My ‘decide’ method is score-based: I pre-assigned scores for each characteristic and kept two variables, a passenger score and a pedestrian score, to track the respective total scores given the scenario characteristics and persona characteristics. The group with the higher score survives.

While assigning the scores, I ran into some moral dilemma which left me with a pretty unsettling feeling. For instance, I decided to assign a score of 0 to the ‘overweight’ characteristic because overweight personas generally have a low life expectancy as compared to average and athletic personas. Yet, this decision made me ponder since I do not feel that it is fair to ‘discriminate’ people by their body type.

Further, I gave the ‘pet’ characteristic a score of 3 since I am personally a big pet-person. This demonstrates that bias is purely subjective to a person’s preferences; someone who does not like animals would not attribute such a high score to pets. A similar case can be argued for the ‘you’ or ‘criminal’ characteristic. The question I have then been asking myself is: what would be the most ethical choice?

My audit results seem to be consistent with the expectations set by my scores. On average, pedestrians and passengers have the highest and lowest survival rates respectively, which I believe is a fair ratio since the autonomous-driving car should ‘know’ that the brakes failing is arguably more the passengers’ (owners of the car) fault that it is the pedestrians’ fault.