

Assignment 3 is due no later than 5pm Friday, 28th of May, 2021. In submitting your work, you are consenting that it may be copied and transmitted by the University for the detection of plagiarism. Submission is your guarantee that the below statement of originality is correct.

“This is my own work. I have not copied any of it from anyone else.”

NAME: **Replace this text with your name.**

STUDENT NUMBER: **Replace this text with your student number.**

Instructions for assignment: You will need to submit two documents for this assignment. The first document is a pdf document named `Assign3_StNo.pdf` which will provide all your analysis and solutions for this assignment. To produce this pdf document you will need to use LaTeX. The LaTeX document which was used to produce this assignment is named `Assign3_StNo.tex` and is located in the Assignment 3 folder in the Topic 9 section of LMS. You can use LaTeX online via Overleaf which is a website dedicated to producing documents from LaTeX. To use LaTeX, follow the instructions in the `Overleaf.pdf` document located in the Assignment 3 folder. The second document that you will need to submit is an R document named `Assign3_R_StNo.R` which is located in the Assignment 3 folder. This document should provide the R code you used to perform all your data manipulation and analysis.

Assessment information for assignment: There are a total of 80 marks for this assignment.

Description of assignment: The data presented in this assignment is adapted from Ocampo (2005)¹ and is stored in the file named `Veneer.csv` located in the Assignment 3 folder in the Topic 9 section of LMS. A study was conducted on fifty-five adult teeth to examine the impact of veneer placement treatment on gingival (gum) health. The variables of interest for Assignment 3 are:

- *Tooth*: This is a factor variable that identifies the tooth.
- *MA*: This is a continuous variable that measures the age of the tooth less the mean tooth age (where the mean tooth age is calculated by averaging over the teeth in the sample).
- *MBG*: This is a continuous variable that measures the pre-treatment gingival crevicular fluid of a tooth less the mean pre-treatment gingival crevicular fluid of teeth (where the mean pre-treatment gingival crevicular fluid of teeth is calculated by averaging over the teeth in the sample).
- *MC*: This is a continuous variable that measures the post-treatment contour difference of a tooth less the mean post-treatment contour difference of teeth (where the mean post-treatment contour difference of teeth is calculated by averaging over the teeth in the sample).
- *T*: This is a factor variable that measures two post-treatment time points. It has two levels (0 = 3-months post-treatment, 1 = 6-months post-treatment).
- *G*: This is the response variable. It's a continuous variable that measures the gingival crevicular fluid of a tooth at 3-months and 6-months post-treatment.

2 marks are allocated for each question that requires the use of the R computer package. These marks are awarded using the following criterion:

- 1. R code that accurately produces the analysis/output required in the question.**

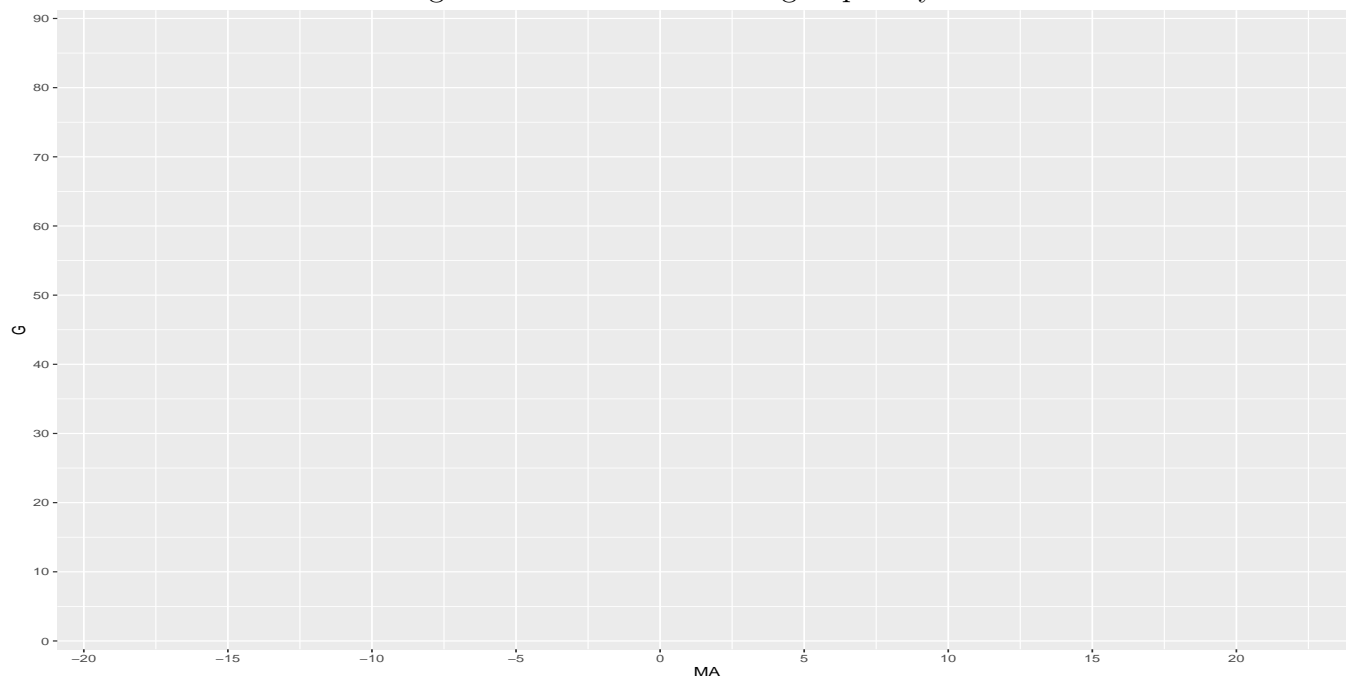
Answer the following questions.

¹OCAMPO, J. (2005). Effect of porcelain laminate contour on gingival inflammation. Master's thesis, University of Michigan School of Dentistry.

1 Graphical analysis

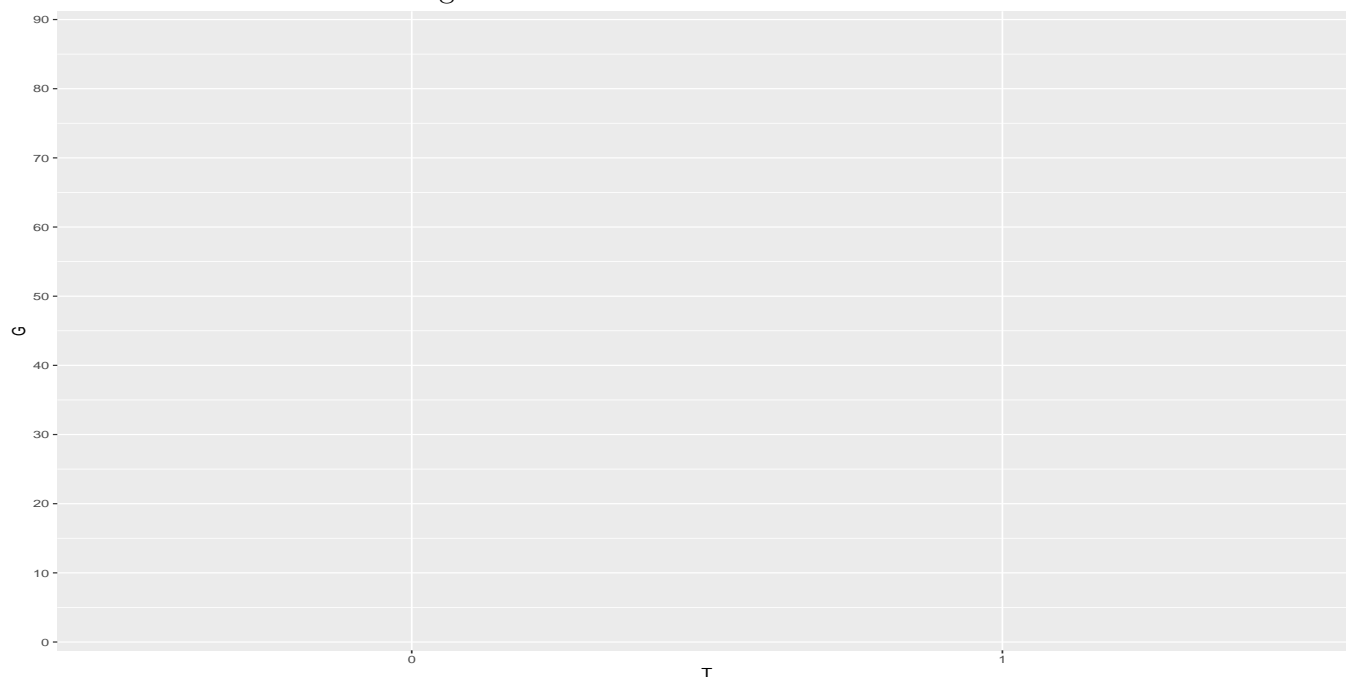
1. Use the R computer package to produce a plot of G vs MA grouped by T . The scale of the vertical and horizontal axes of your figure should be identical to Figure 1 below. (Note that the categories of the T variable need to be represented by colours in Figure 1. Also, for each category of T , you must include a line of best fit which represents the linear relationship between G and MA . Please refer to Figure 3 in the Topics 7 and 8 lectures as a guide.) **(2 marks)**. Do you think that there is an interaction effect between MA and T ? Explain. **(2 marks)**.

Figure 1: Plot of G vs MA grouped by T



2. Use the R computer package to produce a plot of the values of G vs T for each *Tooth*. The scale of the vertical and horizontal axes of your figure should be identical to Figure 2 below **(2 marks)**. Do you think that the random intercept should be included in the linear mixed model? Explain. **(2 marks)** Do you think random effect of T on G should be included in the linear mixed model? Explain. **(2 marks)**

Figure 2: Plot of G vs T for each *Tooth*



2 Describing the model

The researchers in the study set up the following linear mixed model to analyze their research questions.

$$\begin{aligned} G_{ti} = & \beta_0 + \beta_1 MA_i + \beta_2 MBG_i + \beta_3 MC_i + \beta_4 T_{ti} \\ & + \beta_5 MA_i \times T_{ti} + \beta_6 MBG_i \times T_{ti} + \beta_7 MC_i \times T_{ti} \\ & + \mu_{0i} + \mu_{1i} T_{ti} + \varepsilon_{ti}, \end{aligned} \tag{1}$$

- where G_{ti} is the gingival crevicular fluid of for tooth i ($i = 1, 2, \dots, 55$) at occasion t ($t = 1, 2$),
- MA_i is the age of tooth i less the mean tooth age,
- MBG_i is the pre-treatment gingival crevicular fluid of tooth i less the mean pre-treatment gingival crevicular fluid of teeth,
- MC_i is the post-treatment contour difference of tooth i less the mean post-treatment contour difference of teeth,
- $T_{ti} = 1$ if the post-treatment time at occasion t for tooth i is 6 months, and $T_{ti} = 0$ if the post-treatment time at occasion t for tooth i is 3 months,
- β_0 is the fixed intercept,
- $\beta_1, \beta_2, \beta_3$ and β_4 are the fixed simple effects of MA , MBG , MC and T respectively,
- β_5, β_6 and β_7 are the fixed two-way interaction effects of $MA \times T$, $MBG \times T$ and $MC \times T$ respectively,
- μ_{0i} is the random intercept specific to tooth i ,
- μ_{1i} is the random effect of T on G specific to tooth i ,
- ε_{ti} is the random error associated with measuring G for tooth i , on occasion t .
- For model (1), the researchers choose an *unstructured* structure for the variance-covariance matrix of the random effect vector, $\boldsymbol{\mu}_i$. That is, the variance-covariance matrix of the random effect vector, $\boldsymbol{\mu}_i$, is

$$\mathbf{D} = \begin{bmatrix} \theta_0 & \theta_{01} \\ \theta_{01} & \theta_1 \end{bmatrix},$$

- where θ_0 and θ_1 denotes the variance of the random effects μ_{0i} and μ_{1i} , respectively,
- θ_{01} denotes the covariance between the random effects μ_{0i} and μ_{1i} .
- Also for model (1), the researchers choose the following structure for the variance-covariance matrix of the random error vector, $\boldsymbol{\varepsilon}_i$,

$$\mathbf{R} = \begin{bmatrix} \tau & 0 \\ 0 & \tau \end{bmatrix}$$

- where $\tau = Var(\varepsilon_{1i}) = Var(\varepsilon_{2i})$.

3. The researchers would like to express model (1) in matrix form, $\mathbf{Y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\boldsymbol{\mu}_i + \boldsymbol{\varepsilon}_i$, where \mathbf{Y}_i represents the response vector for tooth i , \mathbf{X}_i represents a matrix, for tooth i , that contains the values of the predictors associated with the fixed effects of model (1), $\boldsymbol{\beta}$ is the fixed effect vector, \mathbf{Z}_i is a matrix, for tooth i , that contains the values of the predictors associated with the random effects of model (1), $\boldsymbol{\mu}_i$ is the random effect vector for tooth i and $\boldsymbol{\varepsilon}_i$ is the random error vector for tooth i . Answer the following questions.
 - (a) Write down the observed response vector, \mathbf{Y}_i , of model (1), for tooth $i = 7$ (the 7th tooth in the data set). **(2 marks)**
 - (b) Write down the observed matrix, \mathbf{X}_i , of model (1), for tooth $i = 7$ (the 7th tooth in the data set). **(4 marks)**
 - (c) Write down the fixed effect vector, $\boldsymbol{\beta}$, of model (1). **(1 mark)**
 - (d) Write down the matrix, \mathbf{Z}_i , of model (1), for tooth i . **(2 marks)**
 - (e) Write down the random effect vector, $\boldsymbol{\mu}_i$, of model (1), for tooth i . **(2 marks)**
 - (f) Write down the random error vector, $\boldsymbol{\varepsilon}_i$, of model (1), for tooth i . **(2 marks)**
 - (g) For model (1), derive the variance-covariance matrix of the response vector, \mathbf{Y}_i , for tooth i . Show all workings. **(5 marks)**
4. Interpret the intercept β_0 . **(3 marks)**
5. Interpret the simple effects β_1 , β_2 , β_3 and β_4 . **(12 marks)**
6. Interpret the interaction effects β_5 , β_6 and β_7 . **(9 marks)**

3 Variance-covariance estimates of the final linear mixed model

As their final linear mixed model the researchers choose model (1) which has variance-covariance matrices, \mathbf{D} and \mathbf{R} , defined in section 2. Use this model to answer the questions in this section and in section 4.

7. Use the R computer package to calculate the estimate of the \mathbf{R} matrix of the final linear mixed model. Present this estimate below. Note, round all the elements in the matrix to two decimal places. **(2 marks)**
8. Use the R computer package to calculate the estimate of the variance-covariance matrix of the response vector of the final linear mixed model. Note, round all the elements in the matrix to two decimal places. **(2 marks)**
9. Use your solutions for questions 3(g), 7 and 8 to calculate the estimates of θ_0 , θ_{01} and θ_1 , respectively. Show all your workings. **(4 marks)**

4 Predicted values and residuals of the final linear mixed model

10. Use the R computer package to produce a table that lists the estimates of the fixed effects in the final linear mixed model, together with their corresponding standard errors, degrees of freedom, observed test statistics and p -values. Present this table below. Note that each value in the table needs to be rounded to two decimal places. **(2 marks)**
11. Figure 3 presents the random effect predictions of the final linear mixed model, for the first 10 teeth in the study. These predictions were obtained by using the `ranef()` command in R.

Figure 3: Random Effect Predictions

	(Intercept)	T
1	-3.62	43.47
2	-2.48	23.51
3	-1.65	32.21
4	-5.57	37.25
5	-1.82	15.41
6	-3.86	29.03
7	-0.75	-4.61
8	-7.78	-9.29
9	-5.95	4.62
10	5.65	6.81

Calculate by hand the predicted conditional value of G for tooth $i = 3$ at 6-months post-treatment. Show all your workings. (**Note:** To answer this question, you will need to use the appropriate information presented in Figure 3, the fixed effect estimates you computed in question 10 and the raw data presented in the `Veneer.csv` file). (4 marks)

12. Calculate by hand the marginal residual for tooth $i = 9$ at 3-months post-treatment. Show all your workings. (**Note:** To answer this question, you will need to use the appropriate fixed effect estimates you computed in question 10 and the raw data presented in the `Veneer.csv` file). (4 marks)