

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Categorical Variable

Is a variable that takes on values which are but labels.

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Categorical Variable

Is a variable that takes on values which are but labels.

- ***Gender*** is a categorical variable.

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Categorical Variable

Is a variable that takes on values which are but labels.

- *Gender* is a categorical variable.
- *Nationality* is a categorical variable.

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Categorical Variable

Is a variable that takes on values which are but labels.

- *Gender* is a categorical variable.
- *Nationality* is a categorical variable.
- *Race of a person* is a categorical variable.

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Categorical Variable

Is a variable that takes on values which are but labels.

- ❑ *Gender* is a categorical variable.
- ❑ *Nationality* is a categorical variable.
- ❑ *Race of a person* is a categorical variable.
- ❑ *Height of a person ?*

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Categorical Variable

Is a variable that takes on values which are but labels.

- ❑ *Gender* is a categorical variable.
- ❑ *Nationality* is a categorical variable.
- ❑ *Race of a person* is a categorical variable.
- ❑ *Height of a person* is **not** a categorical variable.

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Categorical Variable

Is a variable that takes on values which are but labels.

- ❑ *Gender* is a categorical variable.
- ❑ *Nationality* is a categorical variable.
- ❑ *Race of a person* is a categorical variable.
- ❑ *Height of a person* is **not** a categorical variable.
- ❑ *Quarter of a year* ?

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Categorical Variable

Is a variable that takes on values which are but labels.

- *Gender* is a categorical variable.
- *Nationality* is a categorical variable.
- *Race of a person* is a categorical variable.
- *Height of a person* is **not** a categorical variable.
- *Quarter of a year (1, 2, 3, 4)*

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Categorical Variable

Is a variable that takes on values which are but labels.

- ❑ *Gender* is a categorical variable.
- ❑ *Nationality* is a categorical variable.
- ❑ *Race of a person* is a categorical variable.
- ❑ *Height of a person* is **not** a categorical variable.
- ❑ *Quarter of a year (1, 2, 3, 4) ?*

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Categorical Variable

Is a variable that takes on values which are but labels.

- *Gender* is a categorical variable.
- *Nationality* is a categorical variable.
- *Race of a person* is a categorical variable.
- *Height of a person* is **not** a categorical variable.
- *Quarter of a year (1, 2, 3, 4)* is a categorical variable.

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Examples

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Examples

- **Study on gender discrimination in salaries at workplace.**

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Examples

- **Study on gender discrimination in salaries at workplace.**
- **Sales regression model.**

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Examples

- Study on gender discrimination in salaries at workplace.
- Sales regression model.

A special technique needed,

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Examples

- Study on gender discrimination in salaries at workplace.
- Sales regression model.

A special technique needed,

'Dummy Variable Regression'

'Indicator Variable Regression'

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Examples

- Study on gender discrimination in salaries at workplace.
- Sales regression model.

A special technique needed,

'Dummy Variable Regression'

'Indicator Variable Regression'

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Examples

- Study on gender discrimination in salaries at workplace.
- Sales regression model.

A special technique needed,

'Dummy Variable Regression'

'Indicator Variable Regression'

It involves the use of,

'Dummy Variables' / 'Indicator Variables'

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Two Rules governing 'Dummy Variables'

- A dummy variable can only take a value 0 or 1 .
- the number of dummy variables required in a regression is one less than the number of categories in the categorical variable.

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Two Rules governing 'Dummy Variables'

- A dummy variable can only take a value 0 or 1 .
- the number of dummy variables required in a regression is one less than the number of categories in the categorical variable.

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Two Rules governing 'Dummy Variables'

- A dummy variable can only take a value 0 or 1 .
- the number of dummy variables required in a regression is one less than the number of categories in the categorical variable.

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Two Rules governing 'Dummy Variables'

- A dummy variable can only take a value 0 or 1 .
- the number of dummy variables required in a regression is one less than the number of categories in the categorical variable.

Gender → Male
 → Female

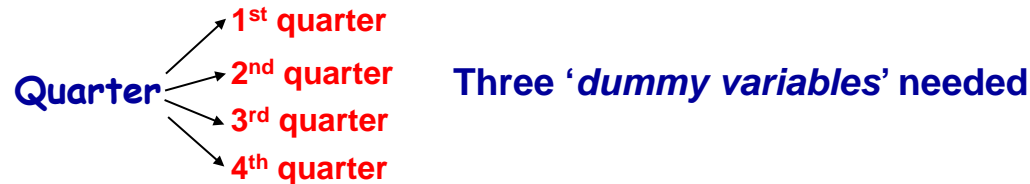
One 'dummy variable' needed

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Two Rules governing 'Dummy Variables'

- A dummy variable can only take a value 0 or 1 .
- the number of dummy variables required in a regression is one less than the number of categories in the categorical variable.



Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Example (delivery1.xlsx)

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Example (delivery1.xlsx)

A parcel delivery service operates in two different regions, region “A” and region “B”. Delivery trucks leave the central warehouse and travel to region A and deliver parcels in that region. Similarly delivery trucks also leave the central warehouse and travel to region B and deliver parcels in that region.

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Example (delivery1.xlsx)

A parcel delivery service operates in two different regions, **region “A”** and **region “B”**. Delivery trucks leave the central warehouse and travel to region A and deliver parcels in that region. Similarly delivery trucks also leave the central warehouse and travel to region B and deliver parcels in that region.

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Example (delivery1.xlsx)

A parcel delivery service operates in two different regions, **region “A”** and **region “B”**. Delivery trucks leave the central warehouse and travel to region A and deliver parcels in that region. Similarly delivery trucks also leave the central warehouse and travel to region B and deliver parcels in that region.

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Minutes

Region

Parcels

TruckAge



Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Minutes

Region

Parcels

TruckAge



Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Minutes

Region

Parcels

TruckAge



Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Minutes

Region

Parcels

TruckAge

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Minutes

Region

Parcels

TruckAge



Y variable

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

Minutes



Y variable

Region



Parcels



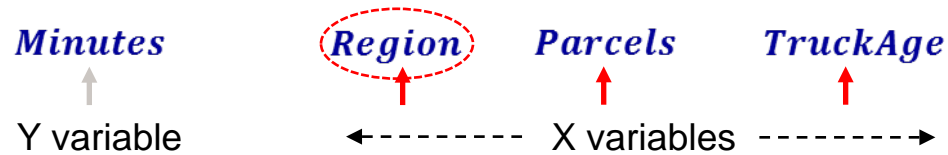
TruckAge



←----- X variables ----->

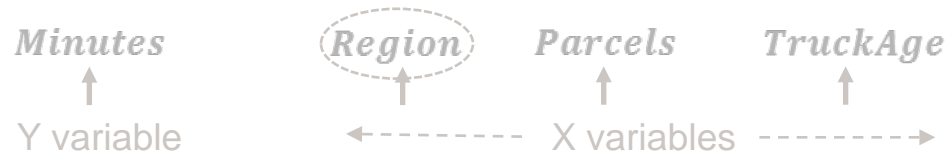
Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.



Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.



Region → **A**
 → **B** One 'dummy variable' needed

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.



Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.



Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.



Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.



Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.



Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.



Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.



Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.



$$\text{Minutes} = \beta_0 + \beta_1 \text{REGA} + \beta_2 \text{Parcels} + \beta_3 \text{TruckAge}$$

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.



$$\text{Minutes} = \beta_0 + \beta_1 \text{REGA} + \beta_2 \text{Parcels} + \beta_3 \text{TruckAge}$$

↑

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

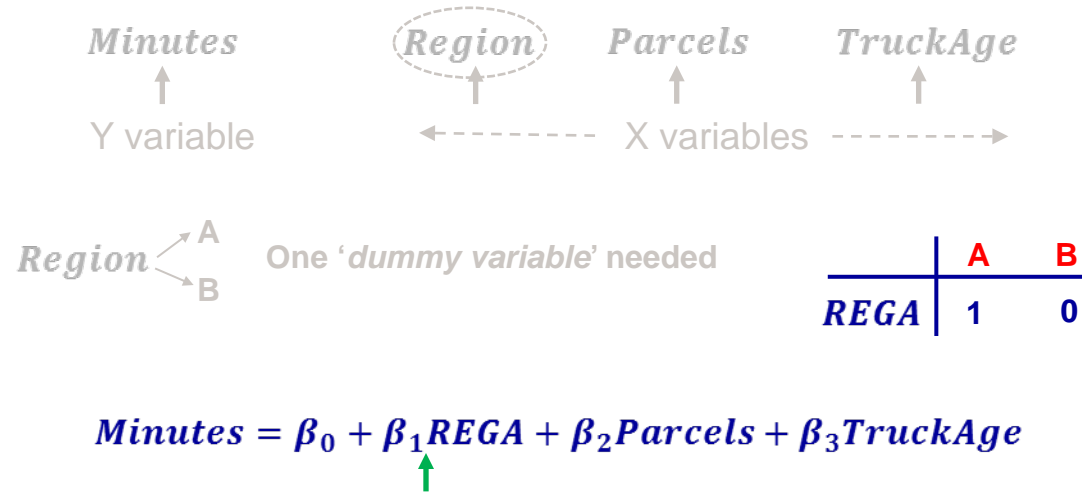


$$Minutes = \beta_0 + \beta_1 REGA + \beta_2 Parcels + \beta_3 TruckAge$$

↑

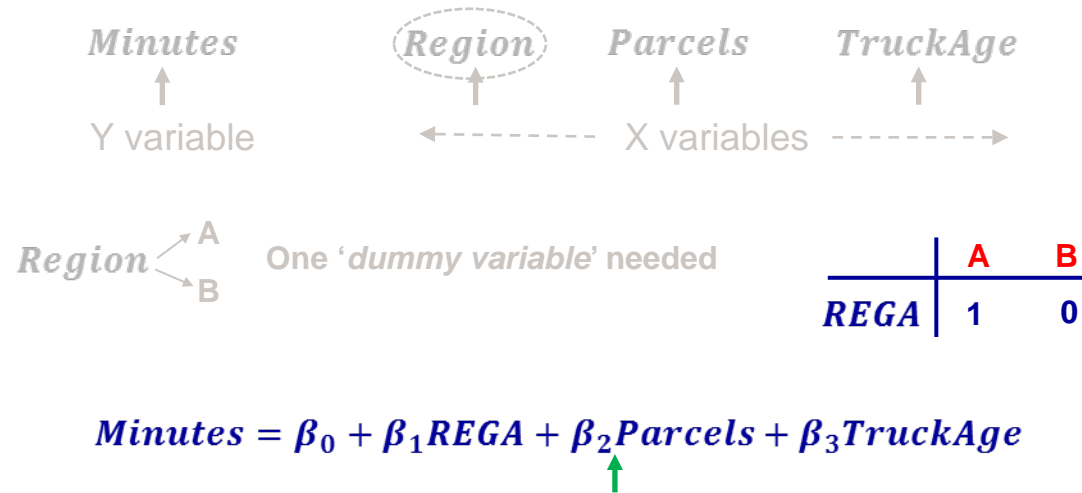
Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.



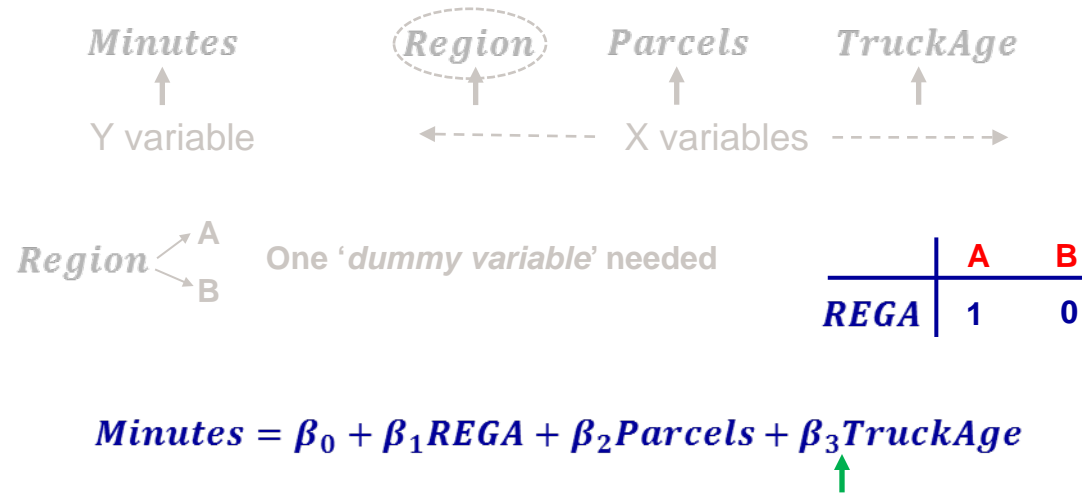
Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.



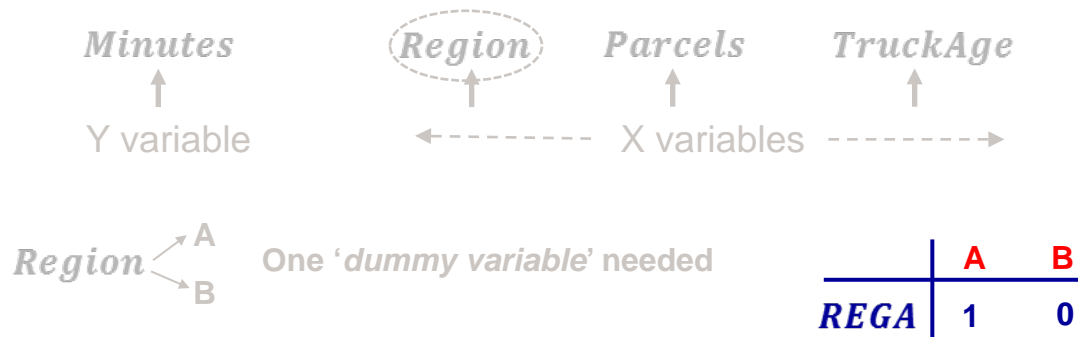
Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.



Linear Regression for Business Statistics

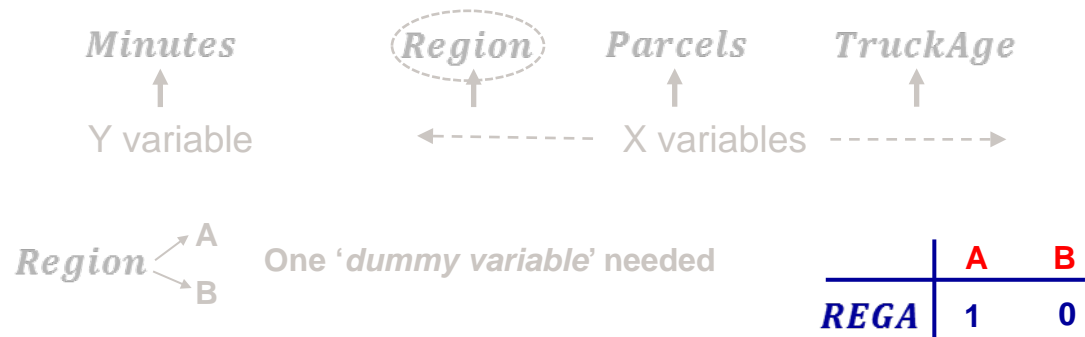
Incorporating 'Categorical Variables' in a regression model.



$$\text{Minutes} = \beta_0 + \beta_1 \text{REGA} + \beta_2 \text{Parcels} + \beta_3 \text{TruckAge}$$

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

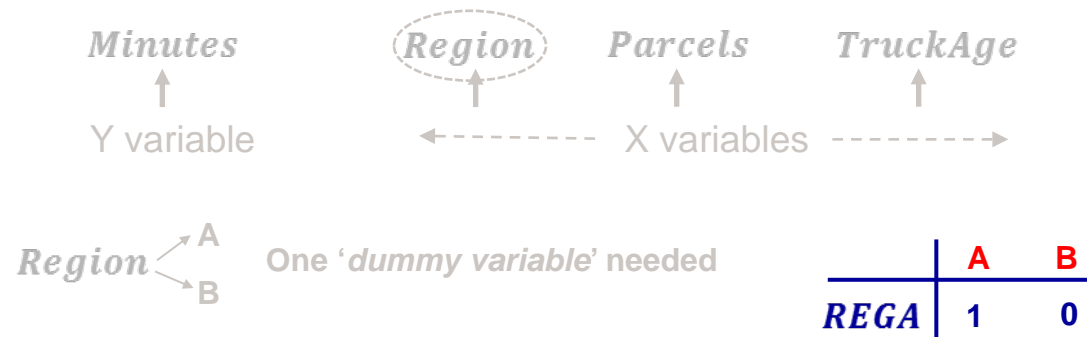


$$\text{Minutes} = \beta_0 + \beta_1 \text{REGA} + \beta_2 \text{Parcels} + \beta_3 \text{TruckAge}$$

$$\text{Minutes} = -33.13 + 106.84 \text{REGA} + 10.02 \text{Parcels} + 3.21 \text{TruckAge}$$

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.



$$\text{Minutes} = \beta_0 + \beta_1 \text{REGA} + \beta_2 \text{Parcels} + \beta_3 \text{TruckAge}$$

$$\text{Minutes} = -33.13 + 106.84 \text{REGA} + 10.02 \text{Parcels} + 3.21 \text{TruckAge}$$

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.



$$\text{Minutes} = \beta_0 + \beta_1 \text{REGA} + \beta_2 \text{Parcels} + \beta_3 \text{TruckAge}$$

$$\text{Minutes} = -33.13 + 106.84 \text{REGA} + 10.02 \text{Parcels} + 3.21 \text{TruckAge}$$

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.



$$\text{Minutes} = \beta_0 + \beta_1 \text{REGA} + \beta_2 \text{Parcels} + \beta_3 \text{TruckAge}$$

$$\text{Minutes} = -33.13 + 106.84 \text{REGA} + 10.02 \text{Parcels} + 3.21 \text{TruckAge}$$

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.

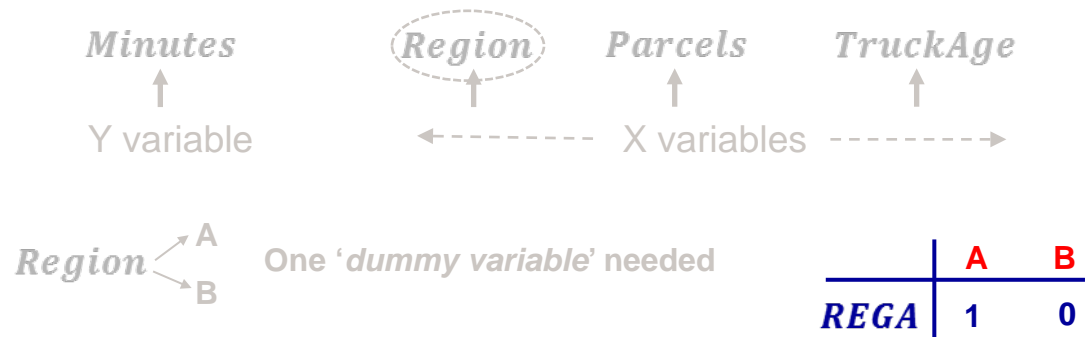


$$\text{Minutes} = \beta_0 + \beta_1 \text{REGA} + \beta_2 \text{Parcels} + \beta_3 \text{TruckAge}$$

$$\text{Minutes} = -33.13 + 106.84 \text{REGA} + 10.02 \text{Parcels} + 3.21 \text{TruckAge}$$

Linear Regression for Business Statistics

Incorporating 'Categorical Variables' in a regression model.



$$\text{Minutes} = \beta_0 + \beta_1 \text{REGA} + \beta_2 \text{Parcels} + \beta_3 \text{TruckAge}$$

$$\text{Minutes} = -33.13 + 106.84 \text{REGA} + 10.02 \text{Parcels} + 3.21 \text{TruckAge}$$