

In the Question 1, we analyze the dataset which were having 256250 number of rows, we drop the rows which had the blank columns and analyzed the different types of entities and observe that the maximum of the entities were falling under the four labels: **component, failure issue, vehicle model and corrective action.**

For analysing the entities in the dataset, we took 4 columns which have the names as mention below

**DEFECT SUMMARY, CONSEQUENCE SUMMARY,
CORRECTIVE SUMMARY, RECALL NOTES**

For further task we used **Mistral 7b and LLaMa 2 7b.**