# BAN 5733 Individual Exercise 8 Solutions (10 points)

## Data Set Description:

Your last report for the ABC Pharmaceutical Company regarding an arthritis treatment trial was so well received by the upper management that they have asked you to look at some additional treatment data. In this study, ADHD children were placed into one of 5 groups: a placebo group, or one of 4 groups that received varying amounts of either Ritalin or Adderall. After the dosage, the children were given instructions to accomplish a task, and were given a score corresponding to how well they followed instructions (a higher score indicated better instruction following). Their internal researchers have determined that the treatment groups are not effective when examined by a one-way ANOVA. However, they have additional variables that they would like considered. The project researchers have moved on to other billable projects so ABC Pharmaceutical deemed it a better investment to have your company look into the additional information. ABC Pharmaceutical would like to know if there is any difference in "instruction score" among the different dosages of drugs and gender groups? They would also like to know if there is any difference in "instruction score" among the treatment groups and gender. Include the "placebo" group as a dosage.

Your answers will be looked at by all the same groups as the last report - the ABC Pharmaceutical managers, researchers and marketing analysts. Write a short section in a manager friendly way with a separate detailed technical section for the researchers and marketing analysts. The managers are interested in primarily non-technical explanation of your final model and recommendations based on that model. The researchers and marketing analysts are interested in knowing all the technical details including but not limited to testing of assumptions of your model. The Researchers/Analysts do not want you to transform any variable for this modeling exercise.

Your analysis may be conducted in JMP or SAS. Your report should be restricted to 6-pages maximum. Please back up your assertions with appropriate statistics and graphs as needed.

Variable Descriptions:

ID is the identification variable

Gender = Male, Female

Treatment = Placebo, Ritalin 10mg, Ritalin 17mg, Adderall 7.5mg, Adderall 12.5mg

TrxGrp = Placebo, Ritalin, Adderall

Score = instruction score; continuous

Manager Report: **(2 points)**

1. Briefly discuss problem, data and model used to solve it
2. Discuss actional insights and recommendations from analysis
3. Make sure you tell them the answer and briefly describe how you got there.

Technical Report: **(6 points)**

1. Go through each of the steps: Background/ Purpose, Data Description, Methods, Results, Insights/Conclusions
2. Define the problem well enough that people could replicate the study.
3. Describe the data including the number of records, number of variables, and where you obtained it.
4. Support your findings with appropriate graphs and descriptive information
5. Provide insights and recommendations based on original problem and analysis

Analytic Code: **(2 points)**

1. Provide 1 set of code from either SAS, R or Python for this ANOVA analysis
2. The remaining analysis may be completed in JMP.

Deliverables:

- As you complete the exercise, create a report in Microsoft Word that includes the managerial report, technical report, and analytic code and is limited to 6 pages.
- Copy and paste supporting tables/diagrams as needed to an appendix to justify any of your answers.
- Make sure you *print your name, student ID#, student email on the cover page* of the report and turn-in the report as communicated by your instructor.
- Please also put a running *header/footer with your name, on each page of your exercise* solution report.
- Failure to follow these instructions will result in deduction of points

# Manager Friendly Report

In this study, attention deficit hyperactivity disorder (ADHD) children were placed into one of 5 groups: a placebo group, or one of 4 groups that received varying amounts of either Ritalin or Adderall. After the dosage, the children were given instructions to accomplish a task, and were given a score corresponding to how well they followed instructions (a higher score indicated better instruction following). When ABC Pharmaceuticals' internal researchers examined the data using one-way ANOVA and the 5 levels of treatment, they did not find any association. However, the management of ABC Pharmaceuticals has asked us to re-evaluate this data with additional information. Questions posed were:

- **If there is any difference in "instruction score" among the different dosages of drugs and gender groups?**
- **If there is any difference in "instruction score" among the treatment groups and gender?**

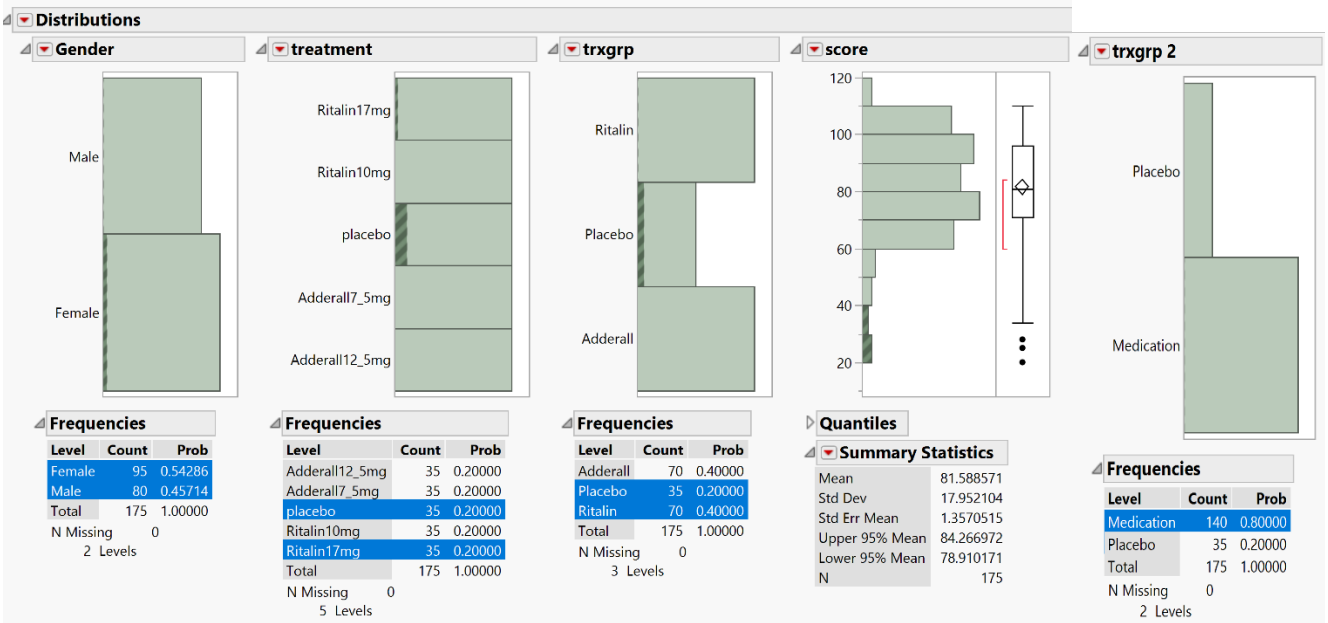
We have explored your data and built several models in an attempt to explain whether prescription control of ADHD is effective. The full model was not statistically significant. This is consistent with the previous work by your internal researchers. However, we did recode the data into placebo and any type of ADHD medication (Adderall and Ritalin). This model showed interesting results. Below you will find highlights of the analyses conducted. For detailed results, please refer to the accompanying technical report that is attached.

- Scores did not differ based on the combination of gender and 5-level treatment or gender and 3-level treatment group.
- Gender alone, 5-level treatment alone, nor 3-level treatment group (placebo, Adderall, Ritalin) were related to final score.
- Even after excluding some potential outliers, non-significant results were obtained.
- We combined the Adderall and Ritalin categories and then compared to placebo. We found a significant relationship between this new 2-level treatment group and the instruction score.
- Either medicine significantly improved the score compared to placebo.
- This indicates that any dosage of Ritalin or Adderall will improve the instruction score compared to the placebo but there was not difference among the dosages.

Technical Report

There are 175 records included in the dataset. The dependent variable "score" is a continuous variable indicating the score for task instruction and completion. In addition, the gender, treatment (5-levels), and trxgrp (3-levels) are categorical variables including the gender information of the patients and the treatment group for which they were assigned.

A quick look at the summary statistics indicates all three nominal variables are evenly distributed with the placebo group having the lowest number of records. Further, the categorical values do not seem to have issues with low cell count or extremely unbalanced data. The continuous variable (score) has some records that are outside of the interquartile range but the mean and median are the same and the variable has low skewness and kurtosis values. None of the variables have missing values. Variable transformations are not indicated in this data. However, we did combine the Ritalin and Adderall dosages groups together and compare with the placebo to look at effects.




A series of models using 5% level of significance were completed in three groupings. The first were a series of One-way ANOVAs to test the main effects of treatment (5-levels), gender, treatment group (3-levels) and treatment group 2 (2-levels). The null and alternative hypotheses being tested were:

$H_0$: All means are equal across the tested variables

$H_a$: A difference occurs among the means within or between variables.

As previously discovered by the internal researchers, the 5-levels of treatment were not statistically significant; neither were gender and treatment group (3-levels). However, the 2-level treatment group 2 variable was found to be significant with the medication group having better task scores than the placebo group (83.2 vs. 75.2, respectively).

Too ensure the variables were not masking effects, a series of two-way ANOVA with interactions were completed as seen in the table below. None of these models were significant so the interaction terms were removed and the main variables were re-examined to ensure adding the control variable didn't alter the results. The treatment group 2 variable remained significant but the gender variable did not change. For simplicity the One-way ANOVA model will be used.

| Model Variables/Parameters | Overall Model Outcome | Number of Significant Parameters | Comments |
|---|---|---|---|
| One-way ANOVA – Main Effects | | | |
| Treatment | F= 1.45<br>p-value = 0.2203 | 0 | Not Significant |
| Gender | F= 0.66<br>p-value = 0.4191 | 0 | Not Significant |
| Treatment Group | F= 2.85<br>p-value = 0.0606 | 0 | Not Significant |
| **Treatment Group 2** | **F= 5.69**<br>**p-value = 0.0181** | **1** | **Significant** |
| Two-way ANOVA with Interactions | | | |
| Gender, Treatment, Gender*Treatment | F= 0.89<br>p-value = 0.5337 | 0 | Not Significant |
| Gender, Treatment Group, Gender*Treatment Group | F= 1.42<br>p-value = 0.2181 | 0 | Not Significant |
| Gender, Treatment Group 2, Gender*Treatment Group 2 | F= 2.25<br>p-value = 0.084 | 0 | Not Significant |
| Two-way ANOVA without Interactions (Main Effect – Control Variable) | | | |
| Gender, Treatment | F= 1.29<br>p-value = 0.271 | 0 | Not Significant |
| Gender, Treatment Group | F= 2.12<br>p-value = 0.0996 | 0 | Not Significant |
| **Gender, Treatment Group 2** | **F= 3.18**<br>**p-value = 0.0442** | **1** | **Significant** |

The Treatment Group 2 model was significant but did not account for that much variation in the score. Only about 3% of the variation in task accomplishment score was accounted for by the treatment group 2 variable.

The medication group in treatment group 2 variable does show promise at increasing the task accomplishment score (85 vs 73 for placebo) but there is not enough information in this data set to determine if it is Adderall or Ritalin that makes the bigger impact as evidenced by the models presented in the Model Table above.

**Summary of Fit**

| | |
|---|---|
| RSquare | 0.031842 |
| RSquare Adj | 0.026246 |
| Root Mean Square Error | 17.71495 |
| Mean of Response | 81.58857 |
| Observations (or Sum Wgts) | 175 |

**Analysis of Variance**

| Source | DF | Sum of Squares | Mean Square | F Ratio |
|---|---|---|---|---|
| Model | 1 | 1785.606 | 1785.61 | 5.6899 |
| Error | 173 | 54290.771 | 313.82 | Prob > F |
| C. Total | 174 | 56076.377 | | 0.0181* |

**Parameter Estimates**

| Term | Estimate | Std Error | t Ratio | Prob>\|t\| |
|---|---|---|---|---|
| Intercept | 79.192857 | 1.673906 | 47.31 | <.0001* |
| trxgrp2[Medication] | 3.9928571 | 1.673906 | 2.39 | 0.0181* |

| Level | Least Sq Mean | Std Error | Mean |
|---|---|---|---|
| Medication | 85.014286 | 1.4933967 | 85.0143 |
| Placebo | 73.142857 | 2.9867933 | 73.1429 |

ANOVA SAS Code:

Creation of final data set – This code created the trxgrp2 variable

```
libname ex6 "C:\Users\miriajm\OneDrive - Oklahoma A and M System\5733 BAN\Fall 2019\Exercises\Exercise 6";

proc contents data = ex6.exercise6;
run;

proc print data = ex6.exercise6 (obs=5);
run;

proc sql;
create table work.ex6 as
  select *, case when trxgrp = 'Placebo' then 'placebo' else 'medication' end as trxgrp2
   from ex6.exercise6;
quit;
```

Creation and running of ANOVA macro to test series of models –

```
%macro anova(class=treatment, inter=);
%if %sysfunc(length("&inter.")) > 2 %then %do;
  title "ANOVA with &class. and &inter. interaction";
  %end;
 %else %do;
   title "ANOVA with &class.";
   %end;
proc anova data = work.ex6;
class &class;
model score = &class &inter;
means &class. /tukey;
run;
title;
%mend;
```

```
%anova;
%anova(class=gender);
%anova(class=trxgrp);
%anova(class=trxgrp2);
%anova(class=gender treatment, inter=gender*treatment);
%anova(class=gender trxgrp, inter=gender*trxgrp);
%anova(class=gender trxgrp2, inter=gender*trxgrp2);
%anova(class=gender treatment, inter=);
%anova(class=gender trxgrp, inter=);
%anova(class=gender trxgrp2, inter=);
```

Verifying ANOVA results with GLM procedure macro –

```
%macro glm(class=treatment, inter=);
%if %sysfunc(length("&inter.")) > 2 %then %do;
  title "GLM with &class. and &inter. interaction";
  %end;
 %else %do;
   title "GLM with &class.";
   %end;
proc glm data = work.ex6;
class &class;
model score = &class &inter;
lsmeans &class. /adjust=tukey;
run;
title;
%mend;

%glm;
%glm(class=gender);
%glm(class=trxgrp);
%glm(class=trxgrp2);
%glm(class=gender treatment, inter=gender*treatment);
%glm(class=gender trxgrp, inter=gender*trxgrp);
%glm(class=gender trxgrp2, inter=gender*trxgrp2);
%glm(class=gender treatment, inter=);
%glm(class=gender trxgrp, inter=);
%glm(class=gender trxgrp2, inter=);
```