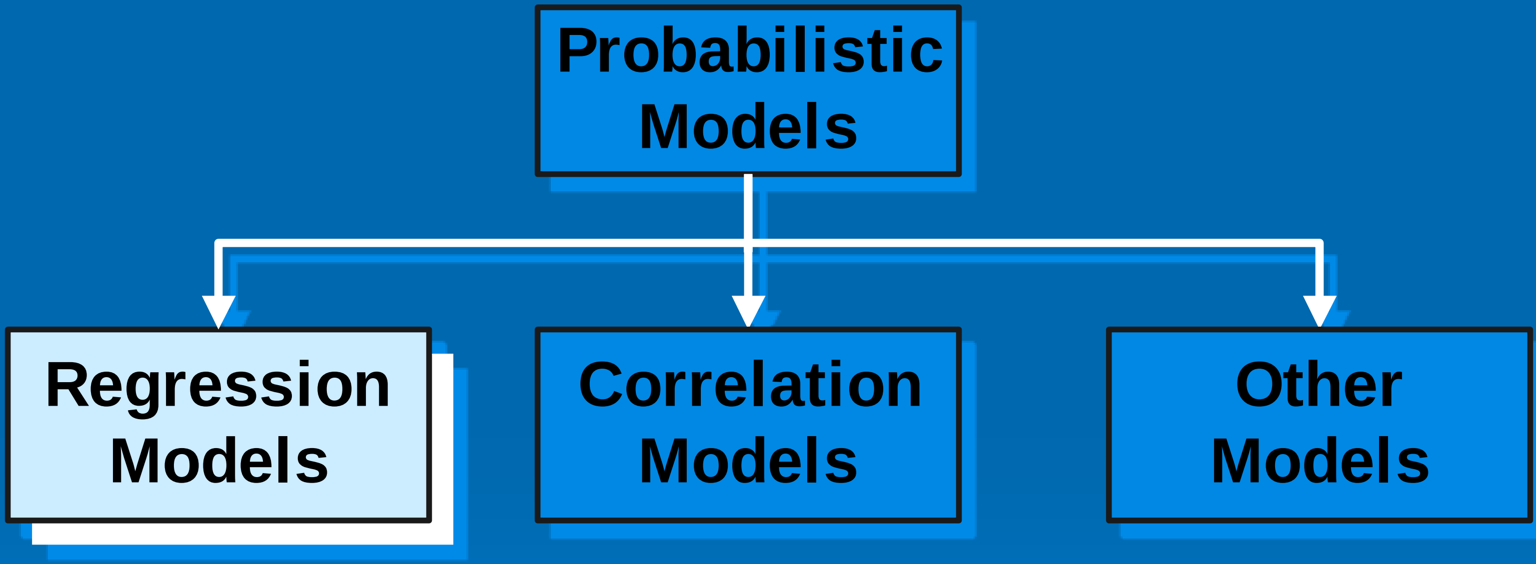
Learning Objectives…

1. Correlation Models

1. Link between a correlation model and a regression model

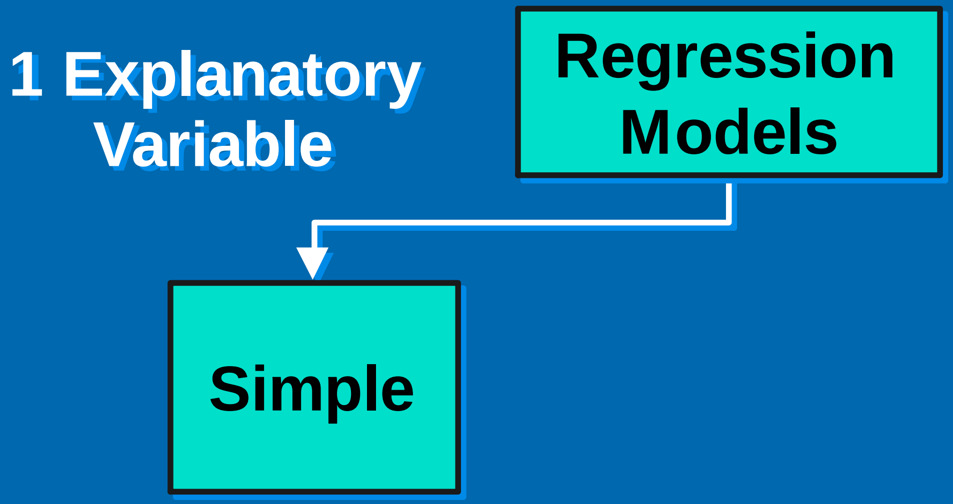
Types of Probabilistic Models



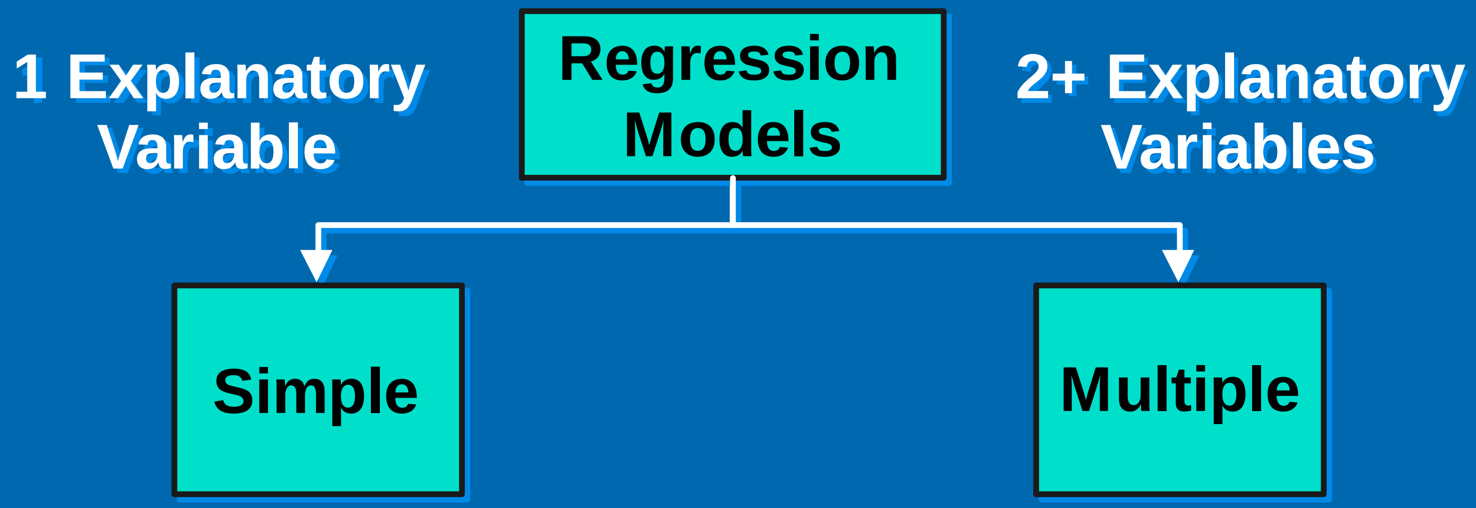
Types of Regression Models

Regression Models

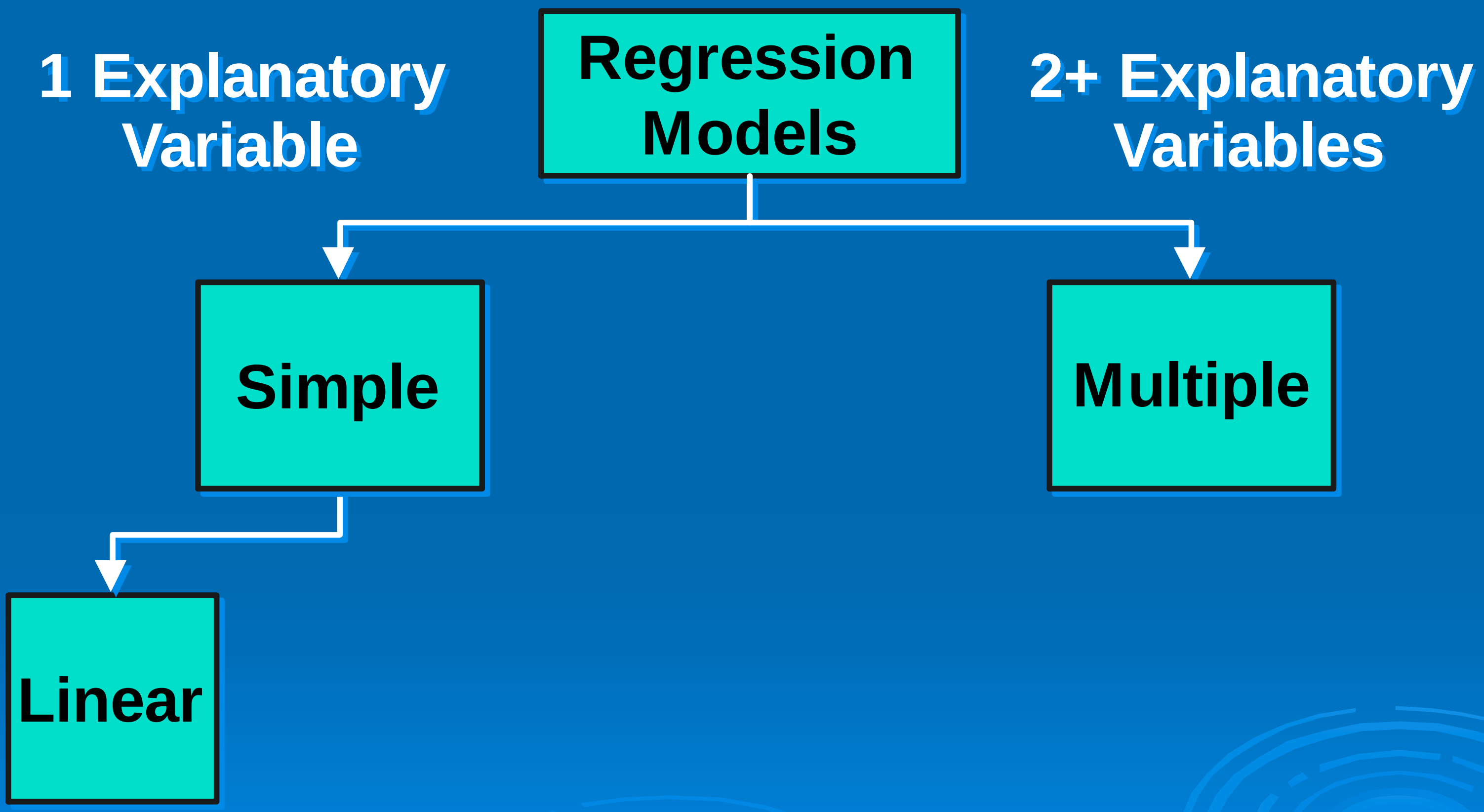
Types of Regression Models



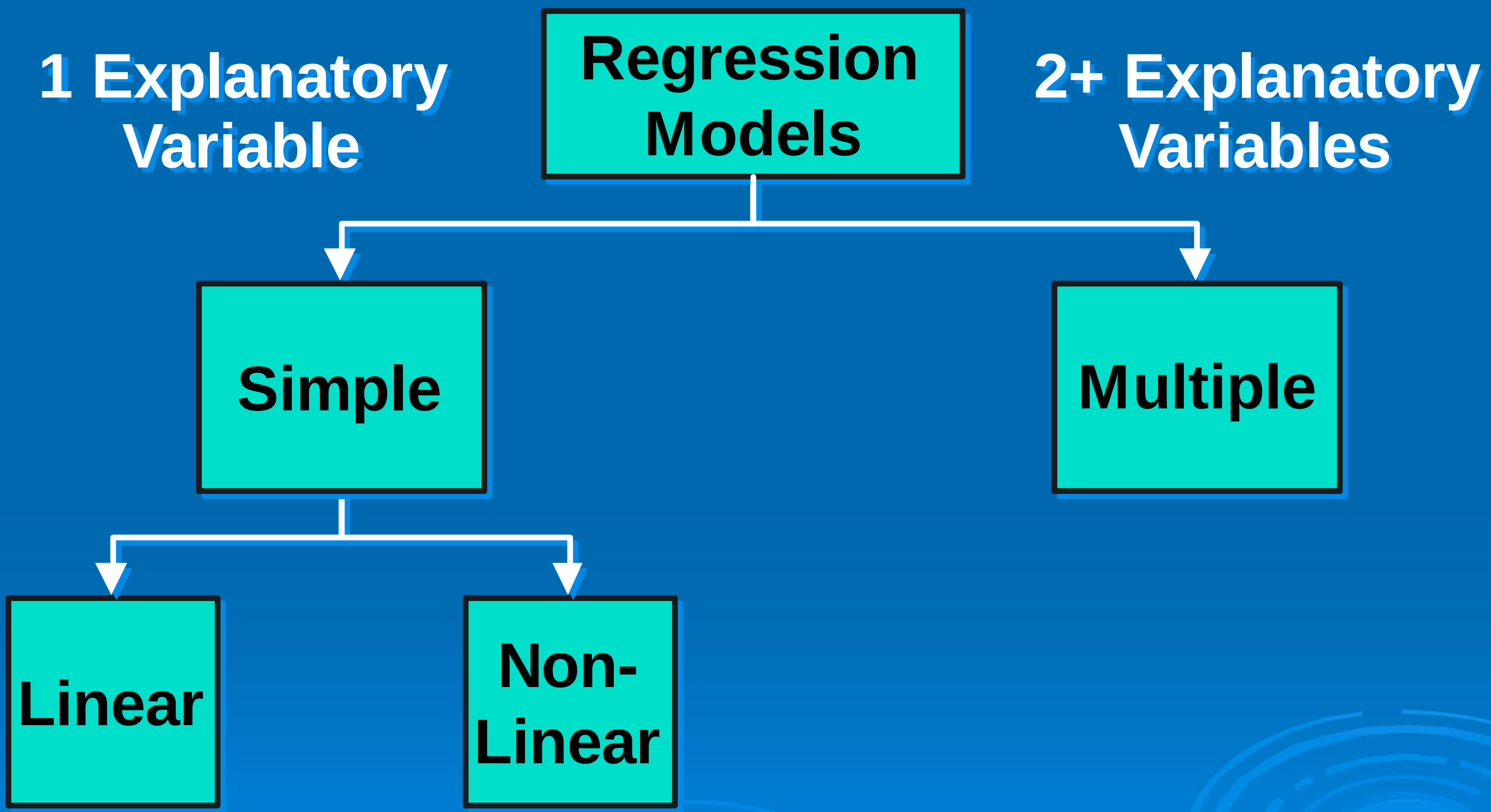
Types of Regression Models



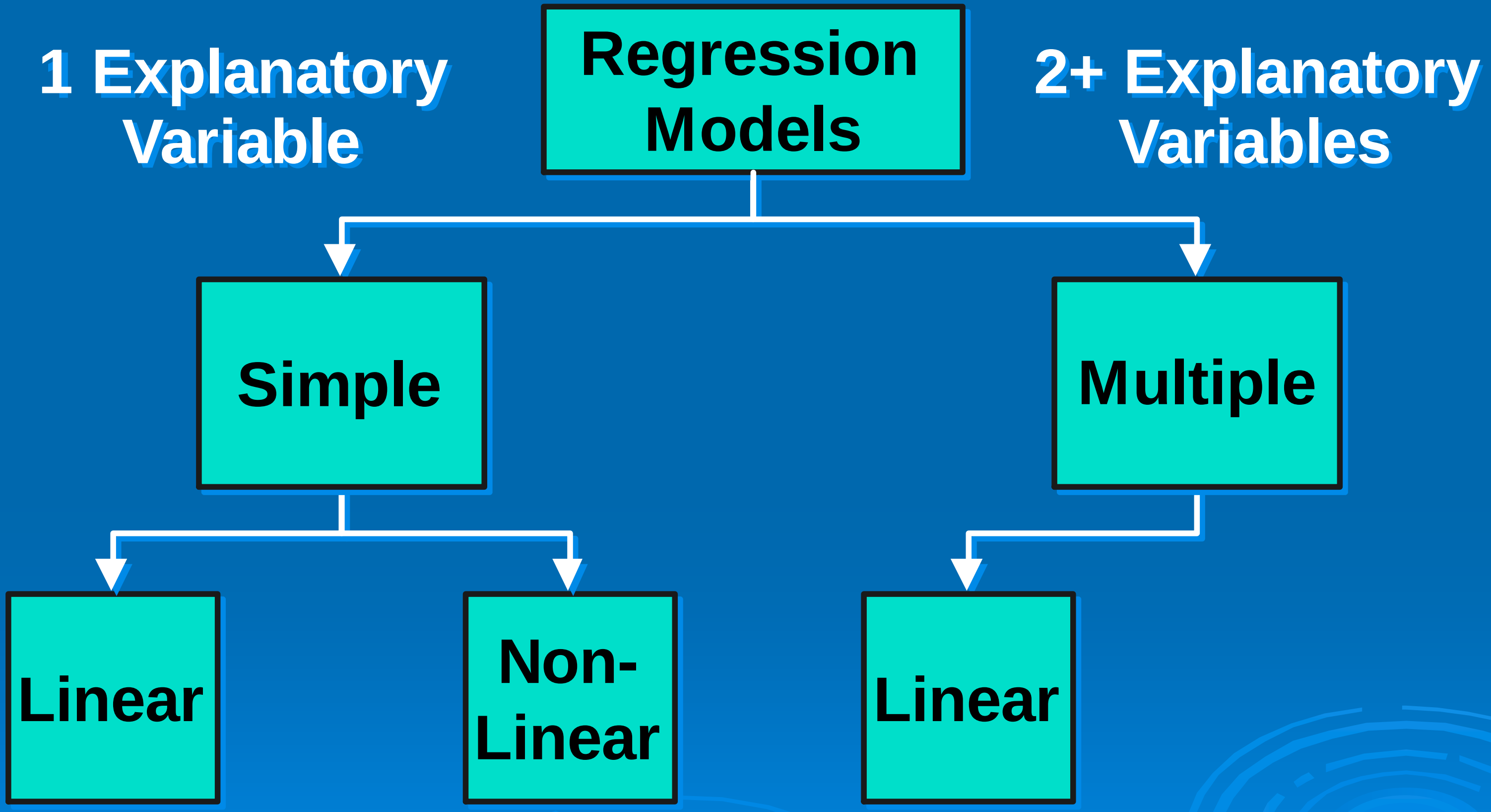
Types of Regression Models



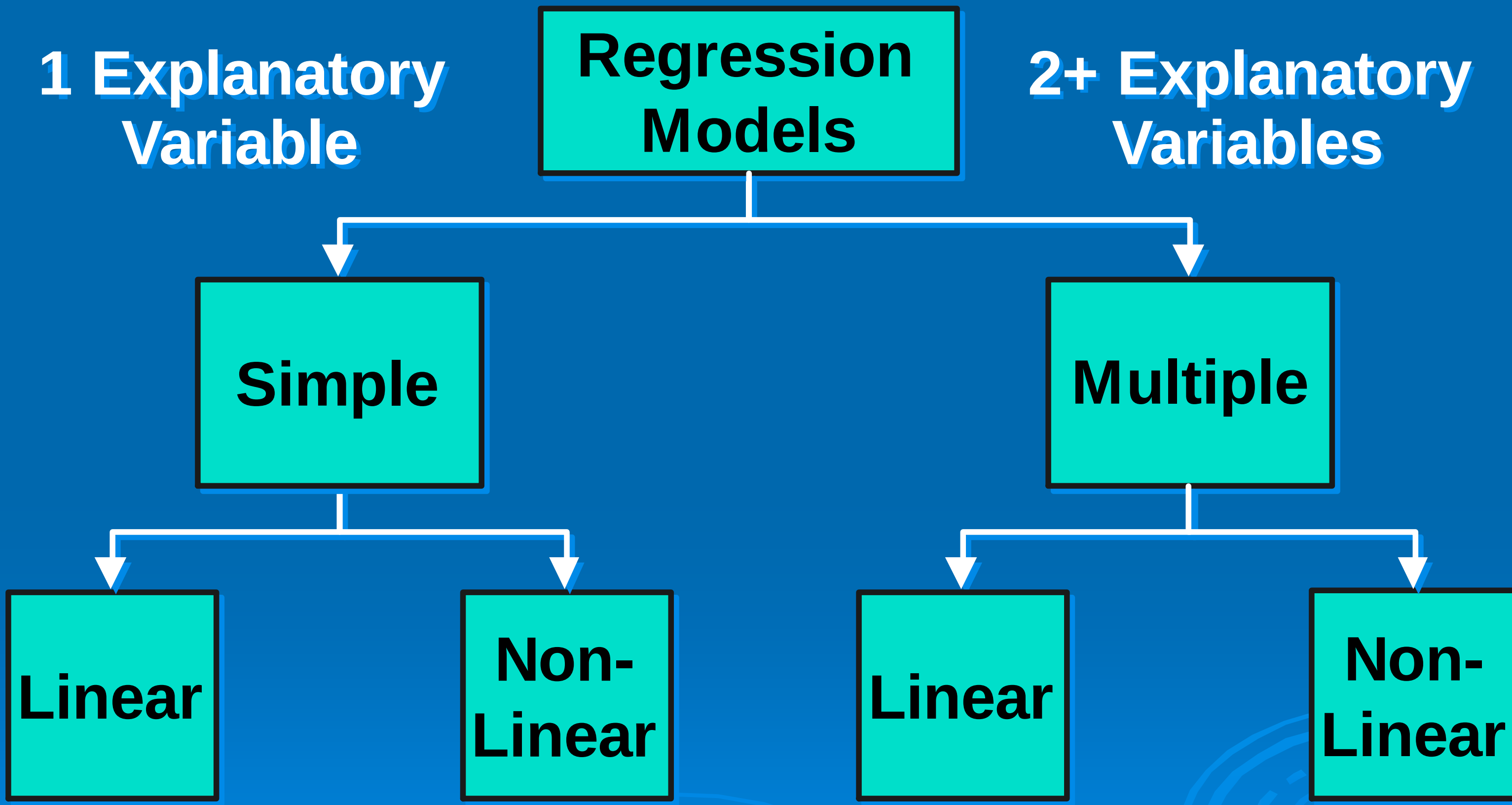
Types of Regression Models



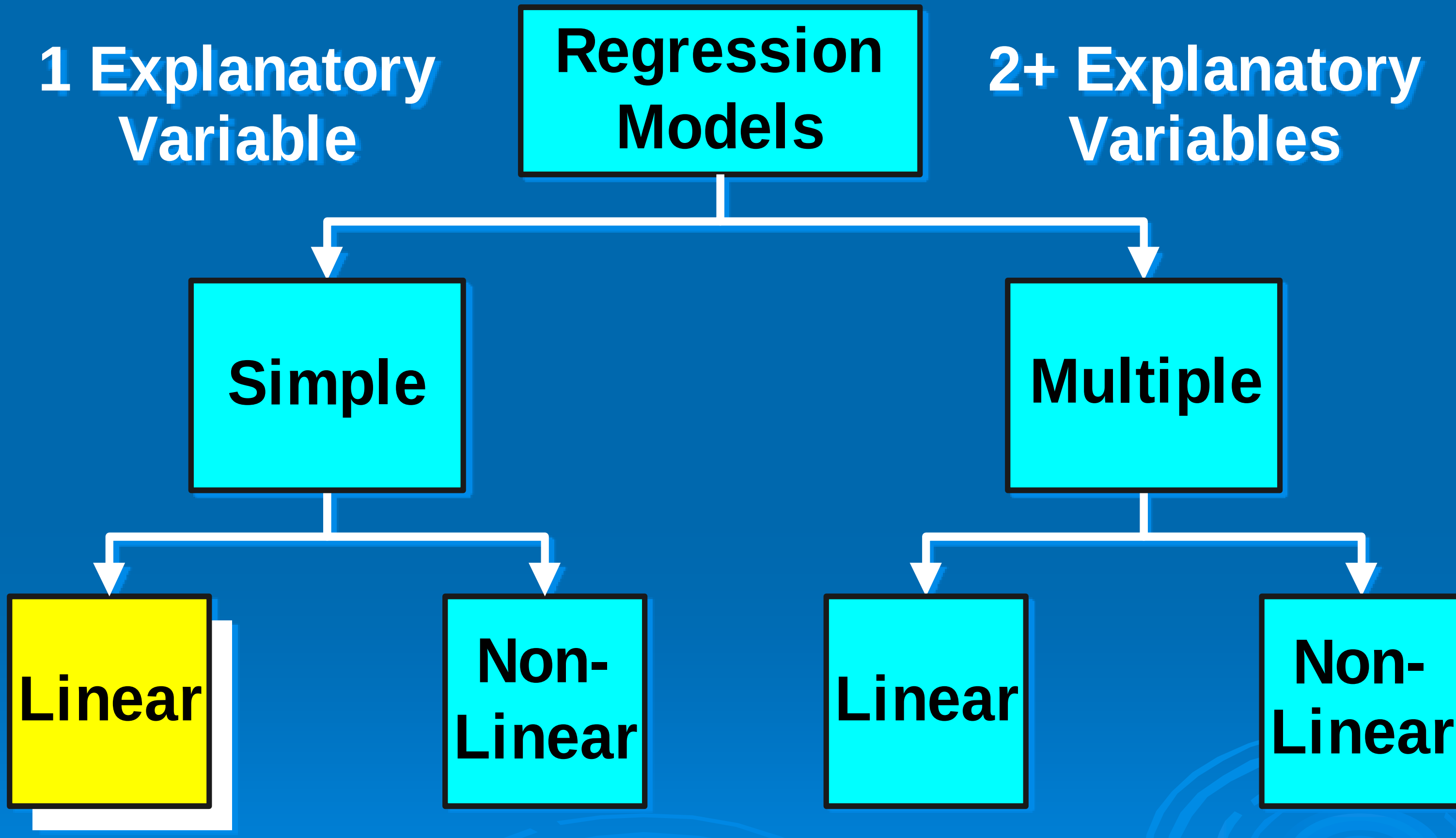
Types of Regression Models



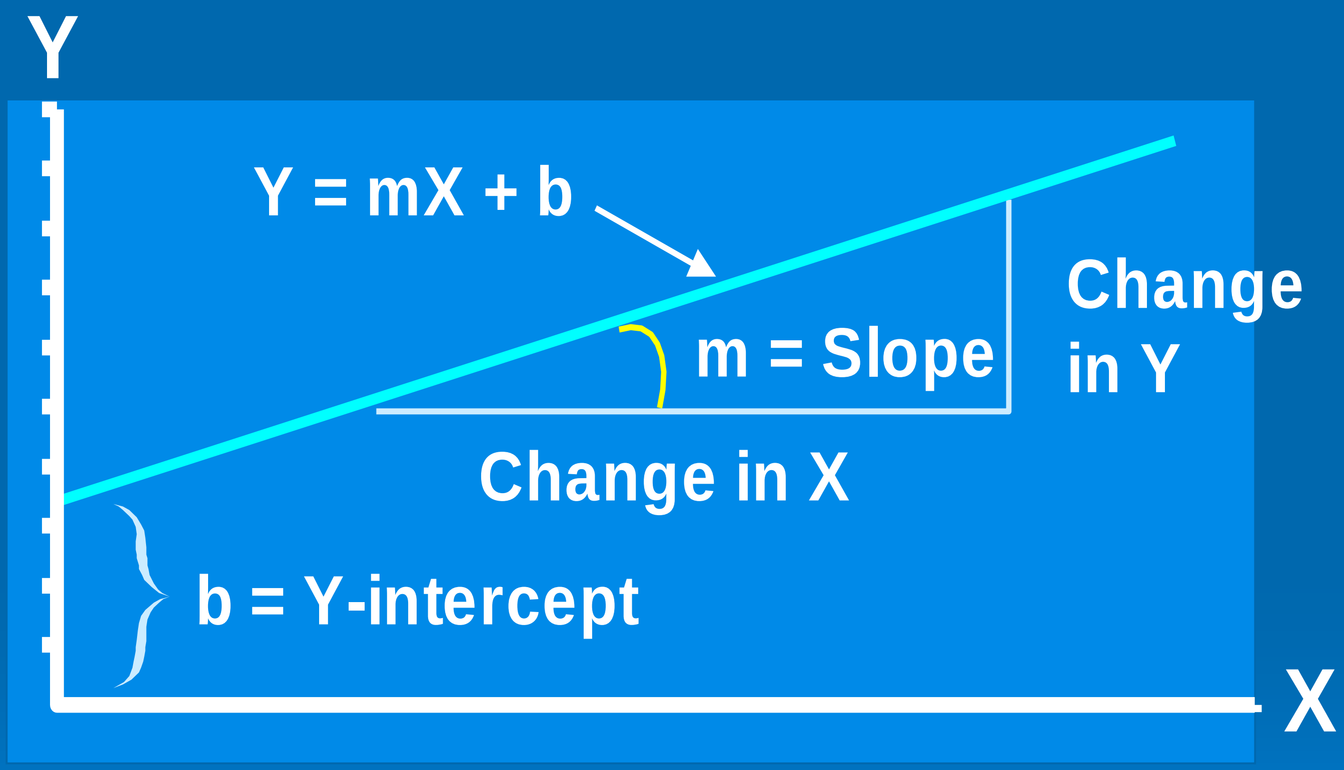
Types of Regression Models



Types of Regression Models



Linear Equations



Linear Regression Model

- 1. Relationship Between Variables is a Linear Function

Population Y-Intercept

Population Slope

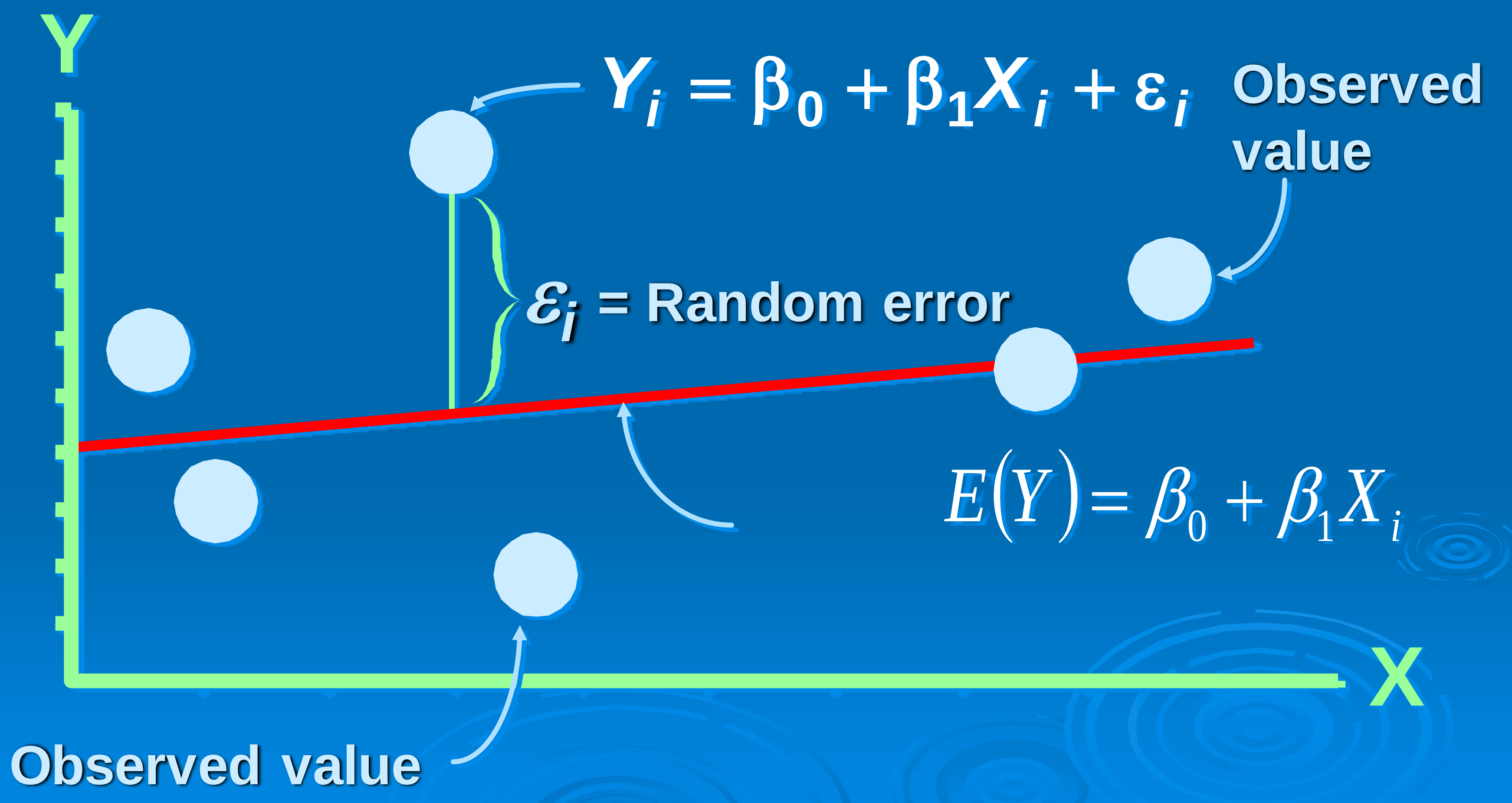
Random Error

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

Dependent (Response) Variable (e.g., CD+ c.)

Independent (Explanatory) Variable (e.g., Years s. serocon.)

Population Linear Regression Model



Example with R

- library("Hmisc")
- mydata <- iris

- str(mydata)

- 'data.frame': 150 obs. of 5 variables:
- $ Sepal.Length: num 5.1 4.9 4.7 4.6 5 5.4 4.6 5 4.4 4.9 ...
- $ Sepal.Width : num 3.5 3 3.2 3.1 3.6 3.9 3.4 3.4 2.9 3.1 ...
- $ Petal.Length: num 1.4 1.4 1.3 1.5 1.4 1.7 1.4 1.5 1.4 1.5 ...
- $ Petal.Width : num 0.2 0.2 0.2 0.2 0.2 0.4 0.3 0.2 0.2 0.1 ...
- $ Species : Factor w/ 3 levels "setosa","versicolor",..: 1 1 1 1 1 1 1 1 1 1 ...

- cor <- rcorr(as.matrix(mydata[1:4]))

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width
Sepal.Length	1.00	-0.12	0.87	0.82
Sepal.Width	-0.12	1.00	-0.43	-0.37
Petal.Length	0.87	-0.43	1.00	0.96
Petal.Width	0.82	-0.37	**0.96**	1.00

- n= 150

- P

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width
Sepal.Length		0.1519	0.0000	0.0000
Sepal.Width	0.1519		0.0000	0.0000
Petal.Length	0.0000	0.0000		0.0000
Petal.Width	0.0000	0.0000	**0.0000**	

- **Fit <- lm (mydata$Petal.Width ~ mydata$Petal.Length)**

- **Summary** (**Fit**)
- Call:
- lm(formula = mydata$Petal.Width ~ mydata$Petal.Length)

- Residuals:
- Min 1Q Median 3Q Max
- -0.56515 -0.12358 -0.01898 0.13288 0.64272

- **Coefficients**:

	Estimate	Std. Error	t value	**Pr(>|t|)**
(Intercept)	-0.363076	0.039762	-9.131	**4.7e-16 *** **
mydata$Petal.Length	0.415755	0.009582	43.387	**< 2e-16 *** **

- ---
- Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

- Residual standard error: 0.2065 on 148 degrees of freedom
- Multiple R-squared: 0.9271, **Adjusted R-squared: 0.9266**
- F-statistic: 1882 on 1 and 148 DF, **p-value: < 2.2e-16**

Linear Regression Model

$$\left\{ \begin{array}{l} y = \beta_0 + \beta_1 x_1 + \mathrm{e} \\ \\ B_0 = -0.363 \quad \beta_1 = 0.415 \end{array} \right.$$

$$y = -0.363 + 0.415\, x_1 + \mathrm{e}$$

p-value: < 2.2e-16