Name: Md Allama Ikbal Sijan

**Hierarchical Linear Modeling**

Psyc 741, Spring 2025

**Due Date**: February 12th (by 11:59 PM)

**Session 4 – In-Class Assignment**

For this in-class assignment, you will be combining together job satisfaction data from three different companies to execute a null multilevel regression model.

You will be completing the following questions, some of which involve conducting analyses in R. You will hand in the answers to the questions; enter them into this word document. Some of the answers involve you copying and pasting your R code. However, you also need to submit your *complete* R code too (saved as a .R script file). Submit both this completed document and your R script file to the submission portal for the Session 4 In-Class Assignment on Canvas.

1. Import the **session4\_companyA**, **session4\_companyB**, and **session4\_companyC** data files into RStudio.

company\_A <- read\_excel("session4\_companyA.xlsx")

company\_B <- read\_csv("session4\_companyB.csv")

company\_C <- read\_sav("session4\_companyC.sav")

1. You notice that there is not a company ID variable in the data files. Create unique company ID values in the three data frames.

company\_A$company\_id <- 1

company\_B$company\_id <- 2

company\_C$company\_id <- 3

1. You also notice that the three companies used individual ID values that overlap with each other. To avoid any potential confusion in the subsequent combined data frame, figure out a way to quickly recode the individual id values so that they are all unique across the three data frames.

# Recode individual id values

company\_B <- company\_B %>%

mutate(id = id + nrow(company\_A))

company\_C <- company\_B %>%

mutate(id = id + nrow(company\_B))

1. Bind/append together the three company data frames into a single data frame.

# Append the three company data

names(company\_A)

names(company\_B)

names(company\_C)

company\_A <- rename(company\_A, "job\_satisfaction" = "js")

df <- bind\_rows(company\_A, company\_B, company\_C)

1. Execute a null multilevel model with company as the group/cluster variable and job satisfaction as the dependent variable. Make sure that maximum likelihood (ML) is used as the estimator. Paste your R syntax below. \**Obviously you would usually want more than three companies of data for a multilevel model like this, but let’s use this example as training.*

null\_model <- lme4::lmer(job\_satisfaction ~ (1|company\_id), data = df, REML = FALSE)

1. Report the following fit statistics from the null model:
   1. AIC: 463.6
   2. BIC: 475.9
   3. Log-likelihood: -228.8
2. Report the following sample sizes from the data:
   1. Number of individual participants: 456
   2. Number of groups: 3
3. Report the following variance components from the model output:
   1. Between-group variance: 0.9078
   2. Within-group variance: 0.1527
4. Based on your responses from Q8 above, calculate and report the intraclass correlation coefficient (ICC) and decide if multilevel modeling seems necessary for this data.

ICC = 0.8560113 & yes, multilevel modeling is very necessary

1. Output the means and standard deviations for the job satisfaction scores for the three companies by grouping the data by company ID and summarizing the job satisfaction scores. Report the values below. Do these descriptive statistics support your conclusion in Q9 above? Explain your reasoning.

company\_id Mean SD

<dbl> <dbl> <dbl>

1 1 1.77 0.294

2 2 3.13 0.477

3 3 4.09 0.352