**CORRELATION ANALYSIS USING HEATMAP**

**University of Pisa, Pisa Italy**

**Department of Data Science & Business Informatics**



In our heat-map, there is a strong positive correlation by all MMR values i.e. [1, 0.99, 0.9+ or 0.8+]. Vehicle Age has only one strongly positive correlation that is to itself. In other words, Vehicle Age has a moderate negative correlation with all the MMR variables and VehBCost. In VehBCost we observe that their is fairly strong correlation with all the MMR values and a moderately negatively correlation with VehicleAge.

Based on our observation and analysis we therefore conclude that there is strong linear relationship among all the attributes apart from VehicleAge which have a moderate negative relationship.



From our analysis of the first heat-map correlation above we realized the need to keep only one MMR variable plus either VehicleAge or VehBcost.

In this heat-map visualization we observe that there is only strongly positive correlation between two distinct variables of MMR\_Factor\_n and VehBCOst\_n which is 0.8. Though all the variables are perfectly positively correlated to themselves i.e. 1. Our further analysis shows that all the remaining variables are weakly correlated (either positive or negative).

Based on our observation and analysis we conclude that there is little relationship among all the variables in the heat-map correlation except MMR\_Factor\_n and VehBCOst\_n of 0.8 where we see strong relationship.