Alex Tabakian

Q1A)

a) Name of paper and student presenter

Towards Enhancing Road Safety in South Carolina Using Insights from Traffic & Driver-Education Data by Sai Teja Paladi

b) Create a new example exemplifying the main conclusion of your example

Many crashes caused by excessive speeding (20mph over speed limit or more)

c) Describe how the conclusion is supported in your answer

The expansion of Alive@25 would teach drivers the dangers of excessive speeding, thus lowering collision rate.

Q2)

a) Describe your data preparation

I prompted Chat GPT to read each pdf file and output comma separated values of each class's attendance records. I then combined the data of each class into a single txt file that could easily be parsed in python.

b) Describe your steps to create a model - pre trained, your own, manual

I used a pretrained model because it already contains strong language reading capabilities, which worked well enough that I did not have to build my own model.

c) Answer the questions from your analysis

a. What are the number of classes and their dates

Number of classes: 25

Class dates:

- 2025-08-19, - 2025-08-21, - 2025-08-26, - 2025-08-28, - 2025-09-04, - 2025-09-09, - 2025-09-11, - 2025-09-16, - 2025-09-18, - 2025-09-23, - 2025-09-25, - 2025-09-30, - 2025-10-02, - 2025-10-07, - 2025-10-14, - 2025-10-16, - 2025-10-21, - 2025-10-23, - 2025-10-28, - 2025-10-30, - 2025-11-04, - 2025-11-11, - 2025-11-13, - 2025-11-18, - 2025-11-20

b. What is the median attendance per class

Median class attendance: 23

c. What are the dates with the lowest and highest attendance

Lowest attendance: 14

Highest attendance: 49

d. Is there a correlation of high attendance with course evaluation dates? When is the attendance the highest?

Yes, there is a correlation between evaluation days and attendance. The average attendance on an evaluation day is 40.0 students while only 32.36 students on a non-evaluation day. This difference in attendance shows that there is statistical evidence to suggest that there is a higher attendance rate on evaluation days.

d) If you had more time, what more could you have done to improve performance?

I noticed that Chat GPT was only able to collect data for 25/27 of the classes. If I had more time, I would look into it and figure out why this is happening, or possibly create my own model.