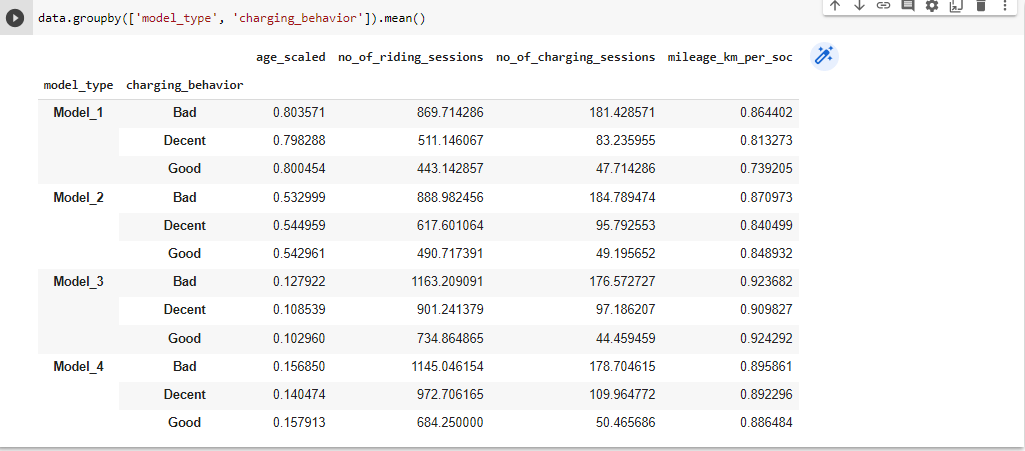
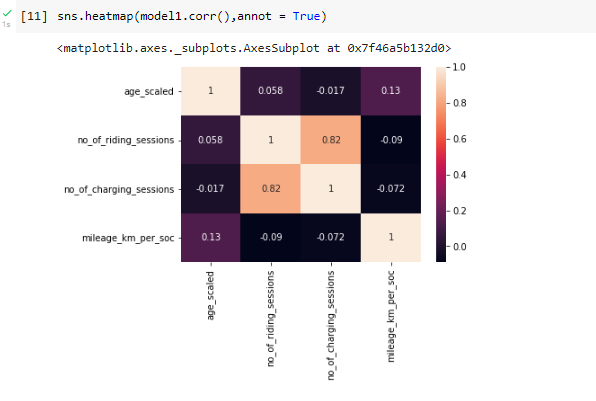
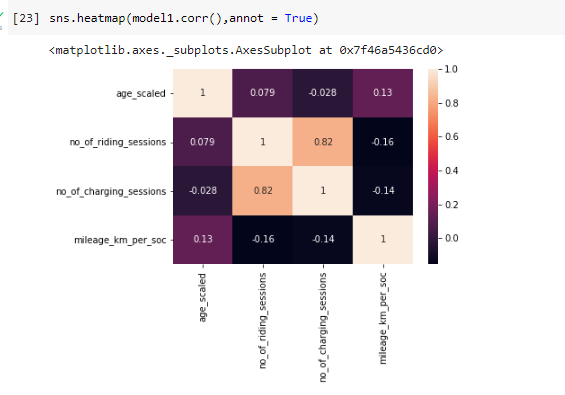
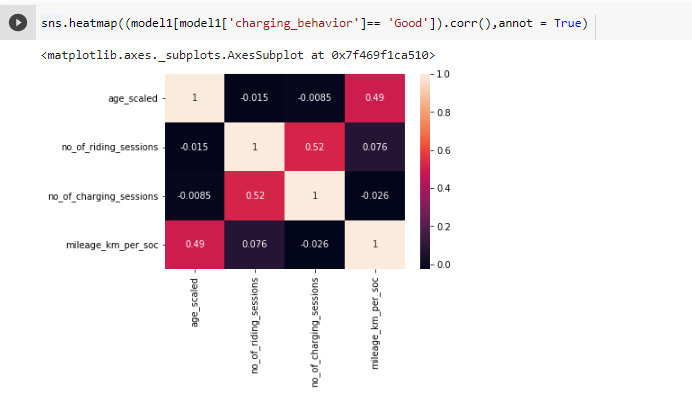
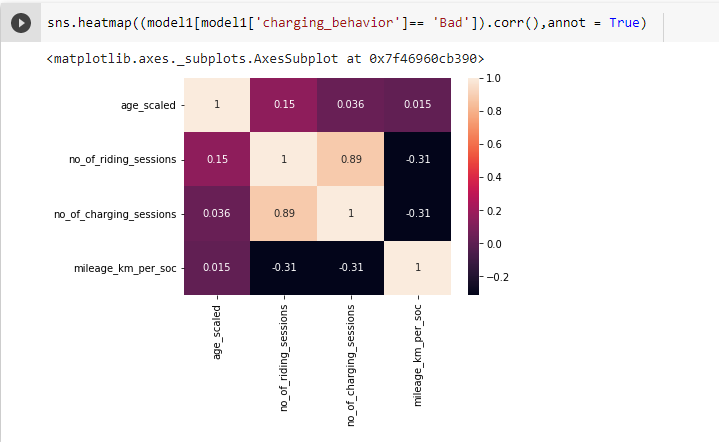
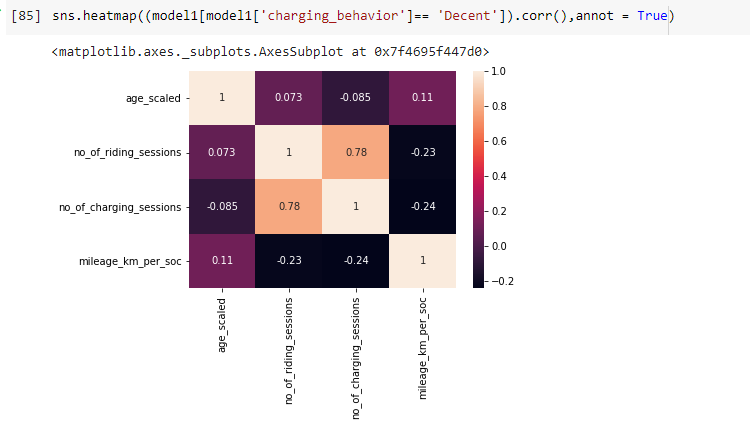
* **Assumption**: It is not fair to come to conclusion regarding influence on mileage by given attributes according to me. Because each vehicle model might have other attributes which are influencing mileage like vehicle weight, battery power, efficiency etc. So, I have decided to make assumptions on mileage by sub-setting data based on Vehicle models and further check the performance of each model with respect to charging behaviour.
* **Additional attributes that could improve the study**: Battery storage power, Vehicle weight, Battery degradation with respect to speed, riding behaviour as riding style is necessary while predicting mileage.
* **The limitations of the approach:** If all the models having same specifications.
* **Results:**



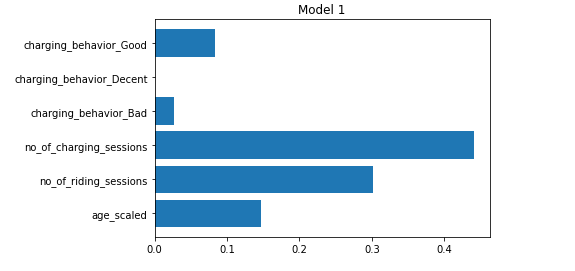
* **Model 1 dataset**
* Using scatter plot against age\_scaled, no\_of\_charging\_sessions and no\_of\_riding\_sessions vs mileage. Removed some data points which I feel they look like outliers. After removing some data points the correlation between mileage and other attributes has been changed with respect to riding sessions and battery sessions.



* We can see there is some correlation between Age scale and Mileage. Let’s subset the model1 dataset based on charging behaviour.
* Subsetting model 1 data charging behaviour == Good, Bad and Decent.
*  **Model 1** vehicles with **Good** charging behaviour age scaled attribute is **positively correlated with the mileage.**
* 

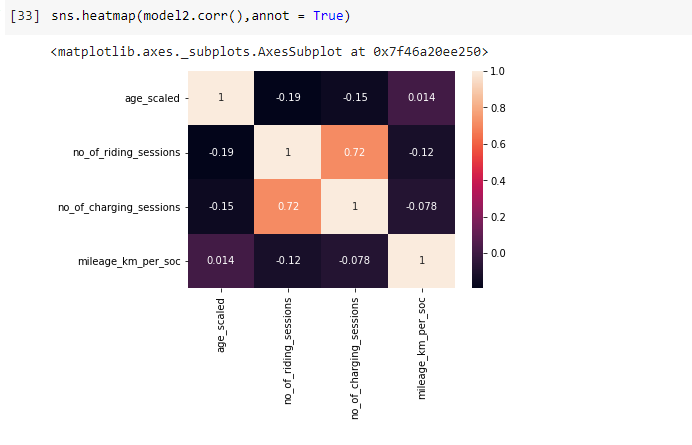


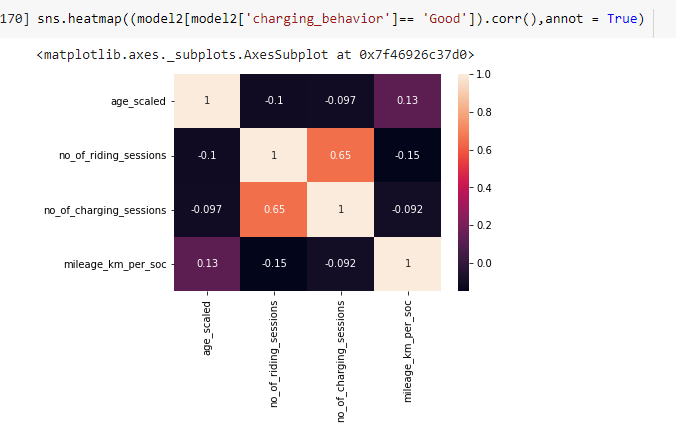
* **Model 1** vehicles with **Bad** charging behaviour with riding and charging attributes are **negatively correlated with the mileage.**
* **Model 1** vehicles with **Decent** charging behaviour with riding and charging attributes **are negatively correlated and age\_scaled is positively correlated with** the **mileage.**
* When the model 1 dataset is trained by Decision Tree regressor and this is the feature importance.



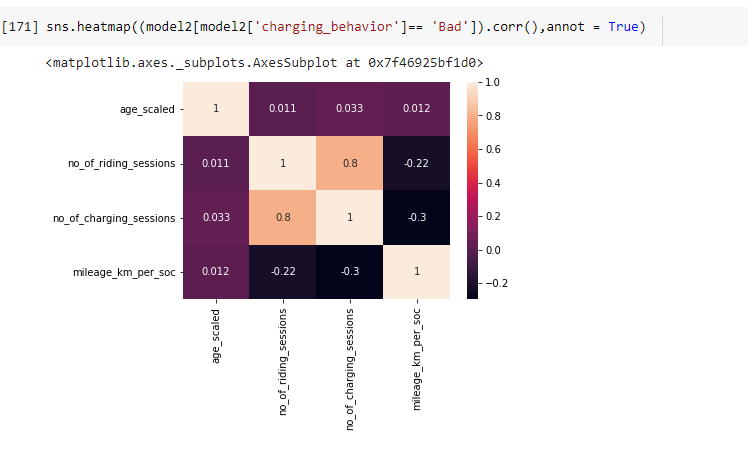
* No\_of\_charging\_session has the most influence on mileage followed by riding sessions and age scaled. Vehicles with good charging behaviour tends to have more mileage than vehicles with bad and decent charging behaviour.

**Model 2**

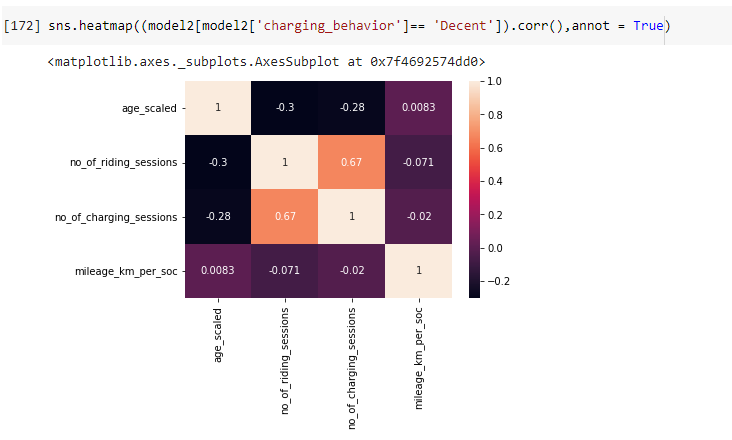
* Using scatter plot against age\_scaled, no\_of\_charging\_sessions and no\_of\_riding\_sessions vs mileage. Removed some data points which I feel they look like outliers



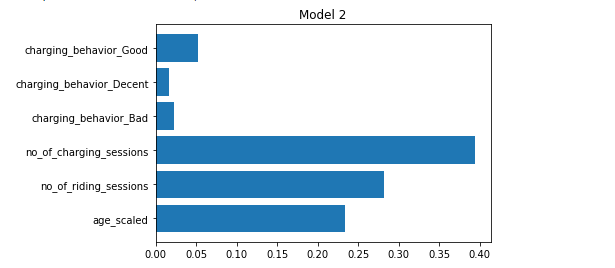
* **Model 2** vehicles with **Good** charging behaviour **age scaled** attribute is **positively correlated** with the **mileage** and **mileage is negatively correlated** with the no of riding sessions**.**



* **Model 2** vehicles with **Bad** charging behaviour **negatively correlated** with the no of **riding sessions.**
* **Model 2** vehicles with **Decent** charging behaviourhasn’t shown any correlation with any of the attributes.

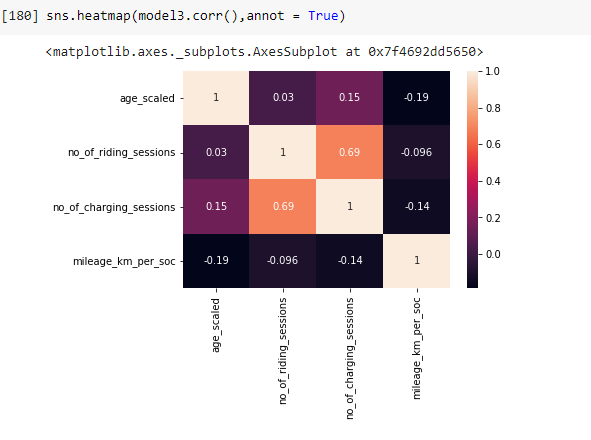


* When the model 2 dataset is trained by Decision Tree regressor and this is the feature importance.

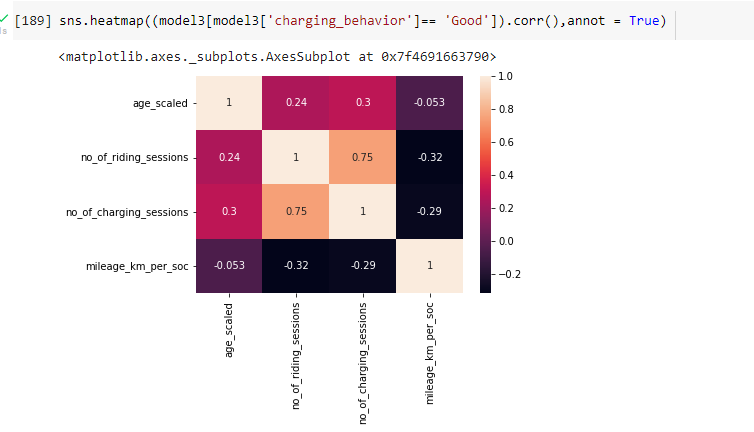


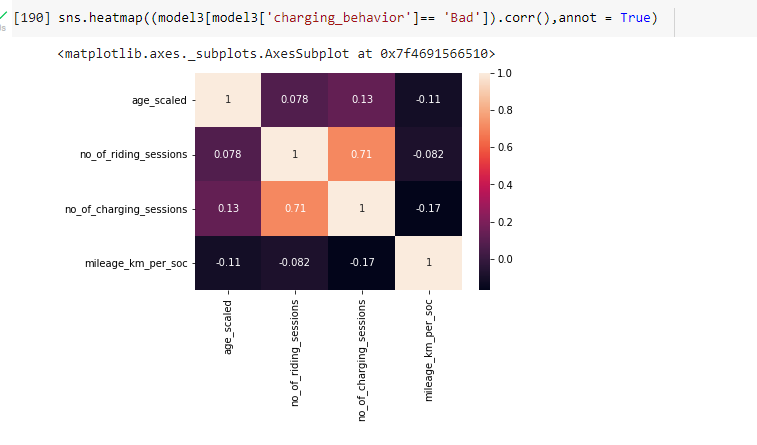
* No\_of\_charging\_session has the most influence on mileage followed by riding sessions and age scaled. Vehicles with good charging behaviour tends to have more mileage than vehicles with bad and decent charging behaviour.

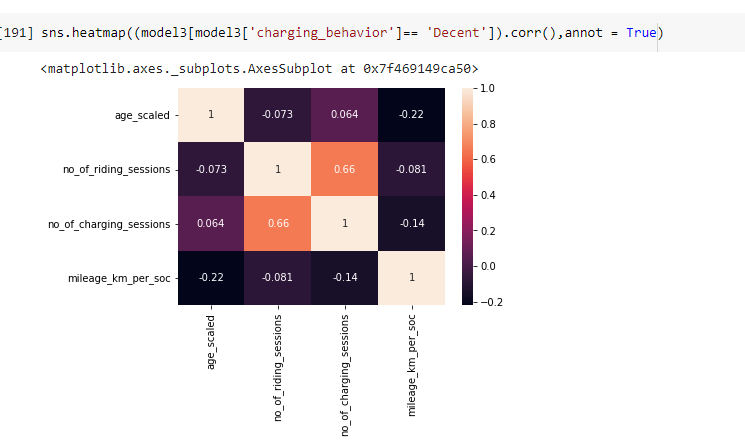
**Model 3**

****

* Using scatter plot against age\_scaled, no\_of\_charging\_sessions and no\_of\_riding\_sessions vs mileage. Removed some data points which I feel they look like outliers.



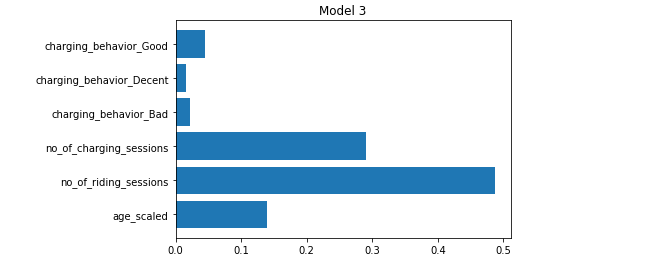
* **Model 3** vehicles with **Good** charging behaviour, **no\_riding\_sessions** and **No\_of\_charging\_sessions** attribute are **negatively correlated** with the **mileage.**
* **Model 3** vehicles with **Bad** charging behaviour, **age\_scaled** and

**No\_of\_charging\_sessions** attribute is **negatively correlated** with the **mileage.**

* **Model 3** vehicles with **Decent** charging behaviour **age\_scaled** and

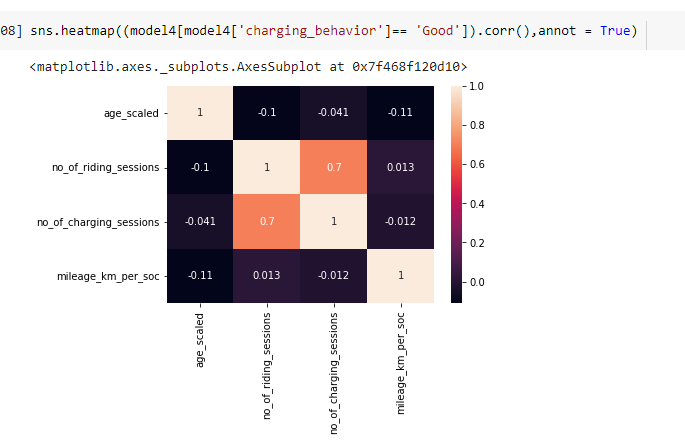
**No\_of\_charging\_sessions** attribute is **negatively correlated** with the **mileage.**

* When the model 3 dataset is trained by Decision Tree regressor and this is the feature importance.

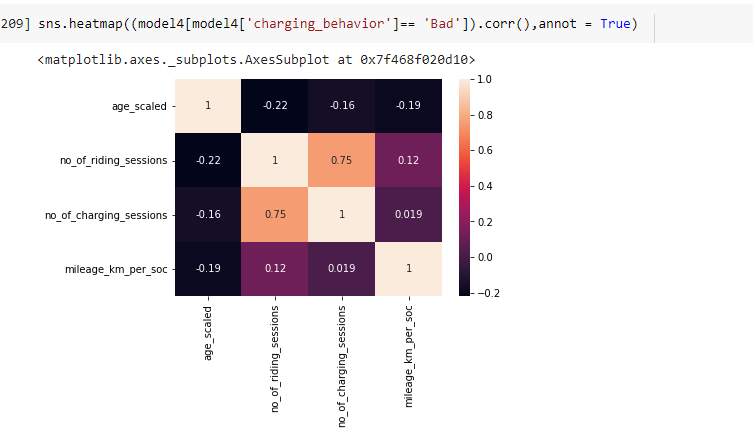
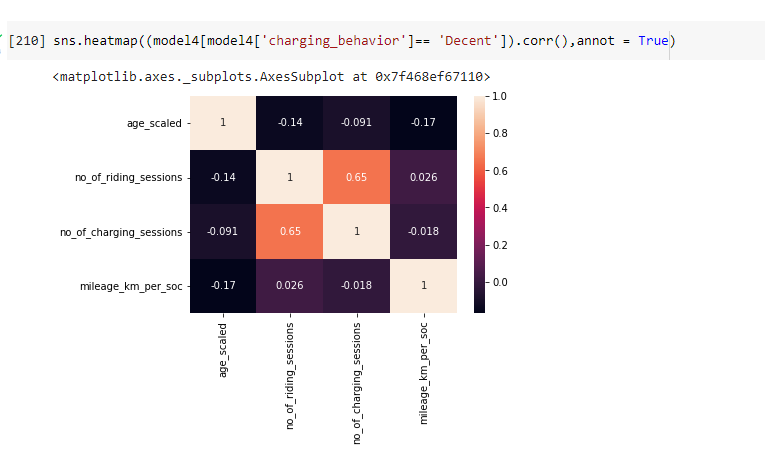


* No\_of\_riding\_session has the most influence on mileage followed by charging sessions and age scaled. Vehicles with good charging behaviour tends to have more mileage than vehicles with bad and decent charging behaviour.

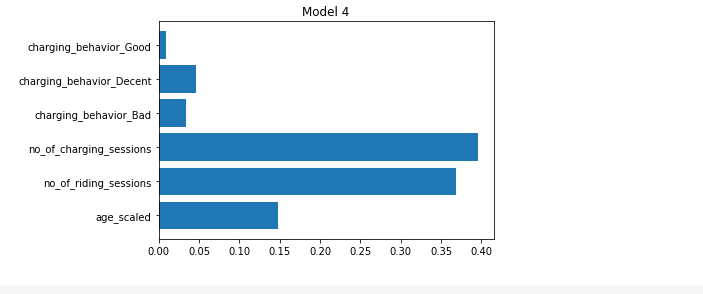
**Model 4**



* **Model 4** vehicles with **Good** charging behaviour, **age scaled** is **negatively** correlated with **mileage**.
* **Model 4** vehicles with **Decent**  charging behaviour, **age scaled** is **negatively** correlated and **no\_of\_riding\_sessions** are **positively** correlated with **mileage**.



* **Model 4** vehicles with **Bad** charging behaviour, **age scaled** is **negatively** correlated and **no\_of\_riding\_sessions** are **positively** correlated with **mileage**.
* When the model 4 dataset is trained by Decision Tree regressor and this is the feature importance.



* No\_of\_charging\_session has the most influence on mileage followed by riding sessions and age scaled. Vehicles with decent charging behaviour tends to have more mileage than vehicles with bad and decent charging behaviour.