

# Attendance Rates in Public Schools

An Analysis of Student Groups and Attendance Rates

# Project Overview

- Attendance rates compared to different student groups
- Attendance rates are an important part of a school's overall **Academic Performance Index (API)**
  - High attendance increases learning and therefore increases a school's API
  - Chronic absenteeism counts against a school's API
- Looking at attendance rates by student groups allows us to know which student groups and areas to provide supports for greater attendance rates.

# The Data Sets

**Data Set 1:** This data set is a collection of the attendance rates of all public schools in the state of Connecticut for the past three schools years. The data is broken down into student count attendance rates by student groups. Through this data set, I investigated the differences between attendance amongst student groups.

**Source:** data.gov

**Data Set 2:** This data set has the demographics of school districts. It includes a list of the zip codes for each district in the state of Connecticut. This allowed me to map the school districts and their populations.

**Source:** Department of Education (.gov)

**Data Set 3:** This data set looks at each districts' API score and test scores. It is broken down by districts and allowed me to compare API scores to attendance.

**Source:** ct.gov

# Data Set 1

# Data Set 1

- Breaks down into the following student groups:
  - Students with Disabilities
  - Free Meal Eligible
  - Reduced Price Meal Eligible
  - Free/Reduced Price Meal Eligible
  - English Learners
  - All Other Races
  - Black/African American
  - Hispanic/Latino of any race
  - White
  - Students with High Needs
  - Students without High Needs
  - All Students
- Looks at the following:
  - District
  - 2022-2023 Attendance Rates
  - 2022-2023 Student Counts
  - 2021-2022 Attendance Rates
  - 2021-2022 Student Counts
  - 2020-2021 Attendance Rates
  - 2020-2021 Student Counts

# Data Set 1 - CT Public School Attendance

| District code | District name           | Category                   | Student group                      | 2021-2022 student count - year to date | 2021-2022 attendance rate - year to date | 2020-2021 student count | 2020-2021 attendance rate | 2019-2020 student count | 2019-2020 attendance rate | Reporting period | Date update |
|---------------|-------------------------|----------------------------|------------------------------------|--|--|-------------------------|---------------------------|-------------------------|---------------------------|------------------|-------------|
| 00000CT       | Connecticut             |                            | All Students                       | 500285                                 | 0.9169                                   | 496092                  | 0.9294                    | 508346                  | 0.9479                    | Jun-22           | 7/22/22     |
| 00000CT       | Connecticut             | Homelessness               | Students Experiencing Homelessness | 1814                                   | 0.8348                                   | 1735                    | 0.8155                    | 3916                    | 0.8884                    | Jun-22           | 7/22/22     |
| 00000CT       | Connecticut             | Students With Disabilities | Students With Disabilities         | 78417                                  | 0.8899                                   | 76487                   | 0.8946                    | 80365                   | 0.9277                    | Jun-22           | 7/22/22     |
| 00000CT       | Connecticut             | Free/Reduced Lunch         | Free Meal Eligible                 | 168984                                 | 0.8851                                   | 176225                  | 0.8861                    | 193706                  | 0.9314                    | Jun-22           | 7/22/22     |
| 00000CT       | Connecticut             | Free/Reduced Lunch         | Reduced Price Meal Eligible        | 29905                                  | 0.9184                                   | 30886                   | 0.9299                    | 27507                   | 0.9518                    | Jun-22           | 7/22/22     |
| 00000CT       | Connecticut             | Free/Reduced Lunch         | Free/Reduced Price Meal Eligible   | 198889                                 | 0.8901                                   | 207111                  | 0.8927                    | 221213                  | 0.934                     | Jun-22           | 7/22/22     |
| 00000CT       | Connecticut             | English Learners           | English Learners                   | 43571                                  | 0.8976                                   | 40619                   | 0.8948                    | 45413                   | 0.9389                    | Jun-22           | 7/22/22     |
| 00000CT       | Connecticut             | Race/Ethnicity             | All other races                    | 48700                                  | 0.9314                                   | 47339                   | 0.9483                    | 47260                   | 0.9559                    | Jun-22           | 7/22/22     |
| 00000CT       | Connecticut             | Race/Ethnicity             | Black or African American          | 63099                                  | 0.8941                                   | 62267                   | 0.8931                    | 64200                   | 0.9401                    | Jun-22           | 7/22/22     |
| 00000CT       | Connecticut             | Race/Ethnicity             | Hispanic/Latino of any race        | 146298                                 | 0.8935                                   | 138260                  | 0.8975                    | 136953                  | 0.9362                    | Jun-22           | 7/22/22     |
| 00000CT       | Connecticut             | Race/Ethnicity             | White                              | 242188                                 | 0.9338                                   | 248226                  | 0.9523                    | 259933                  | 0.9543                    | Jun-22           | 7/22/22     |
| 00000CT       | Connecticut             | High Needs                 | Students Without High Needs        | 241106                                 | 0.9398                                   | 236395                  | 0.9616                    | 241610                  | 0.9606                    | Jun-22           | 7/22/22     |
| 00000CT       | Connecticut             | High Needs                 | Students With High Needs           | 248239                                 | 0.8954                                   | 251220                  | 0.8996                    | 266736                  | 0.9361                    | Jun-22           | 7/22/22     |
| 0010011       | Andover School District |                            | All Students                       | 161                                    | 0.9386                                   | 144                     | 0.968                     | 158                     | 0.9502                    | Jun-22           | 7/22/22     |
| 0010011       | Andover School District | Students With Disabilities | Students With Disabilities         | 23                                     | 0.9315                                   |                         |                           |                         |                           | Jun-22           | 7/22/22     |
| 0010011       | Andover School District | Free/Reduced Lunch         | Free/Reduced Price Meal Eligible   | 30                                     | 0.9274                                   | 31                      | 0.9414                    | 37                      | 0.9401                    | Jun-22           | 7/22/22     |
| 0010011       | Andover School District | Race/Ethnicity             | White                              | 134                                    | 0.9365                                   | 115                     | 0.9699                    | 128                     | 0.9479                    | Jun-22           | 7/22/22     |
| 0010011       | Andover School District | High Needs                 | Students Without High Needs        | 113                                    | 0.941                                    | 100                     | 0.9752                    | 108                     | 0.9552                    | Jun-22           | 7/22/22     |
| 0010011       | Andover School District | High Needs                 | Students With High Needs           | 48                                     | 0.9328                                   | 44                      | 0.9513                    | 50                      | 0.9389                    | Jun-22           | 7/22/22     |
| 0020011       | Ansonia School District |                            | All Students                       | 2139                                   | 0.9045                                   | 2153                    | 0.8923                    | 2185                    | 0.9413                    | Jun-22           | 7/22/22     |
| 0020011       | Ansonia School District | Students With Disabilities | Students With Disabilities         | 369                                    | 0.8897                                   | 394                     | 0.8692                    | 411                     | 0.9337                    | Jun-22           | 7/22/22     |
| 0020011       | Ansonia School District | Free/Reduced Lunch         | Free Meal Eligible                 | 1192                                   | 0.8893                                   | 1217                    | 0.8657                    | 1312                    | 0.9312                    | Jun-22           | 7/22/22     |
| 0020011       | Ansonia School District | Free/Reduced Lunch         | Reduced Price Meal Eligible        | 225                                    | 0.9282                                   | 213                     | 0.9234                    | 147                     | 0.9563                    | Jun-22           | 7/22/22     |
| 0020011       | Ansonia School District | Free/Reduced Lunch         | Free/Reduced Price Meal Eligible   | 1417                                   | 0.8953                                   | 1430                    | 0.8743                    | 1459                    | 0.9338                    | Jun-22           | 7/22/22     |
| 0020011       | Ansonia School District | English Learners           | English Learners                   | 122                                    | 0.918                                    | 119                     | 0.8951                    | 129                     | 0.9506                    | Jun-22           | 7/22/22     |
| 0020011       | Ansonia School District | Race/Ethnicity             | All other races                    | 149                                    | 0.9191                                   | 166                     | 0.9105                    | 170                     | 0.9451                    | Jun-22           | 7/22/22     |
| 0020011       | Ansonia School District | Race/Ethnicity             | Black or African American          | 434                                    | 0.8987                                   | 421                     | 0.8693                    | 434                     | 0.9414                    | Jun-22           | 7/22/22     |
| 0020011       | Ansonia School District | Race/Ethnicity             | Hispanic/Latino of any race        | 987                                    | 0.8994                                   | 946                     | 0.8798                    | 923                     | 0.9366                    | Jun-22           | 7/22/22     |
| 0020011       | Ansonia School District | Race/Ethnicity             | White                              | 569                                    | 0.9136                                   | 620                     | 0.9217                    | 658                     | 0.9468                    | Jun-22           | 7/22/22     |
| 0020011       | Ansonia School District | High Needs                 | Students Without High Needs        | 528                                    | 0.9287                                   | 551                     | 0.9389                    | 587                     | 0.9568                    | Jun-22           | 7/22/22     |
| 0020011       | Ansonia School District | High Needs                 | Students With High Needs           | 1545                                   | 0.8968                                   | 1563                    | 0.8764                    | 1598                    | 0.9355                    | Jun-22           | 7/22/22     |
| 0030011       | Ashford School District |                            | All Students                       | 343                                    | 0.9309                                   | 344                     | 0.955                     | 348                     | 0.9585                    | Jun-22           | 7/22/22     |
| 0030011       | Ashford School District | Students With Disabilities | Students With Disabilities         | 46                                     | 0.9253                                   | 47                      | 0.9573                    | 56                      | 0.9468                    | Jun-22           | 7/22/22     |
| 0030011       | Ashford School District | Free/Reduced Lunch         | Free Meal Eligible                 | 92                                     | 0.9167                                   | 98                      | 0.9281                    |                         |                           | Jun-22           | 7/22/22     |
| 0030011       | Ashford School District | Free/Reduced Lunch         | Reduced Price Meal Eligible        | 21                                     | 0.9174                                   | 28                      | 0.9424                    |                         |                           | Jun-22           | 7/22/22     |

# Data Set 2

## Data Set 2

- Breaks down into the following information:
  - School Districts in Connecticut
- Looks at the following:
  - School ID
  - District Name
  - Mailing Address
  - City
  - Zip Code
  - Telephone



# Data Set 2 - CT Public School Demographics

| NCES<br>LEA ID | State ID | District Name                | Mailing Address        | City           | Zip Code | Telephone  | Locale |
|----------------|----------|------------------------------|------------------------|----------------|----------|------------|--------|
| 1              | 2        | 3                            | 4                      | 5              | 6        | 7          |        |
| 900030         | 001      | ANDOVER SCHOOL DISTRICT      | 35 SCHOOL ROAD         | ANDOVER        | 06232    | 8607427339 | 8      |
| 900090         | 003      | ASHFORD SCHOOL DISTRICT      | P.O. BOX 128           | ASHFORD        | 06278    | 8604291927 | 7      |
| 900150         | 005      | BARKHAMSTED SCHOOL DISTRICT  | 65 RIPLEY HILL ROAD    | BARKHAMSTED    | 06063    | 8607384016 | 7      |
| 900240         | 008      | BETHANY SCHOOL DISTRICT      | 44 PECK ROAD           | BETHANY        | 06524    | 2033931170 | 8      |
| 900390         | 013      | BOZRAH SCHOOL DISTRICT       | P.O. BOX 185           | BOZRAH         | 06334    | 8608874873 | 8      |
| 900630         | 021      | CANAAN SCHOOL DISTRICT       | 47 MAIN STREET         | FALLS VILLAGE  | 06031    | 8608247791 | 7      |
| 900660         | 022      | CANTERBURY SCHOOL DISTRICT   | 45 WESTMINSTER ROAD    | CANTERBURY     | 06331    | 8605466950 | 7      |
| 900720         | 024      | CHAPLIN SCHOOL DISTRICT      | P.O. BOX 277           | CHAPLIN        | 06235    | 8604559306 | 7      |
| 900780         | 026      | CHESTER SCHOOL DISTRICT      | P.O. BOX 187           | DEEP RIVER     | 06417    | 8605262417 | 4      |
| 900870         | 029      | COLEBROOK SCHOOL DISTRICT    | 452 SMITH HILL ROAD    | COLEBROOK      | 06021    | 8603792179 | 7      |
| 900900         | 030      | COLUMBIA SCHOOL DISTRICT     | P.O. BOX 166           | COLUMBIA       | 06237    | 8602289493 | 8      |
| 900930         | 031      | CORNWALL SCHOOL DISTRICT     | 246 WARREN TURNPIKE RO | FALLS VILLAGE  | 06031    | 8606726617 | 7      |
| 901080         | 036      | DEEP RIVER SCHOOL DISTRICT   | P.O. BOX 187           | DEEP RIVER     | 06417    | 8605262417 | 4      |
| 901360         | 253      | EASTERN CONNECTICUT REGIONAL | 376 HARTFORD TURNPIKE  | HAMPTON        | 06247    | 8604550707 | 6,7    |
| 901380         | 039      | EASTFORD SCHOOL DISTRICT     | P.O. BOX 158           | EASTFORD       | 06242    | 8609741130 | 7      |
| 903512         | 242      | EDUCATION CONNECTION         | 345 MAIN STREET        | DANBURY        | 06810    | 8605670863 | 7      |
| 901500         | 050      | ESSEX SCHOOL DISTRICT        | P.O. BOX 187           | DEEP RIVER     | 06417    | 8605262417 | 4      |
| 900018         | 272      | EXPLORATIONS DISTRICT        | 286 MAIN STREET        | WINSTED        | 06098    | 8607389070 | 6      |
| 901590         | 053      | FRANKLIN SCHOOL DISTRICT     | 206 POND ROAD          | NORTH FRANKLIN | 06254    | 8606427063 | 8      |
| 901890         | 063      | HAMPTON SCHOOL DISTRICT      | P.O. BOX 277           | CHAPLIN        | 06235    | 8604559306 | 7      |
| 901950         | 065      | HARTLAND SCHOOL DISTRICT     | 30 SOUTH ROAD          | EAST HARTLAND  | 06027    | 8606537207 | 8      |
| 902040         | 068      | KENT SCHOOL DISTRICT         | 246 WARREN TURNPIKE RO | FALLS VILLAGE  | 06031    | 8609273537 | 7      |
| 902190         | 073      | LISBON SCHOOL DISTRICT       | 15 NEWENT ROAD         | LISBON         | 06351    | 8603765565 | 8      |
| 902760         | 092      | NEW HARTFORD SCHOOL DISTRICT | P.O. BOX 315           | NEW HARTFORD   | 06057    | 8603798546 | 7      |
| 902940         | 098      | NORFOLK SCHOOL DISTRICT      | 128 GREENWOODS ROAD EA | NORFOLK        | 06058    | 8605425553 | 7      |
| 903000         | 100      | NORTH CANAAN SCHOOL DISTRICT | 246 WARREN TURNPIKE RO | FALLS VILLAGE  | 06031    | 8608245149 | 7      |
| 903360         | 112      | POMFRET SCHOOL DISTRICT      | 20 POMFRET STREET      | POMFRET CENTER | 06259    | 8609282718 | 7      |
| 903420         | 114      | PRESTON SCHOOL DISTRICT      | 435 SHETUCKET TURNPIKE | PRESTON        | 06365    | 8608896098 | 8      |
| 903600         | 064      | REGIONAL SCHOOL DISTRICT 04  | 246 WARREN TURNPIKE RO | FALLS VILLAGE  | 06031    | 8608245149 | 7      |

# Cleaning/Exploring Process

# Preliminary Cleaning Process

- Imported csv
- Checked counts for each column
- Retrieved all column names
- Checked for data types

```
In [3]: #Count number of districts
name_unique = data_set_df["District name"].value_counts()
name_count = len(name_unique)
name_count
```

```
Out[3]: 201
```

```
In [4]: #List all columns
data_set_df.columns
```

```
Out[4]: Index(['District code', 'District name', 'Category', 'Student group',
              '2021-2022 student count - year to date',
              '2021-2022 attendance rate - year to date', '2020-2021 student c
              ount',
              '2020-2021 attendance rate', '2019-2020 student count',
              '2019-2020 attendance rate', 'Reporting period', 'Date update'],
              dtype='object')
```

```
In [5]: #Check for incomplete rows
data_set_df.count()
```

```
Out[5]: District code      2019
        District name     2019
        Category         1818
        Student group     2019
```

# Cleaning - Resizing Data Frames

Both data sets had more information than needed.

|   | District code | District name | Category                   | Student group                      | 2021-2022 student count - year to date | 2021-2022 attendance rate - year to date | 2020-2021 student count | 2020-2021 attendance rate | 2019-2020 student count | 2019-2020 attendance rate | Reporting period | Date update |
|---|---------------|---------------|----------------------------|------------------------------------|--|--|-------------------------|---------------------------|-------------------------|---------------------------|------------------|-------------|
| 0 | 00000CT       | Connecticut   | NaN                        | All Students                       | 500285                                 | 0.9169                                   | 496092.0                | 0.9294                    | 508346.0                | 0.9479                    | June 2022        | 07/22/2022  |
| 1 | 00000CT       | Connecticut   | Homelessness               | Students Experiencing Homelessness | 1814                                   | 0.8348                                   | 1735.0                  | 0.8155                    | 3916.0                  | 0.8884                    | June 2022        | 07/22/2022  |
| 2 | 00000CT       | Connecticut   | Students With Disabilities | Students With Disabilities         | 78417                                  | 0.8899                                   | 76487.0                 | 0.8946                    | 80365.0                 | 0.9277                    | June 2022        | 07/22/2022  |
| 3 | 00000CT       | Connecticut   | Free/Reduced Lunch         | Free Meal Eligible                 | 168984                                 | 0.8851                                   | 176225.0                | 0.8861                    | 193706.0                | 0.9314                    | June 2022        | 07/22/2022  |
| 4 | 00000CT       | Connecticut   | Free/Reduced Lunch         | Reduced Price Meal Eligible        | 29905                                  | 0.9184                                   | 30886.0                 | 0.9299                    | 27507.0                 | 0.9518                    | June 2022        | 07/22/2022  |
| 5 | 00000CT       | Connecticut   | Free/Reduced Lunch         | Free/Reduced Price Meal Eligible   | 198889                                 | 0.8901                                   | 207111.0                | 0.8927                    | 221213.0                | 0.9340                    | June 2022        | 07/22/2022  |
| 6 | 00000CT       | Connecticut   | English Learners           | English Learners                   | 43571                                  | 0.8976                                   | 40619.0                 | 0.8948                    | 45413.0                 | 0.9389                    | June 2022        | 07/22/2022  |
| 7 | 00000CT       | Connecticut   | Race/Ethnicity             | All other races                    | 48700                                  | 0.9314                                   | 47339.0                 | 0.9483                    | 47260.0                 | 0.9559                    | June 2022        | 07/22/2022  |
| 8 | 00000CT       | Connecticut   | Race/Ethnicity             | Black or African American          | 63099                                  | 0.8941                                   | 62267.0                 | 0.8931                    | 64200.0                 | 0.9401                    | June 2022        | 07/22/2022  |
| 9 | 00000CT       | Connecticut   | Race/Ethnicity             | Hispanic/Latino of any race        | 146298                                 | 0.8935                                   | 138260.0                | 0.8975                    | 136953.0                | 0.9362                    | June 2022        | 07/22/2022  |

# Cleaning - Filtering and Merging Data Frames

- .loc to filter by groups
- Merged Data Frames

```
#Filter by English learners
ell_df = data_set_df.loc[(data_set_df["Student group"]=="English Learners")]
all_df = data_set_df.loc[(data_set_df["Student group"]=="All Students")]

#Limit data frame again
ell_df = ell_df[["District name", "Student group", "2021-2022 student count - year to date"]]
all_df = all_df[["District name", "Student group", "2021-2022 student count - year to date",
                "2021-2022 attendance rate - year to date"]]

#Merge Data sets right
merged_ell_df = pd.merge(all_df, ell_df, on="District name", how="right")

#Check counts
merged_ell_df.head()
```

|   | District name           | Student group_x | 2021-2022 student count - year to date_x | 2021-2022 attendance rate - year to date | Student group_y  | 2021-2022 student count - year to date_y |
|---|-------------------------|-----------------|--|--|------------------|--|
| 0 | Connecticut             | All Students    | 500285                                   | 0.9169                                   | English Learners | 43571                                    |
| 1 | Ansonia School District | All Students    | 2139                                     | 0.9045                                   | English Learners | 122                                      |
| 2 | Avon School District    | All Students    | 3057                                     | 0.9457                                   | English Learners | 76                                       |
| 3 | Berlin School District  | All Students    | 2613                                     | 0.9341                                   | English Learners | 95                                       |
| 4 | Bethel School District  | All Students    | 3102                                     | 0.9375                                   | English Learners | 167                                      |

# Exploring - Calculate Total Percentages

```
#Calculate ELL percents of total
merged_ell_df["Percentage ELL"]=merged_ell_df["2021-2022 student count - year to date_y"]\
/merged_ell_df['2021-2022 student count - year to date_x']

#Check Count
merged_ell_df.head()
```

|   | District name           | Student group_x | 2021-2022 student count - year to date_x | 2021-2022 attendance rate - year to date | Student group_y  | 2021-2022 student count - year to date_y | Percentage ELL |
|---|-------------------------|-----------------|--|--|------------------|--|----------------|
| 0 | Connecticut             | All Students    | 500285                                   | 0.9169                                   | English Learners | 43571                                    | 0.087092       |
| 1 | Ansonia School District | All Students    | 2139                                     | 0.9045                                   | English Learners | 122                                      | 0.057036       |
| 2 | Avon School District    | All Students    | 3057                                     | 0.9457                                   | English Learners | 76                                       | 0.024861       |
| 3 | Berlin School District  | All Students    | 2613                                     | 0.9341                                   | English Learners | 95                                       | 0.036357       |
| 4 | Bethel School District  | All Students    | 3102                                     | 0.9375                                   | English Learners | 167                                      | 0.053836       |

- Added new column where calculated percentage of student group

# Plotting



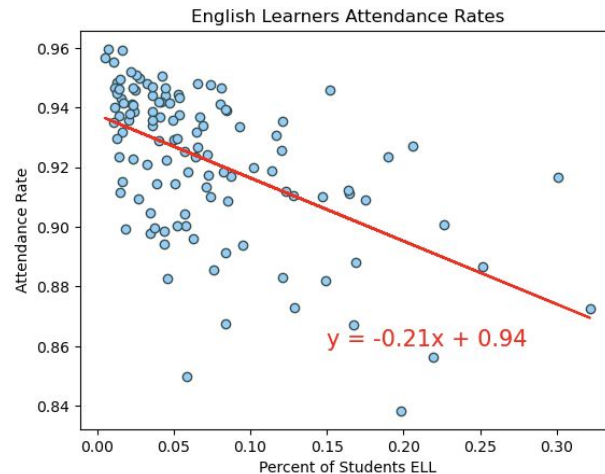
# Plot - Scatter Plot

```
#Plot ELL Attendance Rates
x_values = merged_ell_df["Percentage ELL"]
y_values = merged_ell_df["2021-2022 attendance rate - year to date"]
(slope, intercept, rvalue, pvalue, stderr) = linregress(x_values, y_values)
regress_values = x_values * slope + intercept
line_eq = "y = " + str(round(slope,2)) + "x + " + str(round(intercept,2))
plt.scatter(x_values,y_values, edgecolors="darkslategray", facecolors="lightskyblue")
plt.plot(x_values,regress_values,"r-")
plt.xlabel('Percent of Students ELL')
plt.ylabel('Attendance Rate')
plt.annotate(line_eq,(0.15,0.86), fontsize=15, color = "red")
plt.title("English Learners Attendance Rates")
print(rvalue)

#Save Figure
plt.savefig("Images/ScatterELL.png")

plt.show()
```

-0.5324713816396521





# Plot - Stacked Bar Graph

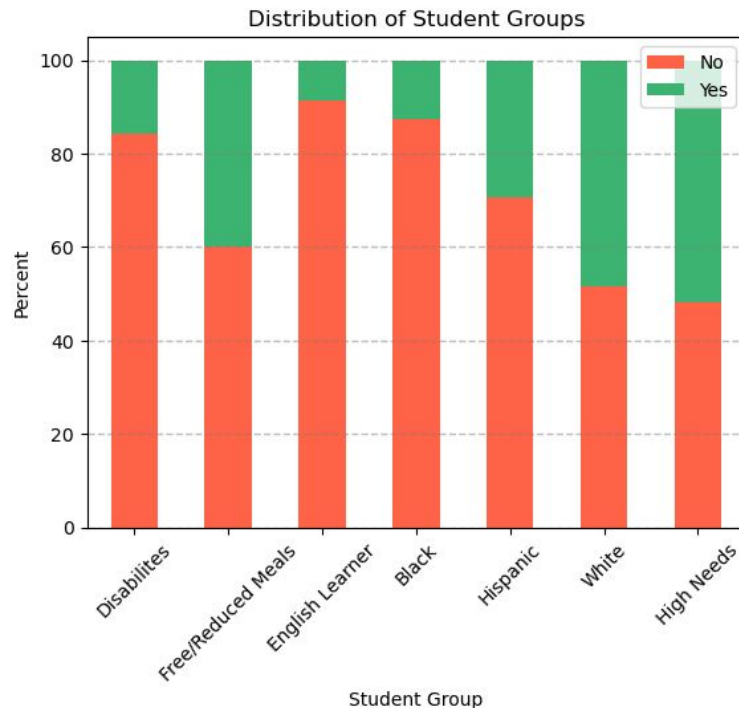
```
#Stacked Bar Graph
plotdata=pd.DataFrame({
    "No":[no_disabilities_percent, no_fr_percent, no_ell_percent, no_black_percent,
        no_his_percent, no_white_percent, no_needs_percent],
    "Yes":[disabilities_percent, fr_percent, ell_percent, black_percent, his_percent,
        white_percent, needs_percent]

    },
    index = ["Disabilites", "Free/Reduced Meals", "English Learner", "Black", "Hispanic",
        "White", "High Needs"])
plotdata.plot(rot=45, kind="bar", stacked = True, color={"No" : "tomato", "Yes":
    "mediumseagreen"})

plt.title("Distribution of Student Groups")
plt.xlabel("Student Group")
plt.ylabel("Percent")
plt.grid(color="gray", linestyle="--", linewidth=1, axis="y", alpha=0.5)

#Save image
plt.savefig("Images/StackedBar.png", bbox_inches="tight")

#Plot
plt.show()
```



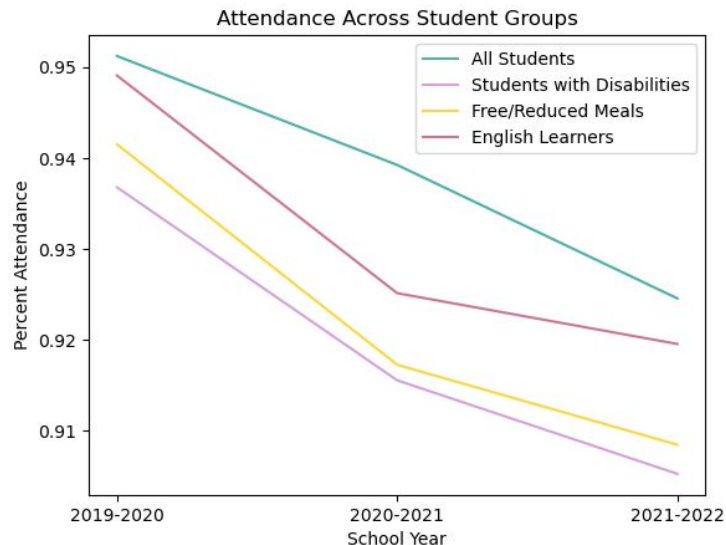
# Plot - Line Graphs

```
#Student Groups Line Graph
x = ["2019-2020", "2020-2021", "2021-2022"]
all_student = [attendance_2020_mean, attendance_2021_mean, attendance_2022_mean]
dis = [dis_2020_mean, dis_2021_mean, dis_2022_mean]
fr = [fr_2020_mean, fr_2021_mean, fr_2022_mean]
ell = [ell_2020_mean, ell_2021_mean, ell_2022_mean]

plt.plot(x, all_student, color="lightseagreen", label = "All Students")
plt.plot(x, dis, color="plum", label = "Students with Disabilities")
plt.plot(x, fr, color="gold", label = "Free/Reduced Meals")
plt.plot(x, ell, color = "palevioletred", label="English Learners")

plt.xlabel("School Year")
plt.ylabel("Percent Attendance")
plt.legend()
plt.title("Attendance Across Student Groups")

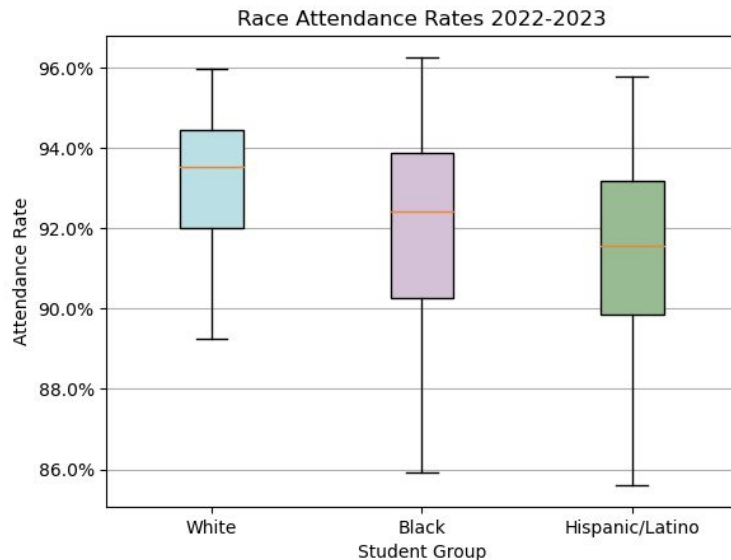
plt.savefig("Images/StudentGroupsLinePlot.png")
plt.show()
```



# Plot - Box and Whisker

```
#Races Plot No Outliers
labels = ["White", "Black", "Hispanic/Latino"]
#Plot
x_axis = [rates_2022_white, rates_2022_black, rates_2022_his]
fig, ax1 = plt.subplots()
ax1.set_ylabel("Attendance Rate")
ax1.set_xlabel("Student Group")
bp_white = ax1.boxplot(x_axis, patch_artist = True, labels = labels, showliers=False)

colors = ["powderblue", "thistle", "darkseagreen"]
for patch, color in zip(bp_white['boxes'], colors):
    patch.set_facecolor(color)
ax1.set_title("Race Attendance Rates 2022-2023")
ax1.yaxis.grid(True)
ax1.yaxis.set_major_formatter(mtick.PercentFormatter(xmax=1, decimals=None, symbol='%', is_latex=False))
plt.savefig("Images/RacesNoOutBoxPlot.png")
plt.show()
```



# Plot - Pie Charts

```
#Races Piechart

#Labels
labels = student_group

#Values
values = student_count

#Colors
colors = ["powderblue", "plum", "lightpink", "lightgreen"]

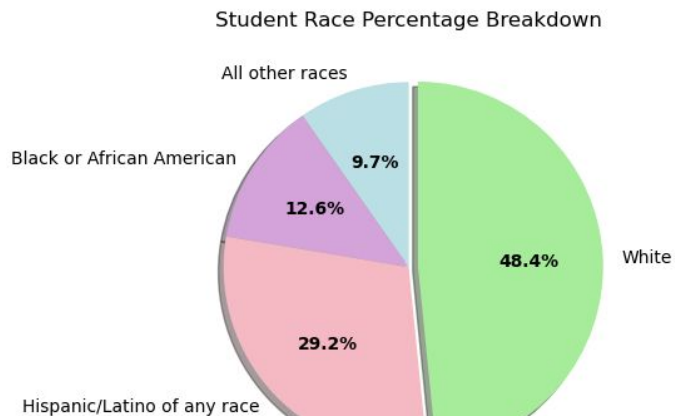
#Explode
explode = (0, 0, 0, 0.05)

#Plot with percents
fig1, ax1 = plt.subplots()
_, _ = ax1.pie(values, explode=explode, labels=labels, autopct='%1.1f%%',
              shadow=True, startangle=90, colors=colors)
plt.setp(ax1.patches, color='black', fontweight='bold')

#plt.pie(values, explode=explode, labels=labels, colors=colors,
#        # autopct="%1.1f%%", shadow=True, startangle=90)
plt.legend(labels, loc="lower right")
plt.title("Student Race Percentage Breakdown")

plt.savefig("Images/RacesPie.png", bbox_inches="tight")

plt.show()
```



# API – Geoapify

# Cleaning - Reducing Data Frame

*#Attempt to clean out populations*

```
all_students_df = data_set_df.loc[(data_set_df["Student group"]=="All Students")]
all_students_df = all_students_df[["District name", "2021-2022 student count - year to date"]]
all_students_df["District name"] = all_students_df["District name"].str.upper()
all_students_df
```

|      | District name                                     | 2021-2022 student count - year to date |
|------|---|--|
| 0    | CONNECTICUT                                       | 500285                                 |
| 13   | ANDOVER SCHOOL DISTRICT                           | 161                                    |
| 19   | ANSONIA SCHOOL DISTRICT                           | 2139                                   |
| 31   | ASHFORD SCHOOL DISTRICT                           | 343                                    |
| 40   | AVON SCHOOL DISTRICT                              | 3057                                   |
| ...  | ...   | ...                                    |
| 1972 | UNIFIED SCHOOL DISTRICT #2                        | 78                                     |
| 1977 | CONNECTICUT TECHNICAL EDUCATION AND CAREER SYSTEM | 11138                                  |
| 1990 | NORWICH FREE ACADEMY DISTRICT                     | 2068                                   |
| 2002 | THE GILBERT SCHOOL DISTRICT                       | 421                                    |
| 2011 | THE WOODSTOCK ACADEMY DISTRICT                    | 1060                                   |

201 rows x 2 columns

- Reduce to columns needed
- Make all uppercase District Name so can merge

# Exploring - Requests JSON

- Build URL and test zipcode to view request

```
# Import API key
from config3 import geoapify_key
```

```
#Set up base API URL
zip_code = "06473"
country = "United States"
url = f"https://api.geoapify.com/v1/geocode/search?text={zip_code}&format=json&apiKey={geoapify_key}"
geo_data = requests.get(url).json()

pprint(geo_data)
```

```
{'query': {'parsed': {'expected_type': 'unknown', 'postcode': '06473'},
  'text': '06473'},
  'results': [{'address_line1': 'North Haven',
    'address_line2': 'North Haven, CT 06473, United States of '
      'America',
    'bbox': {'lat1': 41.221443578114,
      'lat2': 41.541443578114,
      'lon1': -73.019669765131,
      'lon2': -72.699669765131},
    'city': 'North Haven',
    'country': 'United States',
    'country_code': 'us',
    'county': 'South Central Connecticut Planning Region',
    'datasource': {'attribution': '© OpenStreetMap contributors',
      'license': 'Open Database License',
```

# Exploring - Building Lists

```
: zip_code = []
  districts = []
  for item in merged_df["Zip Code"]:
      zip_code.append(item)
  for district in merged_df["2021-2022 student count - year to date"]:
      districts.append(district)
  districts
343,
3057,
199,
2613,
391,
3102,
1734,
745,
159,
2527,
18482,
7439,
2545,
804,
77,
429,
1462,
143,
4119,
214,
```

```
: district_data = []
  country_code = "us"

  i=0
  for code in zip_code:
      try:
          #zip_code_url = f"https://api.geoapify.com/v1/geocode/search?text={code}&format=json&apiKey={api_key}"
          zip_code_url = f"https://api.geoapify.com/v1/geocode/search?text={code}%20{country_code}&format=json&apiKey={api_key}"
          zip_code_requests = requests.get(zip_code_url).json()
          print(f"Zip Code URL: {zip_code_url}")

          #Find Lat and Long
          zip_lat = zip_code_requests['results'][0]['lat']
          zip_lng = zip_code_requests['results'][0]['lon']
          zip_code_1 = zip_code_requests["results"][0]["postcode"]

          # Append the City information into city_data list
          district_data.append({"Zip Code": zip_code_1,
                                "Lat": zip_lat,
                                "Lng": zip_lng,
                                "Population": districts[i]/75
                                })

      except:
          print(f"not working")
          i=i+1
```

Zip Code URL: <https://api.geoapify.com/v1/geocode/search?text=06232%20us&format=json&apiKey=c967a2b03b354238ac3aa60aa86eb97e>

Zip Code URL: <https://api.geoapify.com/v1/geocode/search?text=06401%20us&format=json&apiKey=c967a2b03b354238ac3aa60aa86eb97e>

Zip Code URL: <https://api.geoapify.com/v1/geocode/search?text=06278%20us&format=json&apiKey=c967a2b03b354238ac3aa60aa86eb97e>

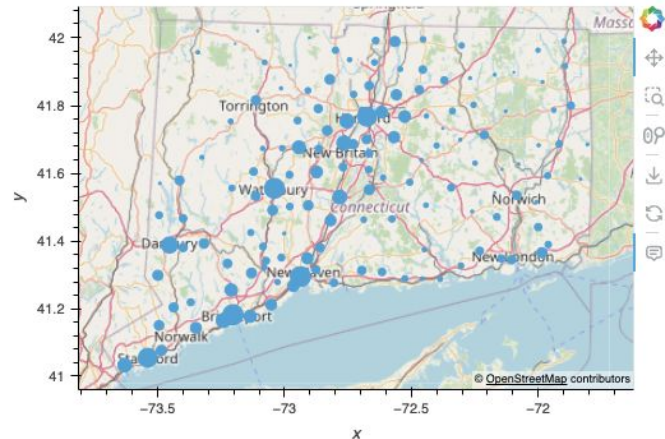


# Exploring - Building Data Frame and Plot

- Turn list of collected data into Data Frame
- Plot Data Frame using hvplot

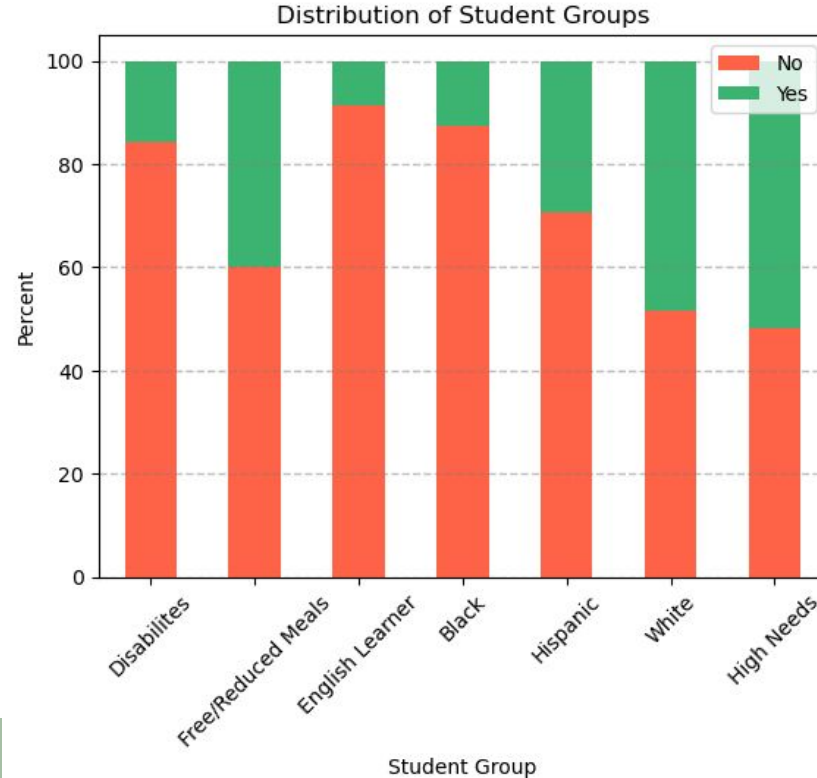
```
map_plot_1 = district_data_df.hvplot.points(  
    "Lng",  
    "Lat",  
    s="Population",  
    geo = True,  
    tiles = "OSM",  
)  
hvplot.save(map_plot_1, "population.png")  
map_plot_1
```

WARNING:bokeh.core.validation.check:W-1005 (FIXED\_SIZING\_MODE): 'fixed' sizing mode requires width and height to be set: figure(id='p1951', ...)

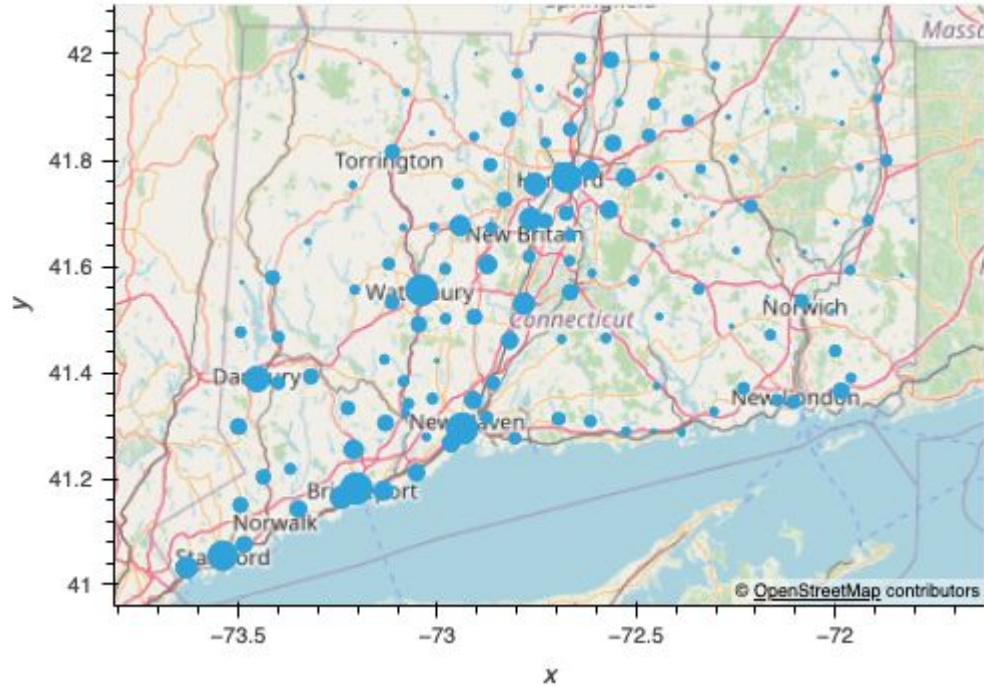


# Results

# Breakdown of Student Groups

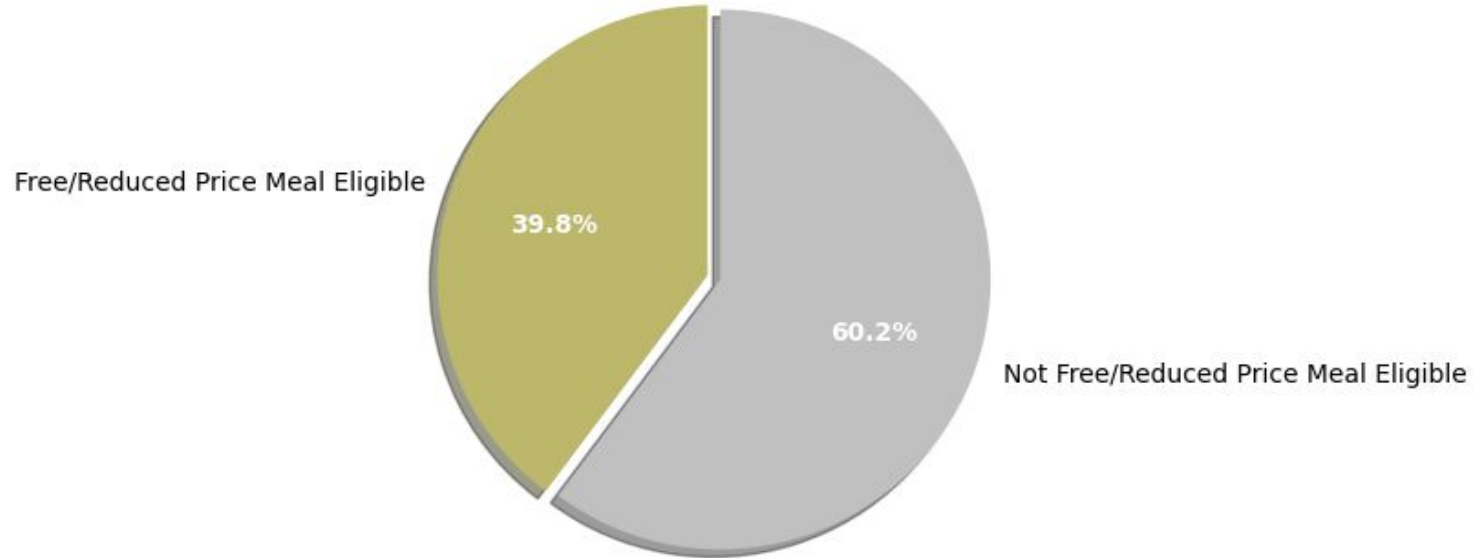


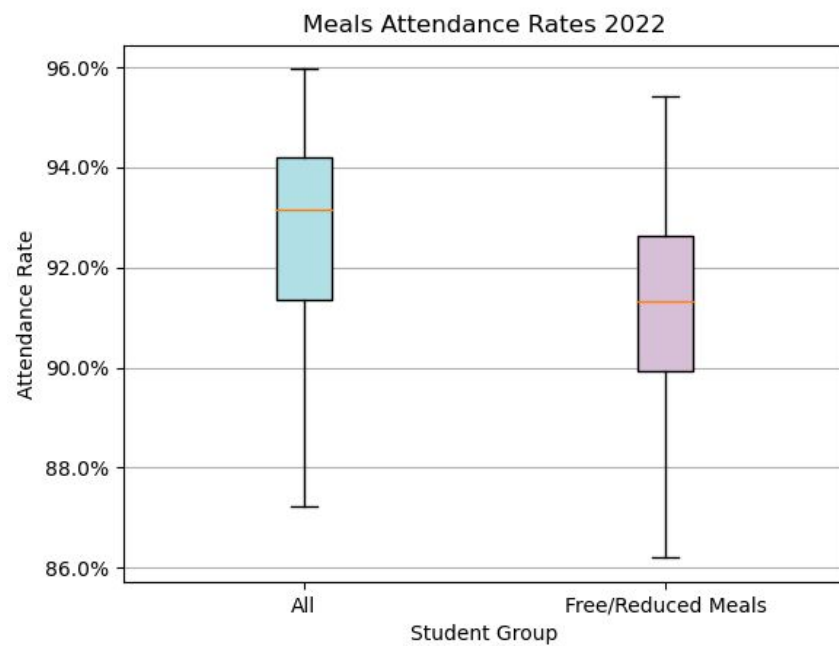
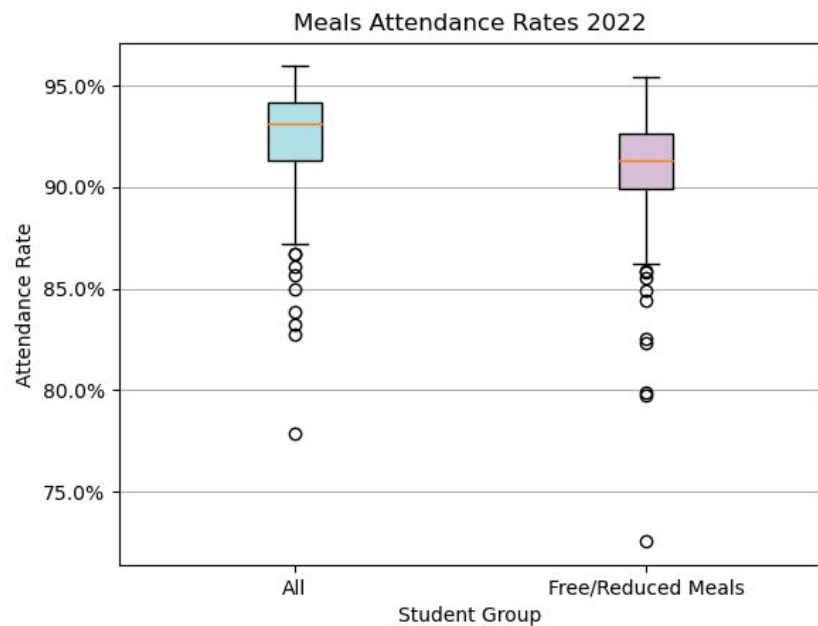
# Map of Student Population

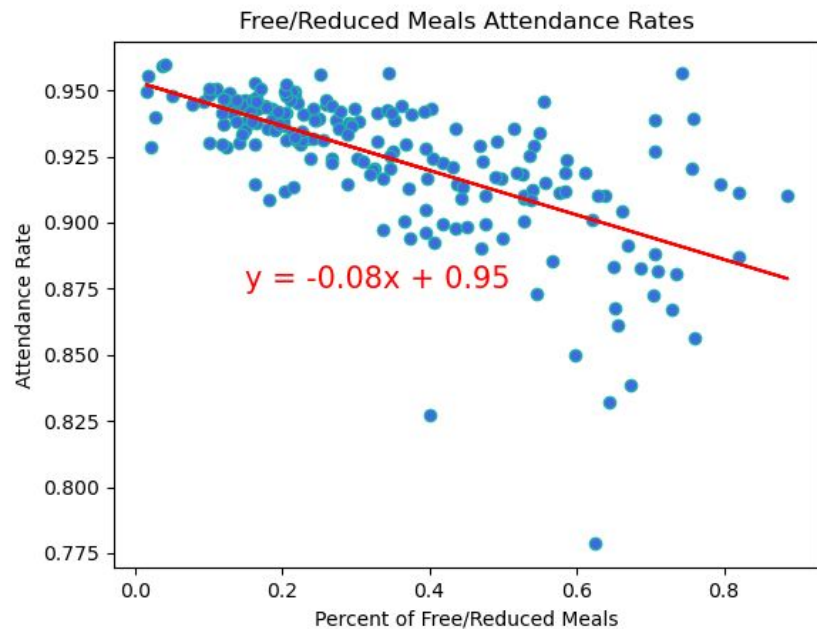


# 1.) Is there a correlation between the percentage of students receiving free or reduced meals in a district with the overall district attendance rate?

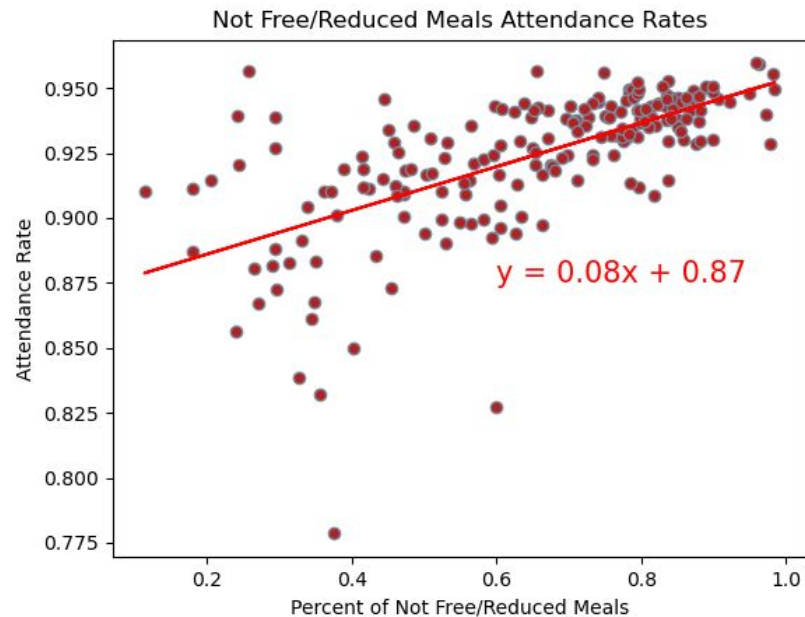
Free/Reduced Meals Percentage Breakdown



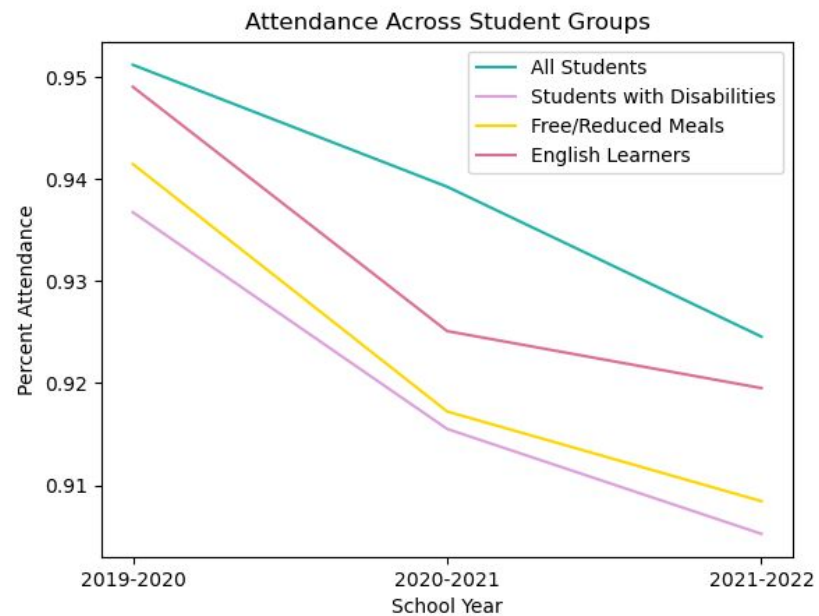
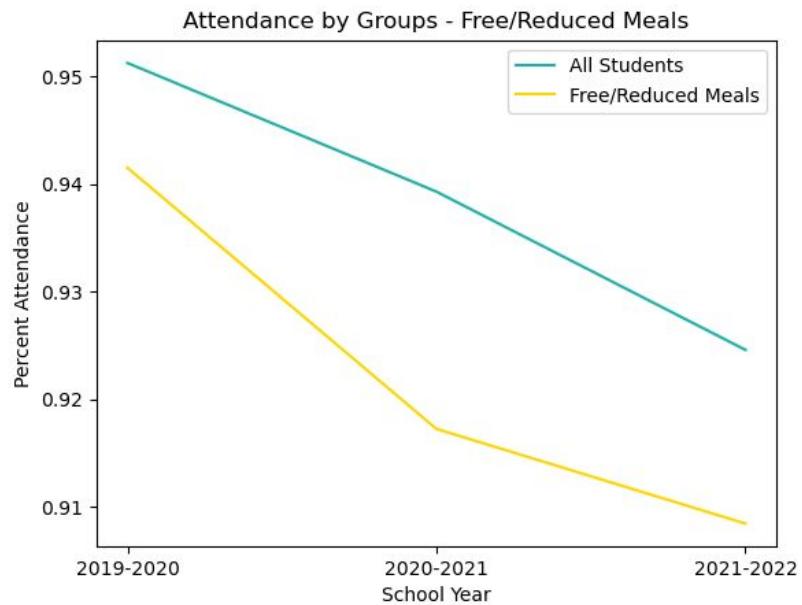




PCC = -0.64



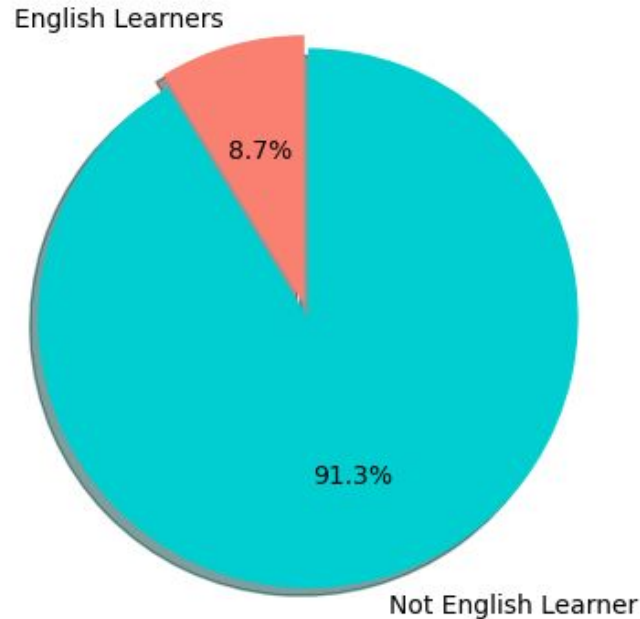
PCC = 0.64

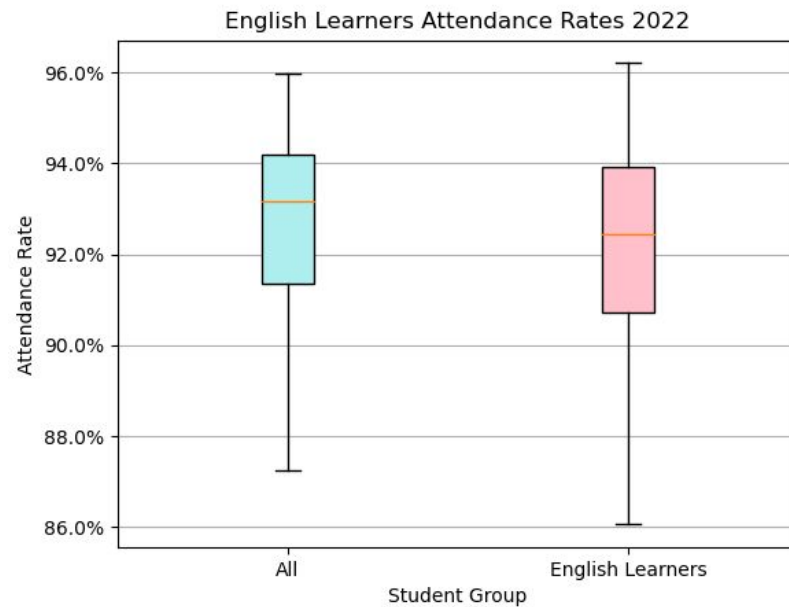
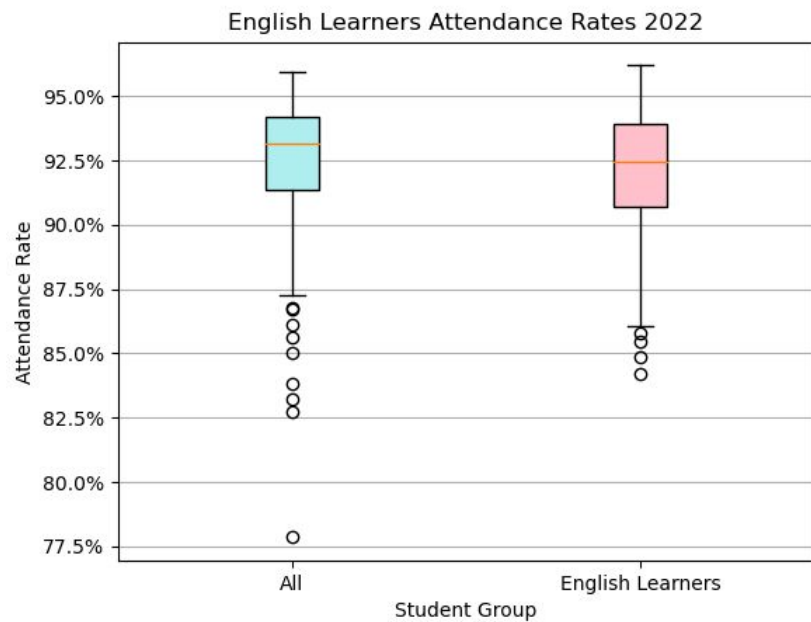


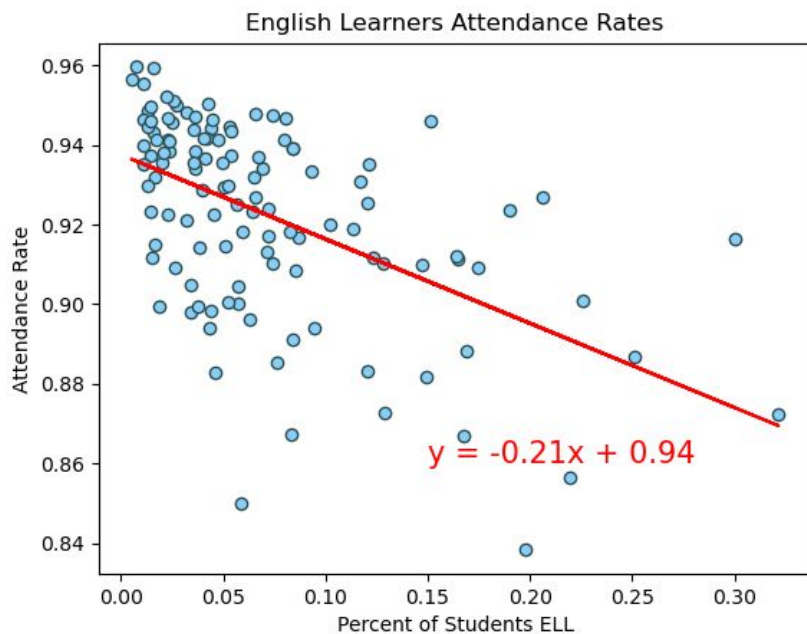


## 2.) Is there a correlation between the percentage of students who are English Language Learners in a district with the overall district attendance rate?

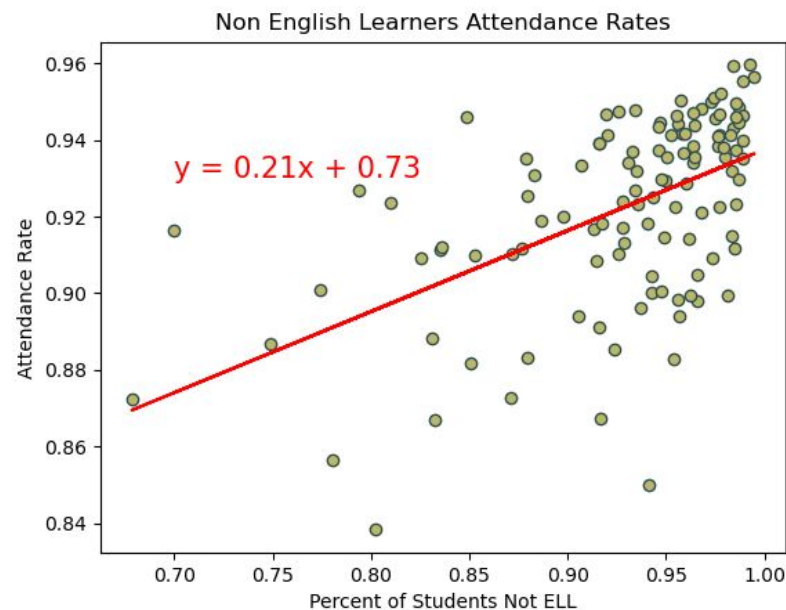
English Learners Percentage Breakdown



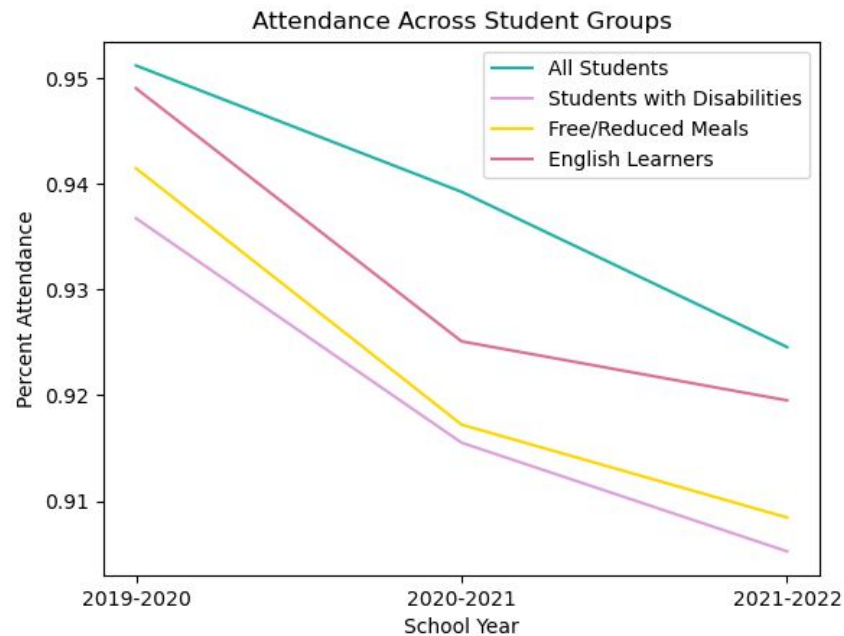
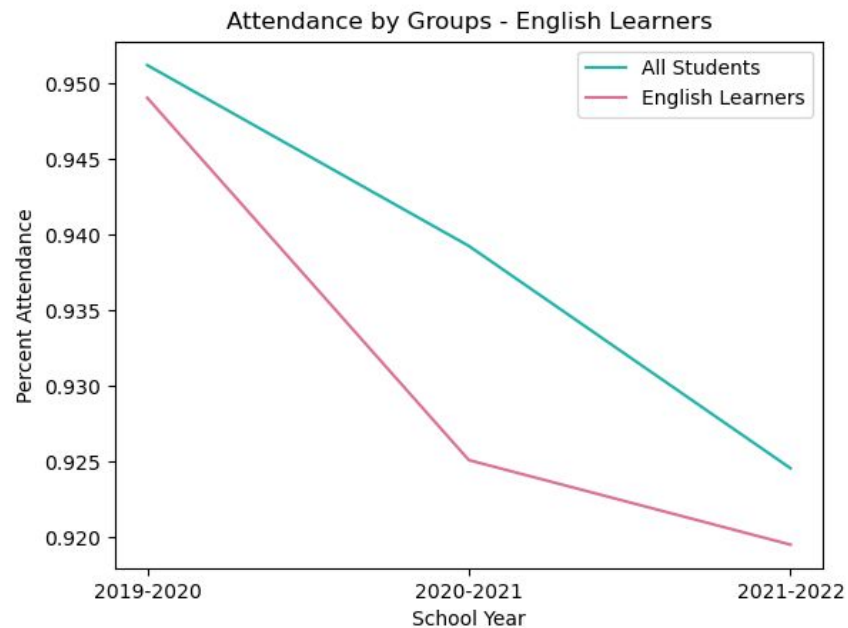




PCC = -0.53

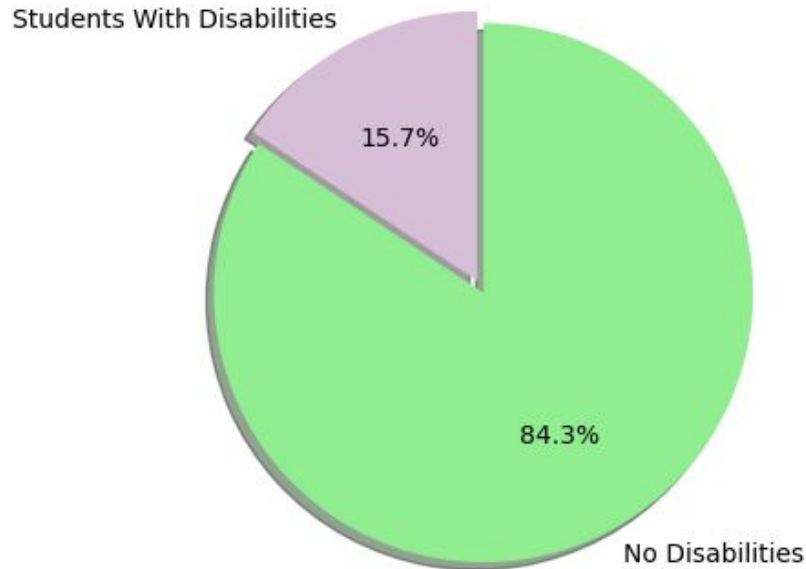


PCC = 0.53

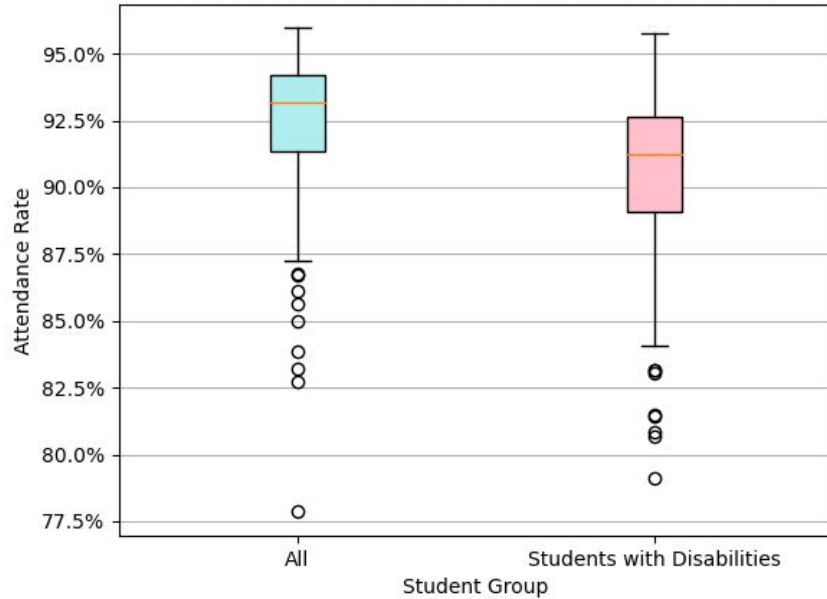


### 3.) Is there a correlation between the percentage of students with disabilities in a district with the overall district attendance rate?

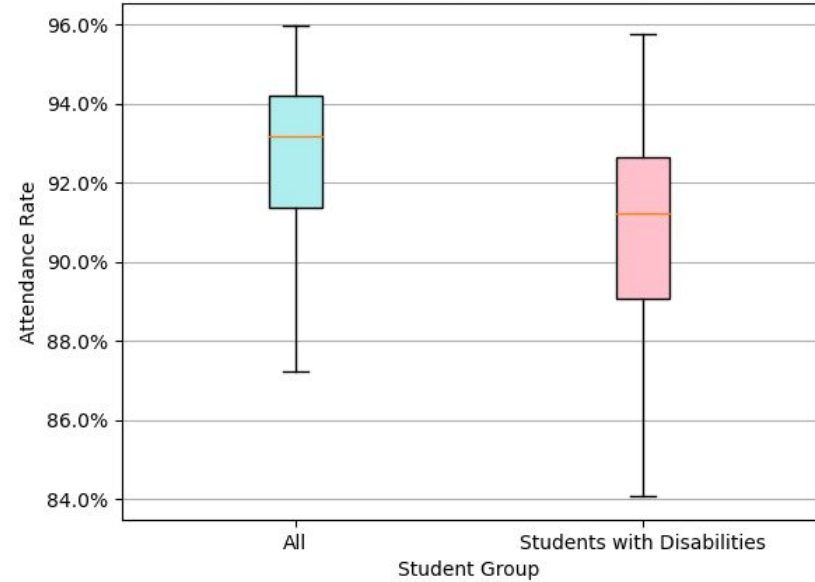
Students with Disabilities Percentage Breakdown

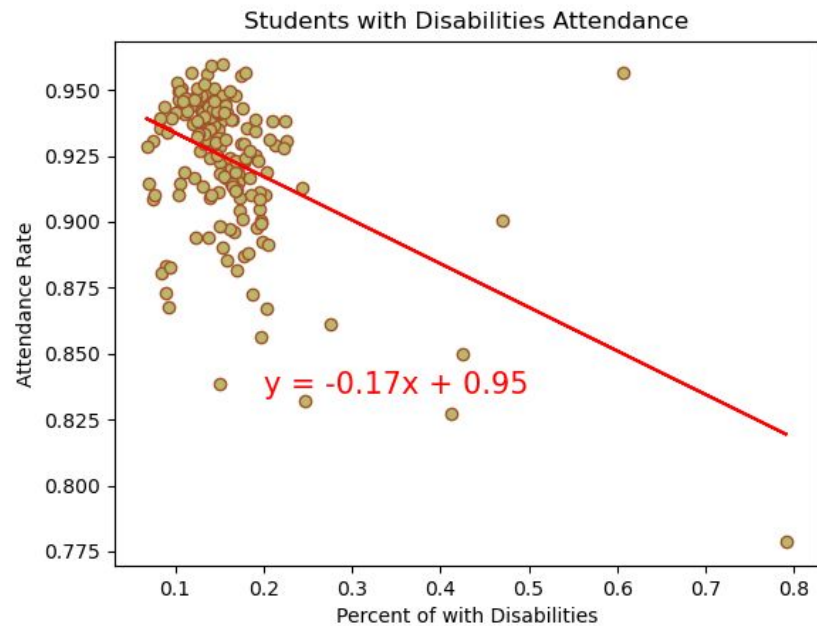


Students with Disabilities Attendance Rates 2022

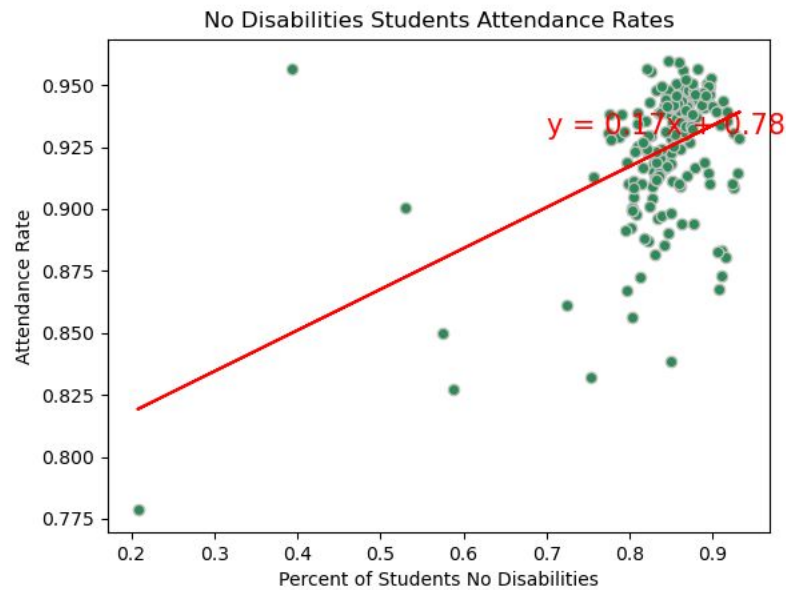


Students with Disabilities Attendance Rates 2022

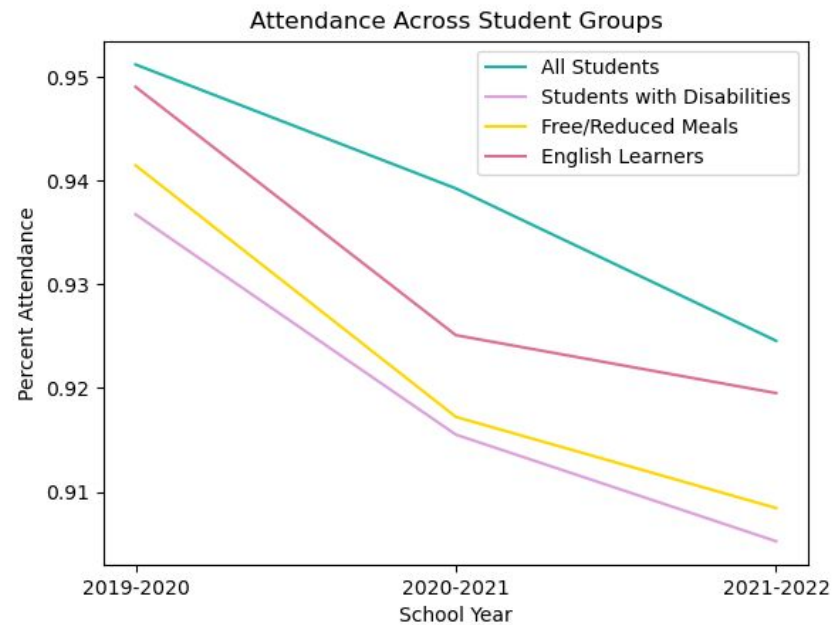
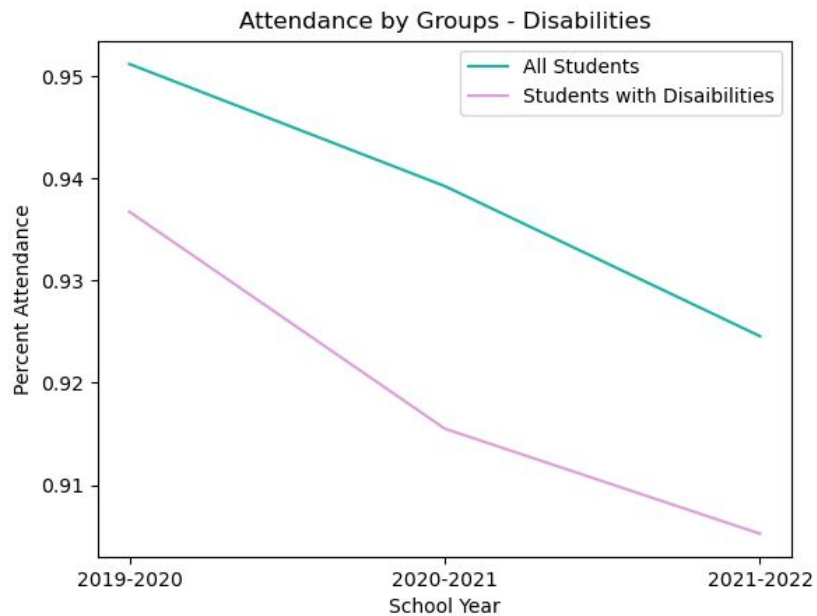




PCC = -0.47



PCC = 0.46



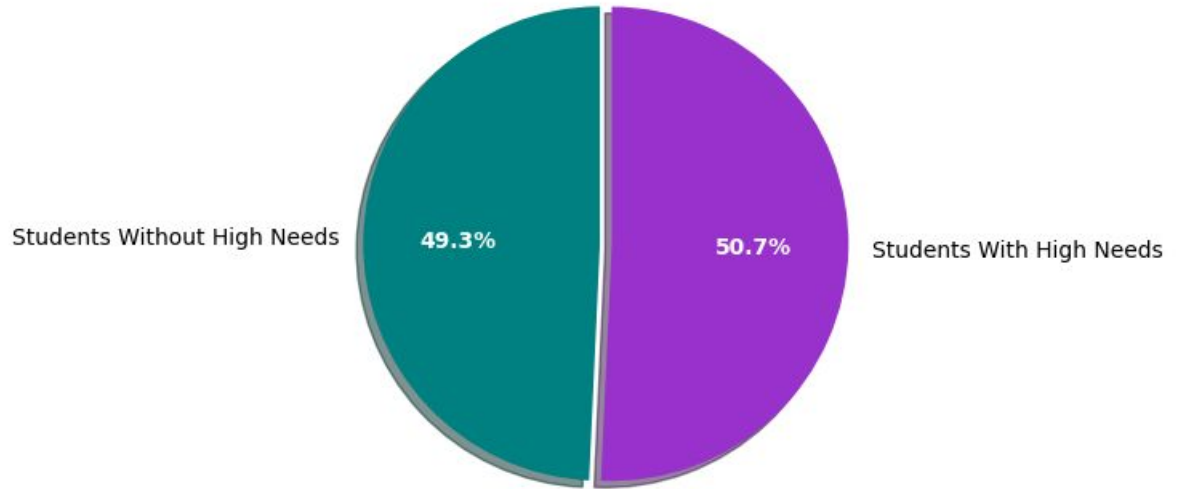


#### 4.) Is there a correlation between the percentage of students with high needs in a district with the overall district attendance rate?

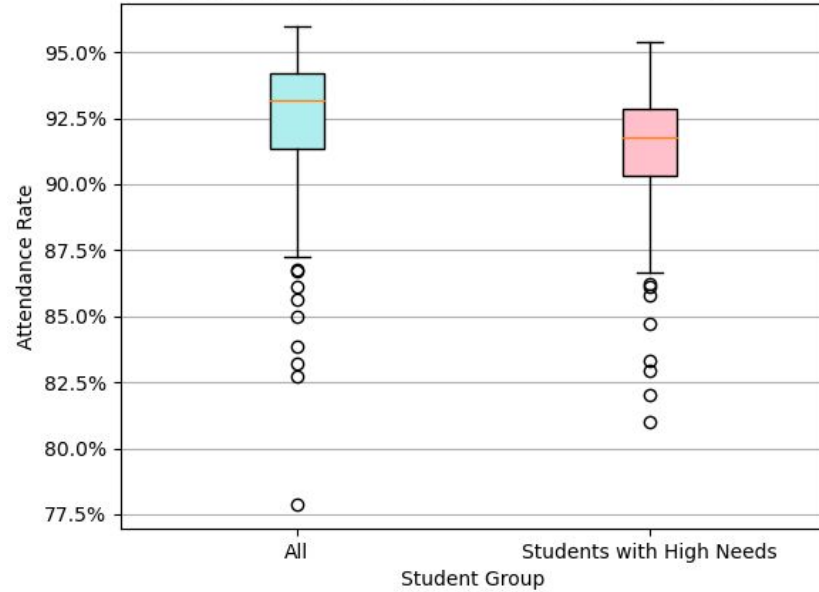
Students are included in the High Needs group if they are a student with a disability, an English Language Learner (ELL), or qualify for Free/Reduced Meals.

Source: ct.gov

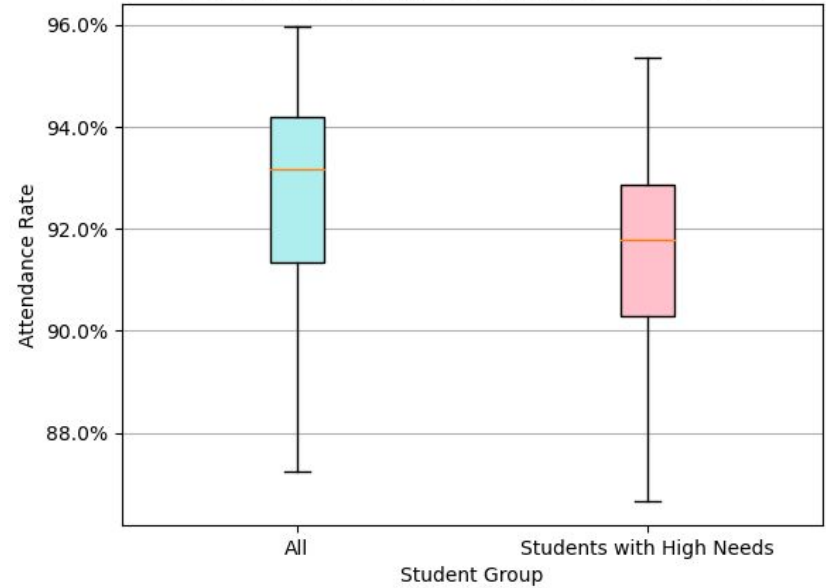
Student Needs Percentage Breakdown

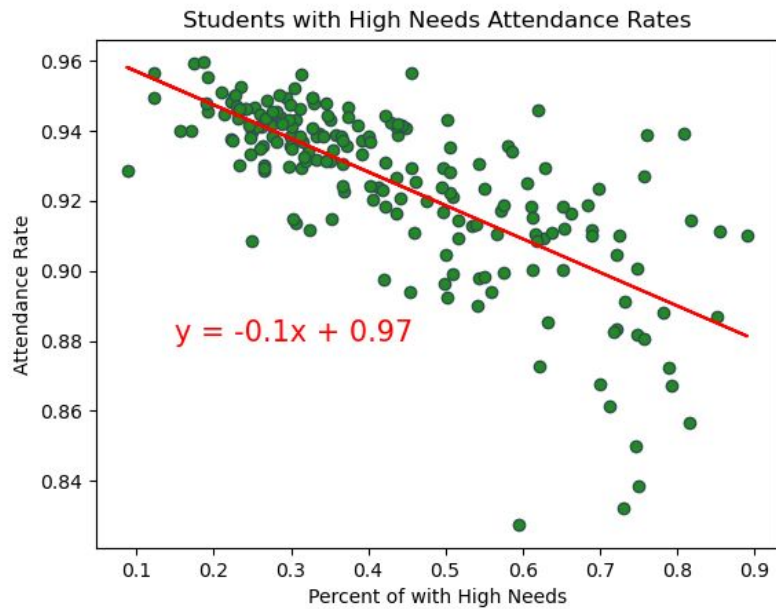


Students with High Needs Attendance Rates 2022

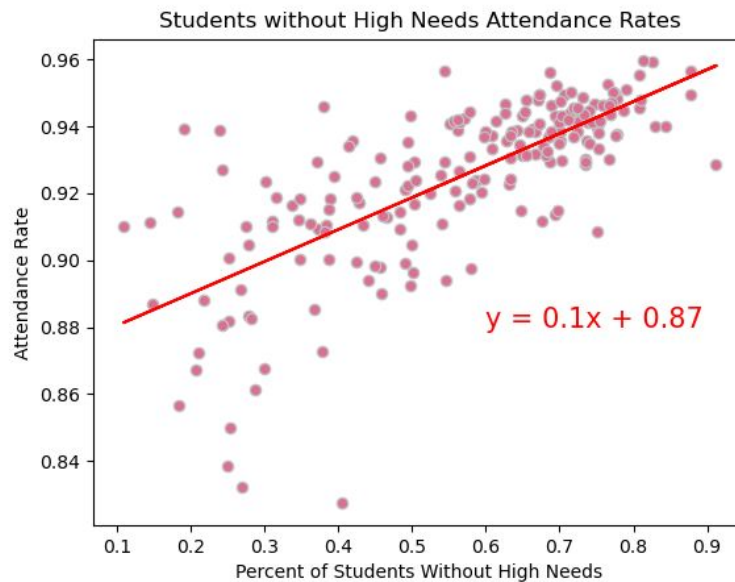


Students with High Needs Attendance Rates 2022

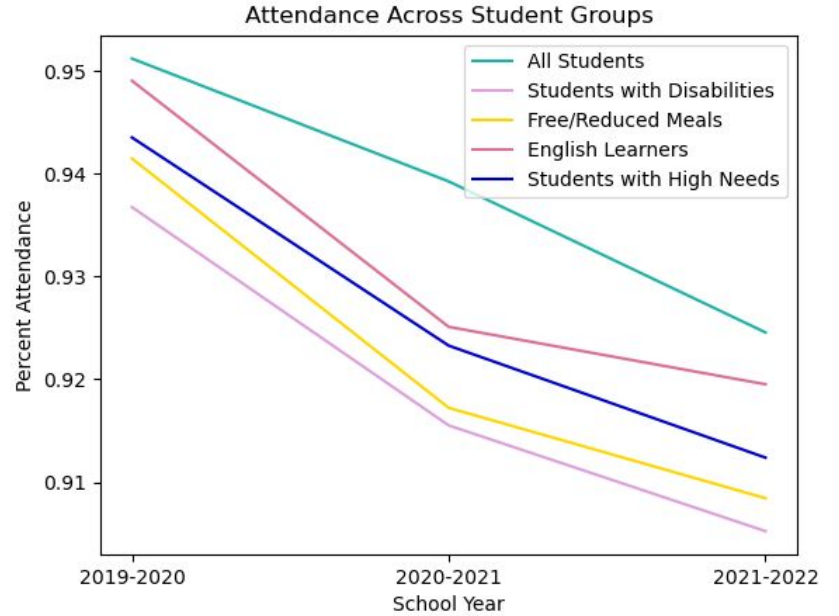
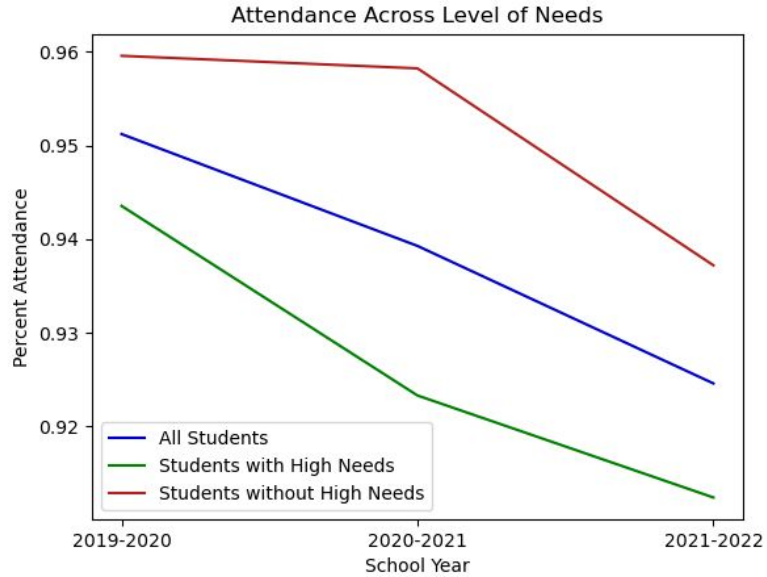




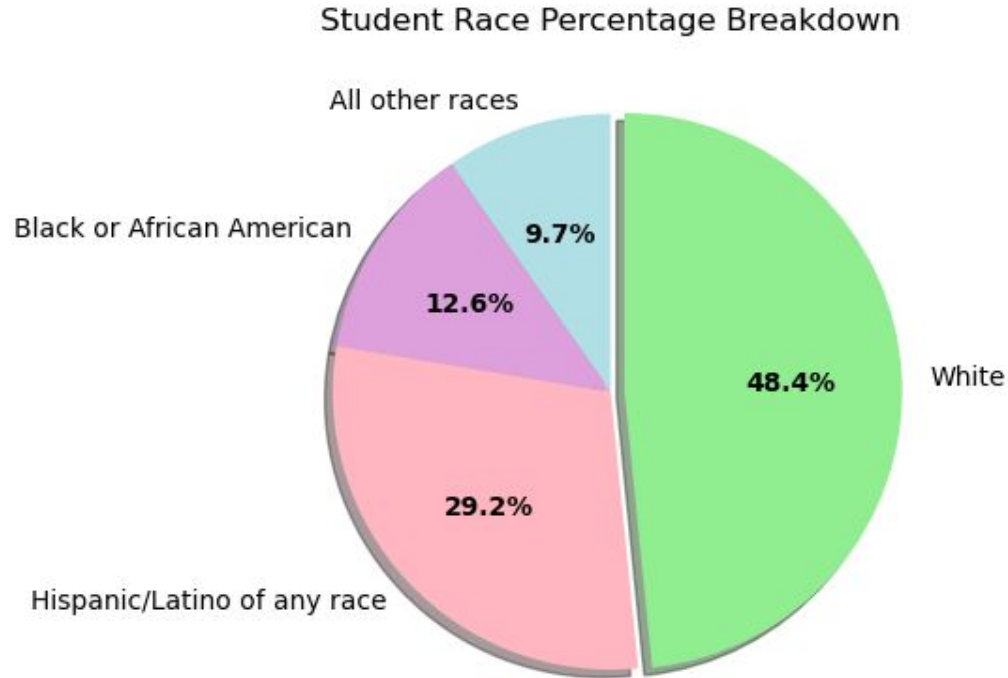
PCC = -0.71

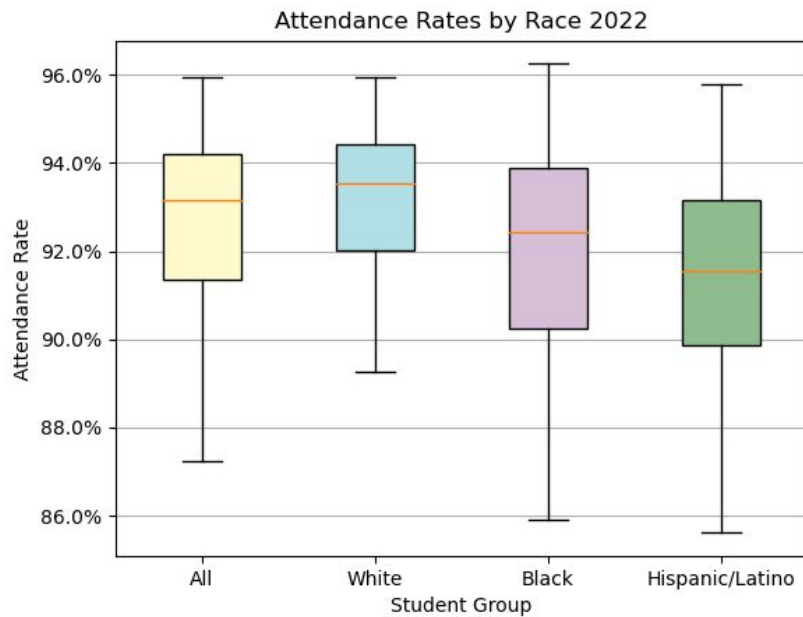
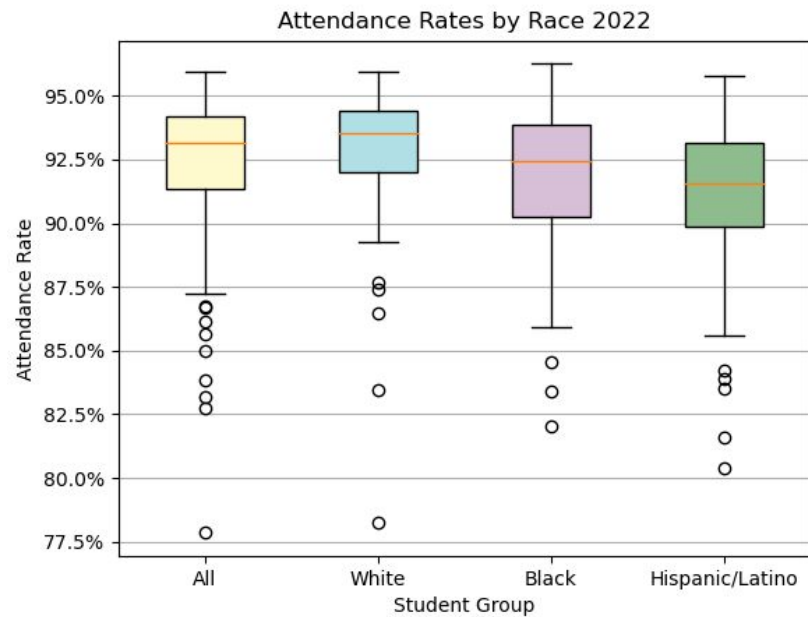


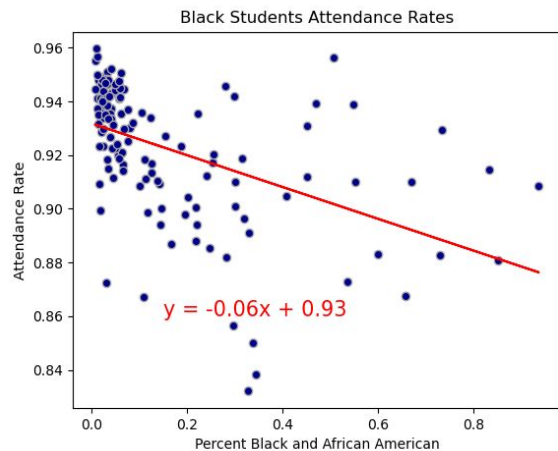
PCC = 0.71



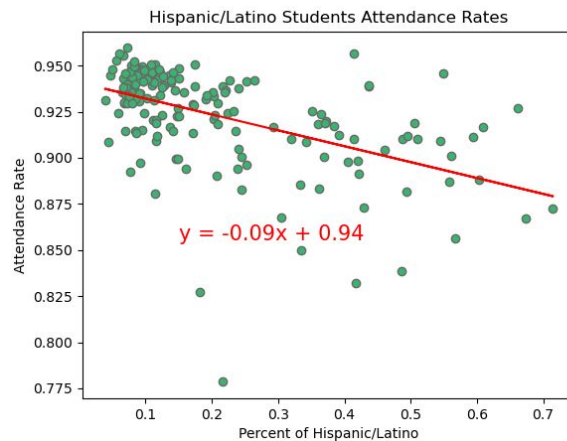
## 5.) Is there a correlation between the percentage of students by race in a district with the overall district attendance rate?



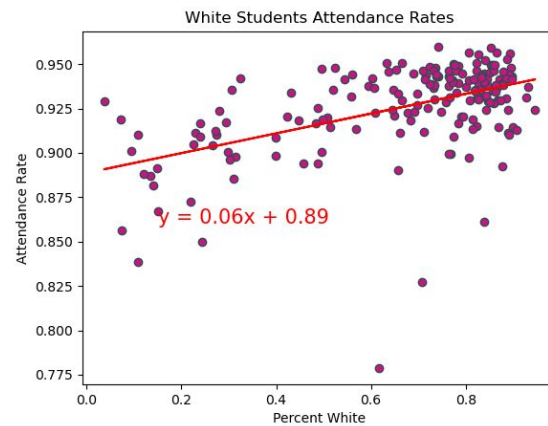




PCC = -0.47

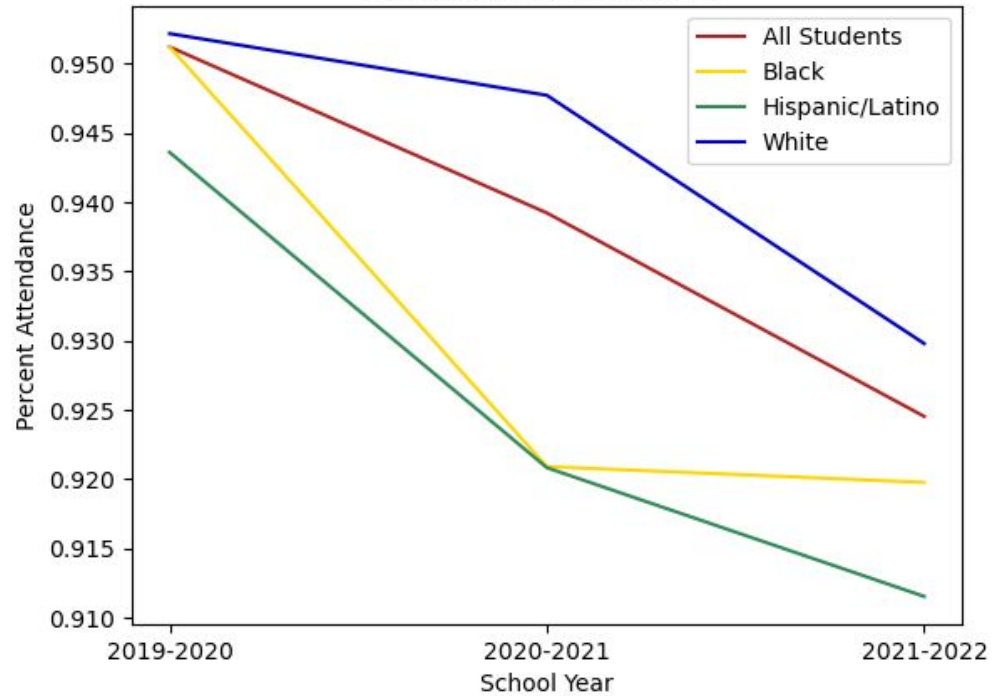


PCC = -0.5



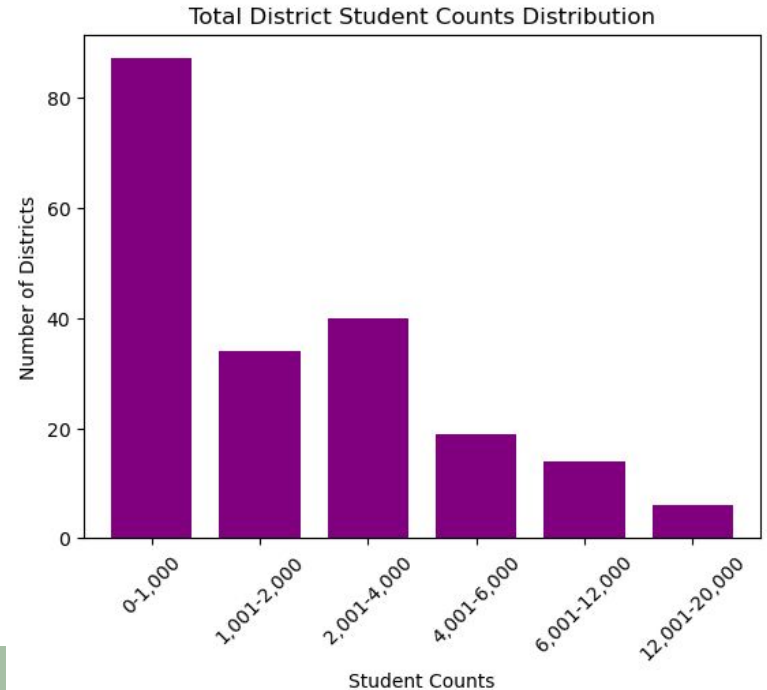
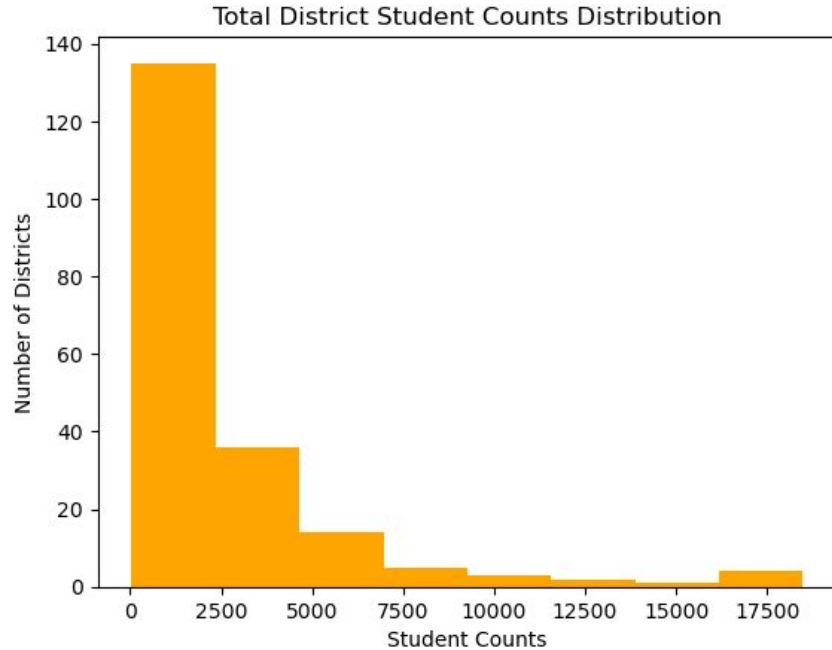
PCC = 0.51

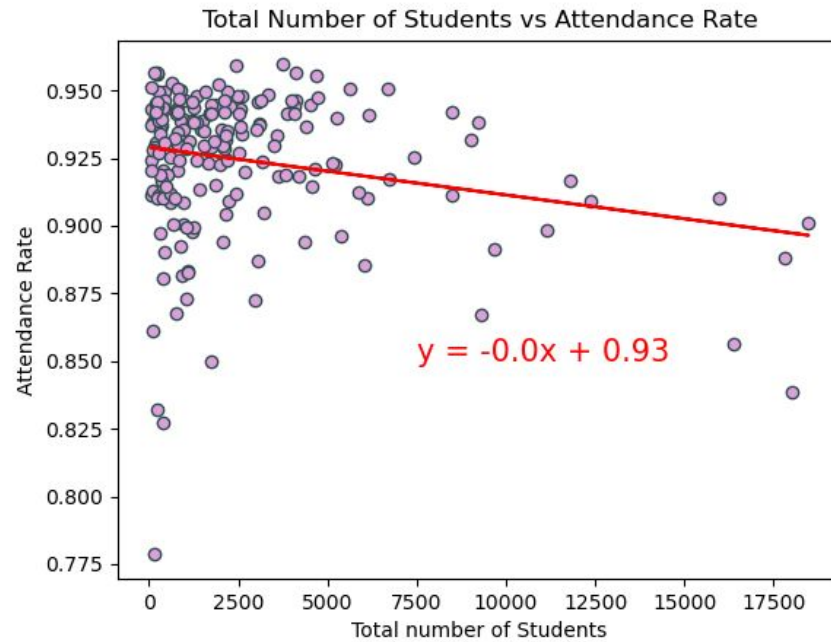
Attendance Across Races





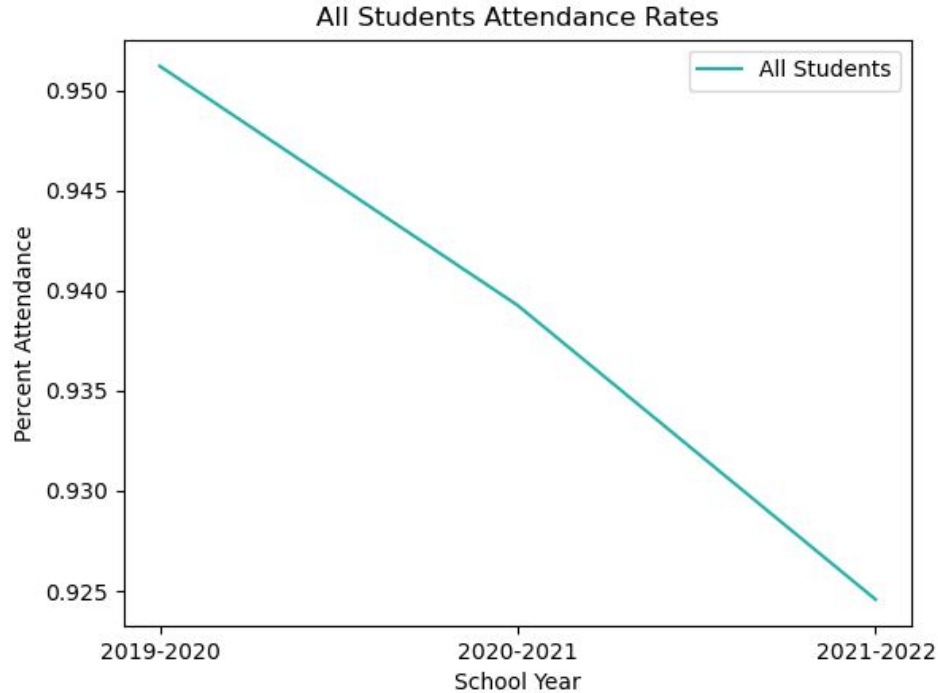
## 6.) Is there a correlation between the size of a district and the attendance rate?



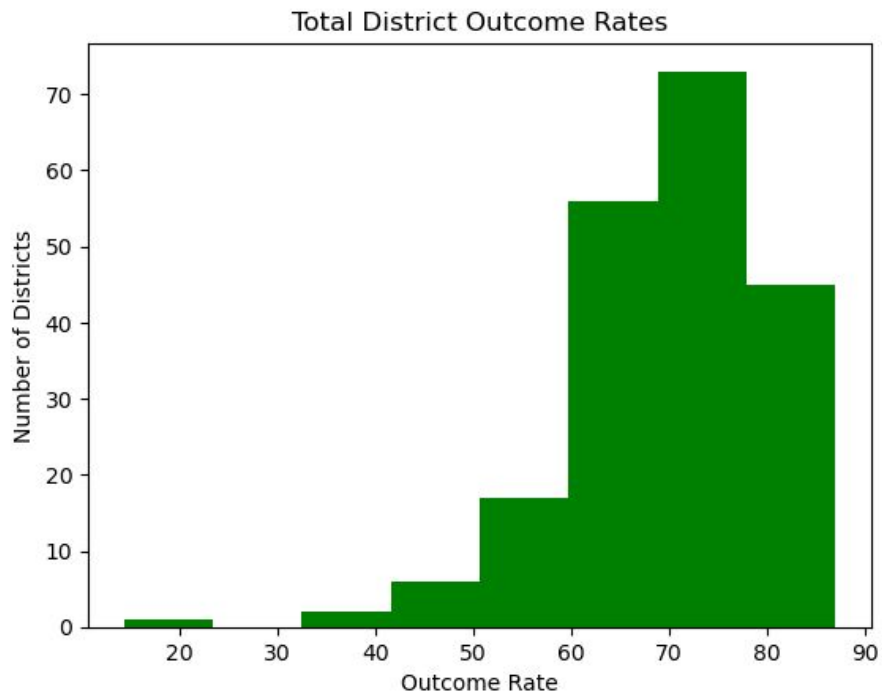
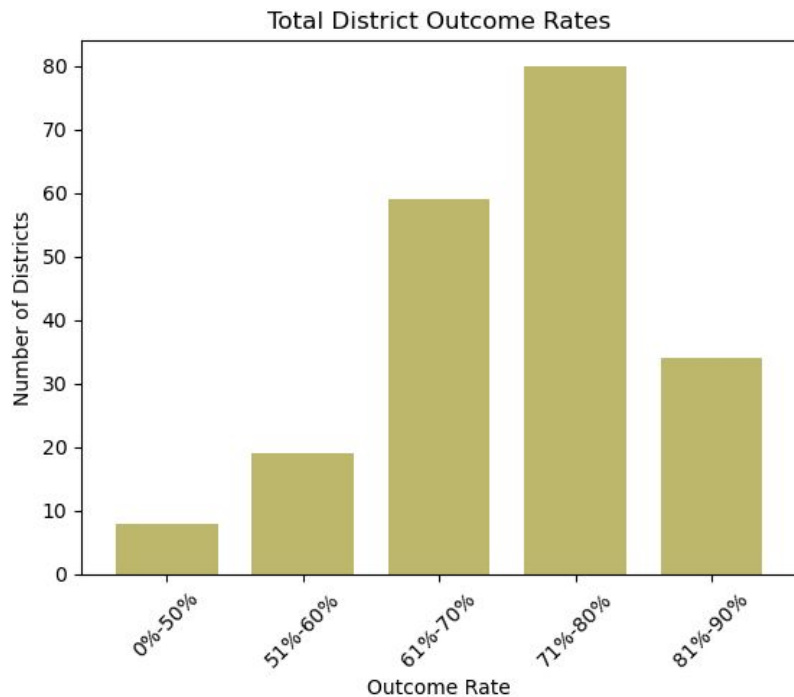


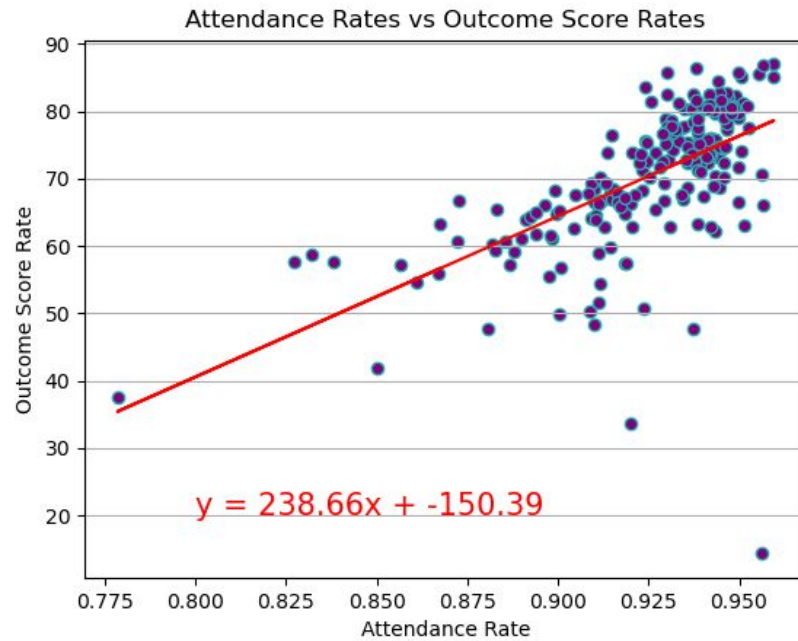
PCC = -0.22

## 7.) Is there a trend in the change in attendance rates over the years?



# Extension: How do API scores relate to student attendance?





PCC = 0.61

# Takeaways

- Strongest correlation between students with high needs and attendance rates
  - The higher percentage of students with high needs, the lower attendance rate
- Strong correlation between students receiving free/reduced meals and attendance.
  - The higher percentage of students receiving free/reduced meals, the lower the attendance rate
- Moderate correlation between race and attendance rates
  - The higher percentage of white students, the higher the attendance rate
- The attendance rate is not affected by the total number of students
- There is a strong correlation between attendance rate and API scores
  - The higher attendance rate, the higher API score

