Name: KEY

1. Below is the R-output from two linear model fits of the number of rideshare trips on a given day (`trips`), to the mean temperature (`temp`) and a binary indicator of rain (`rain`), where `rain=1` indicates rain and `rain=0` indicates no rain.

```
#### ----

lm(formula = trips ~ rain + temp)

             Estimate Std. Error t value Pr(>|t|)
(Intercept) -180.0397    28.6988  -6.273 1.01e-09 ***
rain        -120.2663    10.0154 -12.008  < 2e-16 ***
temp          10.6889     0.4699  22.749  < 2e-16 ***
---
Residual standard error: 91.44 on 362 degrees of freedom
Multiple R-squared:  0.6917,Adjusted R-squared:   0.69

#### ----

lm(formula = trips ~ rain + temp + rain:temp)

             Estimate Std. Error t value Pr(>|t|)
(Intercept) -170.9346    31.6835  -5.395 1.24e-07 ***
rain        -166.3996    68.5258  -2.428   0.0157 *
temp          10.5363     0.5209  20.225  < 2e-16 ***
rain:temp      0.8236     1.2102   0.681   0.4966
---
Residual standard error: 91.5 on 361 degrees of freedom
Multiple R-squared:  0.6921,Adjusted R-squared:  0.6895

#### ----
```

Based on this output, give the quantities below and answer the questions. You may leave answers in numerical form but without doing all of the arithmetic (such as $\hat{\theta} = \frac{\log 4.32}{\sqrt{.74}}$).

(a) The sample size n: 1

$n = dof + p$

$= 362 + 3$

$= 361 + 4$

$= \underline{365}$

(b) Under the first model, the estimated mean of `trips` as a function of `temp` when `rain = 0`, and the estimated mean of `trips` as a function of `temp` when `rain = 1`.

$\widehat{E[\text{trips}|\text{rain}=0, \text{temp}]} \approx -180.04 + 10.69 \times \text{temp}$

$\widehat{E[\text{trips}|\text{rain}=1, \text{temp}]} \approx -180 - 120 + 10.69 \times \text{temp}$

(c) Under the second model, the estimated mean of `trips` as a function of `temp` when `rain = 0`, and the estimated mean of `trips` as a function of `temp` when `rain = 1`.

$\widehat{E[t|r=0, \text{temp}]} : -170.93 + 10.54 \times \text{temp}$

$\widehat{E[t|r=1, \text{temp}]} \approx -170 - 166 + (10.54 + 0.82)$

(d) An estimate of σ^2, under the assumption that the effect of `temp` does not depend on `rain`.

91.44^2

(e) Let a_{r1} be the linear increase in `trips` per unit `temp` when it is raining, and let a_{r0} be the linear increase in `trips` per unit `temp` when it is not raining. Find a 95% CI for $a_{r1} - a_{r0}$.

$\hat{a}_{r1} - \hat{a}_{r0}$ = difference in slopes : $\underline{0.8236}$

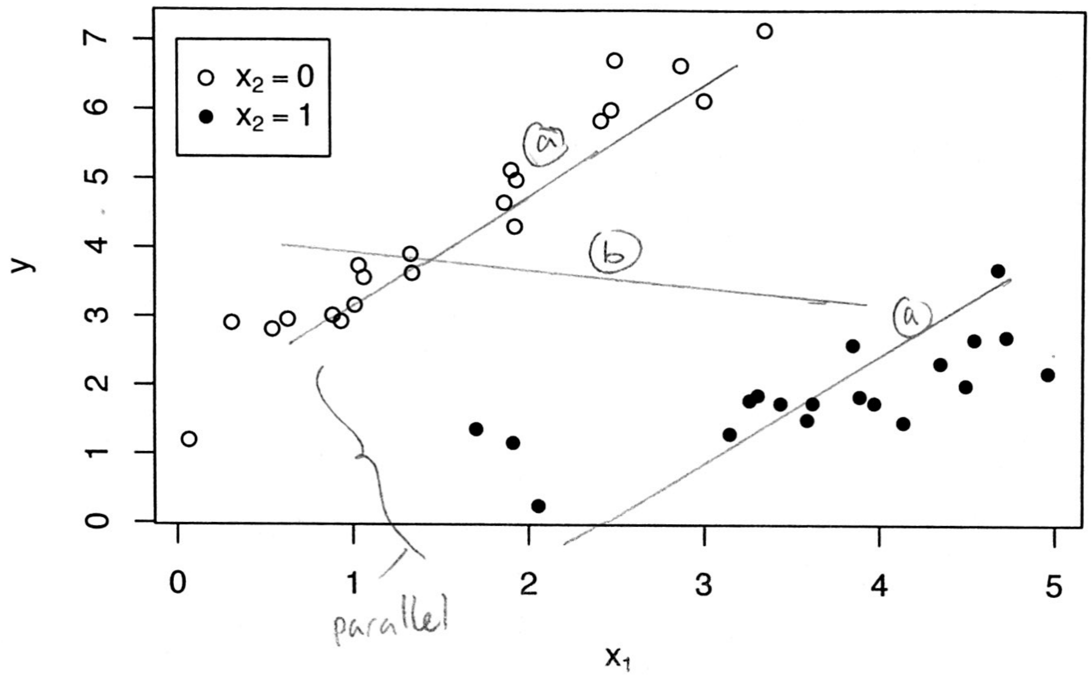
95% CI is $\approx 0.8236 \pm 2 \times 1.2102$

(f) Is there evidence that the effect of `temp` on `trips` depends on whether or not it is raining? Explain your answer.

No, the t-stat for rain:temp is small

p-val. big.

2. The figure below displays data on an outcome y measured under a variety of values of a continuous variable x_1 and two conditions ($x_2 = 0$ and $x_2 = 1$).



(a) Sketch and label on the graph a graphical representation of the estimated mean function for y from the OLS fit to the model `y ~ x1 + x2`.

(b) Sketch and label on the graph a graphical representation of the estimated mean function for y from the OLS fit to the model `y ~ x1`.

(c) Explain why the mean functions in (a) and (b) are similar or different.

The slope in (b) is very different from the slopes in (a) because of the correlation between x_1 and x_2. Removal of x_2 from the model changes the estimated effect of x_1.

(d) Which do you think will be bigger, the RSS under y ~ x1 or the RSS under y ~ x1 + x2? Explain your answer.

RSS always decreases (or in rare cases, stays the same) after adding a term to the model. Therefore

$$RSS(y \sim x_1) \geq RSS(y \sim x_1 + x_2)$$

3. Suppose person 1 runs a regression of a vector $\mathbf{y}$ on two predictors $\mathbf{x}_1, \mathbf{x}_2$ while person 2 runs a regression of $\mathbf{y}$ on $\mathbf{w}_1, \mathbf{w}_2$, where $\mathbf{w}_1 = a\mathbf{x}_1 + b\mathbf{x}_2$ and $\mathbf{w}_2 = c\mathbf{x}_1 + d\mathbf{x}_2$. Both persons include an intercept in their regression model.

(a) Let $\mathbf{X}$ and $\mathbf{W}$ be the design matrices for persons 1 and 2 respectively. Find a 3×3 matrix $\mathbf{G}$ so that the vector $(1, w_{i1}, w_{i2})$ is equal to the vector $(1, x_{i1}, x_{i2})\mathbf{G}$. From this, deduce that $\mathbf{W}$ can be written as $\mathbf{W} = \mathbf{XG}$.

$$(1, w_1, w_2) = (1, ax_1 + bx_2, cx_1 + dx_2)$$

$$= (1, x_1, x_2)\begin{pmatrix} 1 & 0 & 0 \\ 0 & a & c \\ 0 & b & d \end{pmatrix} = (1, x_1, x_2)G$$

$$W = \begin{pmatrix} 1 & w_{11} & w_{12} \\ \vdots & & \\ 1 & w_{n1} & w_{n2} \end{pmatrix} = \begin{pmatrix} (1, x_{11}, x_{12})G \\ \vdots \\ (1, x_{n1}, x_{n2})G \end{pmatrix}$$

$$= XG$$

(b) Recall from the homework that the fitted values for the first person can be written as $\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$. Use this fact to show that the RSS for person 1 and the RSS for person 2 are the same (Hint: Use the fact from matrix algebra that $(\mathbf{A}^T\mathbf{B}^T\mathbf{B}\mathbf{A})^{-1} = \mathbf{A}^{-1}(\mathbf{B}^T\mathbf{B})^{-1}(\mathbf{A}^T)^{-1}$).

RSS will depend on $\hat{Y}$

$$\hat{Y}_{(1)} = X(X'X)^{-1}X^T Y$$

$$\hat{Y}_{(2)} = W(W^TW)^{-1}W^T Y$$

$$= XG(G^TX^TXG)^{-1}G^TX^T Y$$

$$= X^T GG^{-1}(X^TX)^{-1}G^TG^{-T}X^T Y$$

$$= X(X^TX)^{-1}X^T Y$$

$$= \hat{Y}_{(1)} \implies \text{same fitted values} \implies \text{same RSS.}$$