Homework#3 (Total score = 100)

COMPSCI- 5590-0012 Econometrics of Data Science
DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF MISSOURI-KANSAS CITY

Due on Friday, Dec 9, 2022 by11:59 PM (in Canvas)

"It must be an individual submission. Any kind of copying or corroboration will be severely penalized".

The spreadsheet "hw3_car_data.csv" contains the following variables:

Variable name	Variable meaning
foreign	1 if foreign, 0 otherwise
mpg	Gas mileage (miles per gallon)
cylinders	Number of cylinders
displacement	Engine displacement (cubic inches)
hp	Horse power
weight	Vehicle weight (pounds)
acceleration	Maximum acceleration (ft/s2)
modelyr	Model year
origin	Continent of origin. 1=North America; 2=Europe; 3=Asia;
name	Vehicle name
Modyr70-82	Binary variable for model year

I. [75 points total] Using the given data, estimate a logistic regression model to explain "foreign" as a function of selected explanatory variables, including binary variables for model year (Think about your model!).

To earn credit, please solve the following questions using **R**:

1. [15 points] Generate summary statistics for these variables (including min, max, mean, standard deviation, median, 25th and 75th quartiles) and interpret the results;

2. [15 points] State and justify your binary logit model explaining "foreign";

3. [10 points] Estimate a binary logit model that includes binary variables for model year;

4. [25 points] Summarize your results in a table and interpret them (odds ratios are convenient);

5. [10 points] Discuss goodness-of-fit measures;

II. [25 points] Summarize the main ideas in "On the relevance of irrelevant alternatives," by Benson et al. (2016).
Reference: Benson, A. R., Kumar, R., & Tomkins, A. (2016, April). On the relevance of irrelevant alternatives. In *Proceedings of the 25th International Conference on World Wide Web* (pp. 963-973).

The summary should be of one-page (typed, 12 pt, Times New Roman, 1.15 spacing, with 1inch margins).

Submission guidelines:

- In the document, first, write down a question and then write your answer after that.
- In the answer script, write your name and student ID on the right-hand side of the header.
- Please type your response. *No* handwritten submission will be accepted.
- Upload a pdf file containing the results, graphs, reasoning, and interpretations.
- Upload the R file.
- Please write the file name (both pdf and R) in this format: COMPSCI-5590-hw3-your last name
- It is an *individual submission*.