

# Data Analysis with Tableau

---

Name: Jinglong Li

Net ID: jl7152

## 1. Introduction of goal of this analysis

---

In New York over 10,000 parole eligible prisoners are denied release every year. The consequences of these decisions are costly, while the process of how these determinations are made is unclear. Through analyzing this dataset from Nikki Zeichner, I hope to uncover new information about patterns in parole board determinations and useful information about how the parole system works in New York.

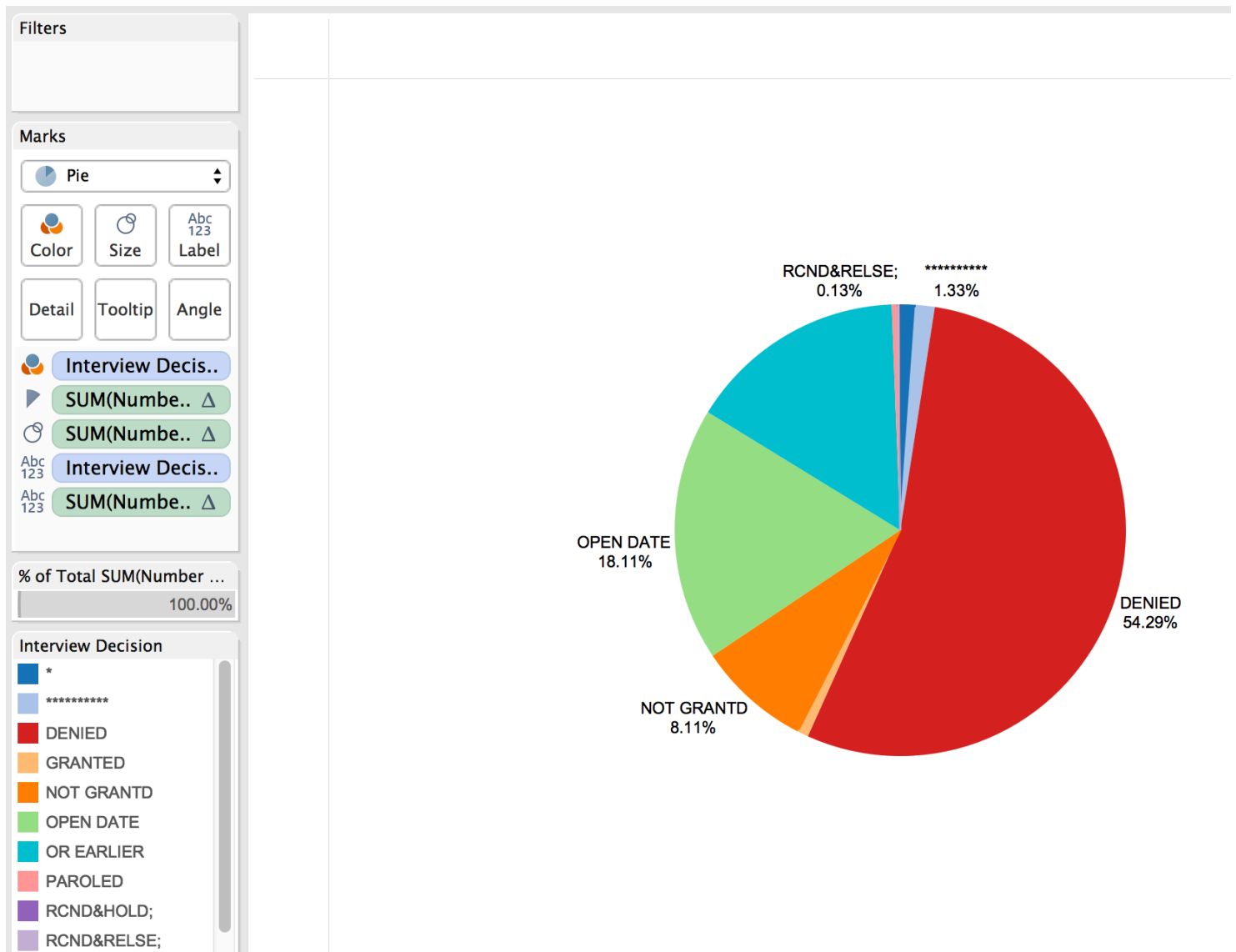
## 2. Analysis

---

Generally speaking, my approach is “overview first, zoom and filter, then details-on-demand”. So let’s start the analytical journey.

① Step 1:

---



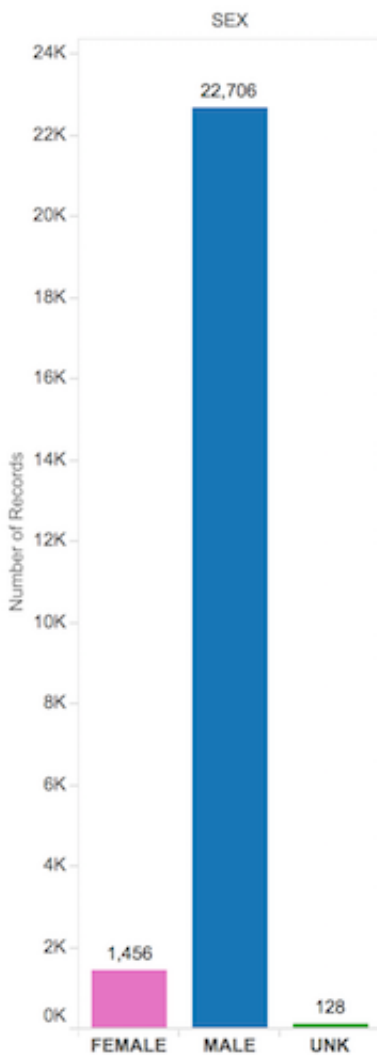
## Analysis:

First, since we are going to find information about how the parole system works in New York State, I import the parole.csv file to Tableau and try to get "a big picture" of the condition of parole interview decision. That is to say, how quantitative values are distributed across their full range?

This chart shows the percentages of each decision type of the whole records. In total 24291 records, we can see like Nikki Zeichner said, most of parole interview decision is "denied"(13186 records, about 54%). The percentage of "open date" is 18.11% overall, the second large. The percentage of "or earlier" is 15.64% overall, the third large. Also, only very little amount of prisoners are given "paroled"(89, about 0.37%).

Then I start to think about that, can we find several factors that may influence the parole interview decision? Then the "sex" dimension came into my mind.

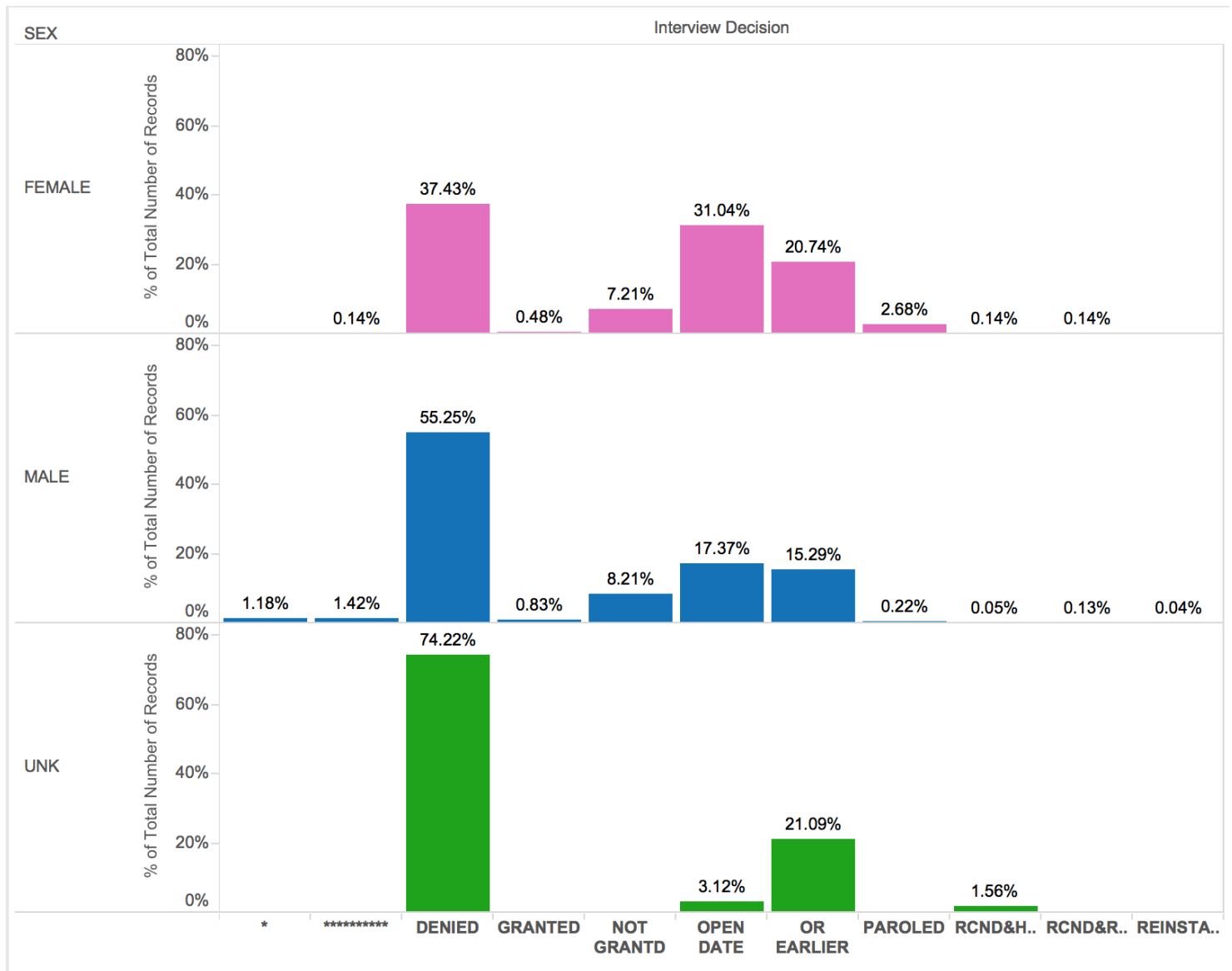
## ② Step 2:



### Analysis:

This chart shows the amount of male's records and female's records and unknown sex type's records. So first, we should look at how many male's and female's records are there. From this chart we can see, male's records amount are far more than female's, which means that, we cannot just consider the amount related to sex in the future research, we should also consider the percentage in each of them. Let's move on and see what interesting findings are.

## ③ Step 3:

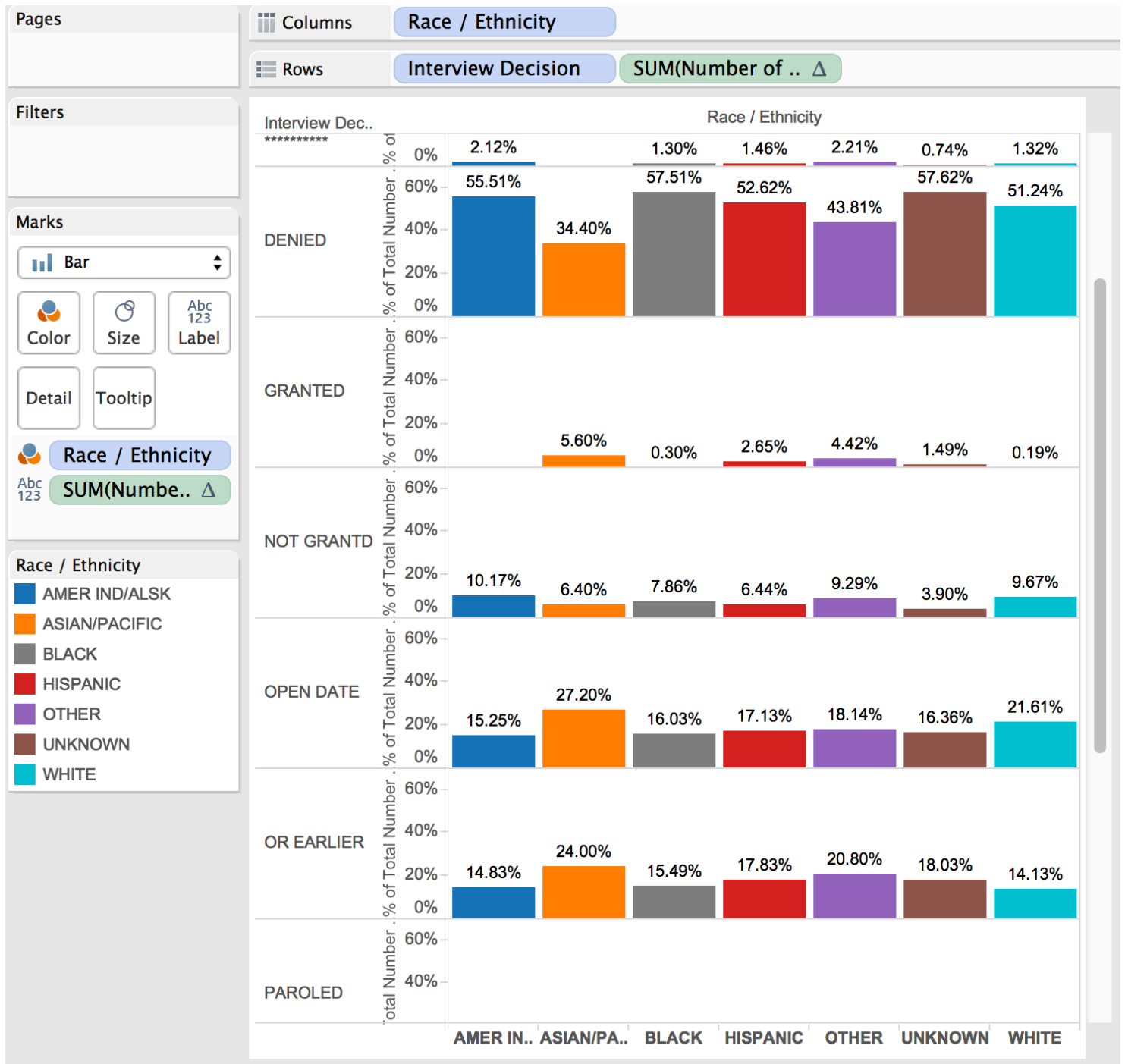


#### Analysis:

This chart display the parole interview decision related to male, female, unknown sex respectively. We can see that 54.46% of female were sentenced "open date" or "paroled" or "or earlier" and only 37.43% were "denied". While in terms of male, the prisoners who were sentenced "denied" is 55.25% and those who were sentenced "open date" or "paroled" or "or earlier" are less than female, which was interesting. It seems that the paroled program gave more mercy to woman instead of man.

Then I begin to think, instead of sex factors, can we find something else from the dataset given? What is the percentage of various race or ethniccity in the records respectively? So then I add "RACE / ETHNICITY" demension to the chart. And for clarity, there is no need to show the sex dimension.

#### ④ Step 4:



### Analysis:

This chart shows what percentage of every the parole interview decision in every race / ethnicity. We can see that generally speaking, the parole interview decision are fair to every race / ethnicity, because in every race / ethnicity, the percentage of "denied", "granted", "not granted", "open date" are almost the same. But we can see that there are more Asians were sentenced "open date", and the percentage of being sentenced to "denied" in Asians is less than other race. So it seems that Asians are more likely not sentenced to "denied" than other race / ethnicity.

Then I begin to think: how about age? I have every prisoner's birthday, so can i find some relation between their ages and the decision?

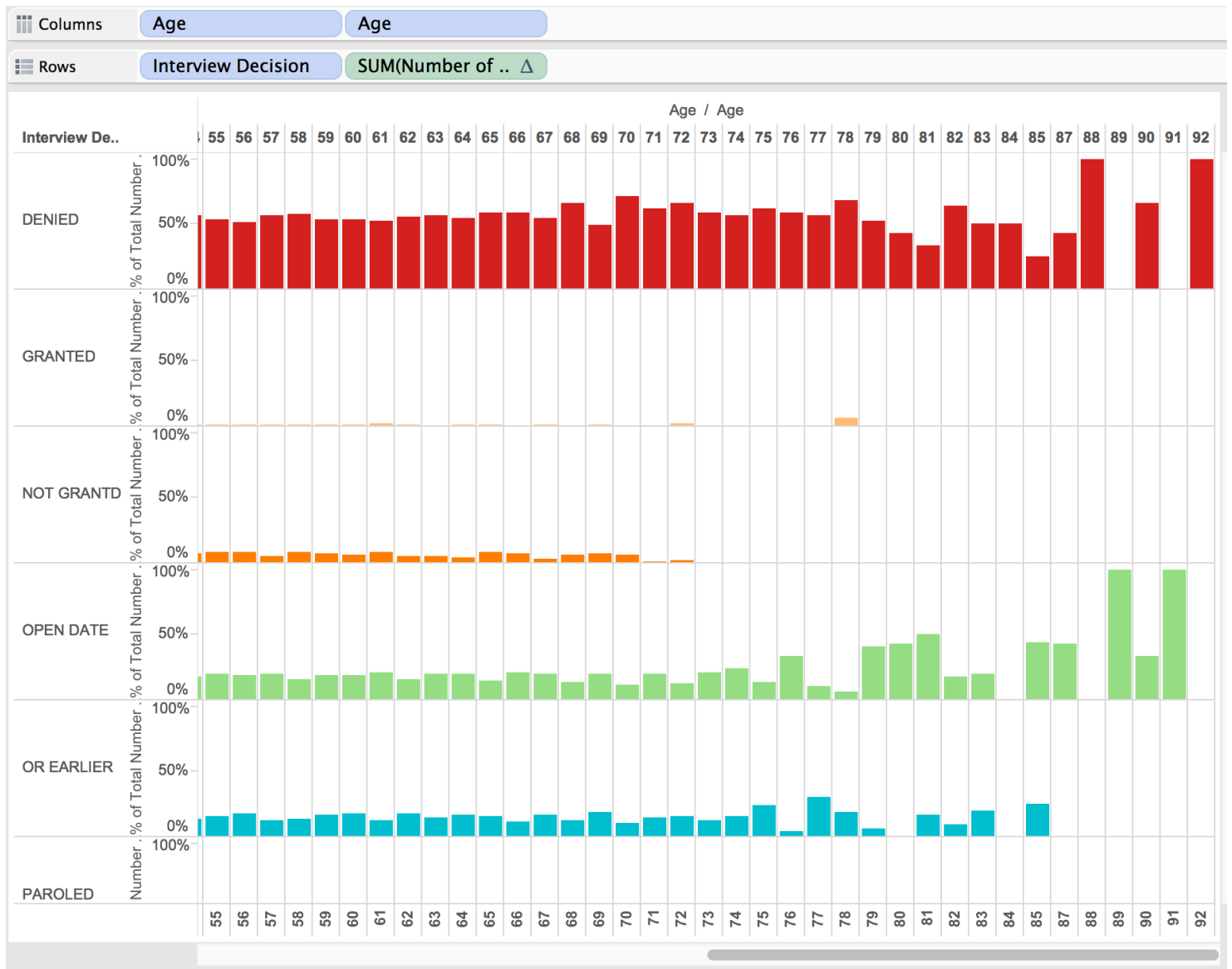
## ⑤ Step 5:

First, I use the code to calculate a new field which is "Age" (based on the prisoner's birthday)

```
1 IF [Birth Date]>DATEADD('year',-DATEDIFF('year',[Birth Date],TODAY()),TODAY()) THE
2     DATEDIFF('year',[Birth Date],TODAY())-1
3 ELSE
4     DATEDIFF('year',[Birth Date],TODAY())
5 END
```

Then we can use "Age" to analyze:





## Analysis:

This chart shows how decision results distribute among different ages(using percentage). We can see that, generally speaking, it is fair. That is to say, in every age, almost same percentage of prisoners were sentenced to "denied", "open date", "granted", "not granted", "or earlier".

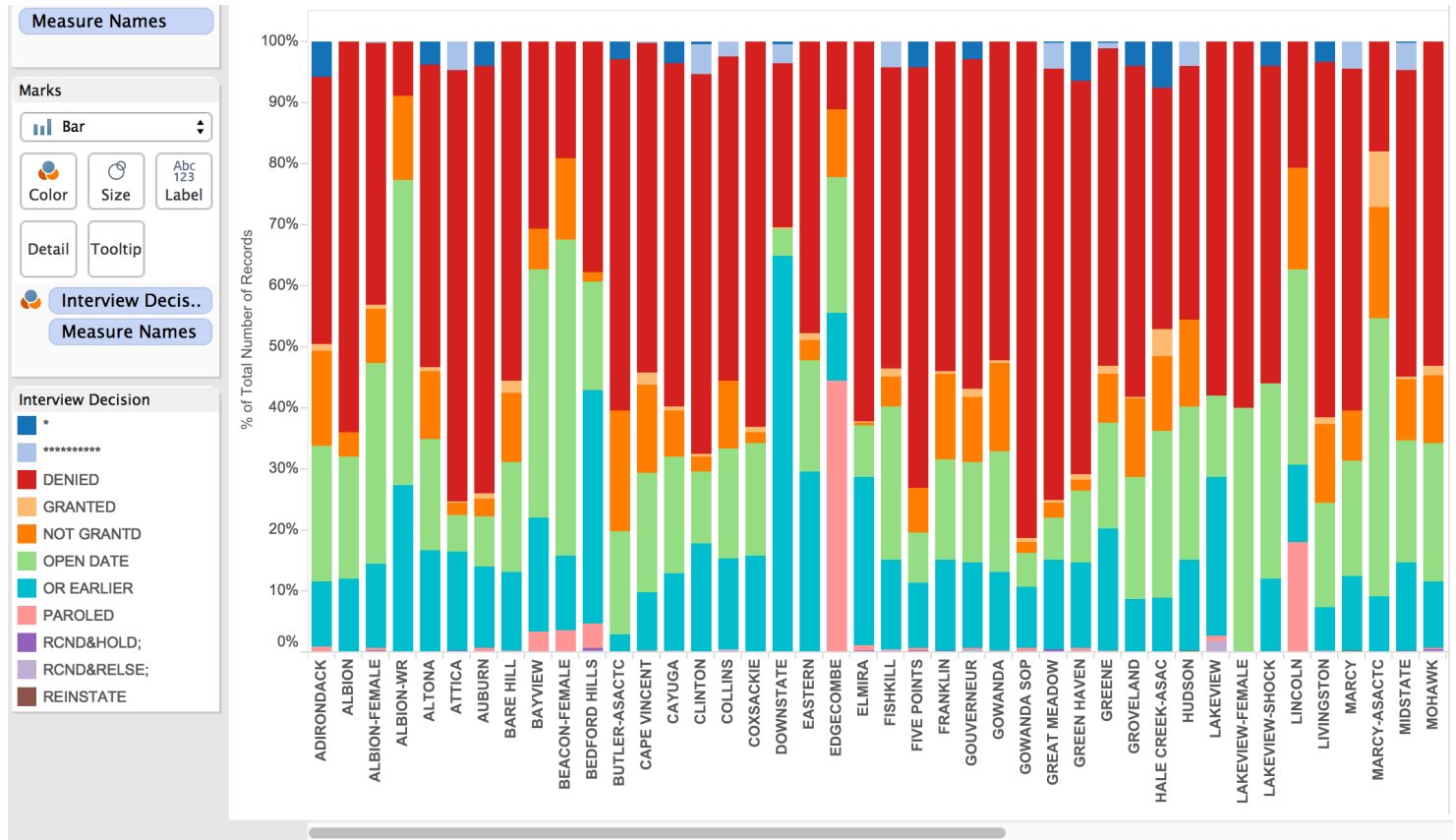
Also we can see, it seems that in age range 88~92, stories are different. But that's not true because if we look at the number of prisoners in that age, we can find the number is very small like 1 or 2. So the amount is too small and we cannot use this small sample to analyze. So we can say, it seems that "Age" are not the factor that influences decision results.

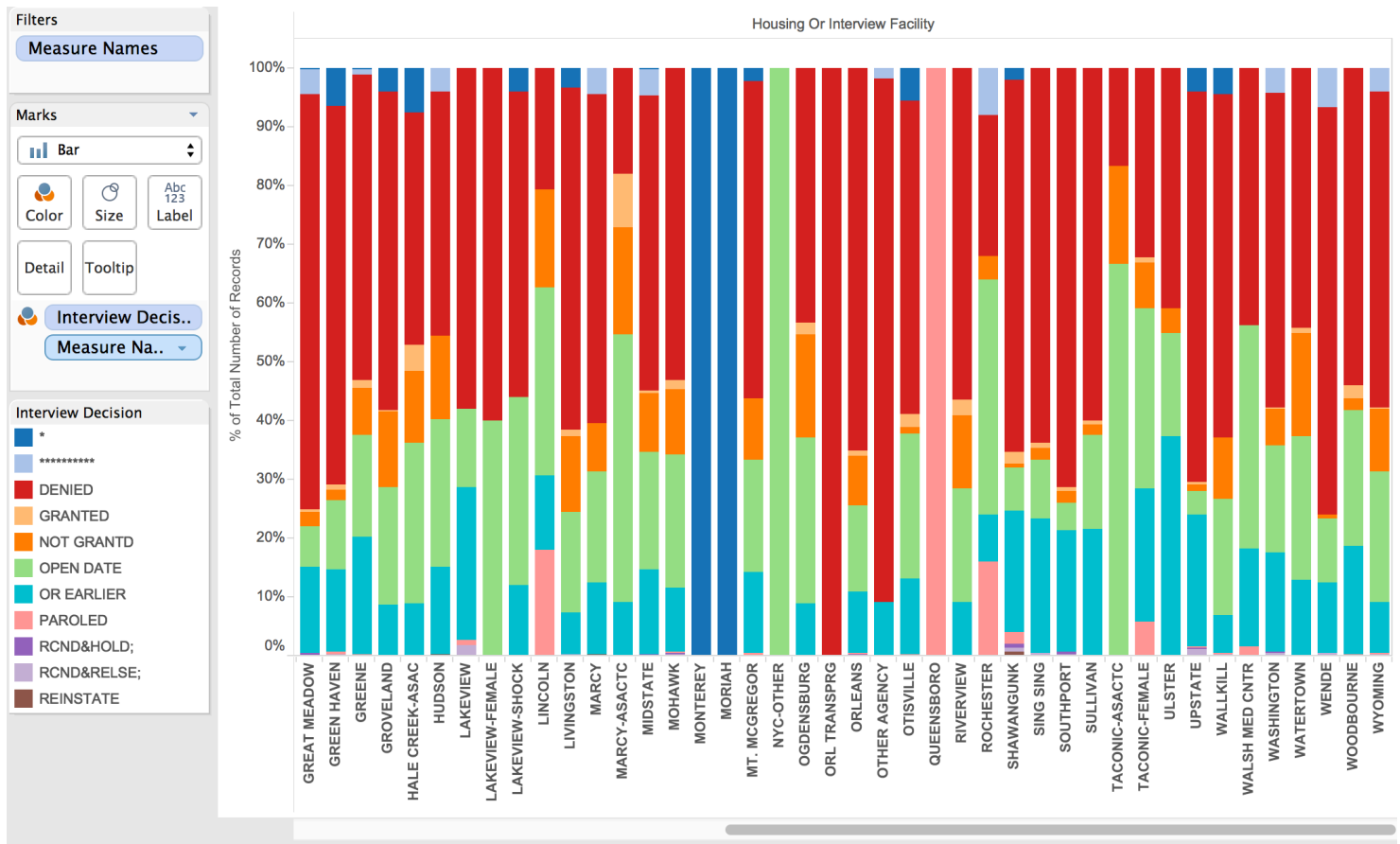
So, since the age seems to have no influence and the decisions are made in different place, are those certain place influence the decision result? Let's find out that.



## ⑥ Step 6:

Now, I begin to notice the "HOUSING OR INTERVIEW FACILITY" attribute. This shows the Correctional Facility where the offender is being housed before his/her Board appearance. I guess maybe this is a key factor that influences what decision result is.





## Analysis:

In this chart, I use percentage to show the how decision results distributed in every Correctional Facility. And then something really interesting happens. We can see, basically, the differences between Correctional Facility really exist. And in the place like "Monterey" and "Moriah", all prisoners, all decisions are "\*". Also in "FIVE POINTS", all the prisoners are denied! So what are the reasons for that? Maybe we need more informations and should go there to find what really happens there.

## 3. Conclusion

From the analysis above, We can see that how sex, race / ethnicity influence the parole interview decision results. Also, it seems that the age of prisoners doesn't really influence the decision result. And the decision result among diffrent "HOUSING OR INTERVIEW FACILITY" have differences and need to be done more real research.