ASSIGNMENT – 3

NAME:MAHASRI V

ROLL NO:CB.EN.U4CSE20235

CASE:

In 2010, the ACLU presented a report on jury selection in Alameda County, California. The report concluded that certain ethnic groups are underrepresented among jury panelists in Alameda County, and suggested some reforms of the process by which eligible jurors are assigned to panels.

GOAL:

To perform analysis of the data and examine whether the panel is selected on random basis.

NULL HYPOTHESIS:

Panels are selected at random from the population of eligible jurors.

ALTERNATE HYPOTHESIS:

Panels were not selected at random.

DATA ANALYSIS:

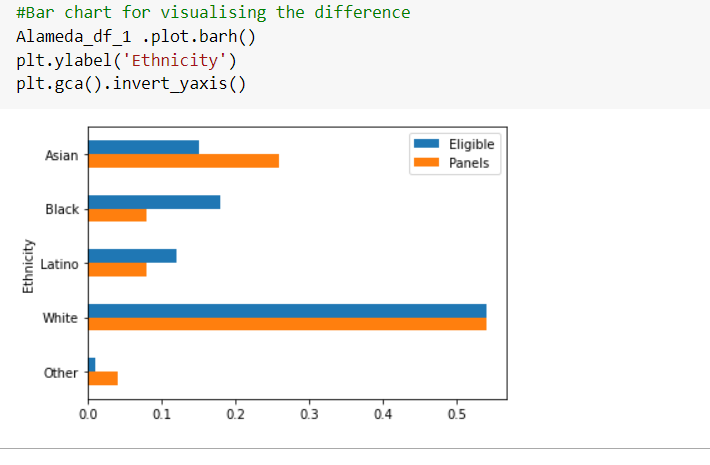
ETHNICITY: Various ethnic groups of Alamenda.

ELIGIBLE: proportion of all eligible juror candidates of that ethnicity.

PANELS: proportion of people of that ethnicity among those who appeared for the process of selection into the jury.

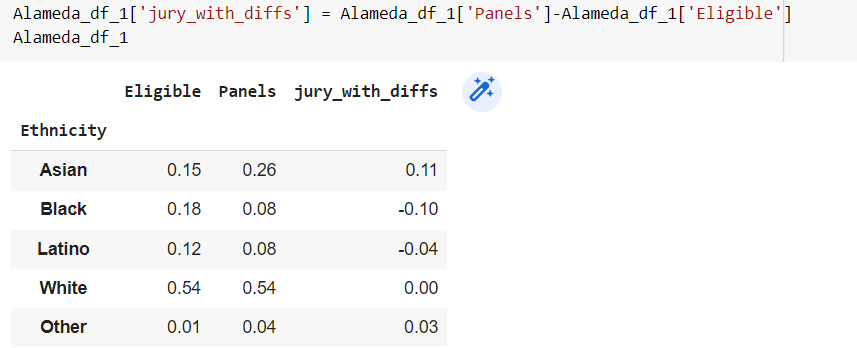


VARIATION IN ELIGIBILITY AND PANELS:



This graph shows that the difference between the panel and eligible candidates. We infer that they are not equally dominating in different ethnicity.

CALCULATING THE DIFFERENCE BETWEEN THE ELIGIBLE AND PANEL:

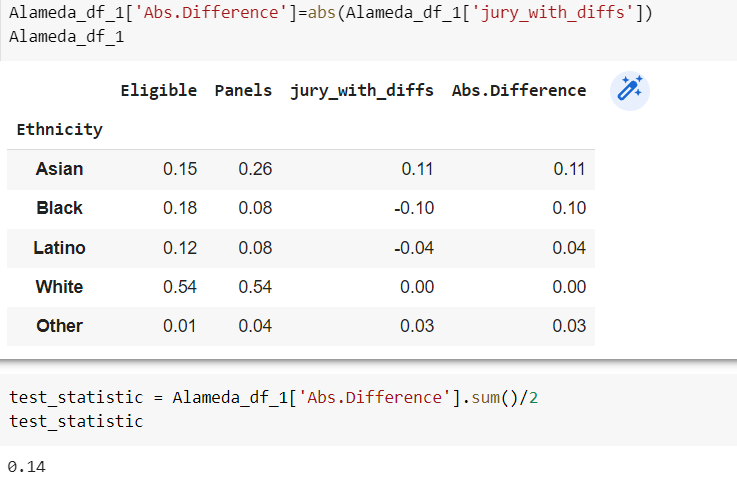


POSITIVE ENTRIES SUM=0.14

NEGATIVE ENTRIES SUM= -0.14

TOTAL SUM=0

To avoid the cancellation, we drop the negative signs and then add all the entries. But this gives us two times the total of the positive entries (equivalently, two times the total of the negative entries, with the sign removed). So we divide the sum by 2.

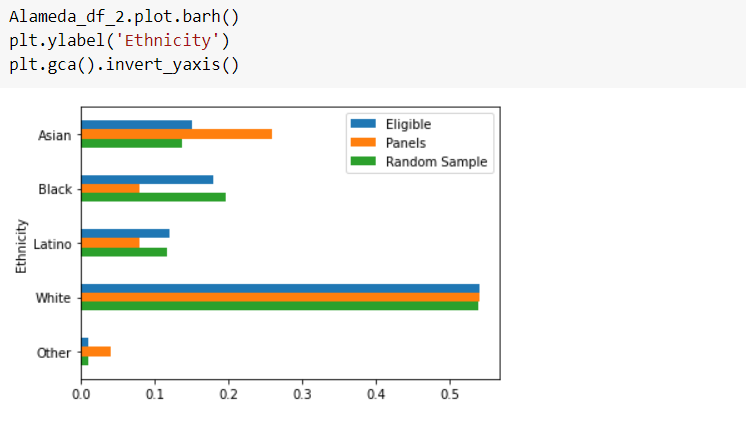


TOTAL VARIATION DISTANCE:0.14

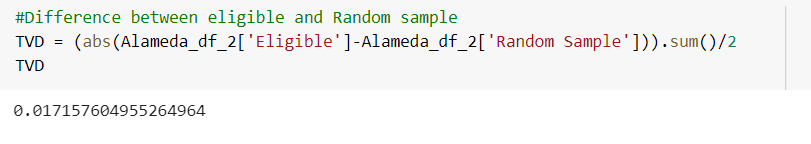
SAMPLING AT RANDOM TO FIND IF THE PANELS ARE REPRESENTATIVE OF THE POPULATION:

Sample size is taken as 1453 and multinomial is used to find the random sample.





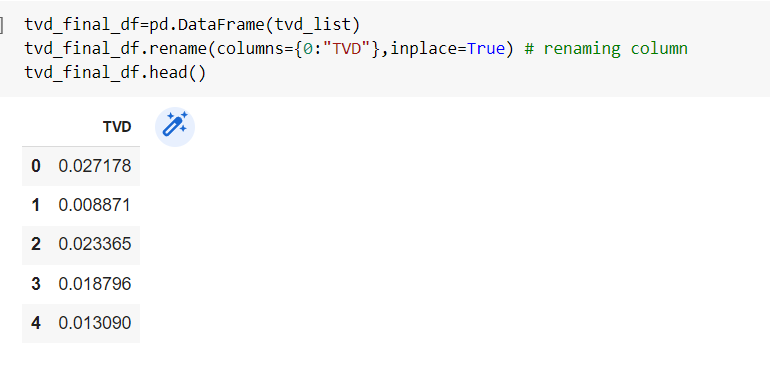
The random sample resembles the eligible population, but the panels don't since the distribution of the random sample is close to the distribution of the eligible population and is different from the distribution of the panels.

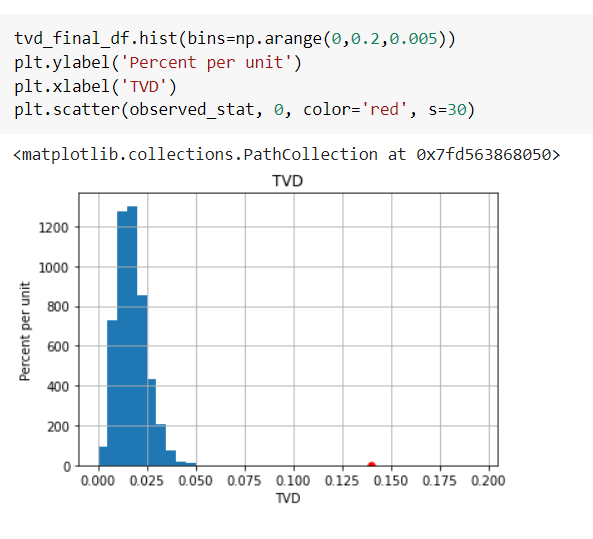


The difference between random sample and eligible is very small that is 0.017 whereas the difference between the random sample and panel is 0.14.



TOTAL VARIATION DISTANCE:





distance between the panels and the population was 0.14,so it is far away from the distribution. panels were not representative of the distribution provided for the eligible jurors.

CONCLUSION:

Null hypothesis is rejected since the panels did not represent the distribution of eligible jurors.