1a) find $\hat{\beta}_2 = \dfrac{\sum (x_i-\bar{x})(y_i-\bar{y})}{\sum (x_i-\bar{x})^2}$

$= \dfrac{46131.6183}{23153.3861}$

$= 1.9924$

find $\hat{\beta}_1 = \bar{y} - \hat{\beta}_2\bar{x}$

$= 69.1478 - (1.9924)(86.0826)$

$= -102.3632$

SRF line will be intercept y at -102.3632 due to $\hat{\beta}_1$ is intercept of SRF and y axis

$\hat{\beta}_2$ is a slope of SRF so if x increase by 1 unit y will increase by 1.9924 unit

1b) $r^2 = 1 - \dfrac{\sum \hat{u}_i^2}{\sum (y_i-\bar{y})^2}$

$= 1 - \dfrac{2610.9211}{94525.1748}$

$= 0.9724$

r^2 is how much SRF fitted the data and x_i explain about 97.24 % of the variation in y_i. If it get near 1, the line will fitted the data

1c) $\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i$

$= -102.3632 + 1.9924(60)$

$= 17.1808$

1d) $var(u_i) = \sigma^2 = \dfrac{\sum \hat{u}_i^2}{n-k}$

$= \dfrac{2610.9211}{46-2}$

$= 59.3391$

$$\sigma^2_{\hat{\beta}_2} = \frac{\sigma^2}{\sum x_i^2} = \frac{59.3391}{364023.30} = 0.000163$$

$$\sigma^2_{\hat{\beta}_1} = \frac{\sum X_i^2}{n \sum x_i^2}\,\sigma^2 = \frac{364023.30}{46\,(23153.3861)}\,(59.3391) = 20.2814$$

1e $CI = 95\% = 1-\alpha$, $\alpha = 0.05$, $\frac{\alpha}{2} = 0.025$, $d.f = 44$

find lower and upper limit

$$= \hat{\beta}_2 \pm t_{\frac{\alpha}{2}} \cdot \sigma_{\hat{\beta}_2}$$

$$= 1.9924 \pm (2.021)(0.0128)$$

lower $= 1.9665$
upper $= 2.0183$

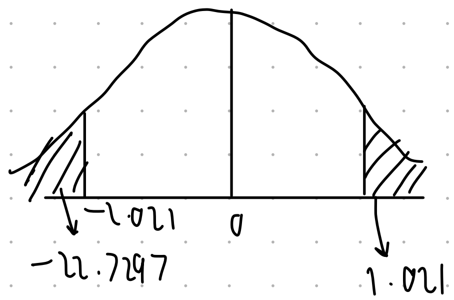
$\therefore P(1.9665 \leq \beta_2 \leq 2.0183) = 0.95$

when the CI is 95%, the value of β_2 will be between 1.9665 (lower bound) and 2.0183 (upper bound)

1f) $\alpha = 0.05$, $\frac{\alpha}{2} = 0.025$

find β_1

1. $H_0 : \beta_1 = 0 \rightarrow$ null hypothesis
   $H_a : \beta_1 \neq 0$
2. $$t_{cal} = \frac{\hat{\beta}_1 - \beta_1}{\sigma_{\hat{\beta}_1}} = \frac{-102.3632 - 0}{4.5035} = -22.7297$$
3. $d.f. = 46 - 2 = 44$
   $t_{\frac{\alpha}{2}} = \pm 2.021$



4. $t_{cal} = -22.7297$ which falls into the rejection area so we reject null hypothesis. we are 95% sure that β_1 is not 0

2a) yes because the sample regression function will be plot to fitted with each data point the most. So if there is only 1 data point the SRF line will pass through the point and make no error term which is the distance between line and the data point.

2b.) β_2 is the true value of the slope of the PRF line, which tell us how much y increase or decrease if x change by 1 unit so β_2 is sufficient enough to tell the relationship between x and y

2c) β_2 significiant different from 0 means the result of t cal fall in the rejection area. So we reject the null hypothesis (Ho) which state that $\beta_2 = 0$. The result will tell how much percentage β_2 will not be 0.

2d) point estimation is when we selected one point in the sample to represent a point in the population but it can not tell the overall estimate of the population value

3a) $lwage = 7.6581 + 0.0318 (main\_hr)$

$lwage = 7.6581$

$wage = 2117.7299$

nominal wage for person who work 0 hour / week is 2117.7299

3b) $lwage = 7.6581 + 0.0318 (main\_hrs)$

$$\frac{d\, lwage}{d\, main\text{-}hr} = 0.0318$$

$$\frac{d\, \hat{wage}}{wage} = 0.0318\, d\, main\text{-}hr$$

$$\%\, \Delta \widehat{wage} = 0.0318\, d\, main\text{-}hr \times 100$$

$$= 3.18\%$$

if person work 1 more hour, we expect wage to increase by 3.18%

3c) $l\hat{wage} = 7.6581 + 0.7632 (main\_days)$

$se = (0.1256)(0.0795)$

the value that will change in main_hr row is Coef and sdd err (SE)

- Coef will change from 0.0318 to 0.7632
- SE will change from 0.003312 to 0.0795