Data Analysis with Tableau

Rahul Ramesh Kumar
rrk310

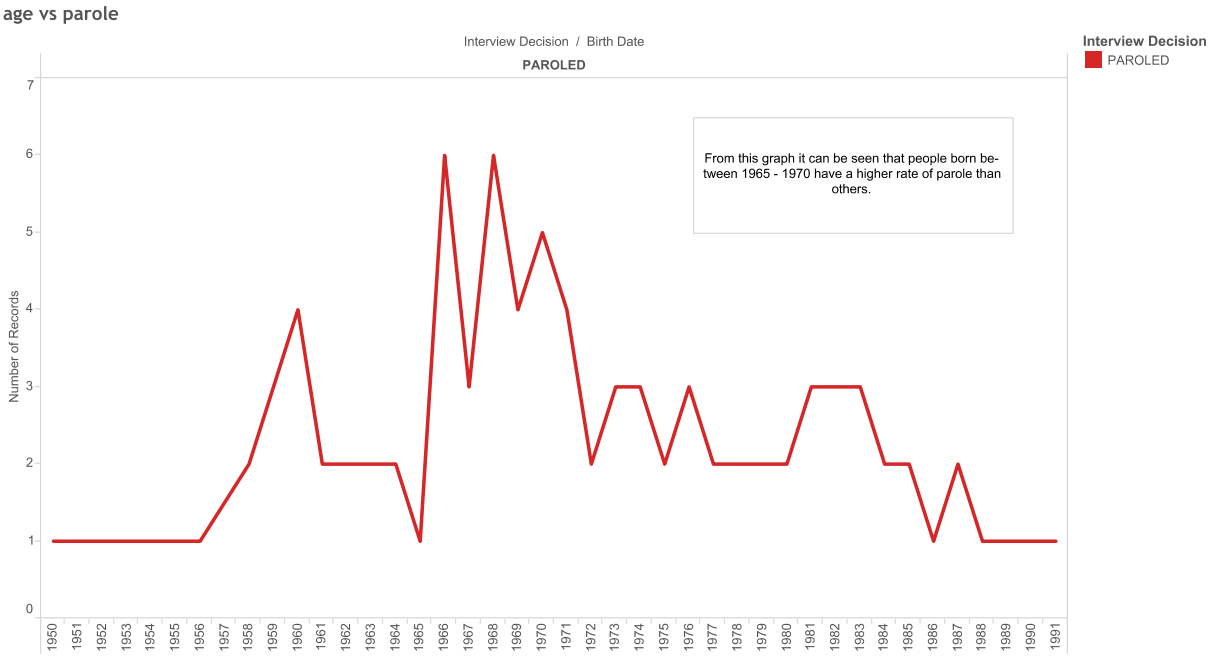
Goal

The following data analysis is of New York State Inmates at different facilities who have either been denied of parole due to various reasons or given parole. The goal of this data analysis is to help identify within the data, a clue to how parole decisions are made by the parole board or obtain a pattern, that will help Nikki Zeichner in her project 'parolehearingdata'. In order to find patterns, I have used information provided from the data set and explore this data using Tableau. One of my basis of analysis is to find a relation between parole decisions and characteristic of each inmate which can be Age, Gender & Ethnicity. I have also used information of interview types to determine who is more likely to get parole and also find patterns of parole hearing decisions across different facilities. Here I present my findings and hope this will help in further development of analysis and insights into the problem.

Analysis

Following information will contain analysis of the data set based on certain criteria's. The dataset is a csv file which is loaded and analyzed using Tableau. Each analysis will present the Visual Image of the data and some text about the analysis presented in the visual image. There are some numeric data which will be presented in key findings in order to support the statements made by me. All analysis is optimized for each page to help the reader.

Analysis 1 -



Analysis -1

Explanation -

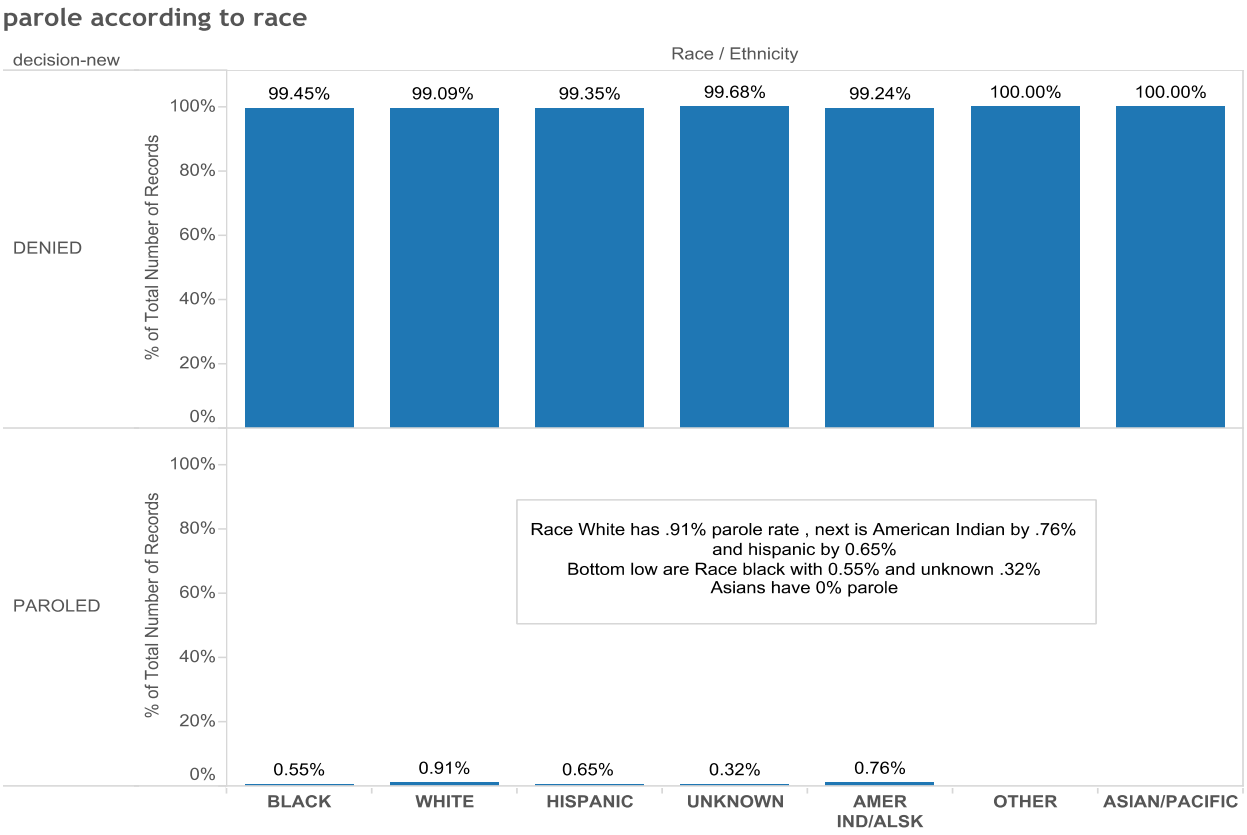
To start off, I wanted to see how well the parole data is distributed based on each inmate's birthdate. Also by using criteria – Interview Decision which is **Paroled**, a pattern emerged which is depicted by the trend line from above figure.

Findings –

1. The trend line shows that people born between 1965 – 1970 inclusive are more likely to get parole than the rest of the others. On basis of latest interview data which is in year 2013, these ages are between 43 - 48.
2. From 1973 – 1994 i.e., people aged between 40 -19 have shown a steady but less parole.
3. Minors have very less chance of getting parole as the trend line approaches 1991
4. Inmates aged between 53 -58 are very less likely to get parole as per this data.
5. The following graph has representations similar to a bell man curve.

From the analysis present there is a pattern in parole across different ages and this information is crucial to determine the success rate of a parole interview. How ever age can't only be one factor. I was also curios to find out how different races of people get parole which I feel is also a factor that I shall present in the next analysis.

Analysis 2 –



% of Total Number of Records for each Race / Ethnicity broken down by decision-new. The view is filtered on decision-new, which keeps DENIED and PAROLED. Percents are based on each column of the table.

Analysis -2

Explanation –

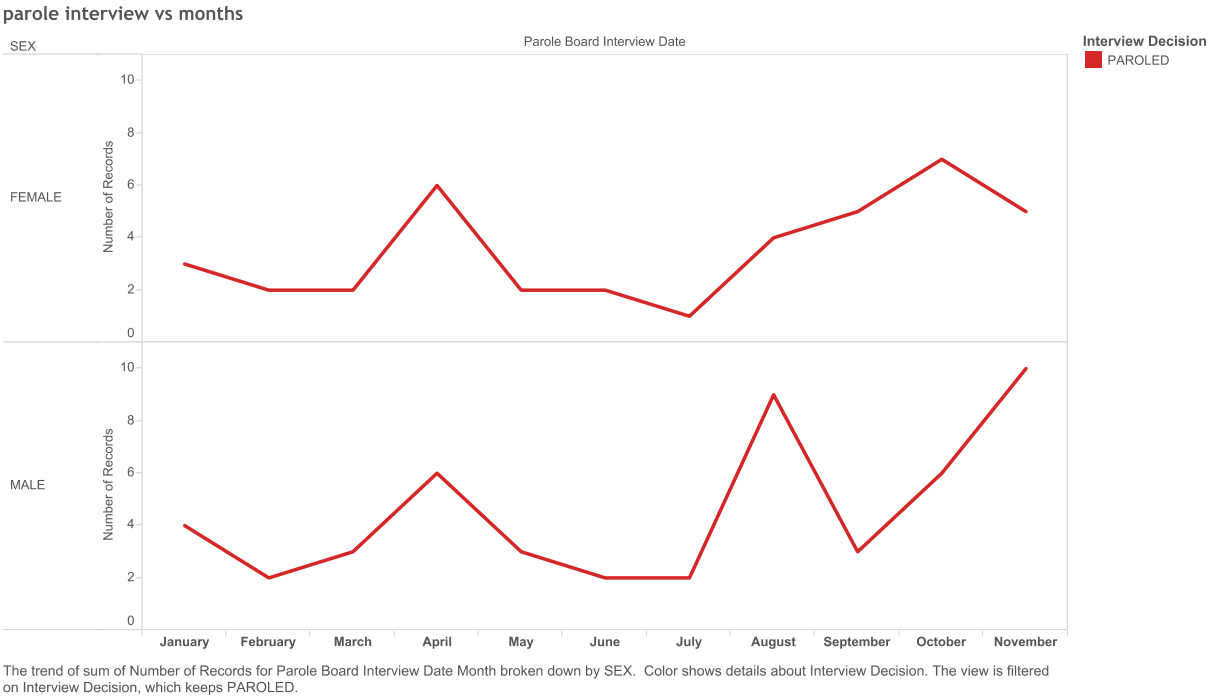
Although I was interested in finding how **Parole** decision is distributed across different races, I was also interested in finding the **Denied** decision distribution. Here in the above image there is a percentage of parole/denied decision on the total number of decisions made for each Race/ Ethnicity.

Here are some striking patterns and observations –

1. Inmates of White Race have the least Denied decisions (99.09%) across all Races.
2. Inmates of Asian and Other Ethnicity have been Denied 100% of parole.
3. From the Paroled graph, its clear that Inmates of White ethnicity have highest Parole decisions of (.91%) along with American Indian with (.76%).
4. Inmates of Black Ethnicity and Hispanic Ethnicity fill the lower level with 0.55% and 0.65% respectively.

From the following graph its clear that Inmates of White Ethnicity have highest rate of Parole (~1%) and Inmates of Asian/Pacific/Other have least rate of Parole (0%). I was next interested in distribution of the data with respect to Gender to find out interesting patterns.

Analysis 3-



Analysis -3

Explanation –

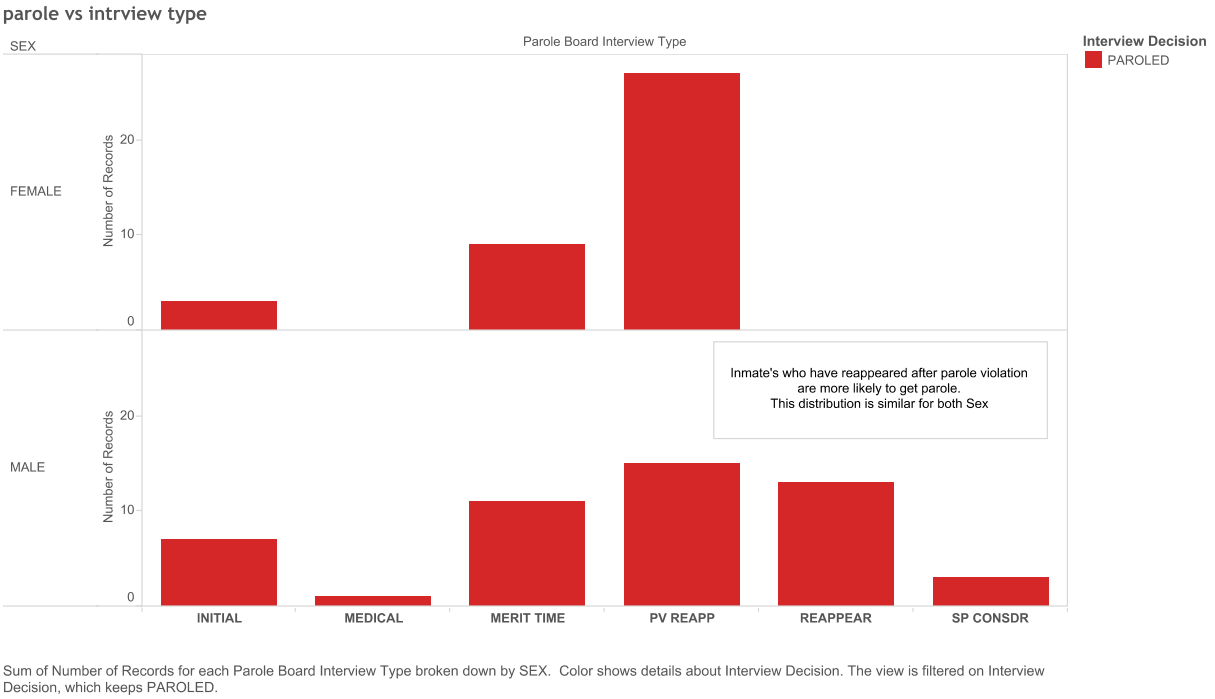
The basic purpose was to show distribution of parole decision for Genders, however I was also interested in knowing if success rate in interviews happen during any period in a year. Hence the chart above shows this distribution for Male and Female.

Following are the key findings-

1. Following line graph shows interesting similarity for both Gender's but this line is not the same in q3 and q4
2. Firstly, Male Gender have higher chances of obtaining Parole with Max(Male) = 10 and Max(Female) = <8
3. None of Genders obtained parole in the month of December and no Interviews on as well. This is probably owing to Christmas and New Year.
4. Parole success is highest in these months - April, August & November for both Gender approximately.
5. An unexplained dip occurs in the month of September for Male gender and there can be several reasons but there is a good chance that this is not a trend.

From these findings its clear that Male and Female obtain parole almost similarly and the parole board makes decisions which is highly successful during q2, q3 and half of q4. From this we can probably refine the findings further since some of them might be re appearing in that case we could account for what kind of interviews helped them achieve highest parole.

Analysis 4-



Analysis -4

Explanation –

Following distribution graph is similar but now accounting for what type of interviews accounted for maximum parole for both Gender?

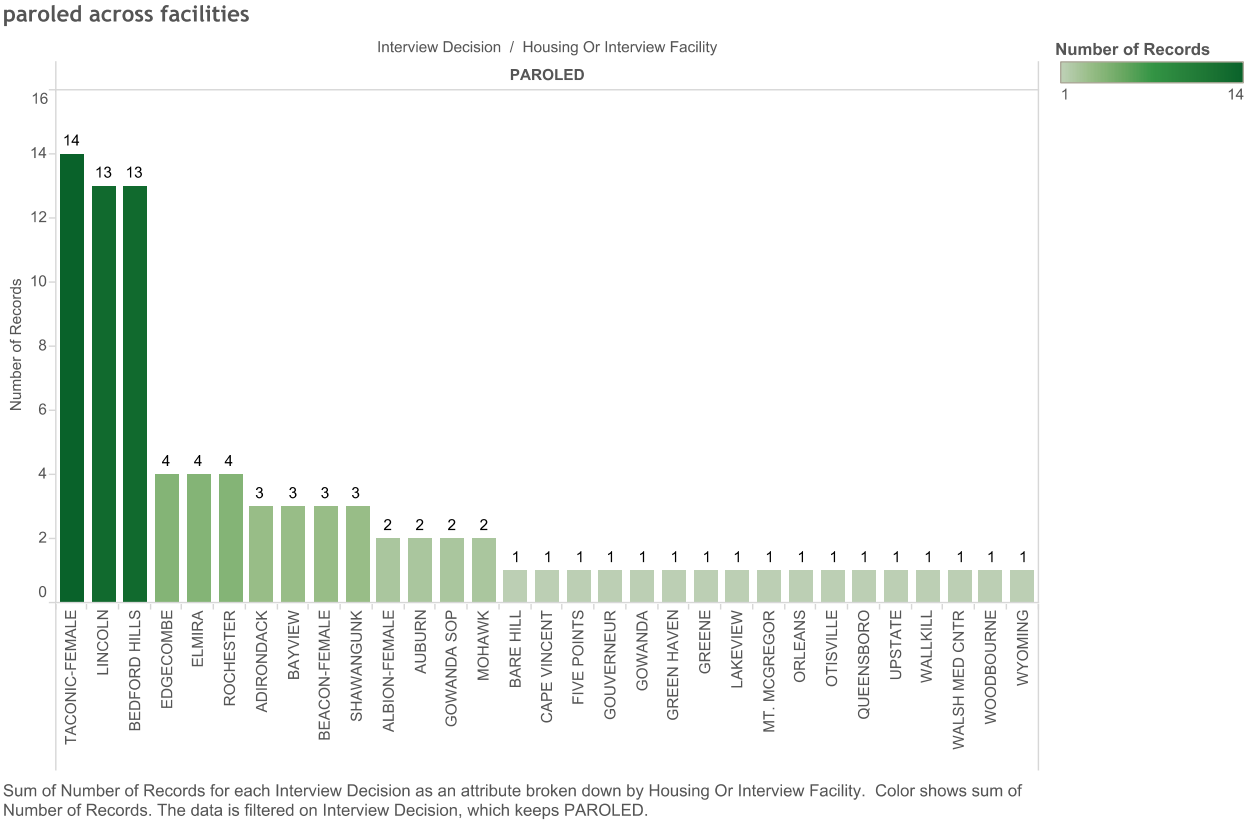
Here are the key findings –

1. From both the graphs its clear that Re appearance after Parole violation has produced the maximum parole for both Gender.
2. This is followed by Re appear which is similar to PV Re appear and Merit time.
3. Another important finding is that Female gender have higher success rate in obtaining Parole than men after Parole Violation Re appear.
4. While doing deeper analysis another interesting pattern appeared only for Female Gender where Females of Black ethnicity had much higher Parole rate than females of Other ethnicity after Parole violation re appearance. This was entirely different for Males.

The findings above have shown that Parole Board Interview Type has lot of effect on Parole decision which is clear from both the Gender distribution.

Some times we also have to take into account how parole decisions happen for inmates across different facility and see if there are any interesting patterns emerge there.

Analysis 5-



Analysis -5

Explanation –

Following graph shows a new distribution that shows the number for Parole decision of inmates that belong to these facilities. Here we want to know interview decision of inmates of different facilities.

These are the key findings –

1. Inmates from Taconic-Female, Lincoln, Bedford Hills have highest rate of Parole decision which is quite large compared to the rest.
2. A Deeper analysis also proves that these correction facilities are largely for Female inmates except for Lincoln.
3. Lincoln is the only Correction Facility for Male with highest Parole decisions and the next being Edgecombe.
4. Lincoln could be the only Correction facility for Male with highest Parole decision because it's a work-release center for drug offenders as per wiki.

From this data its clear that these 3 Facilities have highest parole and also that Female correction facilities have higher parole than Male correction facilities.

Conclusion-

I hope that from this analysis, the reader can get a clue into how parole decisions are made in New York Correction facilities. I also hope that New York correction facility can publish further data on duration of Parole hearing, Parole Interview board and some common questions and answers in interview to further narrow down on analysis.
I wish Nikki Zeichner best success in this project and would like to thank Professor Enrico Bertini for teaching visualization techniques, introducing me to Tableau and providing the data set required for the Data Analysis.

References-

1. http://www.parolehearingdata.org/
2. https://www.parole.ny.gov/calendardatadefinitions.html
3. https://en.wikipedia.org/wiki/Main_Page