



### Welcome to Chicago's Analytics Page!

Here we will be addressing the following Problem Definition

**"Identify strategies to reduce crime rates in Chicago by 30% in the next 5 years by analyzing the correlation between education quality and crime rate to enable strategic actions."**

Plan of Action

- Helping the Chief of Police understand the timeline of crime in Chicago's westside and south side between 2012-2016 (Temporal Analysis Dashboard)
- Help the Mayor understand the proportion of students graduating high schools and dropping out. Which students are more likely to dropout. (DropOutGrid dashboard)
- Map out elementary schools in th city and show a comparison of areas with high/low performance based on race, population, annual income and number of students(Special Dashboard)

### Crime Data Summary

This dataset reflects reported incidents of crime (with the exception of murders where data exists for each victim) that occurred in the City of Chicago from 2012 to 2016.

#### Scope

Time Period(s): 2012-2016

#### Description of Variables:

ID - Unique identifier for the record

Case Number - The Chicago Police Department RC Number (Records Division Number), which is unique to the incident

Date - Date when the incident occurred this is sometimes a best estimate

Block - The partially redacted address where the incident occurred, placing it on the same block as the actual address

UCR - The Uniform Crime Reporting code. This is directly linked to the Primary Type and Description

Primary Type - The primary description of the UCR code

Description - The secondary description of the UCR code, a subcategory of the primary description

Location - A description of the geographic location where the incident occurred

Arrest - Indicate whether an arrest was made

Domestic - Indicates whether the incident was domestic-related as defined by the Illinois Domestic Violence Act

Beat - Indicates the beat where the incident occurred. A beat is the smallest police geographic area - each beat has a dedicated police beat car. Three to five beats make up a police sector, and three sectors make up a police district. The Chicago Police Department has 22 police districts.

District - Indicates the police district where the incident occurred

Ward - The ward (City Council district) where the incident occurred

Community Area - Indicates the community area where the incident occurred. Chicago has 77 community areas.

FB Code - Indicates the crime classification as outlined in the FBI's National Incident-Based Reporting System (NIBRS).

X Coordinate - The x-coordinate of the location where the incident occurred in State Plane (NAD 1983) projection. This location is shifted from the actual location for partial redaction but falls on the same block

Y Coordinate - The y-coordinate of the location where the incident occurred in State Plane (NAD 1983) projection. This location is shifted from the actual location for partial redaction but falls on the same block

Year - Year the incident occurred

Latitude - The latitude of the location where the incident occurred. This location is shifted from the actual location for partial redaction but falls on the same block

Longitude - The longitude of the location where the incident occurred. This location is shifted from the actual location for partial redaction but falls on the same block

Location - The location where the incident occurred in a format that allows for creation of maps and other geographic operations on this data portal. This location is shifted from the actual location for partial redaction but falls on the same block

### Education Data Summary

Chicago Public Schools (CPS) - Report Card Data This dataset shows all school-level performance data used to create CPS School Report Cards for the 2013-2016.

Metrics used are described as follows: for the entire list go to <https://data.chicagopublicschools.org/CPSData/Reports/2013-2016/2013-2016-Report-Cards/2013-2016-Report-Cards>

#### 5 Year Graduation Rate

The five-year cohort graduation rate follows a group of students who enter Chicago Public Schools (CPS) as freshmen and calculates the percent of these students who graduate within five years after their freshman year. For example, the 2006 rate tracks ninth grade students who started in CPS in the 2001-02 school year and graduated from CPS by August of the 2005-06 school year.

#### How is the rate calculated?

The rate is calculated by dividing the number of students from an adjusted ninth grade cohort who graduated over the five year time period by the total number of students in that cohort

#### Example Calculation

School A had 725 ninth grade students enrolled during the 2001-02 school year. Twenty of these were repeating the ninth grade. Twenty-five of them transferred out of CPS and were verified in a new school during the five years following the 2001-02 school year. The adjusted ninth grade cohort for this school was Cohort9th adj. = 725 - 20 - 25 = 680 students. During the five years following the 2001-02 school year, 286 of these 680 students dropped out and 364 graduated. Of the remaining 40 students, 5 were unverified transfers as of August 2006, and twenty-five were still enrolled in CPS as of the end of the 2005-06 school year. School A's Five-Year Cohort Graduation Rate in 2006 was: Graduation Rate=year cohort = 364 / 680 = 53.5%

#### 5 Year Dropout Rate

What is the Five-Year Cohort Dropout Rate? The five-year cohort dropout rate follows a group of students who enter Chicago Public Schools (CPS) as freshmen and calculates the percent of these students who drop out within five years after their freshman year. For example, the 2006 rate tracks ninth grade students who started in CPS in the 2001-02 school year and dropped out of CPS by August of the 2005-06 school year.

#### How is the rate calculated?

The rate is calculated by dividing the number of students from an adjusted ninth grade cohort who dropped out over the five year time period by the total number of students in that cohort.

Example Calculation School A had 725 ninth grade students enrolled during the 2001-02 school year. Twenty of these were repeating the ninth grade. Twenty-five of them transferred out of CPS and were verified in a new school during the five years following the 2001-02 school year. The adjusted ninth grade cohort for this school was Cohort9th adj. = 725 - 20 - 25 = 680 students. During the five years following the 2001-02 school year, 286 of these 680 students dropped out and 364 graduated. Of the remaining 40 students, 5 were unverified transfers as of August 2006, and twenty-five were still enrolled in CPS as of the end of the 2005-06 school year. School A's Five-Year Cohort Dropout Rate in 2006 was: Dropouts = 286 / "year" dropouts) = 5 (unverified transfer students) = 291 (Cohort Dropout Rate)=year cohort = 291 / 680 = 42.8%

#### Student Growth Rating

Overall score of the school based on growth factors

#### Student Attainment Rating

Overall score of the school based on attainment of students

#### Racial Ethnicity of Students -

Based on race, religion and race

#### Population data of Chicago -

for each zip code

#### Avg annual income -

For each zip code





Analyzing Proportions

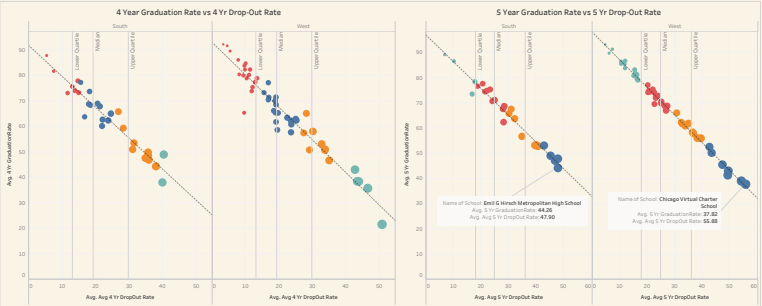


SCATTER PLOTS

Which students are most likely to drop-out?

Here we are trying to understand the proportion of students graduating high schools versus dropping out. Date range for analysis is 2012-16. Avg 4Yr graduation and dropout rate shows inverse linear relationship. Exploring highlighted schools in the respective regions could be a good starting point

Conclusion: Students in the schools above the upper Quartile are most likely to DropOut.



DropOut Rates

1) DropOut Rates - We are taking the ratio of 4 yr and 5 yr dropOut rate vs the 4yr and 5yr Graduation rates and showing the tree map of Community areas from where the students are most likely to DropOut.

2) The data also shows the DropOut rates for each region individually.

3) Englewood in West, Avalon Park in South and North Park in North have the highest dropOut rates in the past 5 yrs and are most likely to produce DropOuts in the future

Dropout Vs Graduation Rates

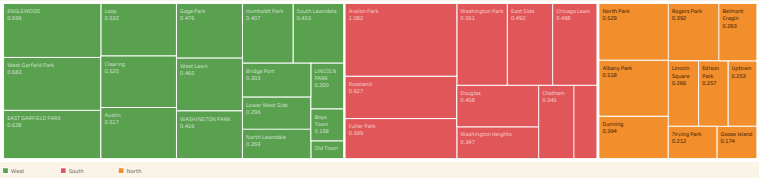
Looking at 4 Yr Graduation rate versus 4 Yr DropOut rate and same for 5 Year DropOut and Graduation rate we can infer that despite of maximum criminal activity prevailing more in west region the school count and schools with higher graduation rates are more when compared to south region.

The probable reason for high dropout rates of these schools are:

- 1. Families involvement is less
- 2. Collaborative Teachers is either Neutral or Weak
- 3. School Safety is weak.

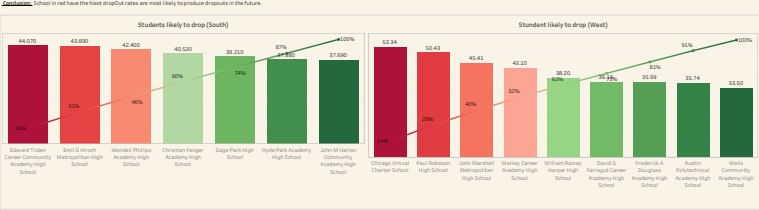
Conclusion: So students from the Community Areas with higher dropout rates are more likely to be affected and likely to drop out in the future too. To justify our analysis we have tried to show this with Pareto chart showing schools where students are more likely to dropout, e.g Englewood, Avalon Park and North Park in West, South and North resp.

DropOut Rate by Region



Using Pareto Analysis

1) We are using the aggregation of 4 and 5 Yr DropOut rates of 2012-2016, find the top schools with highest dropout rates based on their avg dropout rates.



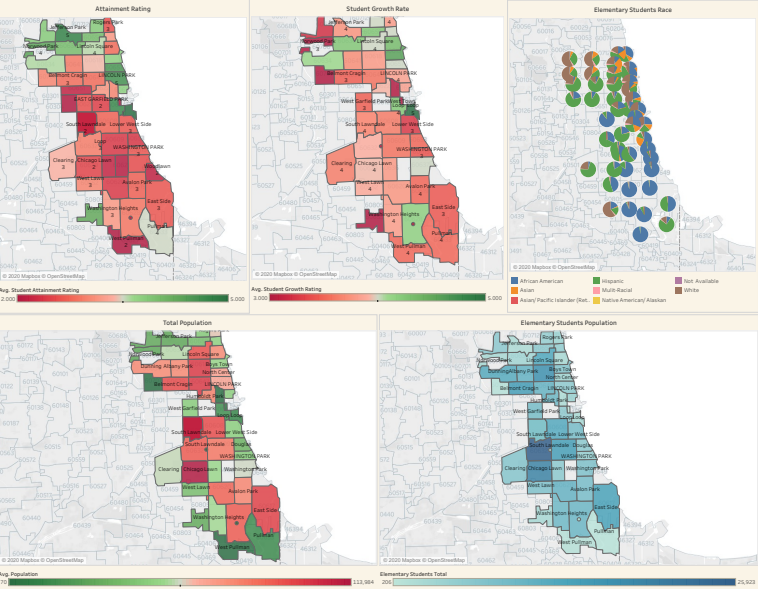


Community Area	Student Attainment Rating	Student Growth Rating	Region
All	2 to 5 and null values	All values	All

**Features:**  
The Dashboard has integrated filters for ease of analysis. Clicking on a menu item applies filters on all the related items. The filters on top like Attainment and Growth applies to all the maps on the dashboard.

- Insights:**
- 1) North, West Pullman and west town have the lowest attainment scores whereas Loop, Rogers park and west town have the lowest growth rate.
  - 2) In northern region students of Hispanic race are more followed by white race students with an average attainment rating of 3 and average growth rate of 2.8. the average population of this region is 70000 with annual salary of 42576 which is highest of all the regions
  - 3) Loop and Rogers park are dominating with Hispanic, African American and White Race.
  - 4) In southern region students of African American race are maximum followed by Hispanic with average student attainment rating of 2.5 and growth rating of 3.8 having an average population of 80000

- Sandboxes:**
- 1) We could see that Racial Student majority does not have a direct effect on the Student Attainment rating or the Student Growth rate.
  - 2) Whereas densely populated areas have high growth rates but lower attainment rates, e.g Washington Park and South Lawndale.





Welcome to Chicago's Analytics Page!

Here we are addressing the following Problem Statement,

"Identify strategies to reduce crime rates in Chicago by 30% in the next 1 Year by analyzing the correlation between education quality and crime rate to enable strategic actions"

#### Plan of Action

- 1) Helping the Chief of Police understand the timeline of crime in Chicago's westside and south side between 2012-2016 (Temporal Analysis Dashboard)
- 2) Help the Mayor understand the proportion of students graduating high schools and dropping out. Which students are more likely to dropout. (DropOutvsGrad dashboard)
- 3) Map out elementary schools in th city and show a comparison of areas with high/poor performance based on race, population, annual income and number of students(Spatial Dashboard)

#### Crime Data Summary:

This dataset reflects reported incidents of crime that occurred in the City of Chicago from 2012 to 2016.

#### Scope:

Time Period(s): 2012-2016

#### Description of Variables :

ID - Unique identifier for the record

Case\_Number - The Chicago Police Department RD Number (Records Division Number), which is unique to the incident

Date - Date when the incident occurred; this is sometimes a best estimate

Block - The partially redacted address where the incident occurred, placing it on the same block as the actual address

IUCR - The Illinois Uniform Crime Reporting code. This is directly linked to the Primary Type and Description.

Primary\_Type - The primary description of the IUCR code

Description - The secondary description of the IUCR code, a subcategory of the primary description

Location\_Description - Description of the location where the incident occurred

Arrest - Indicates whether an arrest was made

Domestic - Indicates whether the incident was domestic-related as defined by the Illinois Domestic Violence Act

Beat - Indicates the beat where the incident occurred. A beat is the smallest police geographic area - each beat has a dedicated police beat car. Three to five beats make up a police sector, and three sectors make up a police district. The Chicago Police Department has 22 police districts.

District - Indicates the police district where the incident occurred.

Ward - The ward (City Council district) where the incident occurred.

Community\_Area - Indicates the community area where the incident occurred. Chicago has 77 community areas.

FBI\_Code - Indicates the crime classification as outlined in the FBI's National Incident-Based Reporting System (NIBRS).

X\_Coordinate - The x coordinate of the location where the incident occurred in State Plane Illinois East NAD 1983 projection. This location is shifted from the actual location for partial redaction but falls on the same block

Y\_Coordinate - The y coordinate of the location where the incident occurred in State Plane Illinois East NAD 1983 projection. This location is shifted from the actual location for partial redaction but falls on the same block

Year - Year the incident occurred

Updated\_On - Date and time the record was last updated

Latitude - The latitude of the location where the incident occurred. This location is shifted from the actual location for partial redaction but falls on the same block

Longitude - The longitude of the location where the incident occurred. This location is shifted from the actual location for partial redaction but falls on the same block

Location - The location where the incident occurred in a format that allows for creation of maps and other geographic operations on this data portal. This location is shifted from the actual location for partial redaction but falls on the same block

#### Education Data Summary

Chicago Public Schools (CPS) – Report Card Data This dataset shows all school-level performance data used to create CPS School Report Cards for the2012-2016.

Metrics used are described as follows: for the entire list got to <https://data.cityofchicago.org/api/assets/AAD41A13-BF8A-4E67-B1F5-86F711E09D5F?download=true>

#### 5 Year Graduation Rate

The five-year cohort graduation rate follows a group of students who enter Chicago Public Schools (CPS) as freshmen and calculates the percent of these students who graduate within five years after their freshman year. For example, the 2016 rate tracks ninth grade students who started in CPS in the 2011-12 school year and graduated from CPS by August of the 2015-16 school year.

#### How is the rate calculated?

The rate is calculated by dividing the number of students from an adjusted ninth grade cohort who graduated over the five year time period by the total number of students in that cohort

#### Example Calculation

School A had 725 ninth grade students enrolled during the 2011-12 school year. Twenty of these were repeating the ninth grade. Twenty-five of them transferred out of CPS and were verified in a new school during the five years following the 2011-12 school year. The adjusted ninth grade cohort for this school was: Cohort9th adj. = 725 – 20 – 25 = 680 students During the five years following the 2011-12 school year, 286 of these 680 students dropped out and 364 graduated. Of the remaining 45 students, 5 were unverified transfers as of August 2016, and twenty-five were still enrolled in CPS as of the end of the 2015-16 school year. School A's Five-Year Cohort Graduation Rate in 2016 was: Graduation Rate5-year cohort = 364 / 680 = 53.5%

#### 5 Year DropOut Rate

What is the Five-Year Cohort Dropout Rate? The five-year cohort dropout rate follows a group of students who enter Chicago Public Schools (CPS) as freshmen and calculates the percent of these students who drop out within five years after their freshman year. For example, the 2016 rate tracks ninth grade students who started in CPS in the 2011-12 school year and dropped out of CPS by August of the 2015-16 school year.

#### How is the rate calculated?

The rate is calculated by dividing the number of students from an adjusted ninth grade cohort who dropped out over the five year time period by the total number of students in that cohort.

Example Calculation School A had 725 ninth grade students enrolled during the 2011-12 school year. Twenty of these were repeating the ninth grade. Twenty-five of them transferred out of CPS and were verified in a new school during the five years following the 2011-12 school year. The adjusted ninth grade cohort for this school was: Cohort9th adj. = 725 – 20 – 25 = 680 students During the five years following the 2011-12 school year, 286 of these 680 students dropped out and 364 graduated. Of the remaining 45 students, 5 were unverified transfers as of August 2016, and twenty-five were still enrolled in CPS in the 2015-16 school year. School A's Five-Year Cohort Dropout Rate in 2016 was: Dropouts = 286 ("true" dropouts) + 5 (unverified transfer students) = 291 Dropout Rate5-year cohort = 291 / 680 = 42.8%

#### Student Growth Rating

Overall score of the school based on growth factors

#### Student Attainment Rating

Overall score of the school based on attainment of students

Racial Ethnicity of Students - Based on cast, religion and race

Population data of Chicago - for each zipcode

Avg annual income - For each zipcode



## Analyzing Proportions

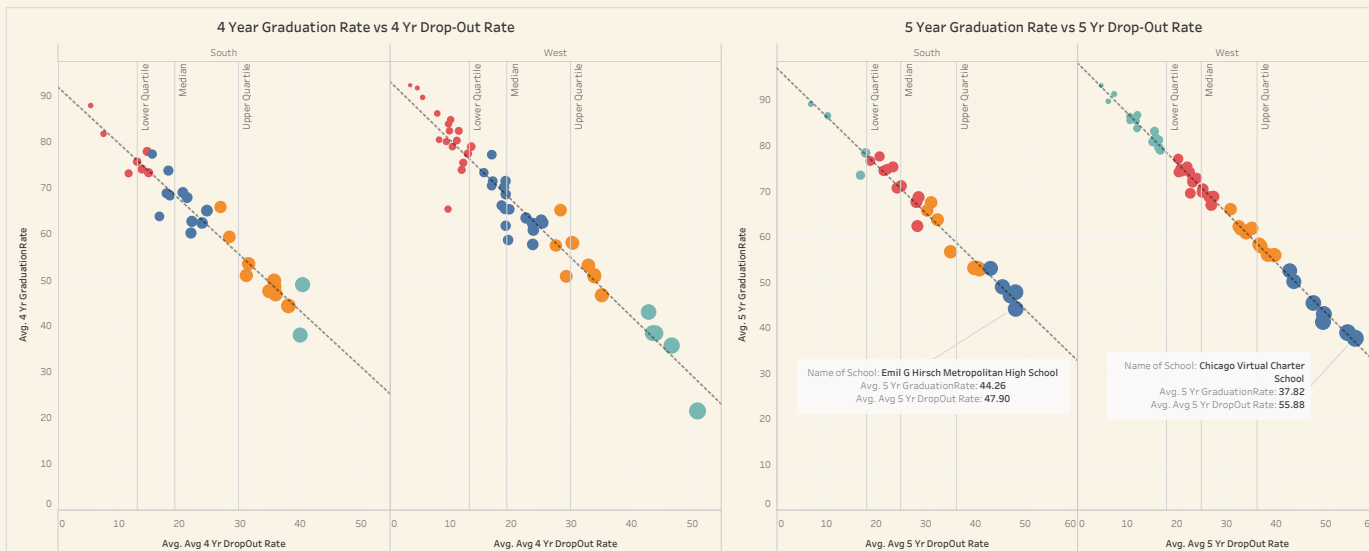


Which students are most likely to drop-out?

Here we are trying to understand the **proportion** of students graduating high schools versus dropping out. Date range for analysis is 2012-16.  
Avg 4/5yr graduation and dropout rate shows inverse linear relationship.  
Exploring highlighted schools in the respective regions could be a good starting point

**Conclusion:**

Students in the schools above the upper Quartile are most likely to DropOut.

**SCATTER PLOTS****Tree Maps**

- DropOutRates - We are taking the ratio of 4 yr and 5 yr dropOut rate vs the 4yr and 5ys Graduation rates and showing the tree map of Community areas from where the students are most likely to DropOut.
- The data also shows the DropOut ratios for each region individually.
- Englewood in West, Avalon Park in South and North Park in North have the highest dropOut rates in the past 5 yrs and are most likely to produce DropOuts in the future

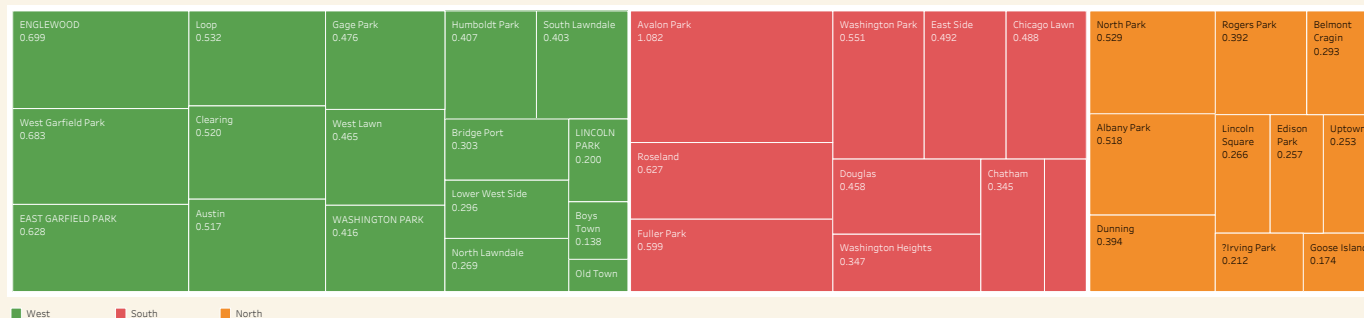
**Dropout Vs Graduation Rates**

Looking at 4 Yrs. Graduation rate versus 4 Yr. Dropout rate and same for 5 Year Dropout and Graduation rate we can infer that despite of maximum criminal activity prevailing more in west region the school count and Schools with higher graduation rates are more when compared to south region.

The probable cause for high dropout rates of those schools are:

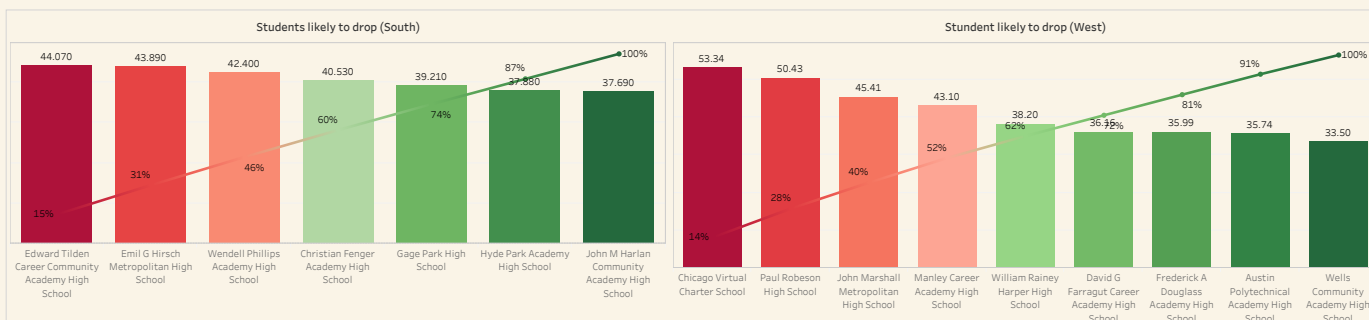
- Families Involvement is less
- Collaborative Teachers is either Neutral or Weak
- School Safety is weak.

**Conclusion:** So students from the Community Areas with higher dropout rates are more likely to be affected and likely to drop out in the future too. To justify our analysis we have tried to show this with Pareto chart showing schools where students are more likely to dropout. e.g Englewood, Avalon Park and North Park in West, South and North resp.

**DropOut Rate by Region****Using Pareto Analysis**

1) We are using the aggregation of 4 and 5 Yr Dropout rates of 2012-1016 find the top schools with highest dropout rates based on their agg dropout rates.

**Conclusion:** School in red have the highest dropout rates are most likely to produce dropouts in the future.



#### Features:

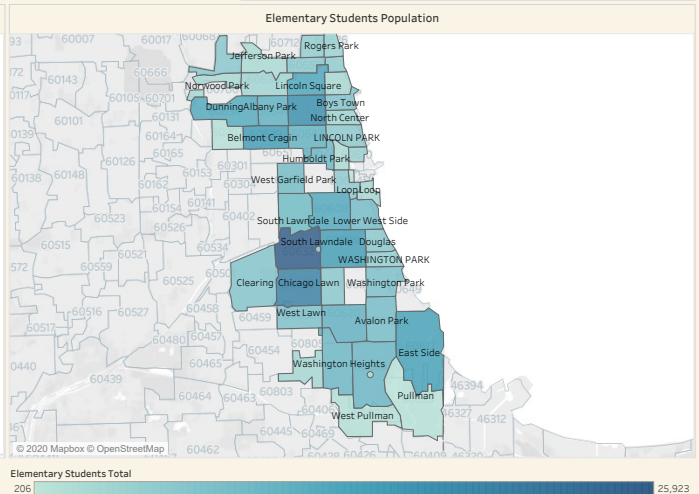
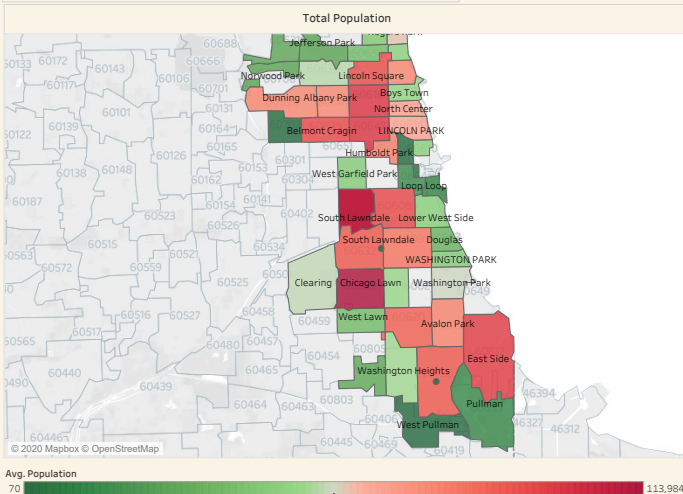
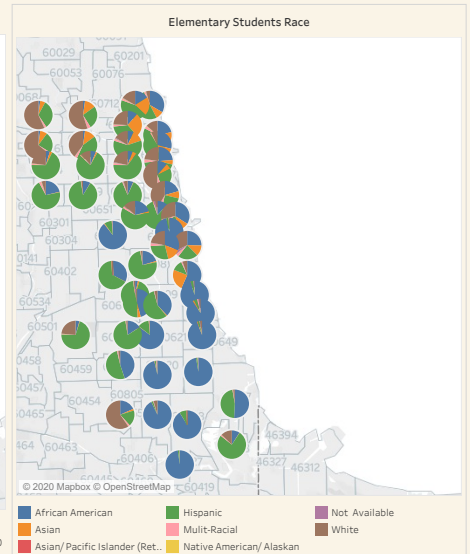
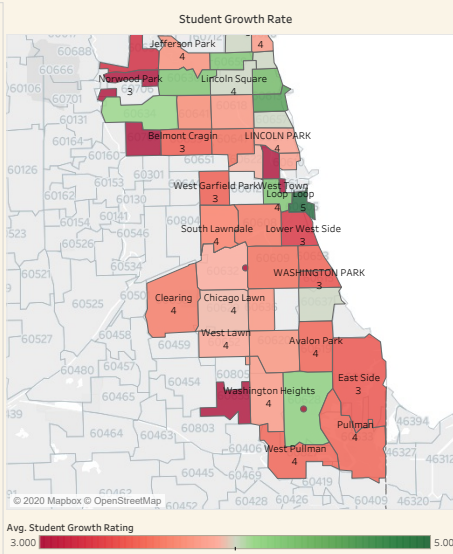
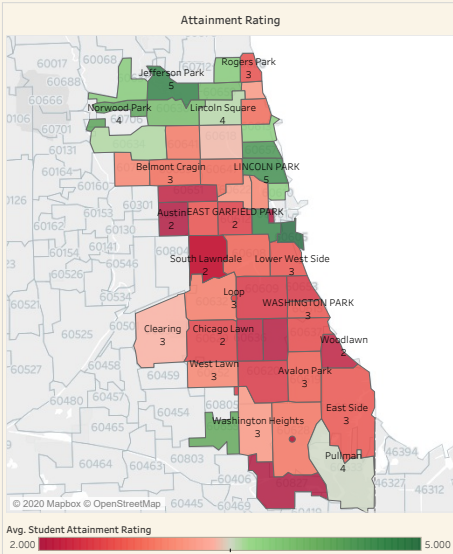
This Dashboard has integrated filters for ease of analysis. Clicking on one item applies filters on all the related items. The filters on top like Attainment and Growth applies to all the maps on the dashboard.

#### Highlights:

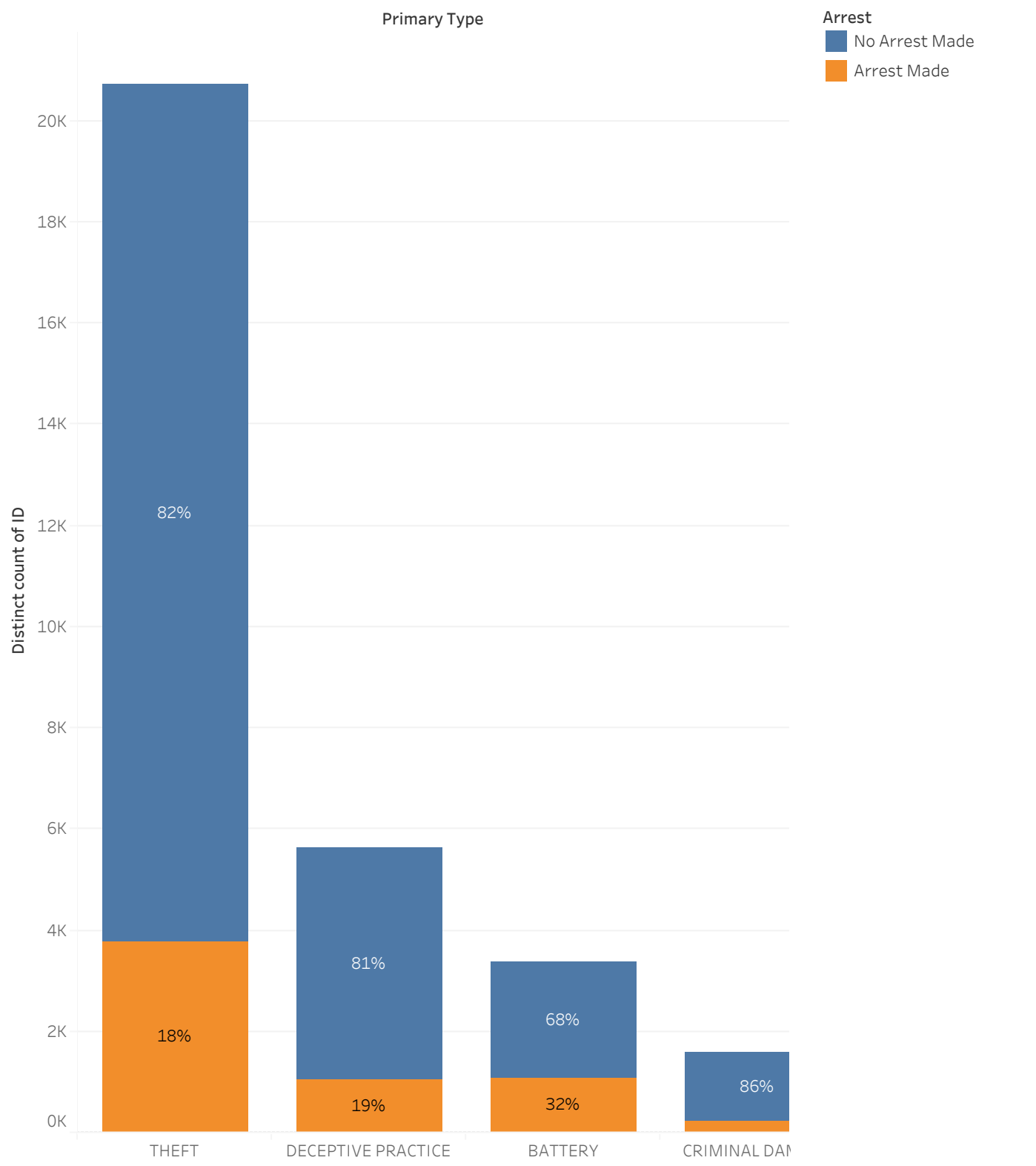
- 1) Austin, west pullman and west town have the lowest attainment score whereas loop, rogers park and west town have the lowest growth rate.
- 2) In northern region students of Hispanic race are more followed by white race students with an average attainment rating of 3 and average growth rate of 2.8. the average population of this region is 75000 with annual salary of 42576 which is highest of all the regions
- 3) Loop and Rogers park are dominating with Hispanic African American and little Asian.
- 4) In southern region students of African American race are maximum followed by Hispanic with average student attainment rating of 2.5 and growth rating of 3.8 having an average population of 80000

#### Conclusion:

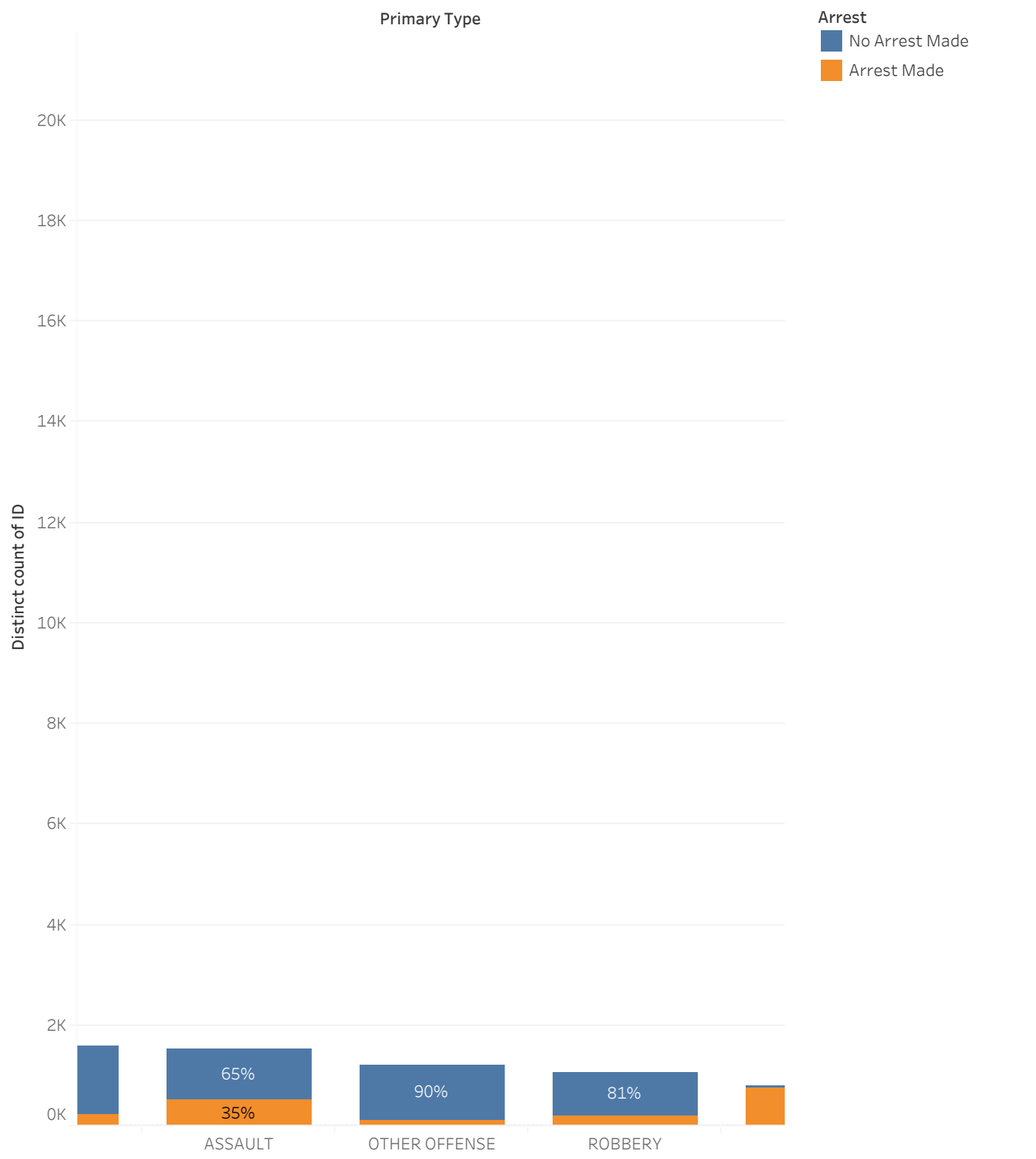
- 1) We could see that Racial Student majority does not have a direct effect on the Student Attainment rating or the Student Growth rate
- 2) Whereas densely populated areas have high growth rates bit lower attainment rates. e.g Washington Park and South Lawndale.





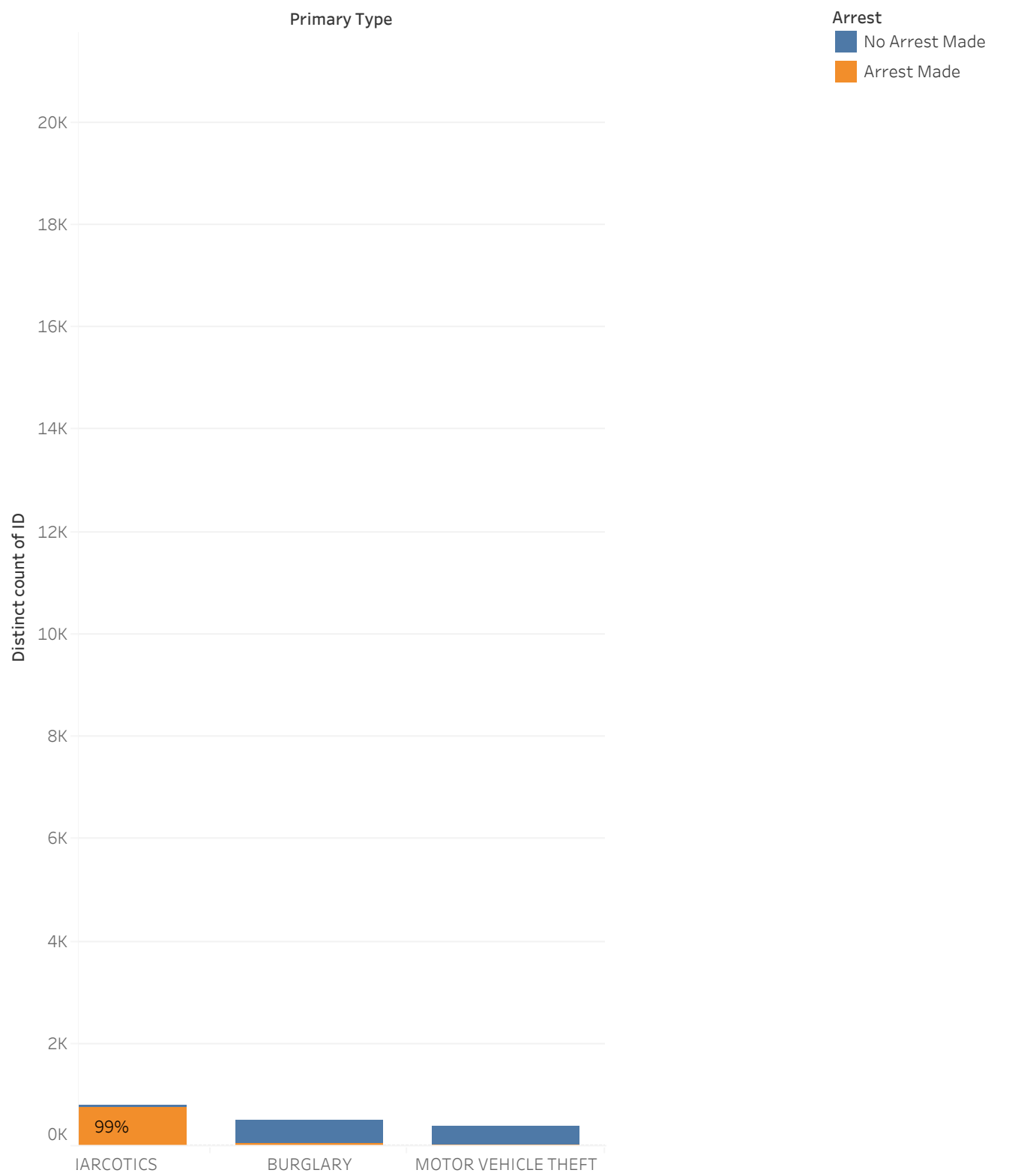


Distinct count of ID for each Primary Type. Color shows details about Arrest. The marks are labeled by % of Total Distinct count of ID. The data is filtered on Community area, which keeps Loop. The view is filtered on Primary Type, which keeps 10 of 33 members.



Distinct count of ID for each Primary Type. Color shows details about Arrest. The marks are labeled by % of Total Distinct count of ID. The data is filtered on Community area, which keeps Loop. The view is filtered on Primary Type, which keeps 10 of 33 members.

Sheet 21



Distinct count of ID for each Primary Type. Color shows details about Arrest. The marks are labeled by % of Total Distinct count of ID. The data is filtered on Community area, which keeps Loop. The view is filtered on Primary Type, which keeps 10 of 33 members.