Concepts:

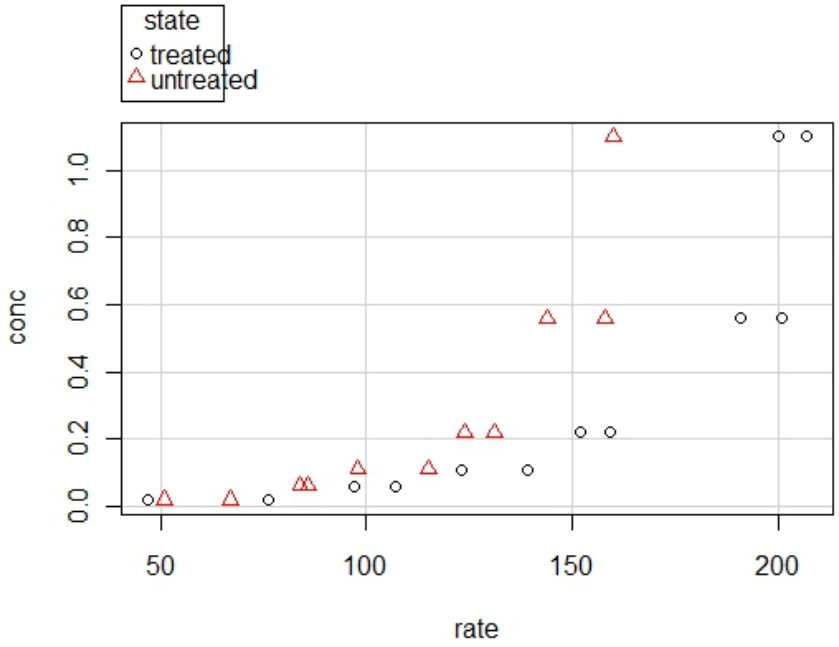
1. () What does it mean to be 95% confident?
2. () What is the purpose of inferential statistics?
3. () What is measured by the P-value?
4. () What is a dummy variable?
5. () What does it mean in mathematical and laymen terms for an estimator to be unbiased?
6. () What are the shortcomings of R^2 as a measure of fit in the MLRM?
7. () What is the hat matrix? Why this name?
8. () What is the difference between $\vec{Y}$ and $\hat{Y}$?
9. () Explain the difference between a prediction interval and a confidence interval for a mean in the MLRM
10. () Prove $E(SS_{Res}) = (n-p)\sigma^2$.
11. () Prove that in the MLRM $\text{Cov}(\hat{\beta}) = \sigma^2(X'X)^{-1}$.
12. () Prove $E(\vec{h}'\vec{Y}) = \vec{h}'E(\vec{Y})$.

Calculations + Interpretations:

13. Data on the velocity of an enzymatic reaction were obtained by Treloar (1974). The number of counts per minute of radioactive product from the reaction was measured as a function of substrate concentration in parts per million (ppm) and from these counts the initial rate (or velocity) of the reaction was calculated (counts/min/min). The experiment was conducted once with the enzyme treated with Puromycin, and once with the enzyme untreated.

    First five rows of data set

	conc	rate	state
1	0.02	76	treated
2	0.02	47	treated
3	0.06	97	treated
4	0.06	107	treated
5	0.11	123	treated



$$(X'X)^{-1} = \begin{pmatrix} 0.15 & -0.20 & -0.15 & 0.20 \\ -0.20 & 0.57 & 0.20 & -0.57 \\ -0.15 & 0.20 & 0.31 & -0.44 \\ 0.20 & -0.57 & -0.44 & 1.46 \end{pmatrix}$$

```
            Coefficients:
                        Estimate Std. Error t value Pr(>|t|)
(Intercept)               103.49      10.53   9.832 6.91e-09 ***
conc                      110.42      20.46   5.397 3.30e-05 ***
stateuntreated            -17.45      15.06  -1.158    0.261
conc:stateuntreated       -21.08      32.69  -0.645    0.527
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 27.05 on 19 degrees of freedom
Multiple R-squared: 0.7201,Adjusted R-squared: 0.6759
F-statistic: 16.29 on 3 and 19 DF,  p-value: 1.742e-05
```

(a) State a MLRM relating the growth rate with the concentration and state (treated versus untreated).

(b) Give the first row of the design matrix to be used in the analysis.

(c) Conduct a model utility test.

(d) Interpret all estimates for the parameters of the model.

(e) Find a confidence interval for β_1.

(f) Find a prediction interval for the rate when the concentration is 100 and the enzyme was treated with Puromycin. Interpret.

(g) Researchers were interested if an increase in the concentration has a positive effect on the growth rate, when correcting for the state. Conduct the appropriate test.

(h) To test if the state has a significant effect on the rate one could do and extra sum of squares test. Outline the procedure completely (state the models and explain how the test statistic can be found from there, include the degree of freedom).

(i) Does the scatterplot support the chosen model?