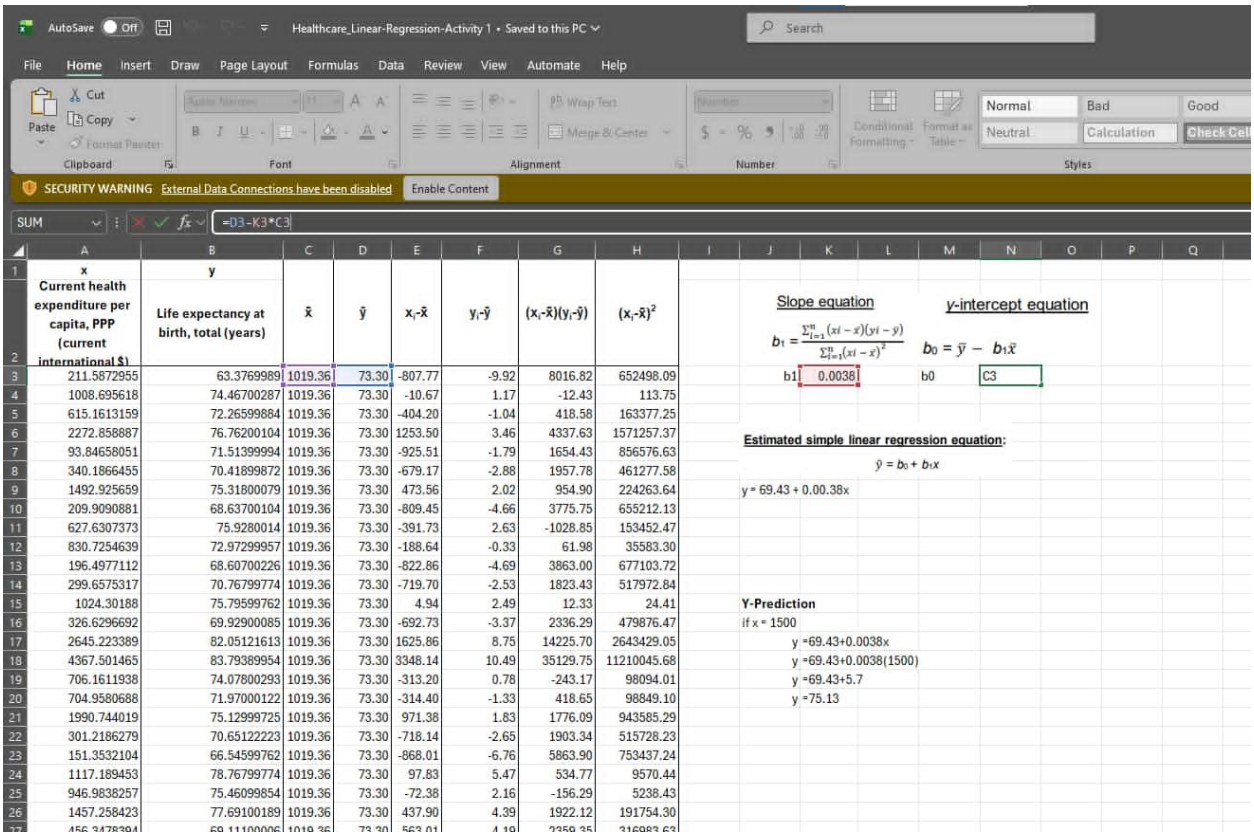

x Current health expenditure per capita, PPP (current international $)	y Life expectancy at birth, total (years)	$\bar{x}$	$\bar{y}$	$x_i-\bar{x}$	$y_i-\bar{y}$	$(x_i-\bar{x})(y_i-\bar{y})$	$(x_i-\bar{x})^2$
211.5872955	63.3769989	1019.36	73.30	-807.77	-9.92	8016.82	652498.09
1008.695618	74.46700287	1019.36	73.30	-10.67	1.17	-12.43	113.75
615.1613159	72.26599884	1019.36	73.30	-404.20	-1.04	418.58	163377.25
2272.858887	76.76200104	1019.36	73.30	1253.50	3.46	4337.63	1571257.37
93.84658051	71.51399994	1019.36	73.30	-925.51	-1.79	1654.43	856576.63
340.1866455	70.41899872	1019.36	73.30	-679.17	-2.88	1957.78	461277.58
1492.925659	75.31800079	1019.36	73.30	473.56	2.02	954.90	224263.64
209.9090881	68.63700104	1019.36	73.30	-809.45	-4.66	3775.75	655212.13
627.6307373	75.9280014	1019.36	73.30	-391.73	2.63	-1028.85	153452.47
830.7254639	72.97299957	1019.36	73.30	-188.64	-0.33	61.98	35583.30
196.4977112	68.60700226	1019.36	73.30	-822.86	-4.69	3863.00	677103.72
299.6575317	70.76799774	1019.36	73.30	-719.70	-2.53	1823.43	517972.84
1024.30188	75.79599762	1019.36	73.30	4.94	2.49	12.33	24.41
326.6296692	69.92900085	1019.36	73.30	-692.73	-3.37	2336.29	479876.47
2645.223389	82.05121613	1019.36	73.30	1625.86	8.75	14225.70	2643429.05
4367.501465	83.79389954	1019.36	73.30	3348.14	10.49	35129.75	11210045.68
706.1611938	74.07800293	1019.36	73.30	-313.20	0.78	-243.17	98094.01
704.9580688	71.97000122	1019.36	73.30	-314.40	-1.33	418.65	98849.10
1990.744019	75.12999725	1019.36	73.30	971.38	1.83	1776.09	943585.29
301.2186279	70.65122223	1019.36	73.30	-718.14	-2.65	1903.34	515728.23
151.3532104	66.54599762	1019.36	73.30	-868.01	-6.76	5863.90	753437.24
1117.189453	78.76799774	1019.36	73.30	97.83	5.47	534.77	9570.44
946.9838257	75.46099854	1019.36	73.30	-72.38	2.16	-156.29	5238.43
1457.258423	77.69100189	1019.36	73.30	437.90	4.39	1922.12	191754.30
456.3478394	69.11100006	1019.36	73.30	-562.01	-4.19	2359.35	316983.63

Slope Equation

$$b1 = \frac{(xi-\bar{x})(yi-\bar{y})}{(x_i-\bar{x})^2}$$

Y-intercept equation

$$b0 = \bar{y} - (b1)(\bar{x})$$

$$= \text{y Mean - Slope*x Mean}$$

Least squares estimates of the regression parameters.

Slope equation	*y*-intercept equation
$b_1 = \frac{\sum_{i=1}^{n}(xi-\bar{x})(yi-\bar{y})}{\sum_{i=1}^{n}(xi-\bar{x})^2}$	$b_0 = \bar{y} - b_1\bar{x}$

x_i = value of the independent variable for the *i*th observation

y_i = value of the dependent variable for the *i*th observation

$\bar{x}$ = mean value for the independent variable

$\bar{y}$ = mean value for the dependent variable

n = total number of observations

- Sum of squares due to regression, SSR:

  $$SSR = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$$

  Measures how much the $\hat{y}$ values on the estimated regression line deviate from $\bar{y}$.

- Relation between SST, SSR, and SSE

  $$SST = SSR + SSE$$

The **coefficient of determination**: The ratio SSR/SST used to evaluate the goodness of fit for the estimated regression equation.

$$r^2 = \frac{SSR}{SST}$$

Take values between zero and one.

Interpreted as the percentage of the total sum of squares that can be explained by using the estimated regression equation.

Square of the correlation between the y_i and $\hat{y}_i$.

Referred to as the simple coefficient of determination in simple regression.

For the Butler Trucking Company example, the coefficient of determination,

$$r^2 = \frac{SSR}{SST} = \frac{15.8712}{23.9} = 0.6641.$$

It can be concluded that 66.41 percent of the total sum of squares can be explained by using the estimated regression equation $\hat{y}_i = 1.2739 + 0.0678x_i$ to predict quarterly sales.

Estimated simple linear regression equation:

$$\hat{y} = b_0 + b_1x$$

$\hat{y}$ = Point estimator of $E(y|x)$

b_0 = Estimated y-intercept

b_1 = Estimated slope

The graph of the estimated simple linear regression equation is called the estimated regression line.

- **Dependent variable** or response: Variable being predicted
- **Independent variables** or predictor variables: Variables being used to predict the value of the dependent variable.
- **Simple regression**: A regression analysis involving one independent variable and one dependent variable.

  In statistical notation:

  y = dependent variable

  x = independent variable

  Linear regression: A regression analysis for which any one unit change in the independent variable, x, is assumed to result in the same change in the dependent variable, y.

  Multiple regression: Regression analysis involving two or more independent variables.

Y-Prediction			
if x = 1500			
	y	=69.43+0.0038x	
	y	=69.43+0.0038(1500)	
	y	=69.43+5.7	
	y	=75.13	
	n	Observations	44
r2		R Square	0.6559129
SSR		Regression	677.55437
SSE		Residual	355.44004
SST		Total	1032.9944
Y-intercept		Intercept	69.433382
slope		X Variable 1	0.0037947