

# LEAN Ohio

**Transforming the Public Sector**

**Green Belt**

**Simple Regression: Reading the  
Graphs**

SIMPLER. FASTER. BETTER. LESS COSTLY.

## Define

1. Identify CTQ

2. Develop PAF

3. Visualize the Process

**Purpose:** To identify and prioritize the business improvement opportunity, define critical customer requirements, document the processes and build effective teams

**Deliverables:**

Project Approval Form; Project Team – TRAIL; CT Flow Down (CTs); SIPOC; Process map (current process) ; Project benefits; Project plan major milestones; Data Collection Plan; Data Integrity Audit (if applicable)

## Measure

4. Understand Metrics

5. Validate measurement System

6. Determine Process Performance

**Purpose:** To determine what to measure, manage the measurement data collection, develop and validate measurement systems and determine process performance

**Deliverables:**

TRAIL Chart – updated; Detailed Process Map; Cause & Effect Matrix; FMEA; Decision Tree; Operational Definitions; MSA Evaluation; Data Collection Plan; Capability Study with Control Charts; Project Status

## Analyze

7. ID potential Sources of Variation

8. Characterize the X's

9. Determine Significant X's

**Purpose:** To determine the root causes, estimate population parameters with confidence intervals and to construct hypothesis about the data and test them to determine significance.

**Deliverables:**

Data Collection Plan- updated; Hypothesis Testing; Decision Tree; MSA Analysis; Capability Analysis- updated; Executive Summary- updated; TRAIL- updated; Analysis Summary; Control Charts- updated

## Improve

10. Establish level for X's

11. Develop Solutions

12. Pilot and Implement

**Purpose:** To develop and quantify potential solutions, improve/optimize the process, evaluate and select final solution and implement the pilot.

**Deliverables:**

DOE; Lean Analysis; Simulation; Optimal settings for X's; Executive Summary- updated; TRAIL- updated; Implementation plan; Control Charts- updated; Capability Analysis- updated

## Control

13. Evaluate Process Performance

14. Develop Control Plan

15. Transition to Project Owner

**Purpose:** Implement final solution, maintain process improvements, ensure new process problems are identified & quickly corrected, disseminate lessons learned. Identify areas for replication & standardization.

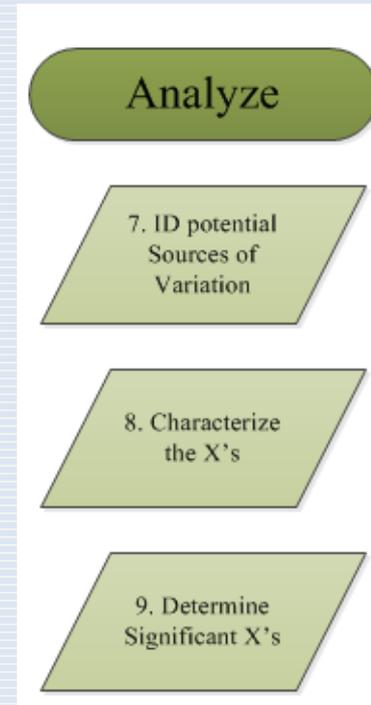
**Deliverables:**

Control Plan; Transition Plan; Capability Analysis- updated; Executive Summary- updated; TRAIL- updated; Control Charts- updated

# DMAIC Flow

- Analyze:
  - ID Potential Sources of Variation
  - Characterize the X's
  - Determine Significant X's

**Analyze Purpose:** To determine the root causes, estimate population parameters with confidence intervals and to construct hypothesis about the data and test them to determine significance.



# Course Agenda

- Introduction of Regression
  - Scatterplots
    - Characteristics of Scatterplots
  - Correlation
  - Simple Regression
- Reading the Graphs

# Regression

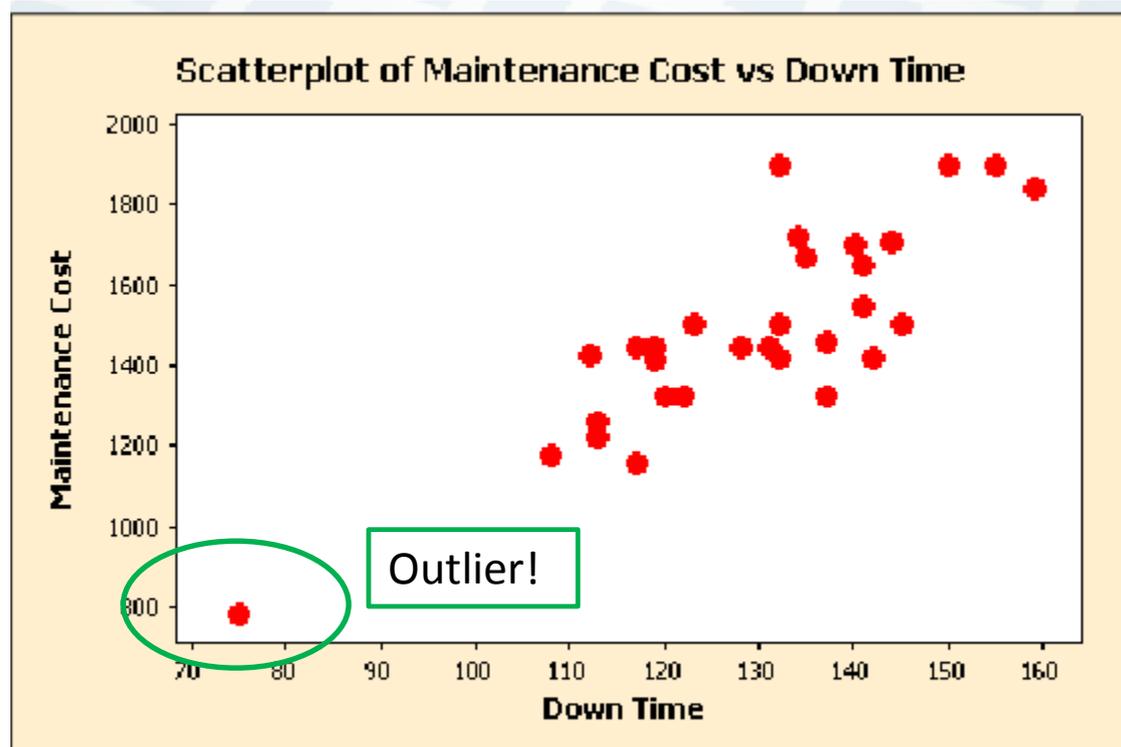
- Regression used to describe the relationship between variables
- Continuous Y and Continuous X
- Can be used to:
  - Make predictions about one variable based on values of another variable
  - Determine what value of X will get you a desired value of Y

# Regression Tools

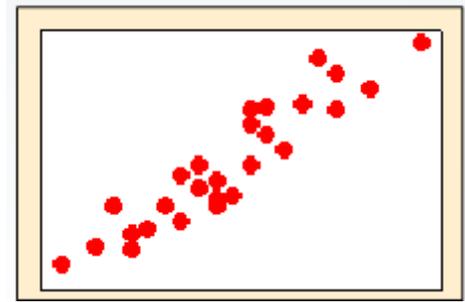
- First Tool: **Scatter plot** used to evaluate these relationships (see patterns)
- Second Tool: **Correlation** – single value to quantify the strength and direction of linear association
- Third Tool: **Simple Regression** – provides an equation between two variables to predict future events – minitab is used for this step

# Scatter plot

- Scatter plot illustrates the relationship between two variables
- Each point on the plot represents one observation
- Scatterplots can point out outliers



# Scatter plots



- Using continuous data – first plot the data
- When we believe one variable may influence another variable, we need to think about which variable is explaining the other (helps determine axis)
  - Mileage explains sales price? **OR**
  - Sales prices explains mileage?
  
  - Mileage = Explanatory Variable
  - Explanatory variable on the X Axis
  
  - Sales Price = Response Variable
  - Response variable on the Y-Axis

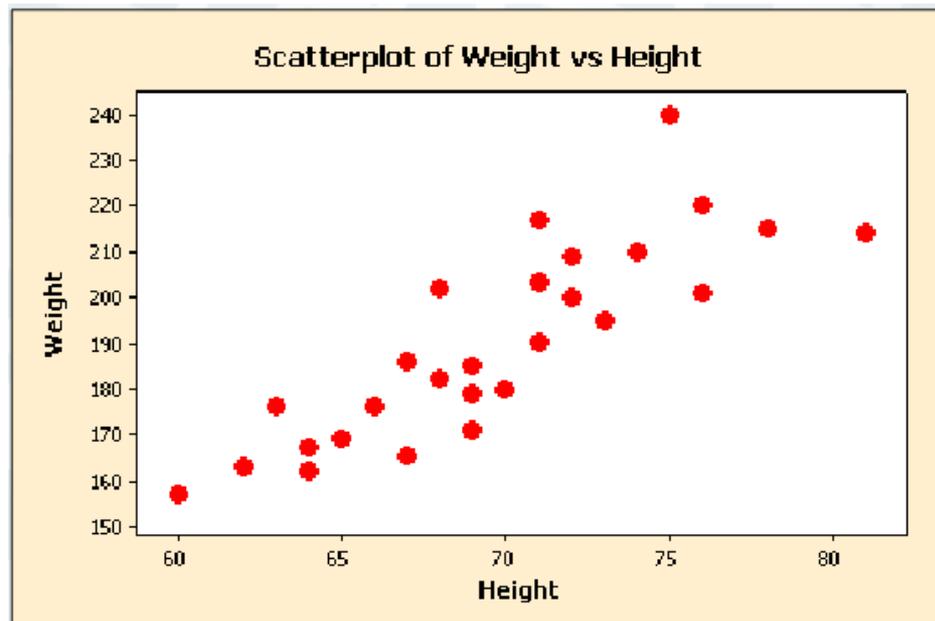
# Scatter Plot Characteristics

- Characteristics of Scatter plot
  - Direction
  - Form
  - Strength

# Scatter Plots: Direction

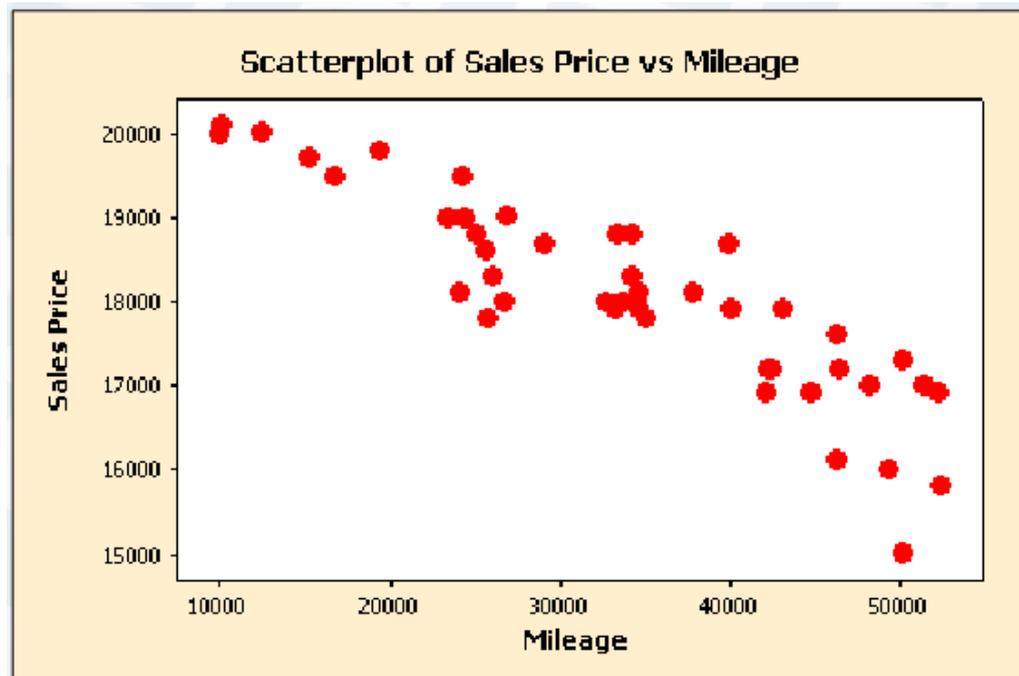
**Direction** of the relationship is the easiest to spot on a scatter plot.

- Consider the relationship between height and weight for men.
- Generally speaking, the taller the man is, the more he weights showing a positive relationship



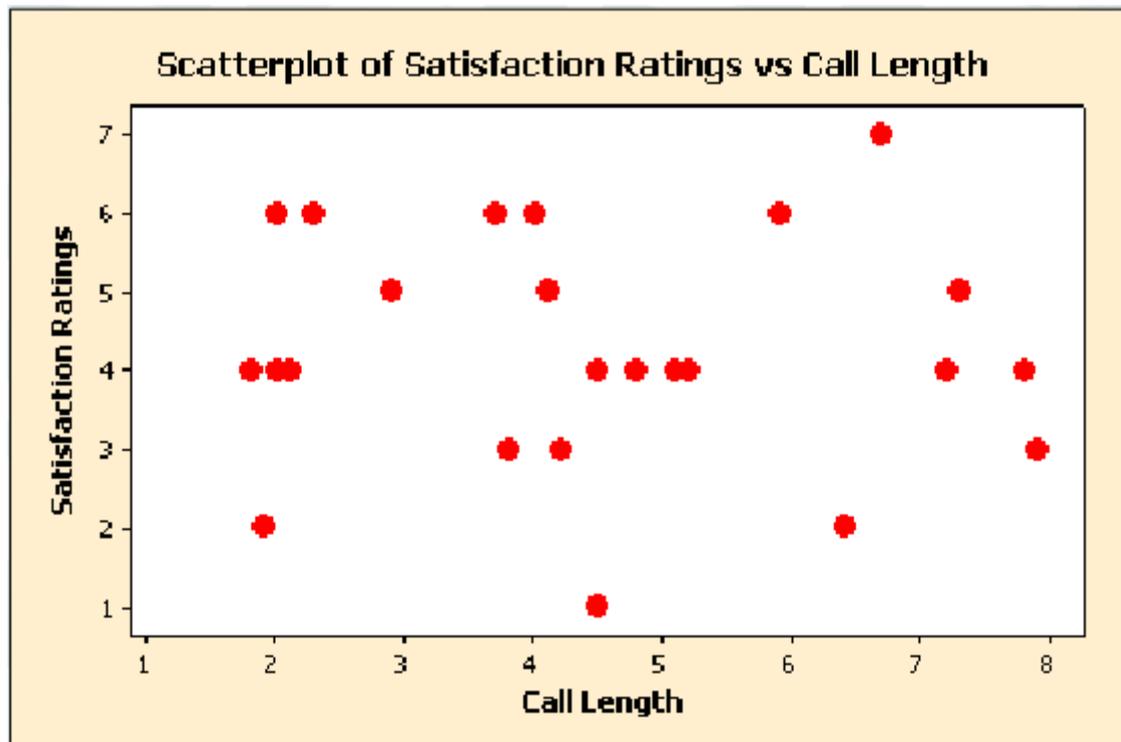
# Scatter Plots: Direction

- Consider plot of mileage and sales price, which demonstrates a **negative relationship** between the variables

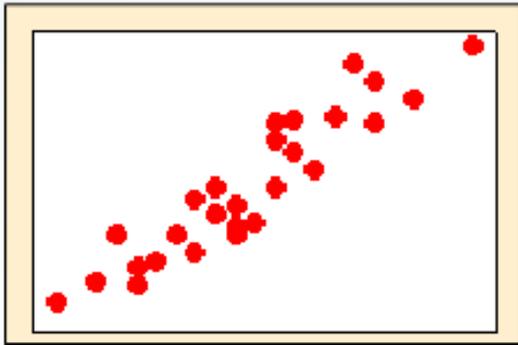


# Scatter Plots: Directions

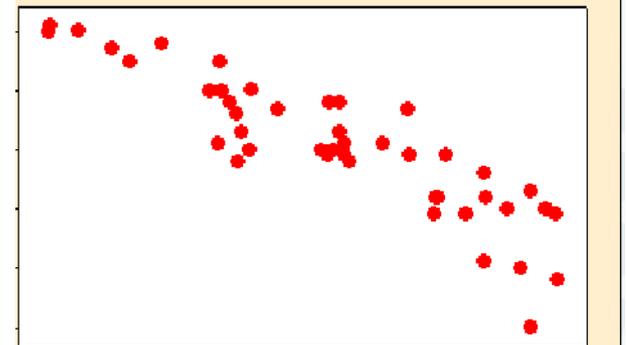
- Sometimes, no clear relationship exists between variables. The points fall randomly and no pattern in either directions.



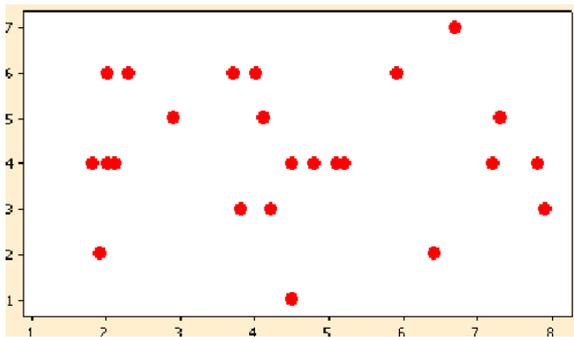
# Scatter plot Direction



- Direction:
  - Positive



- Negative



- or No Relationship

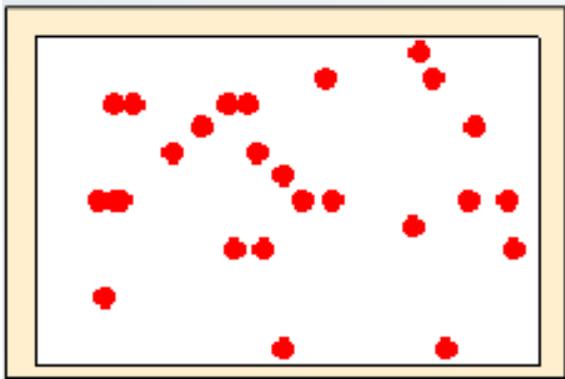
# Scatter plot Characteristics: Form

Form – Pattern of that the data follow when we fit a line to the data (straight line indicates a linear relationship)

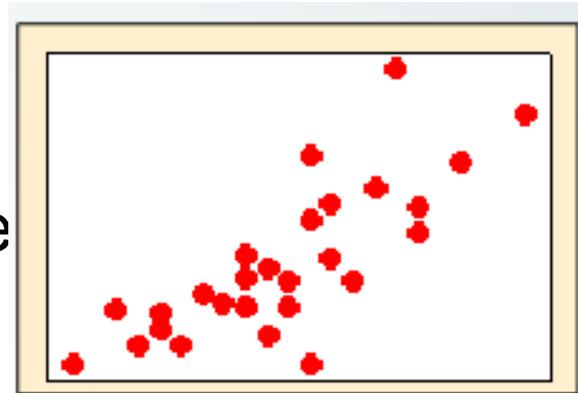
# Scatter plot Characteristics: Strength

- Strength: how closely the data points follow a straight line and the degree of the slope of that line

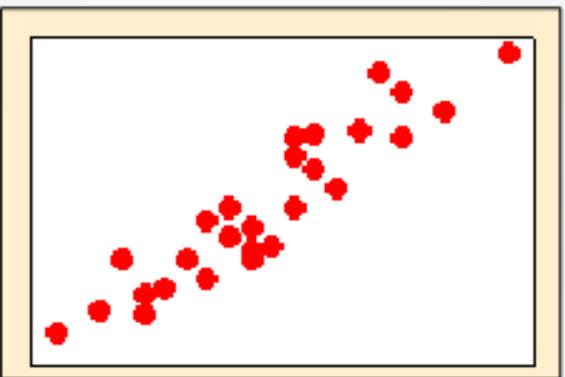
- Weak



- Moderate



- Strong.



# Correlation

- The relationship between the two variables may illustrate: Correlation
- Correlation is the relationship between the two variables

# Correlation

- Correlation:
  - Can indicate whether a linear association exists
  - Quantifies the strength of linear association

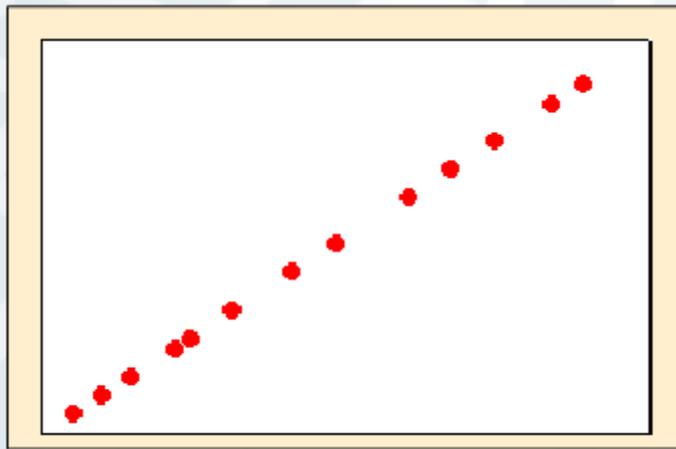
Correlation Coefficient ( $r$ ) quantifies the linear association between variables:

:  $-1 \leq r \leq 1$

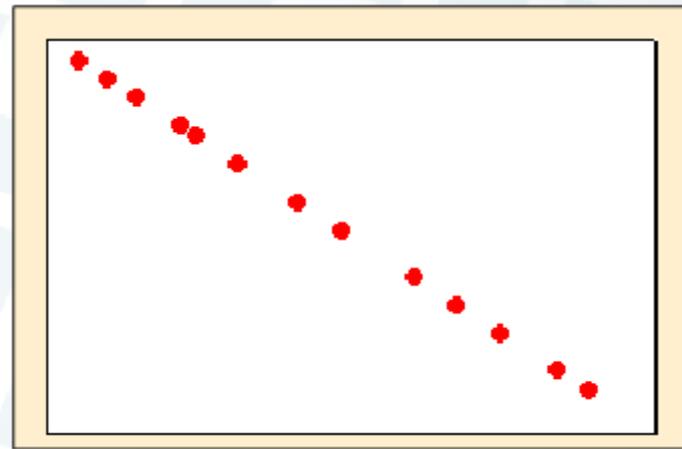
: Value reflects the strength and direction of the relationship.

# Correlation

- The sign of the correlation indicates the direction of the relationship
- Correlation near zero indicates a weak linear relationship
- For a stronger relationship between variables, the correlation is closer to positive or negative one
- Perfect correlation is equal to negative or positive one
- Correlation coefficient is not affected by manipulating the parameters of the scatter plot



$r = 1$



$r = -1$

# Regression Warnings

- DO NOT assume that a relationship means that one variable causes the other factor
- To determine the degree of associations calculate the correlation coefficient
- Correlation is **NOT** Causation

# Simple Regression

- Minitab and other statistical programs can perform **Simple Regression**.
- Simple Regression provides an equation between two variables to predict future events.
- If you have a scenario that would require Minitab's simple regression analysis – contact LeanOhio or your Mentor!