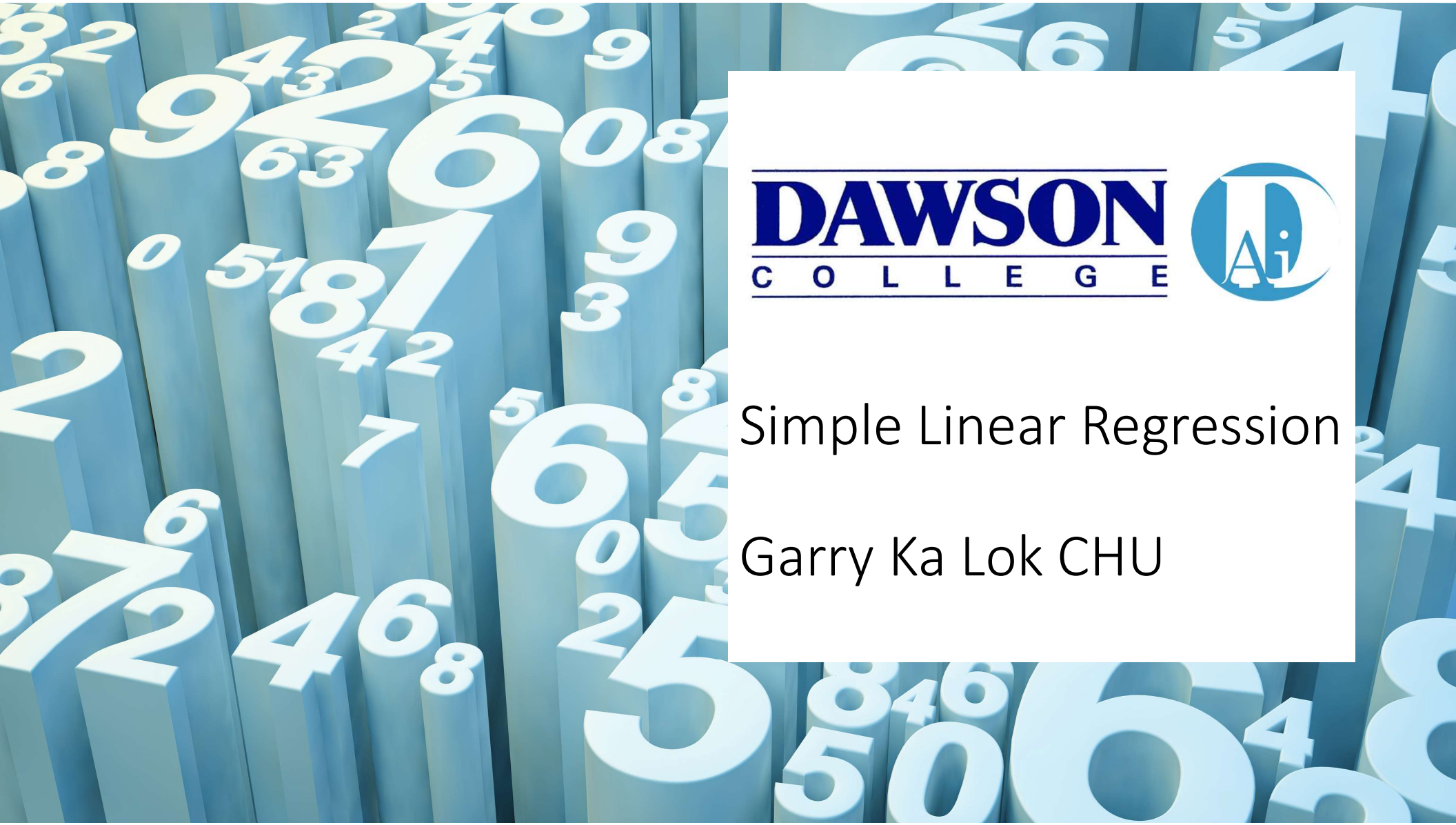


DAWSON COLLEGE

Simple Linear Regression

Garry Ka Lok CHU

Simple Linear Regression

Simple Linear Regression is used to predict the value of a target variable based on the relationship between the input variable and the dependent variable.

Simple Linear Regression

Reference:

S. R. Searle (1971) *Linear Models*, John Wiley & Sons.

Simple Linear Regression

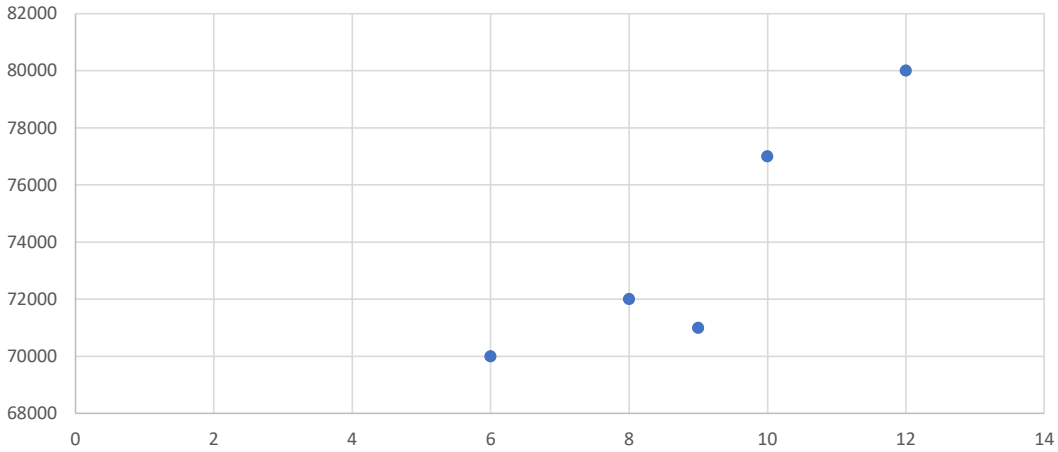
Consider the following dataset:

Person	Annual Income	Years of Schooling
1	70000	6
2	80000	12
3	77000	10
4	72000	8
5	71000	9

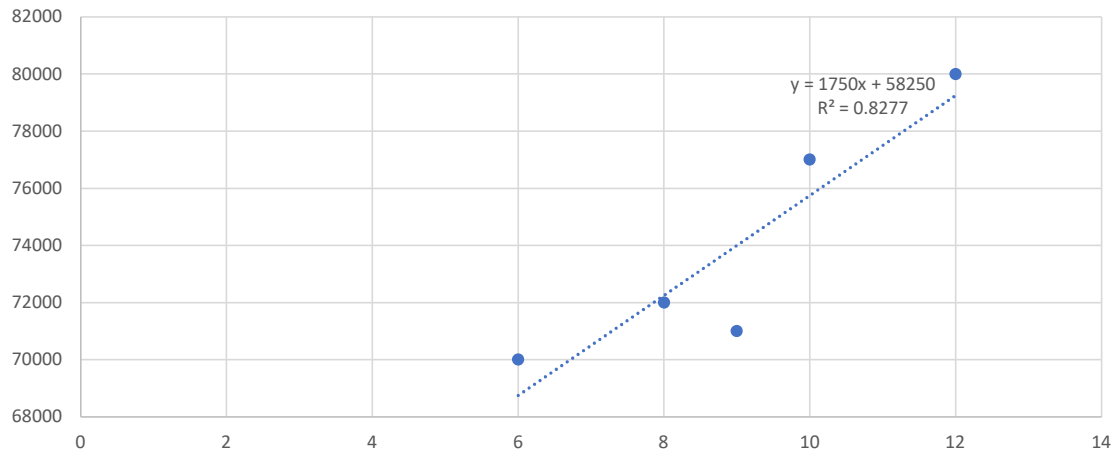
Simple Linear Regression

Use a scatter plot to show the relationship between Annual income (y) and Years of schooling (x)

Person	Annual Income	Years of Schooling
1	70000	6
2	80000	12
3	77000	10
4	72000	8
5	71000	9



Person	Annual Income	Years of Schooling
1	70000	6
2	80000	12
3	77000	10
4	72000	8
5	71000	9



Simple Linear Regression

Use the Trendline function to show the least-squares line. Display the equation of the line and R^2, the coefficient of determination.

Simple Linear Regression

- R, the linear correlation coefficient, measures the strength of the relationship between 2 variables.
- R^2, the coefficient of determination, is a measure to explain how much variability of one factor can be caused by its relationship with the other factor.

Simple Linear Regression

The equation of the least-squares line is used to predict the annual income given the years of schooling.

If the years of schooling is 11, then the annual income is predicted to be:

Y = 1750(11) +58250 = 77500

Simple Linear Regression

Further Questions:

- If the annual income also depends on other factors, may we use regression?
- Other than Linear regression, may we use other relationships to construct the model?

Person	Annual Income	Years of Schooling	Age
1	70000	6	28
2	80000	12	40
3	77000	10	32
4	72000	8	36
5	71000	9	34

Simple Linear Regression

Advanced Topics:

- Multiple Regression
- Nonlinear Regression
- Ridge and Lasso Regression