1. Dataset ‘Doctor.xlsx’ (variables Y=pregnancy outcome (dependent variable), x1, x2, x3, x4, x5 (independent variables, factors)=indicators of the woman under study):

- Construct descriptive statistics using the summary function.

- If necessary, create a categorical variable (factor function).

- Build a basic logistic regression model with as many predictors as possible using the glm function.

- Write a binary regression equation using coefficient estimates.

- Test the significance of the regression coefficients individually.

- Test the significance of the regression as a whole using the Wald and maximum likelihood criteria.

- Construct confidence intervals for the regression coefficients (confint and confint.default functions).

- Conduct comparative analyses of logit and probit models.

- Construct a contingency table with a threshold probability of 0.5 (confusionMatrix function).

- Calculate the specificity and sensitivity of the model (sensitivity and specificity functions).

- Find the optimal threshold value of the prediction probability. Build a contingency table for this probability, calculate the specificity and sensitivity of the model (function optimalCutoff).

- If the sample is large, you can preliminarily split the sample into two parts: training and test samples (function sample\_frac).

- Plot the ROC curve, interpret the results (plotROC function).

- Try to improve the logit or probit model using the AIC coefficient (stepAIC function).