An Example

Example (patient satisfaction **(data)**) Y : patient satisfaction; X_1 : age; X_2 : severity of illness; X_3 : anxiety level. We try a linear regression model

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$$

We have; see **(code)**

Coefficients:

	Estimate	Std. Error	t value	p-value
(Intercept)	158.4913	18.1259	8.744	5.26e-11
x1	-1.1416	0.2148	-5.315	3.81e-06
x2	-0.4420	0.4920	-0.898	0.3741
x3	-13.4702	7.0997	-1.897	0.0647

Residual standard error: 10.06 on 42 degrees of freedom
Multiple R-squared: 0.6822, Adjusted R-squared: 0.6595
F-statistic: 30.05 on 3 and 42 DF, p-value: 1.542e-10

- The estimated model

$\hat{Y}_i$	=	158.4913	-	1.1416 x_{i1}	-	0.4420 x_{i2}	-	13.4702 x_{i3}
(SE)		(18.1259)		(0.2148)		(0.4920)		(7.0997)

$$R^2 = 0.6822, \hat{\sigma} = 10.06, \text{ F-statistic=30.05}$$

- Test $H_0, \beta_2 = 0$.

  Answer: We accept it with $\alpha = 0.05$

- For patients with $X_1 = 50, X_2 = 50$ and $X_3 = 2$, the 95% CI for their average satisfaction is between

  Answer, the predicted value is 52.37012, the CI is [44.3907, 60.34954]

- Since β_0 is likely to be 0, we can simplify the model to

$$Y = \beta_0 + \beta_1 X_1 + \beta_3 X_3 + \varepsilon$$

Coefficients:

	Estimate	Std. Error	t value	p-value
(Intercept)	145.9412	11.5251	12.663	4.21e-16
x1	-1.2005	0.2041	-5.882	5.43e-07
x3	-16.7421	6.0808	-2.753	0.00861

Residual standard error: 10.04 on 43 degrees of freedom
Multiple R-squared: 0.6761, Adjusted R-squared: 0.661
F-statistic: 44.88 on 2 and 43 DF, p-value: 2.98e-11

- The estimated simplified model

$$\begin{array}{lcccccc} \hat{Y}_i & = & 145.9412 & - & 1.2005\ x_{i1} & - & 16.7421\ x_{i3} \\ (\text{SE}) & & (11.5251) & & (0.2041) & & (0.4920) \end{array}$$

$$R^2 = 0.6761, \hat{\sigma} = 10.04, \text{ F-statistic=44.88}$$

- we can make the prediction based on the refined/simplfied model

  Answer: the predicted value is 52.43355, the CI is [44.47879 60.38831]