Neetu Regmi

STAT 930

**Consulting Portfolio**

**Project 1: Visualization of Prognosis and Treatment planning**

**Project Description:** The dataset for this project was survey results. The total sample size was 25. The client wanted to know if there is a relation between years of experience of faculty and treatments plan of choice.

**Data:** There were total 10 questions asked in survey and client was interested in question related to two different types of treatment option maxillary right central incisor and mandibular right central second molar. Both treatments question had several options to choose for the participant.

**Methods:** R was used to perform the fisher exact test and Chi-square test of independence but given the small sample size and the expected frequency in some cells turned out to be less than 5 and zero in some cells leading to non-significant results. After consulting with Kelsey, I did visualization for the client.

**Results:**

Plots shown in the report pertained to the research question of interest between treatment and years of experience. Mostly, we were interested in treatment option selected by the participant and years of experience in dentistry. The highest response for one of treatment choice surgical debriment was selected by 30+years of experience in dentistry whereas for other treatment options response varied among different years of experience. Here is one of the plots from the report.

A picture containing text, plot, diagram, line

Description automatically generated

**Summaries of meetings**

I had an opportunity to be a consultant for client from Dental School. The dataset was survey questions. In our first initial meeting which was held via zoom, she presented her dataset and survey questions that were given to participants. We briefly discussed her research questions that she was interested in analyzing. Her research questions were straightforward related to treatment planning and years of experience. We concluded our meeting by outlining what needs to be done and reaching out (via email or follow up meeting) in case if further questions arise from both parties. I also told client about the policy that nothing is promised until two weeks from the time dataset is received. The client did not have any problem with the policy. I ended up doing fisher exact test and chi-square test of independence and unfortunately, the results were not significant. After seeing the result, I consulted this matter with Kelsey. Following her advice, I reached out to the client (via email) mentioning that sample size was small leading to non-significant result and if she would be interested in visualization. Luckily, she agreed on doing visualization. I did visualization (bar plot) based on her research question and made a report. Afterward, we had second meeting via zoom where I went through the report. Although, the client didn’t have what we discussed in the initial meeting, she seemed happy with overall report and the visualization.

**Reflection:**

Overall, the meetings with the client went well. I had two meeting, and both were via zoom. The first one was to discuss about the dataset and research questions whereas the second one was to talk about the report and the visualization. The client wanted some information from the dataset that she could use for her poster presentation at the end. This worked well because she could use the plots for her poster. Couple of things that went well while working on this project was the client was flexible and understanding for the doing visualization instead of analysis. The communication was clear and there were no issues. Although, the client was fine with bar chart and did not request for any other form of visualization one thing that could be improve upon was having different form of visualization other than bar chart might have been even better.

**Communications:**

Email 1: Set up the meeting.

Email 2: Informed about issue of non-significant result and recommended on doing some visualization or open to any other suggestions or ideas.

Email 3: Set up meeting to talk about the report that had visualization related to research question.

**Project 2: Comparison of different Herbicides for control of Palmer Amaranth in NON- GMO Corn**

**Project Description:** The experimental design for the project was randomized complete block design. The client wanted to test the effect of herbicide on different responses and mainly interested in knowing significant herbicides among different response variables. There were four blocks, and each block was a strip consisting of plots where all herbicides were randomized.

**Data:** There were five responses, and one response was repeated measure with value in percentage (0 to 100) which was modeled using beta distribution and other was count which was modeled using negative binomial distribution. Also, there were 18 herbicides where two of them were control herbicide (Non-treated and Weed-free). They were included for the comparison purpose.

**Model:** Let denote one of the 5 responses for the level of herbicide, block, Then

Where:

is the overall mean.

is the effect (deviation from the mean) of the herbicide, i = 1, 2, …, 18

is the effect of the block, j = 1, 2, 3, 4

is error term, ~ iid N(0,)

**Method:**

The PROC GIMMIX procedure in SAS 9.4 was used to analyze each of the five responses of interest. Residuals and qq-plot were used to assess normality. When difference occurred, LSD’s were reported at the level. Additionally, the least squares mean table showing the significant herbicide denoted by letter were our main interest. See below example for one of the response yield

| **Covariance Parameter Estimates** | | |
| --- | --- | --- |
| **Cov Parm** | **Estimate** | **Standard Error** |
| **Rep** | 0 | . |
| **Residual** | 7104094 | 1367184 |

| **Type III Tests of Fixed Effects** | | | | |
| --- | --- | --- | --- | --- |
| **Effect** | **Num DF** | **Den DF** | **F Value** | **Pr > F** |
| **Herbicide** | 17 | 51 | 2.05 | 0.0251 |



| **Tukey Grouping for Herbicide Least Squares Means (Alpha=0.05)** | | | |
| --- | --- | --- | --- |
| **LS-means with the same letter are not significantly different.** | | | |
| **Herbicide** | **Estimate** |  | |
| Surestart II | 13828 |  | A |
|  |  |  | A |
| Acuron | 12647 | B | A |
|  |  | B | A |
| Weedfree | 11357 | B | A |
|  |  | B | A |
| Atrazine | 11126 | B | A |
|  |  | B | A |
| Resicore | 10103 | B | A |
|  |  | B | A |
| TriVolt | 10100 | B | A |
|  |  | B | A |
| Bicep II Magnum | 9960.55 | B | A |
|  |  | B | A |
| Outlook | 9579.83 | B | A |
|  |  | B | A |
| Balance flexx | 9512.80 | B | A |
|  |  | B | A |
| Harness Max | 9441.78 | B | A |
|  |  | B | A |
| Diflexx | 9095.08 | B | A |
|  |  | B | A |
| Zidua | 8531.62 | B | A |
|  |  | B | A |
| Sharpen | 8095.33 | B | A |
|  |  | B | A |
| Verdict | 8035.45 | B | A |
|  |  | B | A |
| Anthem maxx | 7954.53 | B | A |
|  |  | B | A |
| Corvus | 7622.78 | B | A |
|  |  | B |  |
| Untreated | 6844.28 | B |  |
|  |  | B |  |
| Degree Xtra | 6668.85 | B |  |

**Results:**

For this project, the report turned out to be very long since we had five response variables. Each separate analysis gave covariance parameter of estimates and type III test of fixed effects, least squares mean table and difference of least squares tables and estimated mean plot for herbicides. For yield, there were statistically significant difference between Surestart II and Degree Xtra also Surestart II and untreated.

**Summaries of meetings**

I had an opportunity to be the consultant for a client from the department of Agronomy and Horticulture. This was one of the client Ph.D. projects. The design was randomized complete block design and she was one of the returning clients who had worked with Rachel previously. The first initial meeting was held via zoom. In the meeting, the client discussed about her dataset and outlined what she wanted for the analysis. She wanted to know whether treatments effects were significant for various responses. We concluded our meeting by discussing the expectations and what needs to be done. I also mention the timeline for getting report back according to the sc3l policy and the client seemed to be fine with timeline.

**Reflection:**

Even though the project was long I had only two meetings with her. Mostly all the communications were done via email, and we met via zoom to walk through the report. The analysis took more than two weeks. For some clarification, I ended up asking questions about analysis to Rachel because the design was similar to what she had worked on previously. I had some responses which I wasn’t sure whether it should be proceeded with normality assumption but after bringing that topic to Monday meeting I was able to proceed with normality assumption. The report turned out to be long and took almost a month. Couple of things that went well was I was able to work on analysis that had different distribution such as beta with c-loglog link, negative binomial and repeated measure. Overall, the client was patient throughout the whole analysis. Few things that could be improved upon was I could have more communications with client about the issues regarding analysis. One problem that I faced on this project was client wasn’t very responsive to emails and was not very through on what she wanted regarding analysis.

**Communications:**

Email 1: Send email reconfirming the response variable for analysis and asked about how blocking was done and application.

Email 2: Asked about the plot plan to see how the experiment was laid out.

Email 3: Asked if she would be interested in removing one of the control herbicides for count response because of issues such as convergence.

Email 4: Asked about the contrast statement and what she wanted to look at specifically and she didn’t respond until later.

Email 5: Send a zoom link to go over report and explain the result.

**Project 3:**

**Client 3: Evaluate Calf growth and Cow reproduction when cows are supplemented with methionine hydroxy analog during the last trimester of gestation.**

**Project Description:** The client wanted to know if there was any impact of methionine hydroxy analog on primary variable of interest such as weight, BCS, carcass traits.

**Data:** This study was done for three years and there were 150 cows in each year. There were five treatments. Different variables such as sex, dayseason, pre calf weight, pre breed calf weight and others were of major interest.

**Methods:** The PROC GLIMMIX procedure in SAS 9.4 was used to analyze the variables of interest.

**Model:**

Where:

is the overall mean.

is the effect of the year, i = 1, 2, 3

is the effect of the treatment, j = 1, 2, 3, 4, 5

is the random term where treatment is nested within year.

B is the coefficient indicating dependency of on and B

is the measurement made on the day season covariate corresponding to , where day season is equal to the calf birthdate minus the first calving season plus one.

is the average of the day season values (

is the error term, ~ iid N(0, ) where has a covariance structure

**Results:** for calf wean weight variable

| **Type III Tests of Fixed Effects** | | | | |
| --- | --- | --- | --- | --- |
| **Effect** | **Num DF** | **Den DF** | **F Value** | **Pr > F** |
| **Year** | 2 | 8 | 11.48 | 0.0045 |
| **Trt** | 4 | 8 | 4.73 | 0.0297 |
| **Sex** | 1 | 399 | 33.14 | <.0001 |
| **dayseason** | 1 | 399 | 65.12 | <.0001 |

| **Trt Least Squares Means** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Trt** | **Estimate** | **Standard Error** | **DF** | **t Value** | **Pr > |t|** |
| Hay | 511.54 | 8.9948 | 8 | 56.87 | <.0001 |
| MET | 486.56 | 8.7181 | 8 | 55.81 | <.0001 |
| NS | 464.70 | 8.4115 | 8 | 55.25 | <.0001 |
| T1 | 493.30 | 8.8551 | 8 | 55.71 | <.0001 |
| T2 | 507.98 | 8.8572 | 8 | 57.35 | <.0001 |

**Summaries of meetings**

I had the opportunity to be consultant for Landon from the department of Animal Science. This is one of the Master’s project of client. I was able to work alongside with Dr. Kathy Hanford for his project. The preliminary analysis was done by Dr. Hanford and I was able to get information about the project and variable of interest by meeting with Dr. Hanford in person. I was able to start the analyses for some variable and waited until we had meeting with client to work on the other variables. I had initial meeting with client along with Dr. Hanford. In the meeting, the client explained what the overall project was about and their variables of interest for analysis. He explained the meaning of some variables and we discussed about the covariate that needs to be included in the analysis.

**Reflection:**

Overall, I had one meeting with client and two meeting with Dr. Hanford on the project. Initially, the project was started by Dr. Hanford and I was in need of project to work on so I had a chance to work with her. By meeting with Dr. Hanford, I understood what was going on with project. I encountered some new terminology when we looked at the dataset such as weaning and covariate like dayseason. She explained them to me as well as some preliminary analysis related to variables that she had done already like calf wean weight, calf birth weight. This really helped me to start working on the analysis for other variables. I would say working on the project was smooth and I was able to ask questions in case if I was not understanding anything while working on the analysis or needed help in any way. For the second meeting, we ended up discussing SAS code and few changes in the code whether to have dayseason and sex in the model statement for variables such as bcs and carcass traits. Couple of things that went well I had good understanding of some new terminology and was able to ask question and get response quickly from Dr. Hanford so working on the analysis was less time consuming and more straightforward. Things that could be improve upon would be it would be better if we had more meeting with client.

**Communications:**

**Email 1:** send email asking about variables for analysis such as weaning weights.

**Email 2:** send email asking about terms needed to include in the model.

**Email 2:** email to confirm the meeting to show the code for the analysis.