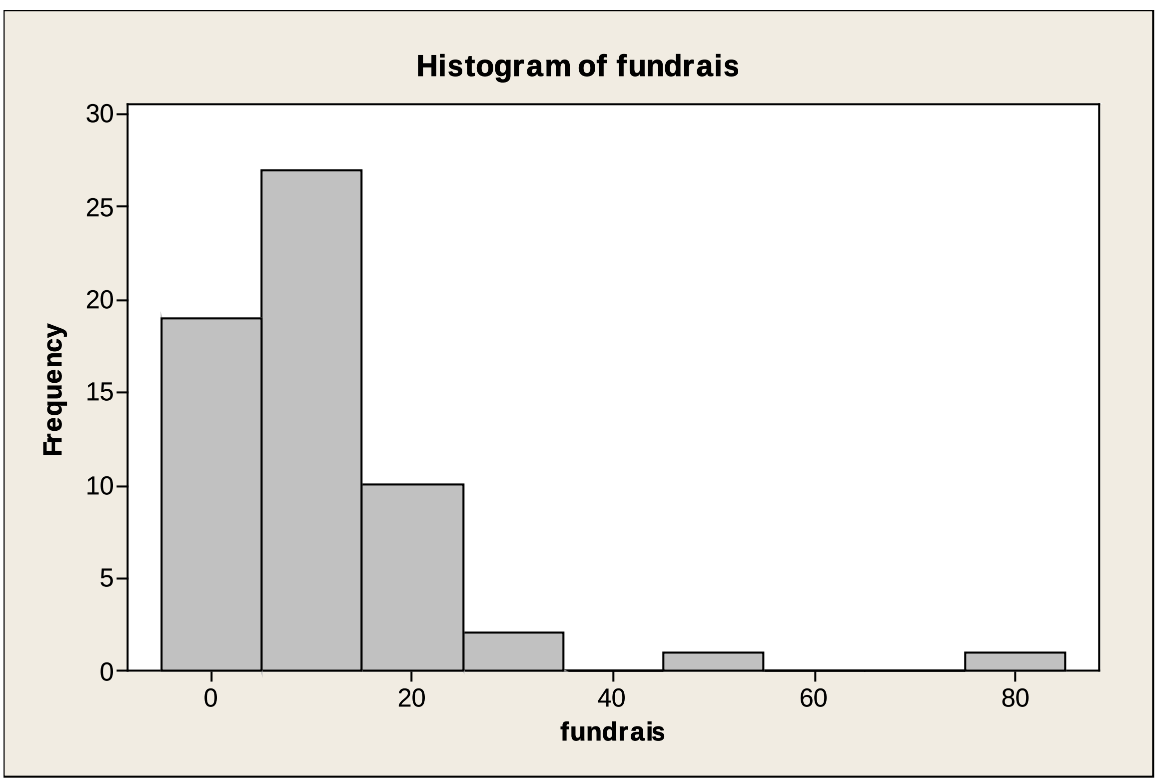
Example 1.1

Graph > Stem-and-Leaf

```
Stem-and-leaf of fundrais  N  = 60
Leaf Unit = 1.0


 19   0  0111112222333333344
(17)  0  55556666666778888
 24   1  0001222244
 14   1  55666789
 6    2  01
 4    2  6
 3    3  4
 2    3
 2    4
 2    4  8
 1    5
 1    5
 1    6
 1    6
 1    7
 1    7
 1    8  3
```

Fundraising expenses (% of total expenses) for organized charities have a unimodal distribution, skewed to the right (+) with two unusual observations at 48 and 83 percent.

Graph > Histogram



Example 1.16

Stat > Basic Statistics > Display Descriptive Statistics (select C1 % Copper)

Results for: exp01-16.mtw

Descriptive Statistics: % Copper

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
% Copper	26	0	3.654	0.303	1.547	2.000	2.700	3.350	3.875	10.100

Conclusion: The average (or mean) copper content is 3.65 %. Half of the Bidri samples had copper content of less than 3.35 % and half had a greater copper content.

Example 1.18 (p. 41)

Stat > Basic Statistics > Display Descriptive Statistics (select C1)

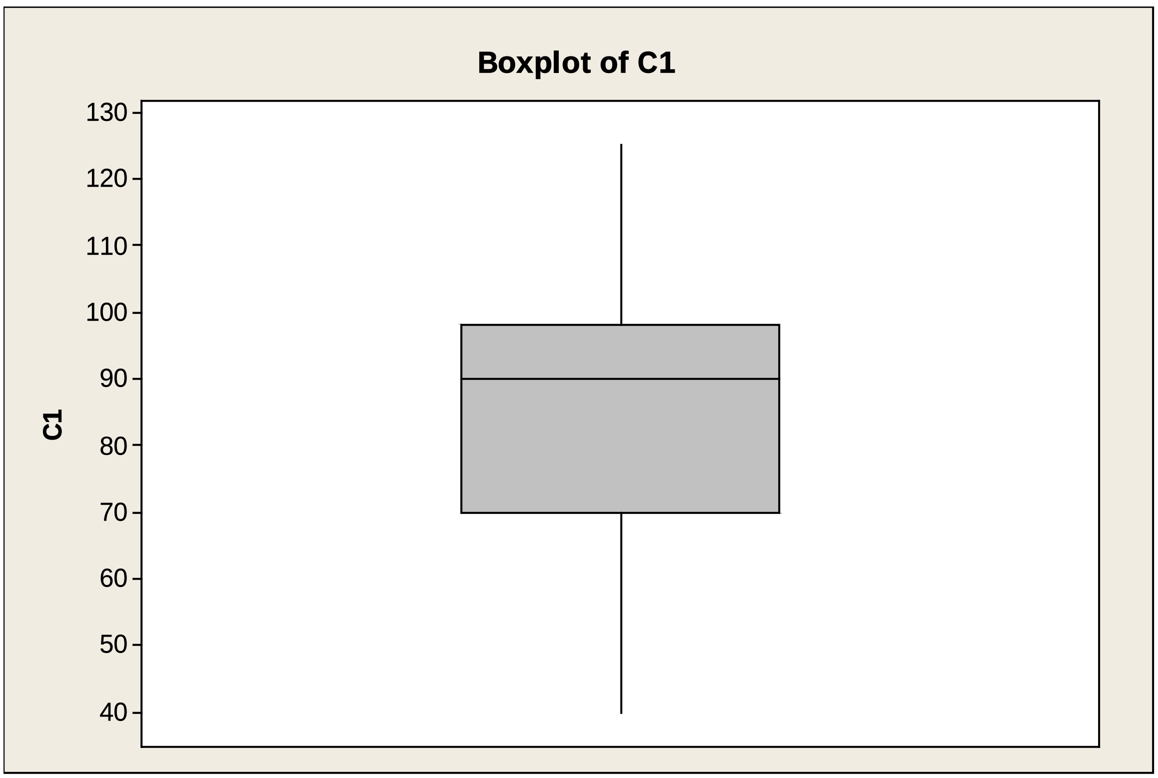
Results for: exp1-18.mtw

Descriptive Statistics: C1

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
C1	19	0	86.32	5.35	23.32	40.00	70.00	90.00	98.00	125.00

Conclusion: A typical pit depth is 86 thousandths of an inch, with a typical spread of 23 thousandths of an inch either way.

Graph > Boxplot (Simple, select C1)



Conclusion: The distribution of pit depths is positively skewed with no apparent outliers.

Example 3.32

Calc > Probability Distributions > Binomial (n=15, p=0.2, CDF, x=8)

Cumulative Distribution Function

```
Binomial with n = 15 and p = 0.2

x  P( X <= x )
8     0.999215
```

Conclusion: Probability that at most 8 fail the test is .9992

Calc > Probability Distributions > Binomial (n=15, p=0.2, pmf, x=8)

Probability Density Function

```
Binomial with n = 15 and p = 0.2

x  P( X = x )
8   0.0034548
```

Conclusion: Probability that exactly 8 fail the test is .003

Calc > Probability Distributions > Binomial (n=15, p=0.2, inverse pmf, prob=.5)

Inverse Cumulative Distribution Function

```
Binomial with n = 15 and p = 0.2

x  P( X <= x )      x  P( X <= x )
2     0.398023      3     0.648162
```

Conclusion: Median is 3. There is at least a 50% chance that <=3 fail the test, and there is at least a 50% chance that >=3 fail the test.

Example 3.40

Calc > Probability Distributions > Poisson (Probability, mean=2, input constant 1)

Probability Density Function

```
Poisson with mean = 2

x  P( X = x )
1    0.270671
```

Conclusion: Probability that exactly one error is found would be 27%

Calc > Probability Distributions > Poisson (Cumulative probability, mean=2, input constant 3)

Cumulative Distribution Function

```
Poisson with mean = 2

x  P( X <= x )
3     0.857123
```

Conclusion: Probability that at most three errors are found would be 85.7%

Calc > Probability Distributions > Poisson (Inverse cumulative probability, mean=2, input constant 0.75)

Inverse Cumulative Distribution Function

```
Poisson with mean = 2

x  P( X <= x )      x  P( X <= x )
2     0.676676      3     0.857123
```

Conclusion: Third quantile Q3 is equal to 3. There is at least a 75% chance that <=3 errors are found, and there is at least a 25% chance that >=3 errors are found.

Example 4.16

Calc > Probability Distributions > Normal (Cumulative probability, mean 1.25, standard deviation 0.46, input constant 1.00)

Cumulative Distribution Function

```
Normal with mean = 1.25 and standard deviation = 0.46

x  P( X <= x )
1     0.293400
```

Calc > Probability Distributions > Normal (Cumulative probability, mean 1.25, standard deviation 0.46, input constant 1.75)

Cumulative Distribution Function

```
Normal with mean = 1.25 and standard deviation = 0.46

   x  P( X <= x )
1.75     0.861472
```

Conclusion: Then P(1.00 < X < 1.75) = .861 - .293 = .568 so the probability that reaction time is between 1.00 and 1.75 seconds is 56.8%

Calc > Probability Distributions > Normal (Cumulative probability, mean 1.25, standard deviation 0.46, input constant 2.00)

Cumulative Distribution Function

```
Normal with mean = 1.25 and standard deviation = 0.46

x  P( X <= x )
2     0.948495
```

Conclusion: Then P(X > 2) = 1-.948 = .052 so the probability that reaction time exceeds 2.0 seconds is 5.2%

Calc > Probability Distributions > Normal (Inverse cumulative probability, mean 1.25, standard deviation 0.46, input constant 0.99)

Inverse Cumulative Distribution Function

```
Normal with mean = 1.25 and standard deviation = 0.46

P( X <= x )        x
       0.99  2.32012
```

Conclusion: 99^{th} percentile is 2.32. There is a 99% chance that reaction time is less than 2.32 seconds

Example 4.24

Calc > Probability Distributions > Gamma (Cumulative probability, shape=8, scale=15, input constant 60)

Cumulative Distribution Function

```
Gamma with shape = 8 and scale = 15

 x  P( X <= x )
60    0.0511336
```

Calc > Probability Distributions > Gamma (Cumulative probability, shape=8, scale=15, input constant 120)

Cumulative Distribution Function

```
Gamma with shape = 8 and scale = 15

  x  P( X <= x )
120     0.547039
```

Conclusion: Then P(60<X<120) = .547 - .051 = .496 so the probability that a mouse survives between 60 and 120 weeks is 49.6%

Calc > Probability Distributions > Gamma (Cumulative probability, shape=8, scale=15, input constant 30)

Cumulative Distribution Function

```
Gamma with shape = 8 and scale = 15

 x  P( X <= x )
30    0.0010967
```

Conclusion: Then P(X>30) = 1-.001 = .999 so the probability that a mouse survives at least 30 weeks is 99.9%

Calc > Probability Distributions > Gamma (Inverse cumulative probability, shape=8, scale=15, input constant 0.25)

Inverse Cumulative Distribution Function

```
Gamma with shape = 8 and scale = 15

P( X <= x )        x
       0.25  89.3416
```

Conclusion: First quantile Q1 is equal to 89.3. There is a 25% chance that a mouse will survive less that 89.3 weeks.

Example 4.25

Calc > Probability Distributions > Weibull (Cumulative probability, shape=2, scale=10, input constant 10)

Cumulative Distribution Function

```
Weibull with shape = 2 and scale = 10

 x  P( X <= x )
10     0.632121
```

Conclusion: There is a 63% chance that nitrous oxide emissions are less than 10.

Calc > Probability Distributions > Weibull (Cumulative probability, shape=2, scale=10, input constant 25)

Cumulative Distribution Function

```
Weibull with shape = 2 and scale = 10

 x  P( X <= x )
25     0.998070
```

Conclusion: There is a 99.8% chance that nitrous oxide emissions are less than 25.

Calc > Probability Distributions > Weibull (Inverse cumulative probability, shape=2, scale=10, input constant 0.95)

Inverse Cumulative Distribution Function

```
Weibull with shape = 2 and scale = 10

P( X <= x )        x
       0.95  17.3082
```

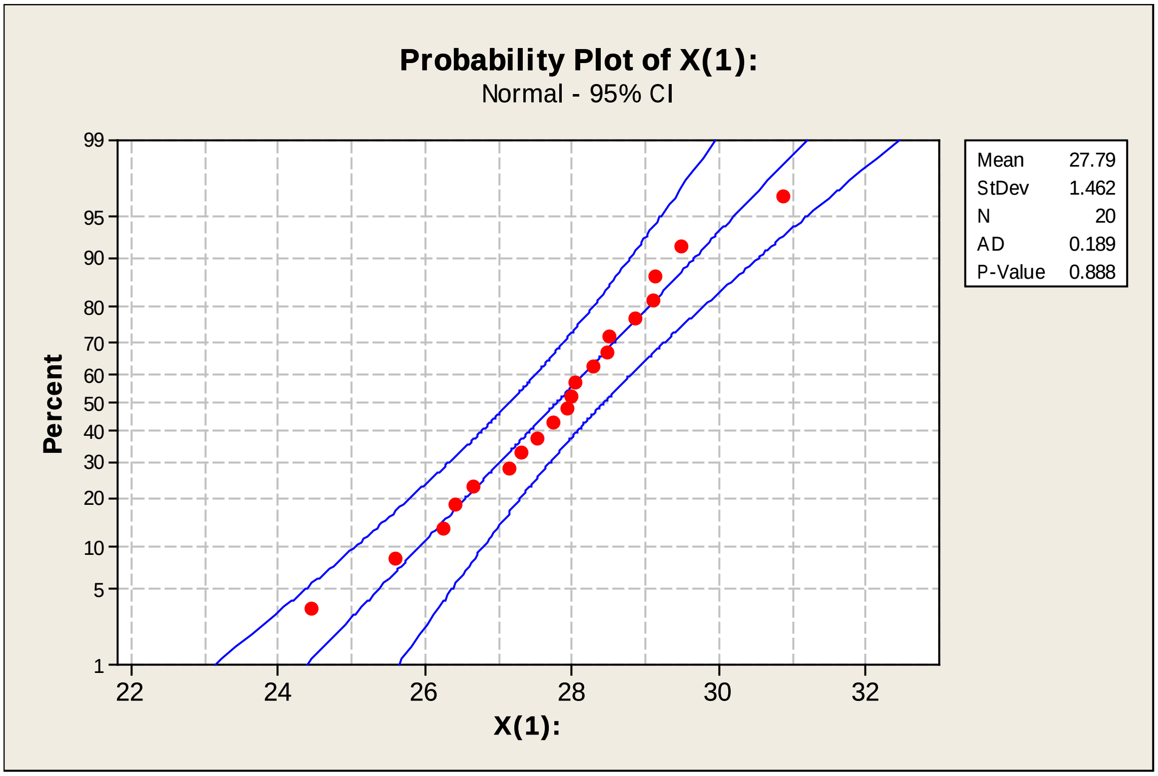
Conclusion: 95% of NOx emissions are less than 17.3.

Example 4.30

Graph > Probability Plot (Single, select data)

Results for: exp4-30.mtw

Probability Plot of X(1):



Conclusion: The distribution of dialectric breakdown voltage data appears to fit a normal distribution with mean 27.79 and standard deviation 1.462.

Example 7.11

Stat > Basic Statistics > One sample-t (select column C1)

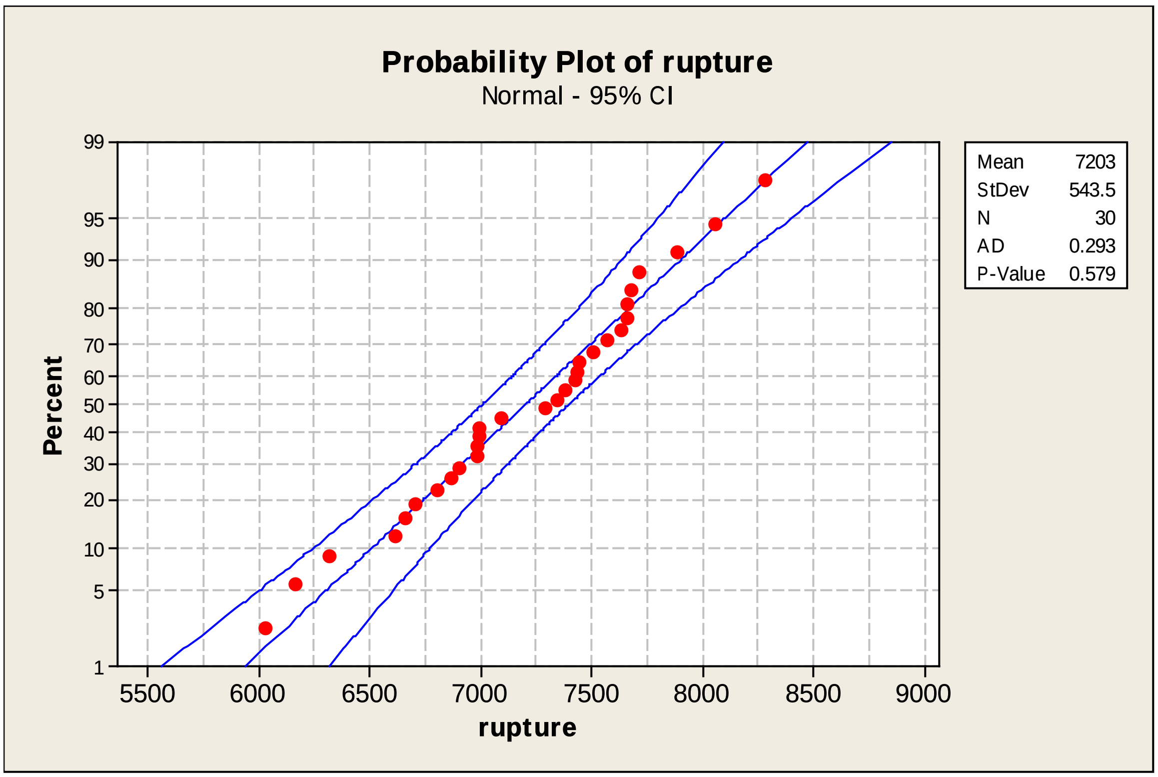
Results for: EXP07-11.MTW

One-Sample T: rupture

```
Variable   N    Mean  StDev  SE Mean        95% CI
rupture   30  7203.2  543.5     99.2  (7000.2, 7406.2)
```

Conclusion: 95% sure that average strength is between 7000 and 7406 psi.

Graph > Probability Plot (Single, select data)



Conclusion: OK, data appears normal, t-test procedure is valid.

Example 8.9

Stat > Basic Statistics > One sample-t > Options (Alternative: not equal)

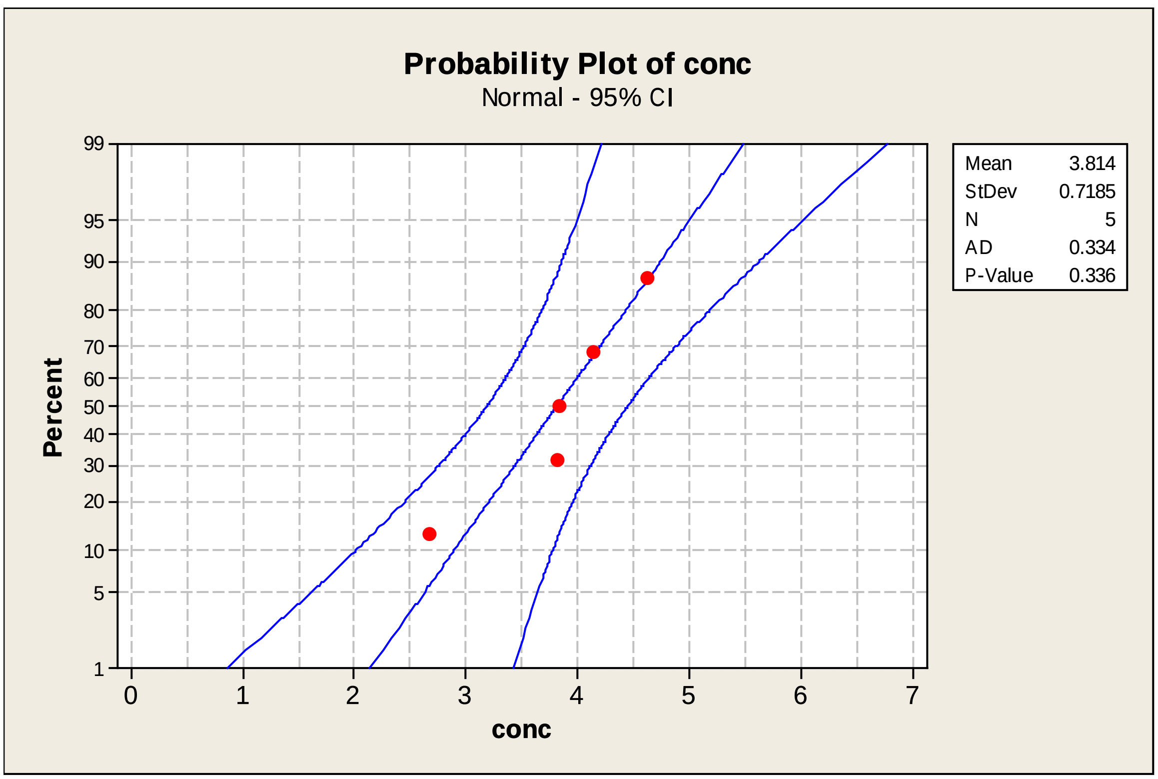
One-Sample T: conc

```
Test of mu = 4 vs not = 4


Variable  N   Mean  StDev  SE Mean       95% CI           T      P
conc      5  3.814  0.718    0.321  (2.922, 4.706)  -0.58  0.594
```

Calculation: Reject H_0 at the 95% level if $T > t_{.025,4} = 2.776$. Since $T = -0.58$, do not reject. [Or, do not reject because the P-value is $P = 0.594 > .05$.]

Conclusion: Insufficient evidence to be 95% sure that average concentration differs from 4 mg/mL.

Graph > Probability Plot (Single, select data)



Conclusion: OK, data appears normal, t-test procedure is valid.

Example 8.11

Stat > Basic Statistics > 1 Proportion > (Number of events: 16, Number of trials: 91, Perform hypothesis test, Hypothesized proportion: 0.15, Options: Confidence level: 90, Alternative: greater than, Use test and interval based on normal distribution)

Test and CI for One Proportion

```
Test of p = 0.15 vs p > 0.15


                          90% Lower
Sample   X   N  Sample p      Bound  Z-Value  P-Value
1       16  91  0.175824   0.124684     0.69    0.245

Using the normal approximation.
```

Since the test statistic z = 0.69 is less than the 90% cut off value of 1.282, we do not reject the null hypothesis.

Conclusion: Insufficient evidence to be 90% sure that more than 15% of corks are bad.

Example 8.18

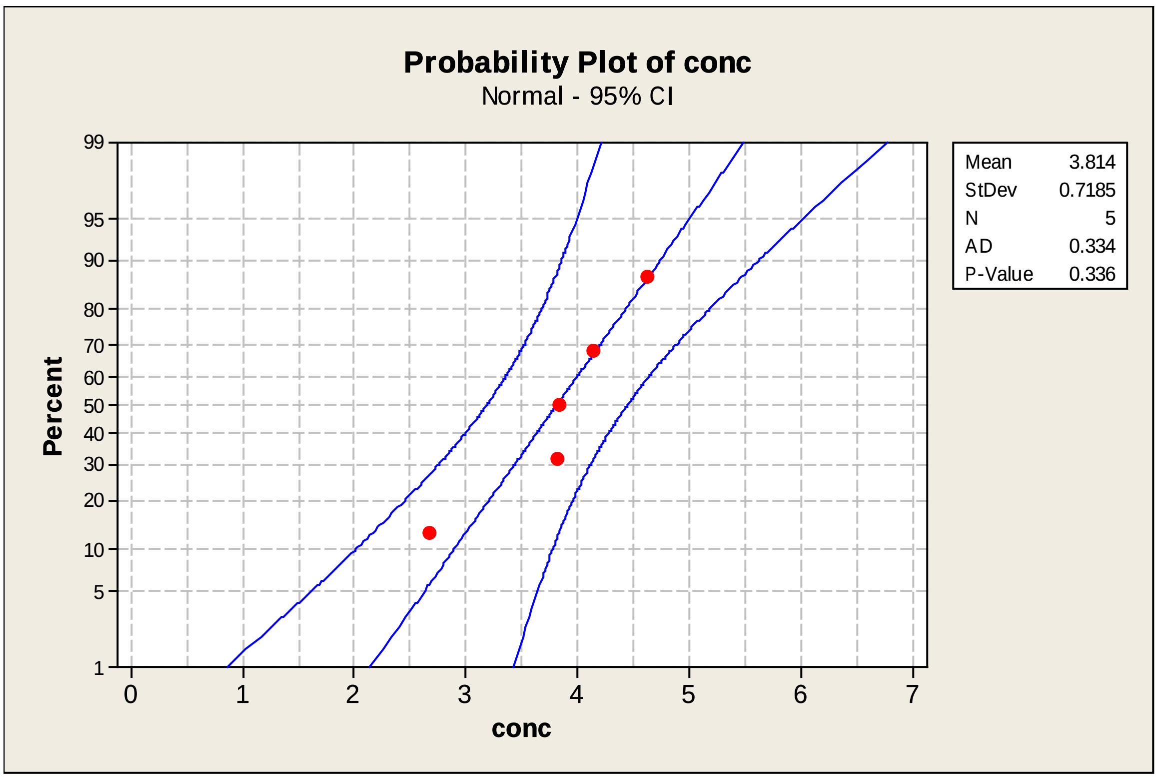
Stat > Basic Statistics > One sample-t > Options (Alternative: not equal)

One-Sample T: conc

```
Test of mu = 4 vs not = 4


Variable  N   Mean  StDev  SE Mean       95% CI            T      P
conc      5  3.814  0.718    0.321  (2.922, 4.706)  -0.58  0.594
```

Conclusion: We are 41.6% sure that average concentration differs from 4 mg/mL. [There is insufficient evidence to conclude that average concentration differs from 4 mg/mL.]

Graph > Probability Plot (Single, select data)



Conclusion: OK, data appears normal, t-test procedure is valid.

Example 9.9

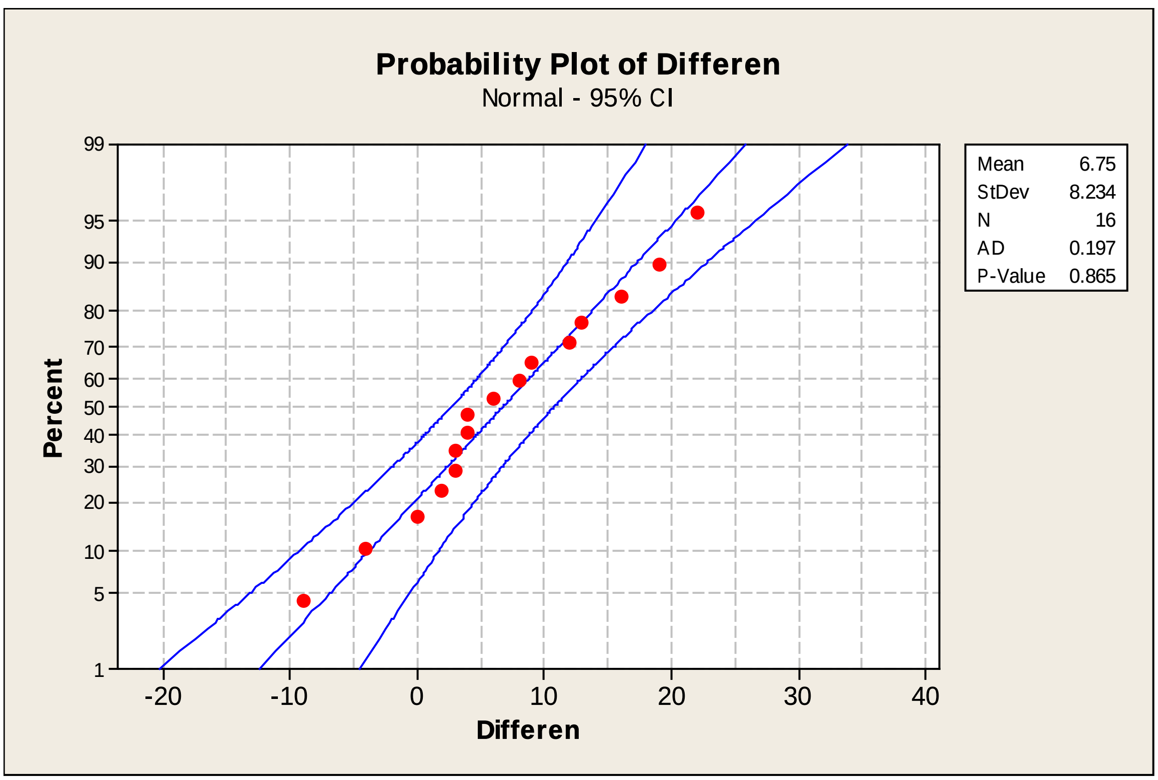
Stat > Basic Statistics > One sample-t > Options (level, alternative)

One-Sample T: Differen

```
Test of mu = 0 vs not = 0


Variable   N     Mean    StDev  SE Mean          95% CI             T      P
Differen  16  6.75000  8.23408  2.05852  (2.36237, 11.13763)  3.28  0.005
```

Conclusion: We are 99.5% sure that the average proportion of time at which arm angle is less than 30 degrees has changed after work conditions were changed (since p=.005). Also, we are 95% sure that the average proportion of time at which arm angle is less than 30 degrees, after the change in working conditions, is between 2% and 11% less than it was before (based on the 95% CI).

Graph > Probability Plot (Single, select data)



Conclusion: OK, data appears normal, t-test procedure is valid.

Alternative procedure based on two-sample t-test

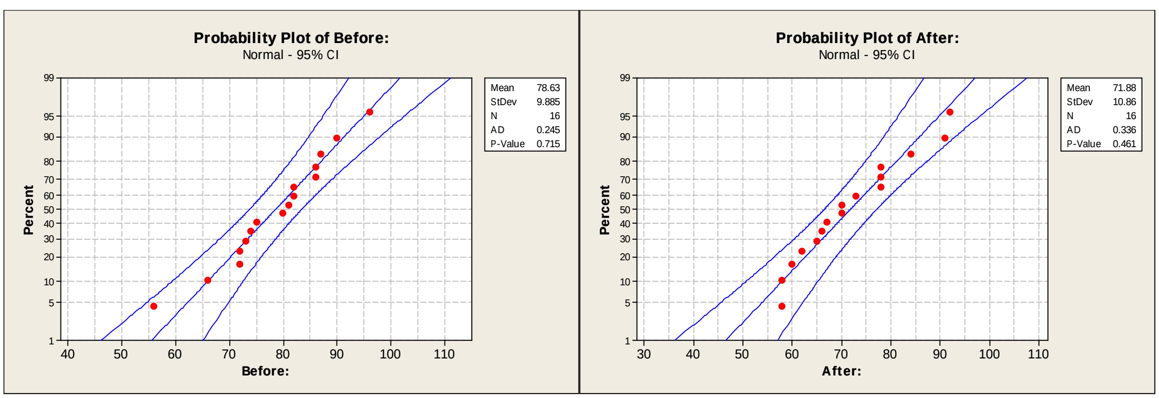
Stat > Basic Statistics > 2 sample-t > Options (level, alternative)

Two-Sample T-Test and CI: Before:, After:

```
Two-sample T for Before: vs After:

          N   Mean  StDev  SE Mean
Before:  16  78.63   9.89      2.5
After:   16   71.9   10.9      2.7


Difference = mu (Before:) - mu (After:)
Estimate for difference:  6.75000
95% CI for difference:  (-0.75711, 14.25711)
T-Test of difference = 0 (vs not =): T-Value = 1.84  P-Value = 0.076  DF = 29
```

Conclusion: We are 92.4% sure that the average proportion of time at which arm angle is less than 30 degrees has changed after work conditions were changed (since p=.076). Also, we are 95% sure that the average amount of this change is between -0.7% and +14.3% (based on the 95% CI). Note that the paired t-test gives more definitive results!

Graph > Probability Plot (Single, select data)



Conclusion: OK, each data set appears normal, 2-sample t-test procedure is valid.

Example 9.11

Stat > Basic Statistics > 2 Proportions (Options: Alternative: greater than, Use pooled estimate for p test)

Test and CI for Two Proportions

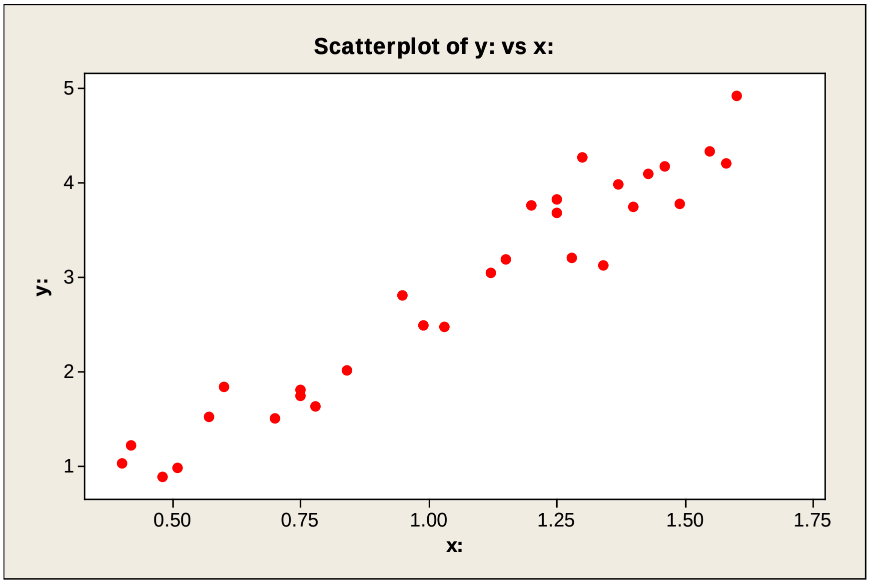
```
Sample    X    N  Sample p
1        81  549  0.147541
2       141  730  0.193151


Difference = p (1) - p (2)
Estimate for difference:  -0.0456097
95% upper bound for difference:  -0.0110060
Test for difference = 0 (vs < 0):  Z = -2.13  P-Value = 0.017

Fisher's exact test: P-Value = 0.019
```

Since p-value is 0.017 we reject the null hypothesis at level 0.01, but not at level 0.01.

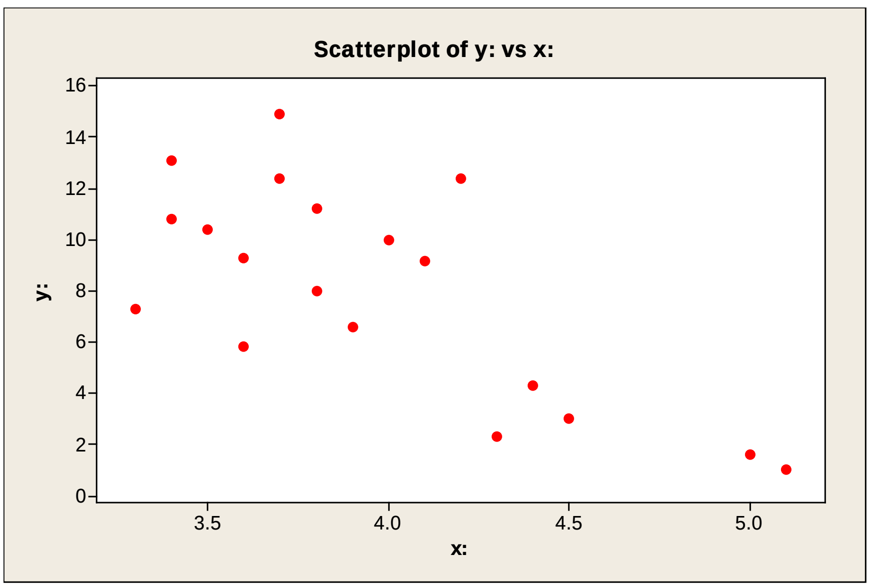
Conclusion: We are 98.3% sure that aspirin use increases the 15 year survival rate for colorectal cancer victims.

Examples 12.1 and 12.2

Graph > Scatterplot



Conclusion: There appears to be a strong positive linear relation between y = OSA and x = palprebal fissure width.



Conclusion: There appears to be a weak negative linear relation between y = mean crown dieback and x = pH.

Example 12.4

Stat > Regression > Regression (Storage: Residuals)

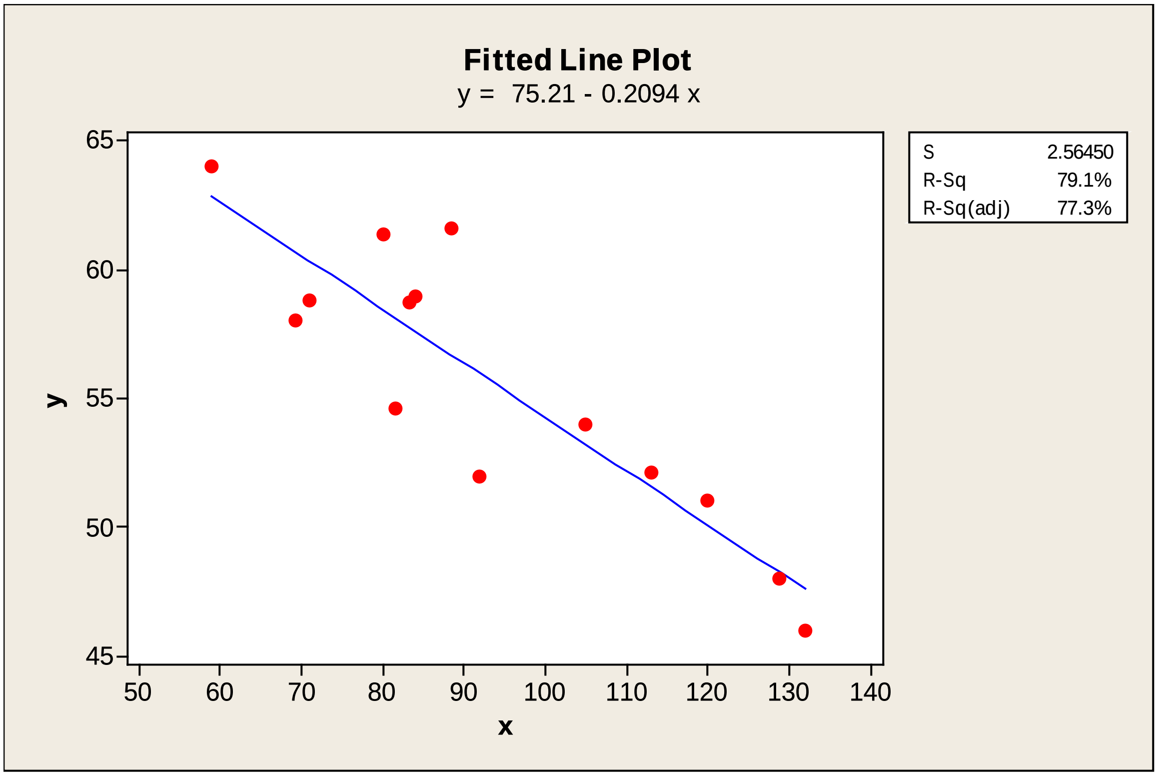
Results for: EXP12-04.MTW

Regression Analysis: y versus x

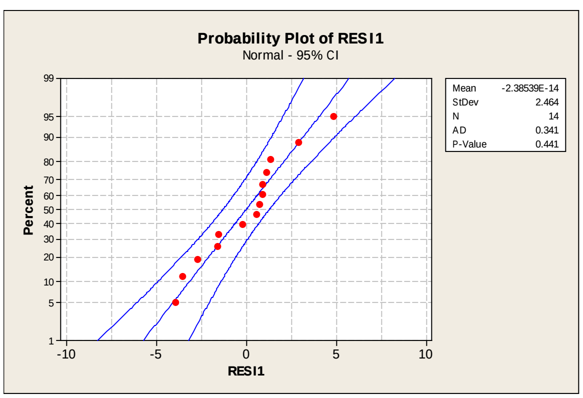
```
The regression equation is
y = 75.2 - 0.209 x


Predictor       Coef  SE Coef      T      P
Constant      75.212    2.984  25.21  0.000
x           -0.20939  0.03109  -6.73  0.000


S = 2.56450   R-Sq = 79.1%   R-Sq(adj) = 77.3%


Analysis of Variance

Source          DF      SS      MS      F      P
Regression       1  298.25  298.25  45.35  0.000
Residual Error  12   78.92    6.58
Total           13  377.17
```

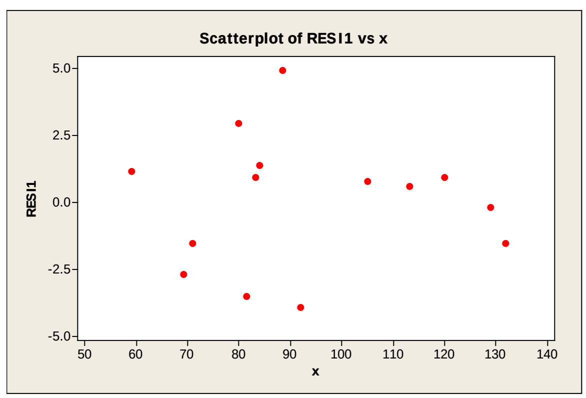
Conclusion: The true mean cetane number y for diesel fuel with iodine value x is estimated to be y = 75.2 – 0.209 x, with a typical spread of 2.6. This regression model explains 77.3% of the variations in cetane number in terms of variations in iodine value.

Stat > Regression > Fitted Line Plot



Conclusion: The regression line provides a reasonably good fit to the data, indicating that x = iodine value can give a useful model to predict y = cetane number.





Conclusion: Probability plot (left) indicates that the residuals fit a normal distribution. Scatterplot (right) indicates the residual seem to be independent and identically distributed. Hence the basic assumptions of the regression model are satisfied.

Example 12.11

Stat > Regression > Regression (Storage: Residuals)

Results for: exp12-11.mtw

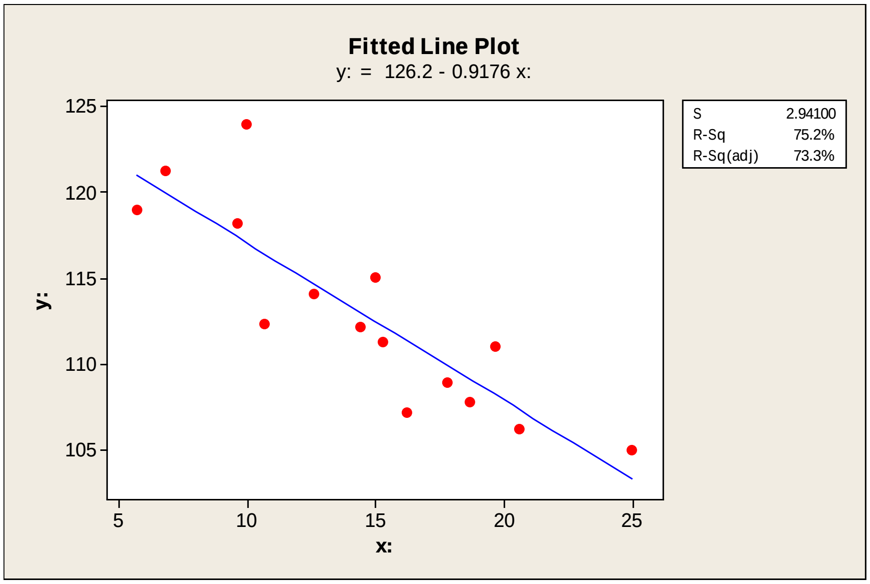
Regression Analysis: y: versus x:

```
The regression equation is
y: = 126 - 0.918 x:


Predictor     Coef  SE Coef      T      P
Constant   126.249    2.254  56.00  0.000
x:         -0.9176   0.1460  -6.29  0.000


S = 2.94100   R-Sq = 75.2%   R-Sq(adj) = 73.3%


Analysis of Variance

Source          DF      SS      MS      F      P
Regression       1  341.73  341.73  39.51  0.000
Residual Error  13  112.44    8.65
Total           14  454.17


Unusual Observations

Obs    x:       y:      Fit  SE Fit  Residual  St Resid
  4  10.0  124.000  117.073   1.008     6.927     2.51R

R denotes an observation with a large standardized residual.
```

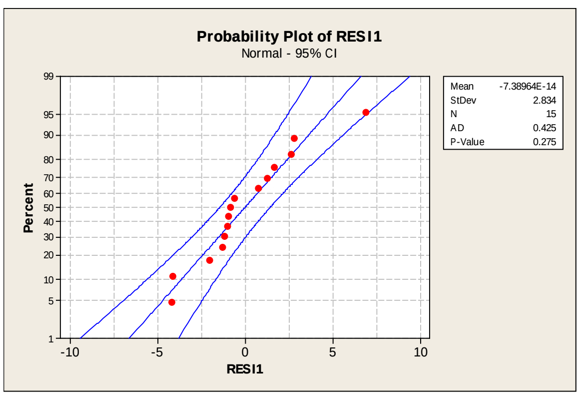
Conclusion: The true mean density for mortar with an air content of x % is estimated to be μ=126-0.918 x lb/cu-ft , with a typical spread of 3 lb/cu-ft. This regression model explains 75.2% of the variations in density in terms of variations in air content.

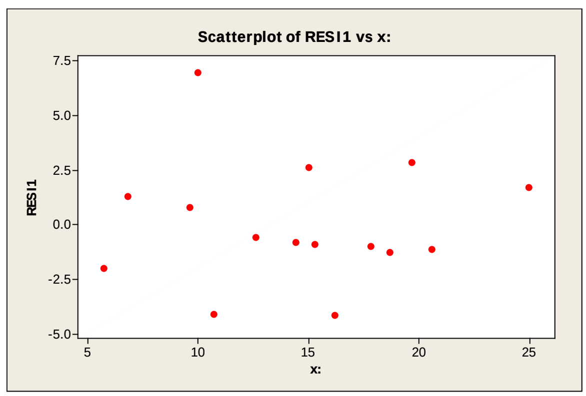
The t statistic is -6.29 with a p-value of 0.000, so we are virtually certain that the regression model provides a useful predictor of mortar density in terms of air content.

Stat > Regression > Fitted Line Plot



Conclusion: The regression line provides a reasonably good fit to the data, indicating that x=air content can give a useful model to predict y=density.





Conclusion: Probability plot (left) indicates that the residuals fit a normal distribution. Scatterplot (right) indicates the residual seem to be independent and identically distributed. Hence the basic assumptions of the regression model are satisfied. Note: The p-value of 0.275 in the probability plot indicates that there is insufficient evidence to reject the null hypothesis of a normal fit.

Example 12.13

Stat > Regression > Regression (Storage: Residuals, Options: Prediction intervals for new observation: 45)

Results for: exp12-13.mtw

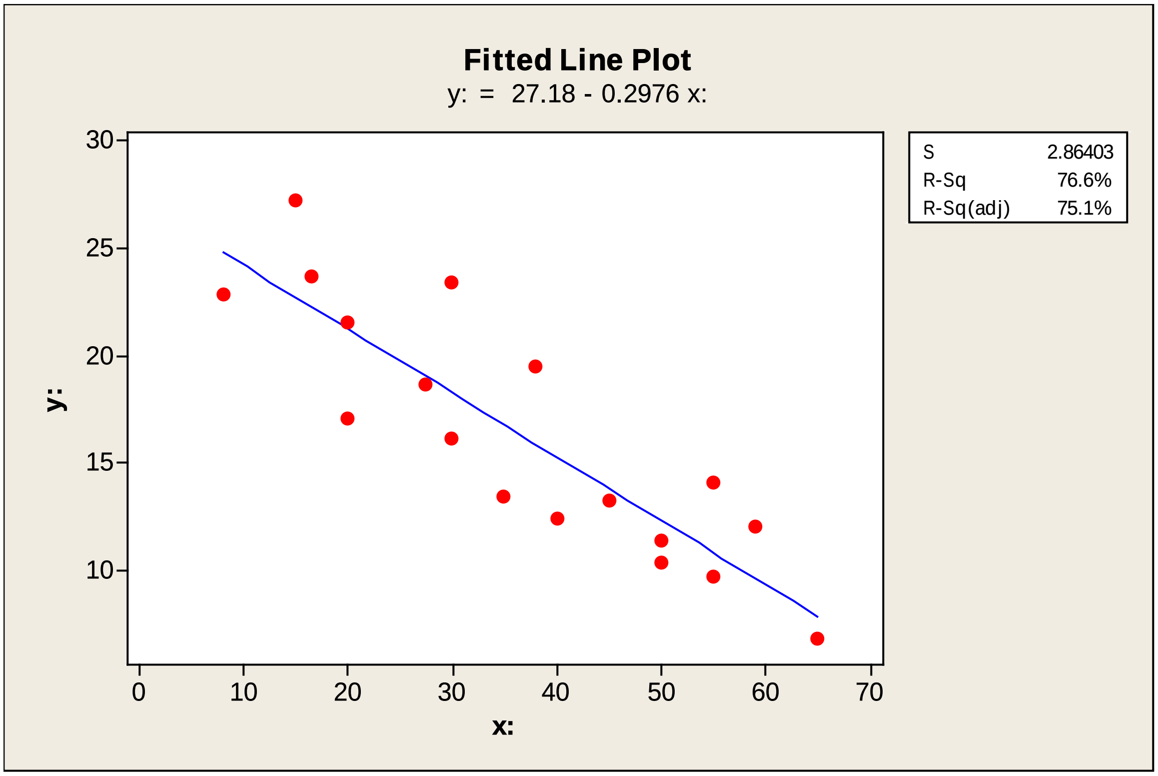
Regression Analysis: y: versus x:

```
The regression equation is
y: = 27.2 - 0.298 x:


Predictor      Coef  SE Coef      T      P
Constant     27.183    1.651  16.46  0.000
x:         -0.29756  0.04116  -7.23  0.000


S = 2.86403   R-Sq = 76.6%   R-Sq(adj) = 75.1%


Analysis of Variance

Source          DF      SS      MS      F      P
Regression       1  428.62  428.62  52.25  0.000
Residual Error  16  131.24    8.20
Total           17  559.86


Predicted Values for New Observations

New
Obs     Fit  SE Fit       95% CI             95% PI
  1  13.793   0.758  (12.185, 15.400)  (7.512, 20.073)


Values of Predictors for New Observations

New
Obs    x:
  1  45.0
```

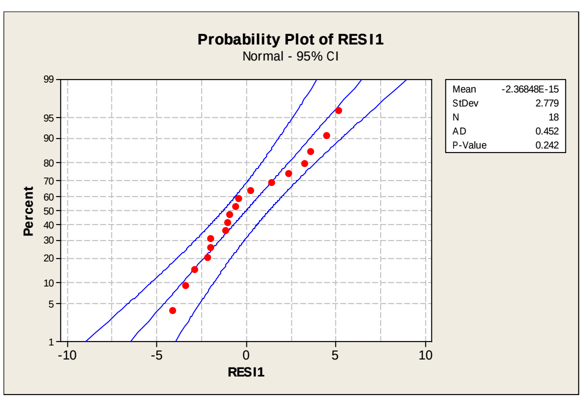
Conclusion: For cement samples with a carbonation depth of 45mm, we are 95% sure that the average strength is between 12.2 PMa and 15.4 MPa. For one sample of cement with a carbonation depth of 45mm, we are 95% sure that the strength of this individual sample is between 7.5 PMa and 20.1 MPa.

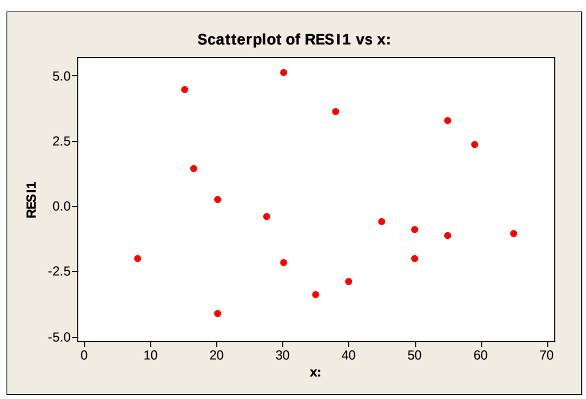
The regression model y = 27.2 - 0.298 x gives the estimated average strength y in MPa for concrete samples with a given carbonation depth x in mm. The t-value of -7.23 and the corresponding p-value of 0.000 indicates strong evidence that there is a positive linear relation between these two variables.

Stat > Regression > Fitted Line Plot



Conclusion: The regression line provides a reasonably good fit to the data, indicating that x=carbonation depth content can give a useful model to predict y=strength with a typical error of 2.86 MPa. 77% of the variations in strength can be attributed to variations in carbonation depth.





Conclusion: Probability plot (left) indicates that the residuals fit a normal distribution. Scatterplot (right) indicates the residuals seem to be independent and identically distributed. Hence the basic assumptions of the regression model are satisfied. Note: The p-value of 0.242 in the probability plot indicates that there is insufficient evidence to reject the null hypothesis of a normal fit.

Example 12.16

Stat > Regression > Regression (Storage: Residuals)

Regression Analysis: y versus x

```
The regression equation is
y = 1.00 + 93.4 x


Predictor   Coef  SE Coef     T      P
Constant   0.998    2.703  0.37  0.717
x          93.38    24.36  3.83  0.002


S = 3.89192   R-Sq = 51.2%   R-Sq(adj) = 47.7%


Analysis of Variance

Source          DF      SS      MS      F      P
Regression       1  222.48  222.48  14.69  0.002
Residual Error  14  212.06   15.15
Total           15  434.54


Unusual Observations

Obs      x       y    Fit  SE Fit  Residual  St Resid
 12  0.074  16.600  7.908   1.210     8.692     2.35R

R denotes an observation with a large standardized residual.
```

Stat > Basic Statistics > Correlation

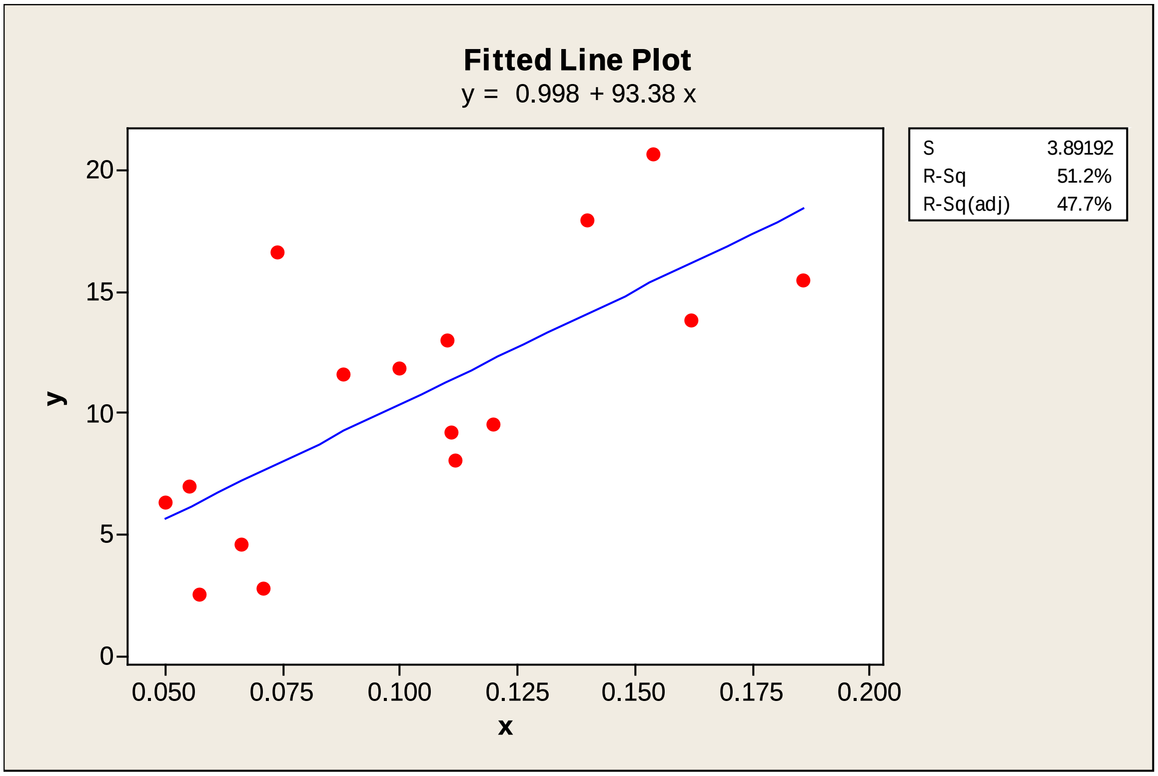
Correlations: x, y

```
Pearson correlation of x and y = 0.716
P-Value = 0.002
```

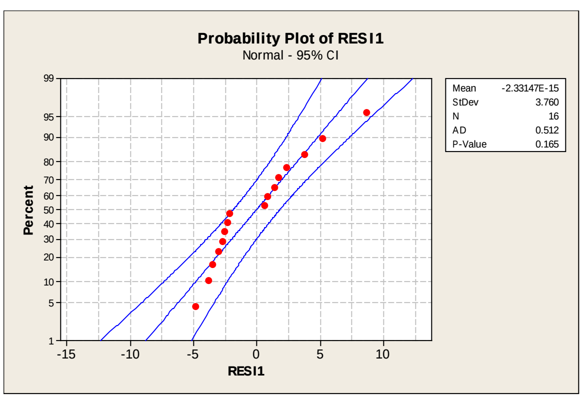
Conclusion: The correlation of .716 indicates a strong positive relation between ozone and carbon concentrations. The p=value of 0.002 indicates we are 99.8% sure there is a linear relation between ozone and carbon concentrations. Note that the p-values for the slope and the correlation are identical, this is always the case!

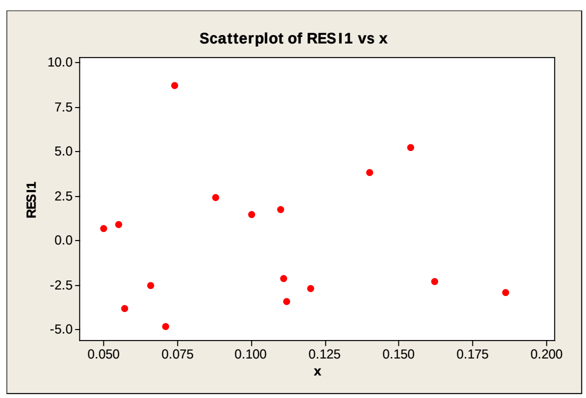
The regression model y = 1.00 + 93.4 x gives the estimated average carbon concentration in $\mu g/mm^3$ for air samples with a given ozone concentration x ppm. Variations in ozone concentration account for 51% of the variations in carbon concentration.

Stat > Regression > Fitted Line Plot



Conclusion: The regression line provides a reasonably good fit to the data, indicating that x=ozone concentration can give a useful model to predict y=carbon concentration with a typical error of 3.89 μg/mm^3.





Conclusion: Probability plot (left) indicates that the residuals fit a normal distribution. Scatterplot (right) indicates the residuals seem to be independent and identically distributed. Hence the basic assumptions of the regression model are satisfied.