P8110 Applied Regression II
Midterm Project

Research Problem:
To determine the effect of parental depression on the risk of depression in their offspring, offspring with depressed index parents were compared to offspring with normal index parents on the age of onset of depression. There were a total of 220 eligible children, 125 children with a depressed index parent and 95 children with a normal index (control) parent.

Diagnostic information, including whether an offspring had ever had depression or ever abused drugs, age of onset of depression and age of initial substance abuse were obtained retrospectively on all 220 children based on a child psychiatrist's blind evaluation of diagnostic information from interviews with the child and/or interviews with the mother about the child, and other corroborating information from school or medical records. The children were between the ages of 6 and 23 years at time of interview.

The table below lists the relevant variables with description for the .csv data file prepared for analysis.

Variable	Definition	Coding
ID	ID of child	
PARDEP	Index parent Depression Status	1 = Ever depressed 0 = never depressed
DSMDEPHR	Depression Status of Child	1 = Ever depressed 0 = Never depressed
PTSEX	Sex of Child	1 = male 2 = female
PTAGE	Age of Child at Interview	Age in years
BEDEPON	Age of onset of depression in offspring	-1: datum unavailable Age in years
DSMSUBHR	Substance abuse status of offspring	1 = substance abuse 0 = no substance abuse
BESUBON	Age of onset of substance abuse in offspring	-1: datum unavailable Age in years
SESCLASS	Categorized social class of parent	Coded as 1, 2, 3, 4, 5; higher scores imply lower SES
PARENTMS	Marital status of index parent	1 = Married with spouse 2 = Separated/Divorced

Examine the data and use appropriate models to assess the following hypotheses,

1. Investigators have hypothesized that offspring of a depressed parent are more likely to have pre-pubertal onset (<13 years of age) depression than offspring of a non-depressed parent, but equally likely to have an onset of adolescent/early

adulthood depression, given demographic and social characteristics. Is this hypothesis supported by this data?

2. Is there evidence for the effect of prior depression in offspring, as well as the effect of parent's depression status on the age of onset of substance abuse/dependence in offspring, given demographic and social characteristics?

using statistical and descriptive justifications including:

i)   Presenting Graphical displays and other descriptive statistics;
ii)   Description of the statistical modeling approach;
iii)   Summary of the results, with appropriate comments to support your answers.

Hint: The outcome needs to be constructed using PTAGE, BEDEPON and DSMDEPHR.

Instructions:

Each group needs to submit two files. File 1 contains the report on data analyses supported with SAS outputs and programs. File 2 summarizes your analytic plan and findings, supported by figures and/or tables, in a manner that could be published in a peer-reviewed paper.

Report 1 is due as a .pdf document via online submission on Canvas by *11:59 pm on Oct 30, 2025*. Only one group member needs to submit the report. The report needs to be organized by sections, including "Data preparation", "Descriptive statistics analyses", "Analyses to address hypothesis 1", and "Analyses to address hypothesis 2". If models are used, they need to be included in the report and each term in the models needs to be defined clearly. SAS program needs to be submitted as a separate .sas file so that TAs can replicate the analyses in the report. The SAS program needs to be well documented. [10 points]

Report 2 is due as a .pdf document via online submission on Canvas by *11:59 pm on Nov 13, 2025*. Only one group member needs to submit the report. The report should be single-spaced and a maximum of two pages excluding any references, tables, and figures. Margins should be at least one inch all around with the font size to be 12 point. We allow a maximum of five tables and/or figures in total (not each). Tables and figures should be self-explanatory (prepare tables and figures in publishable quality; see the paper in HW4). The report should include only three sections: Data Analysis, Results, and Conclusion (one paragraph). [10 points]

Please specify the percentage contribution of each group member. **The final grade each student receives for their midterm project will reflect the scores of the two reports as well as their contribution to the project.**