Appendix 1. A list of variables in this research after the cleaning and clustering steps

vehiclecomfort (low, high)

vehiclesafety (low, high)

brakes (regular brakes, ABS brakes)

qualityroad (low, high)

road (urbanroad, ruralroad)

neuroticism (low, high)

extraversion (low, high)

openness (low, high)

agreeableness (low, high)

conscientiousness (low, high)

sensationseeking (low, high)

normlessness (low, high)

lapse (low, high)

sliperror (low, high)

aggressiveviolation (low, high)

violation (low, high)

lawenforcement: the extent of law enforcement (low, high)

enforcelawtype: how to enforce the law (none, police, both police and camera)

sex (men, women)

age (lessthan29, 30-49, 50-64, morethan65)

marital (singel, married, divorced or widowed)

education (non-university degree, degree university)

income (less than the cost, equal to the cost, more than cost)

motorcycle (No, Yes)

lightvehicle (No, Yes)

heavyvehicle (No, Yes)

workhoursperday: working time in hours per day (4 hours or less, 4 to 8 hours, more than 8 hours)

drivehrsperday: driving time in hours per day (drive 2 hours or less, drive 2 to 4 hours, drive more than 4 hours)

nonstopdrive: duration of non-stop driving in hours per day (drive non-stop 2 hours or less, drive non-stop 2 to 4 hours, drive non-stop more than 4 hours)

foreigncarornot (non-foreign (Iranian-made) car, foreign car)

fatiguedriving: driving while tired and sleepy (low, high)

cellphone: using a cellphone while driving (low, high)

alcoholcannabis: alcohol or cannabis use (No, Yes)

anydrug: taking any psychoactive and stimulant drug (No, Yes)

accident: occurrence or non-occurrence of an accident (No, Yes)

accidenttype: the severity of the accident (Damage, Wounded or Fatal)

The variables accident and accidenttype are dependent and target variables.