1a.) $\hat{\beta}_2 = \dfrac{\Sigma(X_i-\bar{X})(Y_i-\bar{Y})}{\Sigma(X_i-\bar{X})^2} = \dfrac{20.58}{211}$

$= 0.0975$ #

$\bar{X} = 21.5556$, $\bar{Y} = 2.8278$

$\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2\bar{X} = 2.8278 - (0.0975)(21.5556)$

$= 0.7261$ #

$\therefore$ SRF is $\hat{Y}_i = 0.7261 + 0.0975X$, so the intercept is 0.7261 and the slope is 0.0975.

1b.) $r^2 = 1 - \dfrac{\Sigma\hat{u}_i^2}{\Sigma(Y_i-\bar{Y})^2} = 1 - \dfrac{0.5781}{2.5844}$

$= 0.7763$ #

r^2 is a measurement of goodness of fit, 77.63% of Y is explained by X.

1c.) SRF ; $\hat{Y}_i = 0.7261 + 0.0975X$, $X = 30$

$\hat{Y}_i = 0.7261 + 0.0975(30)$

$= 3.6511$ #

The average of $\hat{Y}_i$ is 3.6511 when $X = 30$.

1d.) $var(u_i) = \dfrac{\Sigma\hat{u}_i^2}{n-k} = \dfrac{0.5781}{18-2} = 0.0361$ #

$var(\hat{\beta}_1) = \dfrac{\Sigma X_i^2}{n\Sigma(X_i-\bar{X})^2}\hat{\sigma}^2 = \dfrac{9620}{18 \times 211}(0.0361)$

$= 0.0914$ #

$var(\hat{\beta}_2) = \dfrac{\hat{\sigma}^2}{\Sigma(X_i-\bar{X})^2} = \dfrac{0.0361}{211} = 0.0002$ #

1e.) $P_r\left[\hat{\beta}_2 - (t_{\frac{\alpha}{2}} \cdot se_{\hat{\beta}_2}) \le \beta_2 \le \hat{\beta}_2 + (t_{\frac{\alpha}{2}} \cdot se_{\hat{\beta}_2})\right]$, 90% CI

$\alpha = 0.1$, $se_{\hat{\beta}_2} = \sqrt{0.0002} = 0.0141$, $t_{\frac{0.1}{2}} = t_{0.05} = 1.746$

$P_r[\,0.0975 - (1.746 \times 0.0141) \le \beta_2 \le 0.0975 + (1.746 \times 0.0141)\,] = 0.90$

$P_r(\,0.0729 \le \beta_2 \le 0.1221\,) = 0.90$

It means that 90% of all probabilities, real β_2 will be in the confident interval.

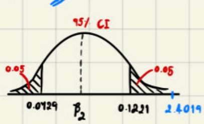
1f.) $H_0 : \beta_2 = 0$ | $se_{\hat{\beta}_2} = \sqrt{0.0914} = 0.3023$

$H_1 : \beta_2 \ne 0$

- $t_{cal} = \dfrac{\hat{\beta}_2 - \beta_2}{se_{\hat{\beta}_2}} = \dfrac{0.7261 - 0}{0.3023} = 2.4019$ //

- upper = 0.1221

  lower = 0.0729



Conclusion : Since $t_{cal} = 2.4019$, greater than 0.1221.
So there is enough evidence claim to support that β_2 is not zero at $\alpha = 0.1$.
$\therefore$ Reject $H_0 : \beta_2 = 0$.

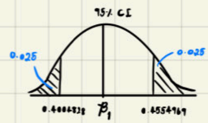
2a) $H_0 : \beta_1 = 0$ $H_1 : \beta_1 \ne 0$; $\alpha = 0.05$, $t_{\frac{0.05}{2},\infty} = 1.96$

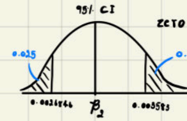
- $t_{cal} = \dfrac{\hat{\beta}_1 - \beta_1}{se_{\hat{\beta}_1}} = \dfrac{0.4279898 - 0}{0.0140339} = 30.4969$ //

$H_0 : \beta_2 = 0$ $H_1 : \beta_2 \ne 0$

- $t_{cal} = \dfrac{\hat{\beta}_2 - \beta_2}{se_{\hat{\beta}_2}} = \dfrac{0.0031338 - 0}{0.0002292} = 13.6728$ //

$\therefore$ At $\alpha = 0.05$, β_1 & β_2 are not zero. So we reject H_0.





2b.) From the results, we would expect that people will increase their visit per year for 0.0031338 time as their ages increase 1 year, this makes economic sense that their health will become weaker when they get older.

2c.) $\widehat{\ln outp_i} = \hat{\beta}_1 + \hat{\beta}_2 age_i$; approx. measure : $0.0031338 \times 100 = 0.31338$ #

If people get older 1 year, it can be expected that their visit will increase by 0.31338 %.

2d.) Coefficient $= 0.0031338 \times 10 = 0.031338$ #

Std. err. $= 0.0002292 \times 10 = 0.002292$ #

Confident Interval $= (\,0.0026846 \times 10\,,\, 0.003583 \times 10\,)$

$= (\,0.026846\,,\, 0.03583\,)$ #

It increases the coefficient, se, and CI by 10 times of its old value.

2e.) $\hat{Y}_0 = 0.4279898 + 0.0031338(50) = 0.5847$

$se_{\hat{Y}_0} = \sqrt{0.00002} = 0.0045$

$t_{\frac{0.01}{2}} = 2.576$

$P_r[\,\hat{Y}_0 - (t_{0.005} \times 0.0045) \le Y_0 \le \hat{Y}_0 + (t_{0.005} \times 0.0045)\,] = 0.99$

$P_r[\,0.5847 - (2.576 \times 0.0045) \le Y_0 \le 0.5847 + (2.576 \times 0.0045)\,] = 0.99$

$P_r(\,0.5731 \le Y_0 \le 0.5963\,) = 0.99$ #

3.) $var(\hat{Y}_0) = \sigma^2\left[\dfrac{1}{n} + \dfrac{(X_0-\bar{X})^2}{\Sigma(X_i-\bar{X})^2}\right]$

As we can see from the equation, the larger X_0 is, the bigger gap from $\bar{X}$. So the variance will be higher, meaning that $se_{\hat{Y}_0}$ will also be higher.