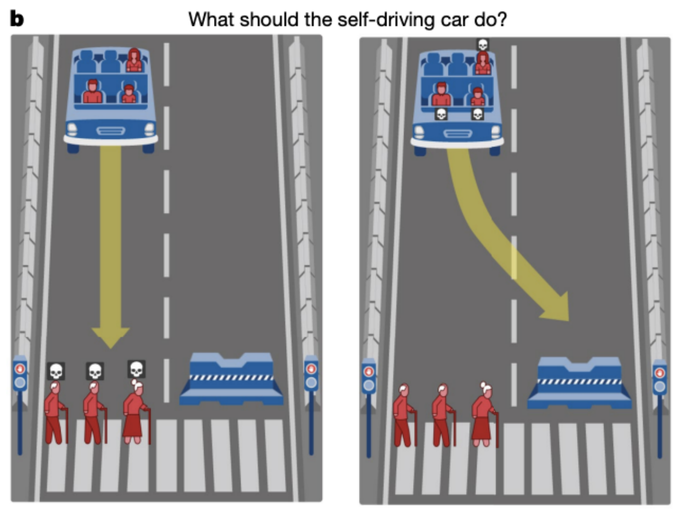
Autonomous Vehicles: Taking the Mean

By Paul Lestz

I. Introduction

In the American panopticon, elected representatives are the prisoners and their constituents are the jailors. Politics is a game of averages, while society bubbles as an ocean of extremes. As we progress deeper into the Information Era, the behavior of autonomous vehicles is destined to be a major player in the continued exacerbation of political contention. This subject is immensely divisive, as it is a matter that is heavily intertwined with a tripartite set of political influences: morality, mortality, and practicality.



Resemblant of the signature Trolley Problem, the potential ethical dilemmas raised by autonomous vehicles (see https://www.moralmachine.net/ for more details) concern human lives. Given this, it is more critical than ever that political systems around the world begin to *develop and ratify* regulations for autonomous vehicles (specifically pertaining to their driving behavior in irregular circumstances). Traditionally, ratification is the greatest obstacle of any legislative agenda, given the relatively inefficient and polarized nature of government and bureaucracy. Autonomous vehicle legislation is atypical in the sense that personalization is infeasible and universal agreement is impossible, leaving agreement to be scavenged in the valley in between.

II. Implementation

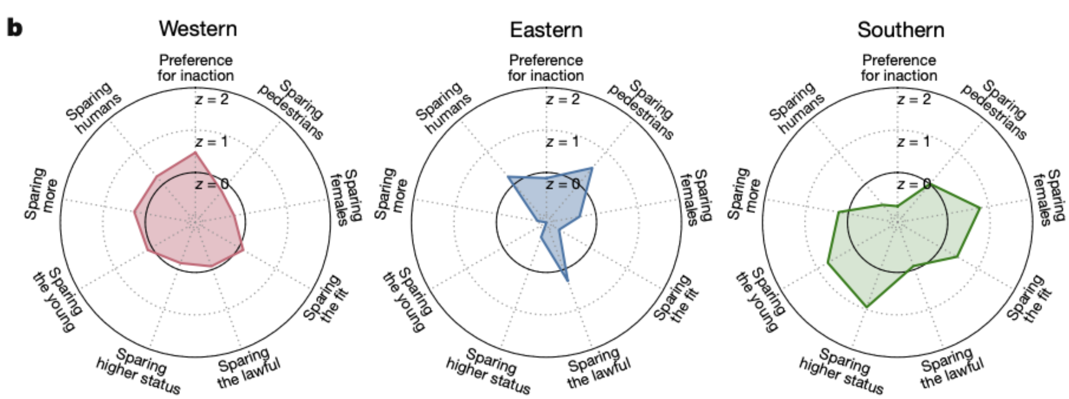
In an effort to model the decision-making process within an autonomous vehicle, I have designed an architectural web in Python to demonstrate (albeit in a simplified manner) the methodology through which a smart car might decide what group of people to save in a hazardous situation. The critical section of such a process would be the calculatory methods within the SmartCar.py class, which scan both the vehicle and the surrounding environment for data in order to make the most efficient, ethical determination as to a course of action. Now, this determination has very little to do with the environment (aside from perhaps vehicle intervention and one's relation to the law) and much more to do with the people (i.e. the passengers and pedestrians) involved.

Within the ‘web/vehicle_architecture’ directory, one will find a Passenger class and a Pedestrian class (both of which inherit from the Person class). Whether a passenger or a pedestrian, these modular representations of people have the same attributes (e.g. an age, a weight, a family, a geographical location, etc.). While collecting the data about these people (especially those not in the vehicle) in real-time is certainly an important issue, that is not the subject of this piece; we will assume that, on average, the self-driving vehicle will collect a proportion of real data that is similar to its maximum collection capacity (which would, in theory, result in the same decision). The true obstacle is, given full information, what constitutes an ‘ethical’ and ‘acceptable’ decision in the face of inevitable danger.

In ‘SmartCar.py’, within ‘web/vehicle_architecture’, there are three methods that perform the crucial decision-making work of the vehicle, those being *calculate_human_value*, *get_regional_bias*, and *calculate_mean_decision*.

The first of these methods is arguably the most controversial, as it essentially performs a calculus on the value of a human life and outputs a single number, which is an inherently disconnected and emotionless way to derive value. However, for now, we as a society are limited by computers that function numerically, not emotionally, at their core; given this, it is necessary that any results we use to make conclusions can be realistically converted to binary. The current weightings applied to each of (what I have deemed are) relevant human attributes to consider are likely based on a fusion between my current beliefs and Western beliefs (as I live in the Western world and thus subconsciously adopt such beliefs). They serve merely as a sample set of decision-making influences to demonstrate potential key outcomes.

The second of these methods exists given the demographic divergences shown below.



Western, Eastern, and Southern nations and cultures undoubtedly have drastically different perspectives on what constitutes a feature of legitimate value. For example, while Southern nations and Western nations (to a lesser degree) value social status, Eastern nations could not care less. In contrast, while Eastern nations prioritize sparing the lawful, Western and Southern nations do not see this as a particularly valuable point of analysis. Within *get_regional_bias*, I attempt to replicate some of the priorities above by assigning scaled values to different bias terms (which are utilized in *calculate_human_value* to adjust one’s worth based on the region of the vehicle itself).

Finally, the last of these three methods sums together the calculated values of the passengers and pedestrians separately, and then decides which party to save based on the higher *mean* score plus an additional bias for every person in the vehicle (in line with the principle of Utilitarianism, which advocates for saving the group for more individuals in order to maximize welfare). In order to not leave the decision solely to the quantity of people on both sides, I instead opt for an additive bias (of 300 units) for every person in the vehicle, such that a high enough mean value on one side could outweigh another side with relatively more individuals. For example, perhaps a scenario pitting one doctor against two active criminals *could* be worth a deontological debate (to save the individual), given one's feature weights.

The crux of our value determination is taking the mean of our calculated human values (with reasonable bias for the number of humans present), with respect to their region of origin, simply because there is no fairer way to make this determination but on a small number of geographic averages. In order to earn support from users, autonomous vehicles must be predictable and encourage faith within their drivers that buying the car is not a death sentence. On top of this, in order to earn support from politicians, there must be a prudential argument for a collective majority buy-in. There is a very narrow crevice between the closing walls of practicality, fairness, safety, and security that must be taken in order for autonomous vehicles to be both safe and successful.

I believe that the key to living within this space is simply to take the mean, as much as possible. When I provide weights in the *calculate_human_value* method, these are only a placeholder for what I hope would be regionally (e.g. Western, Eastern, and Southern) agreed upon feature weights. Thus, in total, there would be two averages being combined and compared: the regional feature average (for each human) and the human value average (to compare passengers to pedestrians).

Note: While named pedestrians, technically this group could be in another car. The code structure allows for this with no alteration.

Some of the results of my simulation can be found in Appendix 1, Appendix 2, and Appendix 3. The people within these scenarios are not real. They were created using Python's *random*.

III. Consultation

Having spoken to drivers of different ages (in addition to engaging in personal reflection), the scope of autonomous vehicle regulation (as expected) is not a unanimously agreed upon matter. Now, I would advocate that most living drivers on the road are, in one way or another, experts in the subject; given the safety risks of the endeavor, a sufficient level of mastery is required (both to earn a license and survive). While the levels of expertise may differ, I would insist that most drivers (of different ages) can provide valuable insights regarding proper programming of autonomous vehicles (especially given that they are all likely to operate one in the near future).

To begin, I spoke with my father – a man who has spent countless hours driving across the country for work, trips, and emergencies. While a lengthy conversation, it revolved around one main question: what would it take for you to let an autonomous vehicle drive for you? In other words, I wanted to know what kind of assurances that he wanted from this technology. In this conversation, I served as the voice of realism. My father wanted a vehicle that would always prioritize and keep him and his family safe; however, I had to remind him that, ultimately, every family member everywhere feels this way.

Having grown up in and lived through a world full of polarization for over fifty years, my father is eerily aware of two things: (1) different regions of the world never tend to get along, and (2) satisfying all people is a goal that lives in the ideal, not the real. Ultimately, this was a preface for suggesting a regional system, with some sort of averaging to balance out extreme opinions.

While my father is closer to my grandfather in age than I am, even they grew up in vastly different societies. The world of technology that surrounded my father was nonexistent when my grandfather was a child. The thought of a car containing portable technology, let alone being able to drive itself, would exist only in the realm of science fiction. When I asked my grandfather the same question as my father, he expressed reluctance at even answering my question. He said, “I have built my life on confidence in that which I know, not faith in that which I do not.” Unfortunately, as I told him, technological innovation works in a different time frame than typical inventions. New technology is a virus that spreads throughout the population. At first, no one wants the scary new device; however, once its prowess is shown, it spreads through the population like wildfire.

A significant amount of debate later, my grandfather was able to concur with me that we must be able to reach a balance between fear (a feeling that he is intimately familiar with given all that he has been through) and excitement (a feeling that can often get lost to time). My grandfather wanted assurances that these vehicles would not disrupt his way of life, which is his way of saying that he wants to ensure that the technology is safe for his loved ones, who are destined to use it in the future. To that, my father, my grandfather, and myself are very alike.

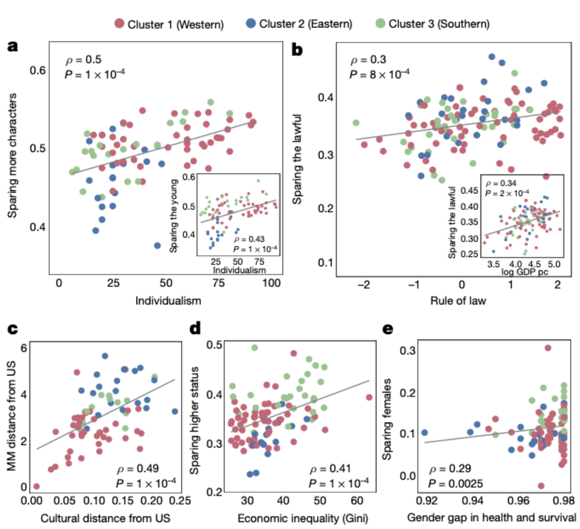


Fig. 4 | Association between Moral Machine preferences and other variables at the country level. Each panel shows Spearman's ρ and P value for the correlation test between the relevant pair of variables. **a,** Association between individualism and the preference for sparing more characters ($n = 87$), or the preference for sparing the young (inset; $n = 87$). **b,** Association between the preference for sparing the lawful and each of rule of law ($n = 122$) and log GDP per capita (pc) (inset; $n = 110$). **c,** Association between cultural distance from the United States and MM distance (distance in terms of the moral preferences extracted from the Moral Machine) from the United States ($n = 72$). **d,** Association between economic inequality (Gini coefficient) and the preference for sparing higher status ($n = 98$). **e,** Association between the gender gap in health and survival and the preference for sparing females ($n = 104$).

IV. Confronting Reality

Humanity was not designed for agreement. As you can see in the diagram to the left, even splitting us into three groups brings an incredible amount of diversity of opinions. Personalization and universality are no better though; one lets chaos run rampant and the other stifles freedom. So, all that can be done is to adopt this tripartite regional system, founded on a base of built-in limitations, to serve as guarantees to an erratic public and wary politicians.

For example, these vehicles should always be prioritizing a good: human life. Allowing unrestricted access to the code could allow those with malicious intentions to abuse such privilege. Another example, we must have detection mechanisms for detecting sampling bias and action bias among active vehicles. It is imperative that we ensure that autonomous vehicles are not only trained properly, but also make decisions fairly and justly. In other words, they should not be using characteristics that have no relevance to whether one should live (e.g. gender, nationality, ethnicity, sexual orientation, etc).

In my implementation, I attempt to build in reasonable safeguards for some of the above characteristics by not allowing the vehicle to know such information about its passengers (or pedestrians that it may scan). In future versions, I would also seek to implement self-checks, where the vehicle can scan its own decision history to ensure that it is not unjustly saving a higher proportion of a certain group. These types of safety features are crucial to not only the success of autonomous vehicles on the road, but also for getting such legislation passed in the first place, as they allow for assurances in a world of technology where there is sometimes very little.

V. Conclusion

To abstain from attempting to regulate autonomous vehicles would be analogous to holding off from creating digital privacy legislation at the advent of the Digital Age. We either sacrifice our time now, or our lives later, and I believe that the latter is simply not a viable option. However, I cannot guarantee that all people feel the same. Many politicians are influenced by their own advancement, and private companies are burdened with stakeholders demanding monetary success at every turn.

We require a solution that works for as many people as possible in the present. I see no fairer cause then but developing a system of averages, where no one can question its fairness because it is the unbridled definition of fairness.

The alternatives to taking the mean are weak. If autonomous vehicles always prioritized their passengers, then no one would ever be able to walk without immense risk. If pedestrians are always prioritized, no one would drive the car. If all people were able to customize vehicles to operate according to their specific beliefs, then cars would be crashing into each other left and right, trying to satisfy their user. The middle ground lies vacant in our legislation, and it is critical that we rally, as constituents, in order to fill this void before a death toll does.

Thank you to Edmond Awad, Sohan Dsouza, Richard Kim, Jonathan Schulz, Joseph Henrich, Azim Shariff, Jean-François Bonnefon, and Iyad Rahwan for their work in this Moral Machine analysis. All credit provided.

Appendix 1

In the diagram below, we see detailed scenarios in which the Autonomous Vehicle always decides to save the pedestrians in the scenario, despite its geographical location.

```
passengers

Name: James Martinez
Age: 44
Gender: M
Weight: 114.4 lbs
Education Level: 3
Occupation: Journalist
Is Doctor: False
Employed: True
Wealth Class: MIDDLE
Married: True
Number of Children: 0
Number of Living Relatives: 11
Actively Breaking Law: False
---------------------------------------
Name: Oliver Williams
Age: 24
Gender: M
Weight: 165.5 lbs
Education Level: 12
Occupation: Police Officer
Is Doctor: False
Employed: True
Wealth Class: LOWER_MIDDLE
Married: False
Number of Children: 5
Number of Living Relatives: 15
Actively Breaking Law: False
---------------------------------------
pedestrians

Name: Amelia Davis
Age: 77
Gender: M
Weight: 207.2 lbs
Education Level: 20
Occupation: Journalist
```

```
Is Doctor: False
Employed: True
Wealth Class: MIDDLE
Married: False
Number of Children: 3
Number of Living Relatives: 15
Actively Breaking Law: False
---------------------------------------
Name: Charlotte Garcia
Age: 78
Gender: M
Weight: 172.7 lbs
Education Level: 14
Occupation: Police Officer
Is Doctor: False
Employed: True
Wealth Class: UPPER_MIDDLE
Married: False
Number of Children: 5
Number of Living Relatives: 4
Actively Breaking Law: False
---------------------------------------
Name: Elijah Anderson
Age: 8
Gender: F
Weight: 56.0 lbs
Education Level: 12
Occupation: Student
Is Doctor: False
Employed: True
Wealth Class: UPPER_MIDDLE
Married: False
Number of Children: 0
Number of Living Relatives: 5
Actively Breaking Law: False
---------------------------------------
tot_passenger_value 3691
avg_passenger_value 1845.5
tot_pedestrian_value 5124
avg_pedestrian_value 1708.0
tot_passenger_value 2491
avg_passenger_value 1245.5
```

```
tot_pedestrian_value 3184
avg_pedestrian_value 1061.3333333333333
tot_passenger_value 4331
avg_passenger_value 2165.5
tot_pedestrian_value 6364
avg_pedestrian_value 2121.3333333333335
num_passengers:  2
num_pedestrians:  3
Western Decision:  Decision.PROTECT_PEDESTRIANS
Eastern Decision:  Decision.PROTECT_PEDESTRIANS
Southern Decision:  Decision.PROTECT_PEDESTRIANS


(base) @plestz ➜ .../moral-machine-analysis/av-final/web/vehicle_architecture
(main) $ python3 smart_car.py

passengers

Name: Charlotte Martinez
Age: 1
Gender: M
Weight: 11.7 lbs
Education Level: 9
Occupation: Student
Is Doctor: False
Employed: True
Wealth Class: UPPER_MIDDLE
Married: False
Number of Children: 5
Number of Living Relatives: 4
Actively Breaking Law: False
---------------------------------------
Name: Amelia Rodriguez
Age: 83
Gender: F
Weight: 163.8 lbs
Education Level: 14
Occupation: Police Officer
Is Doctor: False
Employed: True
Wealth Class: IMPOVERISHED
Married: True
```

```
Number of Children: 0
Number of Living Relatives: 0
Actively Breaking Law: False
---------------------------------------
pedestrians

Name: Evelyn Jackson
Age: 98
Gender: Non-binary
Weight: 128.1 lbs
Education Level: 12
Occupation: Scientist
Is Doctor: False
Employed: True
Wealth Class: IMPOVERISHED
Married: False
Number of Children: 2
Number of Living Relatives: 6
Actively Breaking Law: False
---------------------------------------
Name: Isabella Davis
Age: 50
Gender: F
Weight: 226.0 lbs
Education Level: 16
Occupation: Police Officer
Is Doctor: False
Employed: True
Wealth Class: MIDDLE
Married: True
Number of Children: 5
Number of Living Relatives: 7
Actively Breaking Law: False
---------------------------------------
tot_passenger_value 3145
avg_passenger_value 1572.5
tot_pedestrian_value 3558
avg_pedestrian_value 1779.0
tot_passenger_value 1985
avg_passenger_value 992.5
tot_pedestrian_value 2438
avg_pedestrian_value 1219.0
```

```
tot_passenger_value 3705
avg_passenger_value 1852.5
tot_pedestrian_value 4038
avg_pedestrian_value 2019.0
num_passengers:  2
num_pedestrians:  2
Western Decision:  Decision.PROTECT_PEDESTRIANS
Eastern Decision:  Decision.PROTECT_PEDESTRIANS
Southern Decision:  Decision.PROTECT_PEDESTRIANS
```

Appendix 2

In the diagram below, we see a detailed scenario in which the Autonomous Vehicle opts to save the pedestrians, even though there are more passengers. This is a unique edge case, which arises when the calculated human value of the smaller group is significantly higher than the calculated human value of the larger group.

```
passengers

Name: Sophia Gonzalez
Age: 41
Gender: F
Weight: 218.4 lbs
Education Level: 6
Occupation: Scientist
Is Doctor: False
Employed: True
Wealth Class: MIDDLE
Married: True
Number of Children: 3
Number of Living Relatives: 4
Actively Breaking Law: False
---------------------------------------
Name: Charlotte Hernandez
Age: 62
Gender: Non-binary
Weight: 102.0 lbs
Education Level: 10
Occupation: Engineer
Is Doctor: False
Employed: True
Wealth Class: LOWER
Married: False
Number of Children: 0
Number of Living Relatives: 11
Actively Breaking Law: False
---------------------------------------
Name: Emma Miller
Age: 22
Gender: Non-binary
Weight: 152.4 lbs
Education Level: 17
Occupation: Salesperson
```

```
Is Doctor: False
Employed: True
Wealth Class: LOWER
Married: False
Number of Children: 4
Number of Living Relatives: 5
Actively Breaking Law: True
---------------------------------------
pedestrians

Name: Noah Martin
Age: 40
Gender: F
Weight: 217.7 lbs
Education Level: 18
Occupation: Doctor
Is Doctor: True
Employed: True
Wealth Class: LOWER_MIDDLE
Married: False
Number of Children: 5
Number of Living Relatives: 11
Actively Breaking Law: False
---------------------------------------
Name: Liam Miller
Age: 74
Gender: F
Weight: 168.3 lbs
Education Level: 11
Occupation: Unemployed
Is Doctor: False
Employed: False
Wealth Class: MIDDLE
Married: False
Number of Children: 3
Number of Living Relatives: 11
Actively Breaking Law: False
---------------------------------------
tot_passenger_value 4647
avg_passenger_value 1549.0
tot_pedestrian_value 3831
avg_pedestrian_value 1915.5
```

```
tot_passenger_value 2947
avg_passenger_value 982.3333333333334
tot_pedestrian_value 2631
avg_pedestrian_value 1315.5
tot_passenger_value 5407
avg_passenger_value 1802.3333333333333
tot_pedestrian_value 4471
avg_pedestrian_value 2235.5
num_passengers:  3
num_pedestrians:  2
Western Decision:  Decision.PROTECT_PEDESTRIANS
Eastern Decision:  Decision.PROTECT_PEDESTRIANS
Southern Decision:  Decision.PROTECT_PEDESTRIANS
```

Appendix 3

In the diagram below, we see detailed scenarios in which the Autonomous Vehicle differs in its decision on who to save based on its geographical location. Such divergences are important to understand, especially for those who (1) program vehicles in these different regions, and (2) travel in different regions.

```
passengers

Name: Sophia Johnson
Age: 59
Gender: F
Weight: 124.3 lbs
Education Level: 14
Occupation: Salesperson
Is Doctor: False
Employed: True
Wealth Class: UPPER_MIDDLE
Married: True
Number of Children: 2
Number of Living Relatives: 15
Actively Breaking Law: False
---------------------------------------
pedestrians

Name: Logan Miller
Age: 85
Gender: M
Weight: 220.8 lbs
Education Level: 5
Occupation: Journalist
Is Doctor: False
Employed: True
Wealth Class: LOWER_MIDDLE
Married: False
Number of Children: 5
Number of Living Relatives: 10
Actively Breaking Law: False
---------------------------------------
Name: Noah Garcia
Age: 79
Gender: F
Weight: 111.6 lbs
```

```
Education Level: 10
Occupation: Scientist
Is Doctor: False
Employed: True
Wealth Class: LOWER_MIDDLE
Married: False
Number of Children: 5
Number of Living Relatives: 5
Actively Breaking Law: False
---------------------------------------
tot_passenger_value 2067
avg_passenger_value 2067.0
tot_pedestrian_value 3615
avg_pedestrian_value 1807.5
tot_passenger_value 1407
avg_passenger_value 1407.0
tot_pedestrian_value 2455
avg_pedestrian_value 1227.5
tot_passenger_value 2507
avg_passenger_value 2507.0
tot_pedestrian_value 4175
avg_pedestrian_value 2087.5
num_passengers:  1
num_pedestrians:  2
Western Decision:  Decision.PROTECT_PEDESTRIANS
Eastern Decision:  Decision.PROTECT_PEDESTRIANS
Southern Decision:  Decision.PROTECT_PASSENGERS
```

```
passengers

Name: Lucas Smith
Age: 53
Gender: F
Weight: 215.6 lbs
Education Level: 2
Occupation: Journalist
Is Doctor: False
Employed: True
Wealth Class: MIDDLE
Married: False
```

```
Number of Children: 1
Number of Living Relatives: 8
Actively Breaking Law: True
---------------------------------------
Name: Olivia Jackson
Age: 34
Gender: Non-binary
Weight: 220.3 lbs
Education Level: 16
Occupation: Programmer
Is Doctor: False
Employed: True
Wealth Class: LOWER_MIDDLE
Married: False
Number of Children: 3
Number of Living Relatives: 6
Actively Breaking Law: False
---------------------------------------
Name: Emma Moore
Age: 87
Gender: M
Weight: 150.5 lbs
Education Level: 20
Occupation: Manager
Is Doctor: False
Employed: True
Wealth Class: LOWER
Married: False
Number of Children: 2
Number of Living Relatives: 3
Actively Breaking Law: False
---------------------------------------
Name: Amelia Jones
Age: 91
Gender: F
Weight: 168.5 lbs
Education Level: 17
Occupation: Police Officer
Is Doctor: False
Employed: True
Wealth Class: MIDDLE
Married: True
```

```
Number of Children: 2
Number of Living Relatives: 14
Actively Breaking Law: False
---------------------------------------
pedestrians

Name: William Jackson
Age: 26
Gender: M
Weight: 168.3 lbs
Education Level: 1
Occupation: Scientist
Is Doctor: False
Employed: True
Wealth Class: LOWER_MIDDLE
Married: True
Number of Children: 2
Number of Living Relatives: 7
Actively Breaking Law: False
---------------------------------------
Name: Lucas Jones
Age: 56
Gender: F
Weight: 170.1 lbs
Education Level: 13
Occupation: Scientist
Is Doctor: False
Employed: True
Wealth Class: IMPOVERISHED
Married: False
Number of Children: 4
Number of Living Relatives: 7
Actively Breaking Law: False
---------------------------------------
Name: Emma Williams
Age: 27
Gender: M
Weight: 137.8 lbs
Education Level: 3
Occupation: Doctor
Is Doctor: True
Employed: True
```

```
Wealth Class: MIDDLE
Married: False
Number of Children: 4
Number of Living Relatives: 9
Actively Breaking Law: False
---------------------------------------
tot_passenger_value 6212
avg_passenger_value 1553.0
tot_pedestrian_value 5542
avg_pedestrian_value 1847.3333333333333
tot_passenger_value 3852
avg_passenger_value 963.0
tot_pedestrian_value 3842
avg_pedestrian_value 1280.6666666666667
tot_passenger_value 7412
avg_passenger_value 1853.0
tot_pedestrian_value 6302
avg_pedestrian_value 2100.6666666666665
num_passengers:  4
num_pedestrians:  3
Western Decision:  Decision.PROTECT_PASSENGERS
Eastern Decision:  Decision.PROTECT_PEDESTRIANS
Southern Decision:  Decision.PROTECT_PASSENGERS
```