

Linear Regression for Business Statistics

$$MPG = \beta_0 + \beta_1 Displacement$$

$$MPG = \beta_0 + \beta_1 Cylinders$$

Linear Regression for Business Statistics

$$MPG = \beta_0 + \beta_1 Displacement$$

$$MPG = \beta_0 + \beta_1 Cylinders$$



Linear Regression for Business Statistics

$$MPG = \beta_0 + \beta_1 \text{Displacement}$$

$$MPG = \beta_0 + \beta_1 \text{Cylinders}$$

Linear Regression for Business Statistics

$$MPG = \beta_0 + \beta_1 Displacement$$

(p-value 4.99×10^{-09})

$$MPG = \beta_0 + \beta_1 Cylinders$$

Linear Regression for Business Statistics

$$MPG = \beta_0 + \beta_1 Displacement$$

$$MPG = 35.027 - 0.0579 Displacement$$

(p-value 4.99×10^{-09})

$$MPG = \beta_0 + \beta_1 Cylinders$$

Linear Regression for Business Statistics

$$MPG = \beta_0 + \beta_1 Displacement$$

$$MPG = 35.027 - 0.0579 Displacement$$

(p-value 4.99×10^{-09})

$$MPG = \beta_0 + \beta_1 Cylinders$$

Linear Regression for Business Statistics

$$MPG = \beta_0 + \beta_1 Displacement$$

$$MPG = 35.027 - 0.0579 Displacement$$

(p-value 4.99×10^{-09})

$$MPG = \beta_0 + \beta_1 Cylinders$$

(p-value 1.07×10^{-09})



Linear Regression for Business Statistics

$$MPG = \beta_0 + \beta_1 Displacement$$

$$MPG = 35.027 - 0.0579 Displacement$$

(p-value 4.99×10^{-09})

$$MPG = \beta_0 + \beta_1 Cylinders$$

$$MPG = 42.509 - 3.29 Cylinders$$

(p-value 1.07×10^{-09})



Linear Regression for Business Statistics

$$MPG = \beta_0 + \beta_1 Displacement$$

$$MPG = 35.027 - 0.0579 Displacement$$

(p-value 4.99×10^{-09})

$$MPG = \beta_0 + \beta_1 Cylinders$$

$$MPG = 42.509 - 3.29 Cylinders$$

(p-value 1.07×10^{-09})

Linear Regression for Business Statistics

$$MPG = \beta_0 + \beta_1 Displacement + \beta_2 Cylinders$$

- Both X variables '*insignificant*'
- Reasonable R-square fits

$$MPG = \beta_0 + \beta_1 Displacement$$

$$MPG = \beta_0 + \beta_1 Cylinders$$

Linear Regression for Business Statistics

$$MPG = \beta_0 + \beta_1 Displacement + \beta_2 Cylinders$$

- Both X variables '*insignificant*'
- Reasonable R-square fits

$$MPG = \beta_0 + \beta_1 Displacement$$

$$MPG = \beta_0 + \beta_1 Cylinders$$

Linear Regression for Business Statistics

$$MPG = \beta_0 + \beta_1 Displacement + \beta_2 Cylinders$$

- Both X variables '*insignificant*'
- Reasonable R-square fits

$$MPG = \beta_0 + \beta_1 Displacement$$

$$MPG = \beta_0 + \beta_1 Cylinders$$

Linear Regression for Business Statistics

$$MPG = \beta_0 + \beta_1 Displacement + \beta_2 Cylinders$$

- Both X variables '*insignificant*'
- Reasonable R-square fits

$$MPG = \beta_0 + \beta_1 Displacement$$

- X variable '*significant*'.

$$MPG = \beta_0 + \beta_1 Cylinders$$

- X variable '*significant*'.

Linear Regression for Business Statistics

$$MPG = \beta_0 + \beta_1 Displacement + \beta_2 Cylinders$$

- Both X variables '*insignificant*'
- Reasonable R-square fits



$$MPG = \beta_0 + \beta_1 Displacement$$

- X variable '*significant*'.

$$MPG = \beta_0 + \beta_1 Cylinders$$

- X variable '*significant*'.


Linear Regression for Business Statistics

$$MPG = \beta_0 + \beta_1 Displacement + \beta_2 Cylinders$$

- Both X variables '*insignificant*'
- Reasonable R-square fits

$$MPG = \beta_0 + \beta_1 Displacement$$

- X variable '*significant*'.


$$MPG = \beta_0 + \beta_1 Cylinders$$

- X variable '*significant*'.

Linear Regression for Business Statistics

$$MPG = \beta_0 + \beta_1 Displacement + \beta_2 Cylinders$$

- Both X variables '*insignificant*'
- Reasonable R-square fits

$$MPG = \beta_0 + \beta_1 Displacement$$

- X variable '*significant*'.

$$MPG = \beta_0 + \beta_1 Cylinders$$

- X variable '*significant*'.

Linear Regression for Business Statistics

$$MPG = \beta_0 + \beta_1 Displacement + \beta_2 Cylinders$$

- Both X variables '*insignificant*'
- Reasonable R-square fits

$$MPG = \beta_0 + \beta_1 Displacement$$

- X variable '*significant*'.

$$MPG = \beta_0 + \beta_1 Cylinders$$

- X variable '*significant*'.

Multicollinearity

Linear Regression for Business Statistics

Multicollinearity

- ❑ Occurs when two or more X variables are highly correlated.
- ❑ Variation in Y cannot be apportioned individually across X variables.
- ❑ Model fits may be reasonable and prediction may be ok.
- ❑ Interpretation of individual X variable impact will be suspect.

Linear Regression for Business Statistics

Multicollinearity

- ❑ Occurs when two or more X variables are highly correlated.
- ❑ Variation in Y cannot be apportioned individually across X variables.
- ❑ Model fits may be reasonable and prediction may be ok.
- ❑ Interpretation of individual X variable impact will be suspect.

Linear Regression for Business Statistics

Multicollinearity

- ❑ Occurs when two or more X variables are highly correlated.
- ❑ Variation in Y cannot be apportioned individually across X variables.
- ❑ Model fits may be reasonable and prediction may be ok.
- ❑ Interpretation of individual X variable impact will be suspect.

Linear Regression for Business Statistics

Multicollinearity

- ❑ Occurs when two or more X variables are highly correlated.
- ❑ Variation in Y cannot be apportioned individually across X variables.
- ❑ Model fits may be reasonable and prediction may be ok.
- ❑ Interpretation of individual X variable impact will be suspect.

Linear Regression for Business Statistics

Multicollinearity

- ❑ Occurs when two or more X variables are highly correlated.
- ❑ Variation in Y cannot be apportioned individually across X variables.
- ❑ Model fits may be reasonable and prediction may be ok.
- ❑ Interpretation of individual X variable impact will be suspect.

$$MPG = \beta_0 + \beta_1 Displacement + \beta_2 Cylinders$$

- This model may predict reasonably.
- However, individual impact of *Displacement* and *Cylinders* cannot be discerned.

Linear Regression for Business Statistics

Multicollinearity

- ❑ Occurs when two or more X variables are highly correlated.
- ❑ Variation in Y cannot be apportioned individually across X variables.
- ❑ Model fits may be reasonable and prediction may be ok.
- ❑ Interpretation of individual X variable impact will be suspect.

$$MPG = \beta_0 + \beta_1 Displacement + \beta_2 Cylinders$$

- This model may predict reasonably.
- However, individual impact of *Displacement* and *Cylinders* cannot be discerned.

Linear Regression for Business Statistics

Multicollinearity

- ❑ Occurs when two or more X variables are highly correlated.
- ❑ Variation in Y cannot be apportioned individually across X variables.
- ❑ Model fits may be reasonable and prediction may be ok.
- ❑ Interpretation of individual X variable impact will be suspect.

$$MPG = \beta_0 + \beta_1 Displacement + \beta_2 Cylinders$$

- This model may predict reasonably.
- However, individual impact of *Displacement* and *Cylinders* cannot be discerned.

Linear Regression for Business Statistics

Detecting Multicollinearity

- ❑ Some software produce collinearity statistics such as VIF (Variance Inflation Factor).
- ❑ We can look at correlations across X variables.

Linear Regression for Business Statistics

Detecting Multicollinearity

- ❑ Some software produce collinearity statistics such as VIF (Variance Inflation Factor).
- ❑ We can look at correlations across X variables.

Linear Regression for Business Statistics

Detecting Multicollinearity

- ❑ Some software produce collinearity statistics such as VIF (Variance Inflation Factor).
- ❑ We can look at correlations across X variables.

Correcting Multicollinearity

- ❑ Do you *need* to correct for collinearity ?
 - May not be a *problem* if using regression only to predict.
 - Collinearity is problematic when interpreting coefficient impacts.
- ❑ To remove multicollinearity you may drop variables causing high correlation.

Linear Regression for Business Statistics

Detecting Multicollinearity

- ❑ Some software produce collinearity statistics such as VIF (Variance Inflation Factor).
- ❑ We can look at correlations across X variables.

Correcting Multicollinearity

- ❑ Do you *need* to correct for collinearity ?
 - May not be a *problem* if using regression only to predict.
 - Collinearity is problematic when interpreting coefficient impacts.
- ❑ To remove multicollinearity you may drop variables causing high correlation.

Linear Regression for Business Statistics

Detecting Multicollinearity

- ❑ Some software produce collinearity statistics such as VIF (Variance Inflation Factor).
- ❑ We can look at correlations across X variables.

Correcting Multicollinearity

- ❑ Do you *need* to correct for collinearity ?
 - May not be a *problem* if using regression only to predict.
 - Collinearity is problematic when interpreting coefficient impacts.
- ❑ To remove multicollinearity you may drop variables causing high correlation.

Linear Regression for Business Statistics

Detecting Multicollinearity

- ❑ Some software produce collinearity statistics such as VIF (Variance Inflation Factor).
- ❑ We can look at correlations across X variables.

Correcting Multicollinearity

- ❑ Do you *need* to correct for collinearity ?
 - May not be a *problem* if using regression only to predict.
 - Collinearity is problematic when interpreting coefficient impacts.
- ❑ To remove multicollinearity you may drop variables causing high correlation.

Linear Regression for Business Statistics

Detecting Multicollinearity

- ❑ Some software produce collinearity statistics such as VIF (Variance Inflation Factor).
- ❑ We can look at correlations across X variables.

Correcting Multicollinearity

- ❑ Do you *need* to correct for collinearity ?
 - May not be a *problem* if using regression only to predict.
 - Collinearity is problematic when interpreting coefficient impacts.
- ❑ To remove multicollinearity you may drop variables causing high correlation.