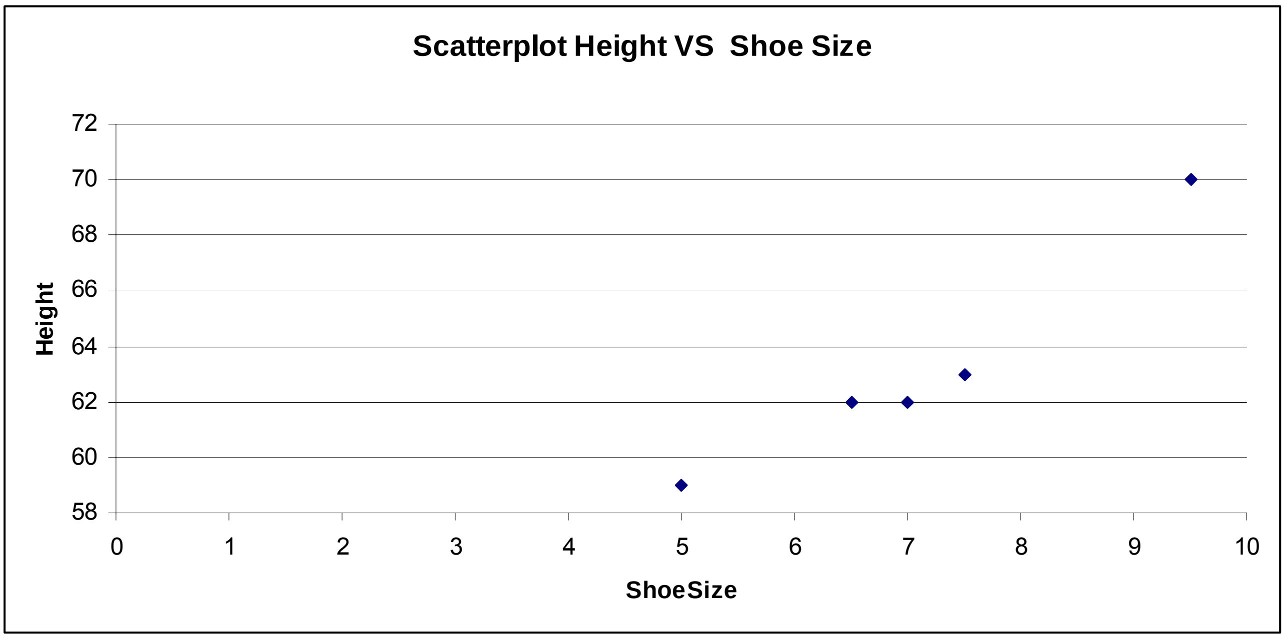
Example: Suppose we want to find the relationship between the show sizes of women and their height. We randomly selected 5 women and recorded their shoe sizes and heights.

Shoe Size	Height in.
5	59
6.5	62
7	62
7.5	63
9	70



Step1

Hypothesize a model to relate the height(y) and shoe size(x): straight line probabilistic model $y = \beta_0 + \beta_1 + \varepsilon$

Or same thing as saying $E(y) = \beta_0 + \beta_1$

Since we expect the error to be 0

Step2

MINITAB RESULTS

Regression Analysis: Height versus ShoeSize

The regression equation is
Height = 46.0 + 2.42 ShoeSize

Predictor	Coef	SE Coef	T	P
Constant	**46.014**	2.593	17.74	0.000
ShoeSize	**2.4206**	0.3577	6.77	0.007

S = 1.17011   R-Sq = 93.9%   R-Sq(adj) = 91.8%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	62.693	62.693	45.79	0.007
Residual Error	3	4.107	1.369		
Total	4	66.800			

Predicted Values for New Observations

New Obs	Fit	SE Fit	95% CI	95% PI
1	65.379	0.614	(63.423, 67.334)	(61.173, 69.584)

Values of Predictors for New Observations

New Obs	ShoeSize
1	8.00

$\hat{\beta}_1 = 2.42$, implies that the estimated mean height(y) increases by 2.42 inches for each additional size of shoe. (Valid only over range of x : 5 to 9.5)

$\hat{\beta}_0 = 46.0$, implies that a shoe size of 0 has an estimated mean height (y) of 46 inches. (no practical interpretation)

The least square line:
$\hat{y} = 46.0 + 2.42x$

Step3

Assume Assumptions of Error are satisfied.

We can find the estimated standard error of the regression model, the estimator of σ ,

$$s = 1.17011$$

Interpretation:

→We expect approximately 95% of the observed height(y) values to lie within 2.34022 of the least square line.

Regression Analysis: Height versus ShoeSize

The regression equation is
Height = 46.0 + 2.42 ShoeSize

Predictor	Coef	SE Coef	T	P
Constant	46.014	2.593	17.74	0.000
ShoeSize	2.4206	0.3577	6.77	0.007

S = 1.17011 R-Sq = 93.9% R-Sq(adj) = 91.8%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	62.693	62.693	45.79	0.007
Residual Error	3	4.107	1.369		
Total	4	66.800			

Predicted Values for New Observations

New Obs	Fit	SE Fit	95% CI	95% PI
1	65.379	0.614	(63.423, 67.334)	(61.173, 69.584)

Values of Predictors for New Observations

New Obs	ShoeSize
2	8.00

Step4

Check usefulness of the hypothesized model.

I) Test for Model Utility: Simple Linear regression

Since we hypothesize from the graph that they are positive linearly related, then just check this hypothesis.

$$H_o : \beta_1 \le 0$$
$$H_A : \beta_1 > 0$$

Test Statistics: t=6.77 with df=3

Rejection Region**:** t>2.353

Decision: Reject Ho at α=0.05

Conclusion: There is enough evidence to conclude that the height and shoe size are positively correlated (height increases as shoe size increases).

Regression Analysis: Height versus ShoeSize

```
The regression equation is
Height = 46.0 + 2.42 ShoeSize


Predictor    Coef  SE Coef      T      P
Constant   46.014    2.593  17.74  0.000
ShoeSize   2.4206   0.3577   6.77  0.007


S = 1.17011   R-Sq = 93.9%   R-Sq(adj) = 91.8%


Analysis of Variance

Source          DF      SS      MS      F      P
Regression       1  62.693  62.693  45.79  0.007
Residual Error   3   4.107   1.369
Total            4  66.800


Predicted Values for New Observations

New
Obs     Fit  SE Fit       95% CI             95% PI
  1  65.379   0.614  (63.423, 67.334)  (61.173, 69.584)


Values of Predictors for New Observations

New
Obs  ShoeSize
1        8.00
```

II) 95% CI for β_1

$$\hat{\beta}_1 \pm t_{\alpha/2} s_{\beta_1} = 2.42 \pm 3.182(0.3577)$$

$$(1.28, 3.56)$$

→ We are with 95% confidence that the true mean increase in the height per additional shoe size is between 1.28 and 3.56 inches.

III) Coefficient of determination

$$r^2 = 0.939$$

→ 93.9% of the sample variation in height(y) is explained by the shoe size(x)

IV) Coefficient of correlation

$$r = +\sqrt{r^2} = \sqrt{0.939}$$

→ r=0.969

→ The height tend to increase as the shoe size increases, strong positive linear relation

Regression Analysis: Height versus ShoeSize

```
The regression equation is
Height = 46.0 + 2.42 ShoeSize


Predictor    Coef  SE Coef      T      P
Constant   46.014    2.593  17.74  0.000
ShoeSize   2.4206   0.3577   6.77  0.007


S = 1.17011   R-Sq = 93.9%   R-Sq(adj) = 91.8%


Analysis of Variance

Source          DF      SS      MS      F      P
Regression       1  62.693  62.693  45.79  0.007
Residual Error   3   4.107   1.369
Total            4  66.800


Predicted Values for New Observations

New
Obs     Fit  SE Fit       95% CI            95% PI
  1  65.379   0.614  (63.423, 67.334)  (61.173, 69.584)


Values of Predictors for New Observations

New
Obs  ShoeSize
   1      8.00
```

All signs in step 4 pointed out a strong linear relationship between the height(y) and the shoe size(x).

→Proceed step 5

Step5

- Suppose we want to predict the height of women if the shoe size is 8

$$\hat{y} = 46.0 + 2.42(8)$$

$$\hat{y} = 65.36$$

→ The predicted height value for a woman with shoe size is 65.36 inches.

- Find and interpret the 95% Confidence Interval for the mean height when the shoe size is 8.

  From the Minitab , we get (63.423, 67.334)

  → We are 95% confident that the average estimated height for all possible subject with shoe size 8 is between 63.423 and 67.334 inches

- Find and interpret the 95% Prediction Interval for the mean height when the shoe size is 8.

  From the Minitab , we get (61.173,69.584))

  → We predict with 95% confidence that the height for a women with shoe size 8 to fall between 61.173 and 69.584 inches

Regression Analysis: Height versus ShoeSize

```
The regression equation is
Height = 46.0 + 2.42 ShoeSize


Predictor    Coef   SE Coef      T      P
Constant   46.014    2.593   17.74  0.000
ShoeSize   2.4206   0.3577    6.77  0.007


S = 1.17011   R-Sq = 93.9%   R-Sq(adj) = 91.8%


Analysis of Variance

Source          DF      SS      MS      F      P
Regression       1  62.693  62.693  45.79  0.007
Residual Error   3   4.107   1.369
Total            4  66.800


Predicted Values for New Observations

New
Obs     Fit  SE Fit        95% CI            95% PI
  1  65.379   0.614  (63.423, 67.334) (61.173,69.584)


Values of Predictors for New Observations

New
Obs  ShoeSize
1    8.00
```