

# FitFuel - Linear Regression with SHAP analysis

Number of influencer posts (x)	Bottles sold (y)
1	120
2	150
3	180
1	130
2	160

Linear regression model

$$y = \beta_0 + \beta_1 x$$

$$\bar{x} = \frac{1+2+3+1+2}{5} = 1.8$$

$$\bar{y} = \frac{120+150+180+130+160}{5} = 148$$

$$\beta_1 = \frac{\sum (x - \bar{x})(y_i - \bar{y})}{\sum (x - \bar{x})^2}$$



analysis

$$(1-1.8)(220-148) + (2-1.8)(150-148) + (3-1.8)(180-148) + (1-1.8)(130-148) + (2-1.8)(160-148)$$

$$(1-1.8)^2 + (2-1.8)^2 + (3-1.8)^2 + (1-1.8)^2 + (2-1.8)^2$$

$$\beta_1 = \frac{17.6}{0.56} = 31.43$$

$$\beta_0 = \bar{y} - \beta_1 \bar{x} = 148 - (31.43)(1.8) = 91.43$$

$$y = 91.43 + 31.43x$$

Baseline (mean of y) = 148

Final Predictions & SHAP values.

X	actual y	predicted y
1	120	$91.43 + 31.43 \times 1 = 122.86$
2	150	$91.43 + 31.43 \times 2 = 154.29$
3	180	$91.43 + 31.43 \times 3 = 185.71$



$$1 \quad 130 \quad 91.43 + 31.43 \times 1 = 122.86$$

$$+ (3A) - 0.01 \times (3.1 - 1.0) + (3A) - 0.01 \times (3.1 - 1.0)$$

$$2 \quad 160 \quad 91.43 + 31.43 \times 2 = 154.29$$

$$(3A) - 0.01 \times (3.1 - 1.0)$$

~~Baseline + SHAP Predicted.~~

SHAP value = Predicted - Baseline

$$= 122.86 - 148 = -25.14$$

$$EA \cdot IP = \frac{2 \cdot P1}{22.0} = 18$$

$$EA \cdot IP = (3.1 - 1.0) \times EA \cdot IP = 154.29 - 148 = 6.29$$

$$= 185.71 - 148 = 37.71$$

$$= 122.86 - 148 = -25.14$$

$$= 154.29 - 148 = 6.29$$

SHAP interpretation.

Record 1 & 4 ( $x = 1$ ) = -25.14 Fewer

Posts  $\rightarrow$  prediction pulled down

From baseline



Record 2 & 5 ( $x=2$ ) : SHAP = +6.29.  
moderate increase from baseline.

Record 3 ( $x=3$ ) SHAP: +37.71  $\rightarrow$  more  
posts strong increase from baseline

Prediction vs actual analysis

actual y	predict y	error	under/over
120	122.86	2.86	over
150	154.29	4.29	over
180	185.71	5.71	over
130	122.86	-7.14	under
160	154.29	-5.71	under

slope = +31.43

1  $\rightarrow$  post = 31.43 bottles (on average).

prediction error is low.