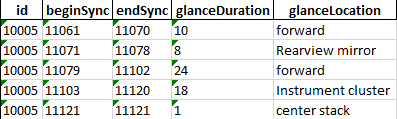
- outlining all the categories that we used

- outlining importance of deriving driver’ behaviour or environment state from the glances

- future work(?)

**Preprocessing part**

The dataset we are using has very detailed description about the glance location and we decided to group them to first have a general understanding of them. For example, we grouped ‘left mirror’, ‘left window’, ‘left forward’ as ‘left’ and some rarely appeared glance location as other. There are five columns in the dataset namely ‘id’, ‘beginSync’, ‘endSync’, ‘glanceDuration’ and ‘glanceLocation’. The columns that we are most interested in is ‘id’ which identifies the user, ‘glanceDuration’ which indicates the time spent glancing at a certain direction and ‘glanceLocation’ indicating where the user is glancing at. There are 61 states for each user and a new row is appended whenever a user changes his or her state as shown below. We decided to combine the separate rows in a single row by representing the glance data in a list with 61 items which represents the 61 states.



**Why the fuck is this important**

In Singapore, there are roughly 7700 traffic accidents that resulted in injuries in 2017. We think that it is extremely important to find a better way to predict accidents. Glance Data is easily obtainable by simply installing few cameras in the vehicle which can be easily implemented. A short period of time is needed as it only requires 61 states to come up with a prediction. We believe that this project can be useful at predicting accidents and hence prevent them.