Name: Samar Ibrahim Antar

Assignment #3: Decision Trees

- **Binary_split function**: for Splitting a dataset into 2 lists (groups) of rows (left and right) given the index of feature and a split value for that feature.
- **Gini_cost function:** calculating the gini and cost function that is minimized in classification for the list of groups & list of known class values.
- **Best_split function:** to obtain best split value for classifying data by checking the value on each feature and calculate the cost until find the best. (greedy search)
- **Child_splits function:** it takes node, max depth , min size or number of rows in a node and current depth to check if it should split again or stop and create a leaf node. (check stop criteria by max depth, min size or if either left or right group of rows is empty).
- **Build CART tree function:** Building the tree involves creating the root node and calling the Child_splits() function that then calls itself recursively to build out the whole tree. It takes train data, max depth and min size. So output of building a tree of max depth=2 and min size=1 for first 7000 rows of train data -because it takes very long time for all train data- showing in following figure.