = 0.009
Only 0.9% of the variation in y (temperature) is explained by least-square regression of y (temperature) on x (year).

# Future Studies

- Base on the data collected in this study, there is not enough evidence to draw conclusion on the global temperature changes. The linear regression model shows the temperature is neither increasing nor decreasing. The hypothesis testing for correlation however suggests a possible correlation do exist. Future studies include looking at more data over a longer period of time, more cities around the world and a better regression model such as the Sine regression model.