dtsa_5304_nyc_math_test_results

December 10, 2024

Interactive elements (selectors) do not work and are not displayed in the pdf file. Use Jupiter Notebook and .ipynb file instead of pdf to evaluate interactive elements.

Final Project for DTSA 5304 - NYC Math Test Results 2013-2023

For this project I choose the Math Test Results 2013-2023 dataset from NYC OpenData. This dataset is a pre-aggregated dataset of math test scores for students in grades 3-8 in New York City public schools from 2013 to 2023. The dataset includes several files, those contain the same data but aggregate data by different attributes. This is convenient on one hand, that data is aggregated and I don't need to handle millions of records, but on other hand it imposes limitations and I will have to use different files for different visualisations.

All files in dataset have similar attributes: Location (borough, or school district, or particular school), Grade, Year, Category (gender or race of students), Number of tested students, Mean scale score, Levels 1 to 4 (quantity and percentages) and sum of levels 3+4 (also quantity and percentages).

I have seen several visualisations of this dataset, but this dataset is mostly used as an additional one, when analyzing, for example, economic factors or unemployment. Among the shortcomings of existing visualizations, one can note the low level of detail in the analysis of individual demographic groups and the absence of combined analysis (for example, the intersection of temporal and demographic sections). The intuitiveness of color schemes and scales is not always taken into account, which can affect the perception of data.

Goals:

Temporal Analysis: I want to examine trends in academic performance over the years. And identify changes in test scores over 10 years.

Demographic Analysis: I plan to visualize changes in performance from lower to higher grades. Compare results between student genders and races.

Geographical Analysis: Comparing performance between boroughs and identifying potential differences between school districts

General trend: Identification of patterns and possible dependencies in academic performance (combined influence of attributes on results).

Tasks:

1. Geographical Analysis of Academic Performance

- **Goal**: Identifying differences in test scores between school districts can help identify low-performing areas and possible reasons for these differences.
- **Means**: Create a heat map to show average test scores by borough. Overlay the data on a map of New York City for more detailed analysis.
- **Characteristics**: Examining how school location influences math test performance. Visual identification of problem areas in a city.
- **Target data**: Dataset files in which data is grouped by borough and school district.
- **Workflow**: In the initial stages of analysis to identify key areas requiring further investigation.
- **Roles**: Data analyst interested in educational trends by district.

2. Temporal Analysis of Academic Performance

- **Goal**: Identifying time trends can help us understand how academic performance has changed over the years and what factors (such as COVID) may have influenced this.
- **Means**: Making line graphs showing changes in average test scores over the years.
- **Characteristics**: Trends and patterns in data throw 2013–2023.
- **Target data**: Data grouped by year.
- **Workflow**: After geographical analysis to study in detail the dynamics in each borough.
- **Roles**: A researcher who analyses educational outcomes in the context of policy or external events.

3. Demographic Analysis of Academic Performance

- **Goal**: Comparing results across different demographic groups (grades and races) can help identify disparities and areas requiring intervention.
- **Means**: Making line graphs across grades and years to compare groups.
- **Characteristics**: Differences in results between groups, such as average scores or distribution of proficiency levels.
- **Target data**: Data grouped by demographic characteristics.
- **Workflow**: In the final stages of analysis, to draw conclusions about the influence of demographic factors.
- **Roles**: A researcher working on the social dimension in education.

Evaluation Strategy

1. Key Assessment Question. How does visualization help us understand geographic, temporal, and demographic trends in math test scores? Including: - Identifying differences in performance between boroughs. - Analyzing changes in performance over the years. - Exploring the impact of demographic factors on scores.

2. Audience for evaluation. Practitioners who analyze data to make decisions. School administrators who make decisions about instructional strategies. Educational researchers interested in identifying general trends.

3. Qualitative metrics for evaluation. - Insight Depth: Measures how effectively users extract new insights from data. - Interface larity: How easily users understand the visualization.

4. Quantitative metrics: - Accuracy: how accurately users interpret data. - Task Completion Time: how quickly users find answers.

5. Evaluation approach: - Journaling study: Participants use a data analysis tool for several days and record insights, problems, and questions.

6. Success Criteria: Most users understand how to work with the interface without significant training. Users note that visualization helps find new insights. And the average time to complete 80% of tasks does not exceed 2 minutes.

7. Summary: Understand how useful visualization is for users. Identify weak points in design and interface and obtain data for improvement and justification of the project's success.

[13]:
```
import altair as alt
import pandas as pd
import csv
import warnings
import json
import numpy as np

warnings.filterwarnings("ignore", category=FutureWarning)

alt.renderers.enable('svg') # uncomment for export
alt.data_transformers.disable_max_rows()

with open('/Users/ilyapetrusenko/Downloads/School Districts.geojson', 'r') as f:
    geojson_data = json.load(f)

data = pd.read_csv('nyc_math_test_results.csv', quoting=csv.QUOTE_NONE,␣
 ↪delimiter=';')
data = data[data['Mean Scale Score'] != 's']
data[['District', 'Borough', 'Number']] = data['DBN'].str.
 ↪extract(r'(\d{2})([A-Z])(\d{3})')
data['District'] = data['District'].astype(int)
data['Borough'] = data['Borough'].map({
    'M': 'Manhattan',
    'K': 'Brooklyn',
    'X': 'Bronx',
    'Q': 'Queens',
    'R': 'Staten Island'
})
data['# Level 1'] = data['# Level 1'].astype(int)
data['# Level 2'] = data['# Level 2'].astype(int)
data['# Level 3'] = data['# Level 3'].astype(int)
data['# Level 4'] = data['# Level 4'].astype(int)
data['# Level 3+4'] = data['# Level 3+4'].astype(int)
data['% Level 1'] = data['% Level 1'].str.replace(',', '.').astype(float)
data['% Level 2'] = data['% Level 2'].str.replace(',', '.').astype(float)
data['% Level 3'] = data['% Level 3'].str.replace(',', '.').astype(float)
```

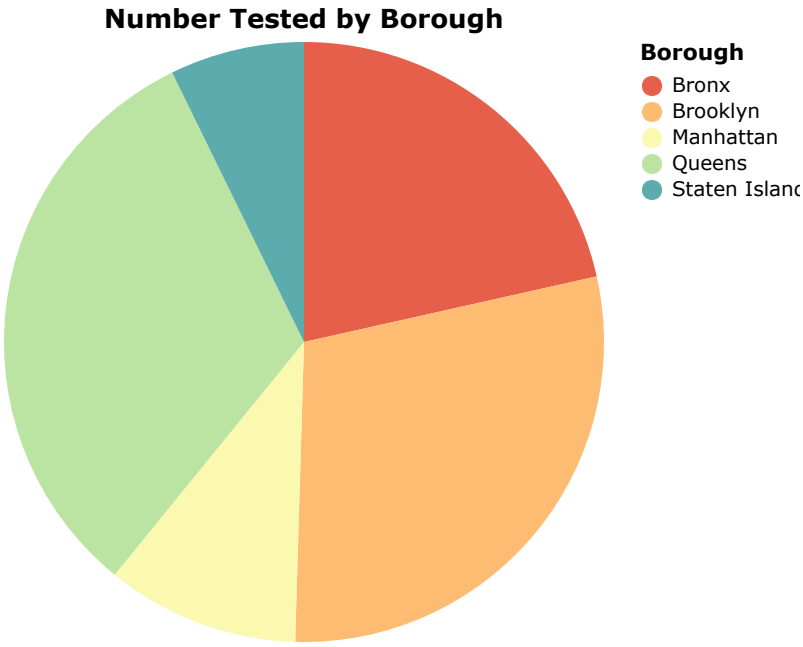
```
data['% Level 4'] = data['% Level 4'].str.replace(',', '.').astype(float)
data['% Level 3+4'] = data['% Level 3+4'].str.replace(',', '.').astype(float)

data.head()
```

[13]:

	DBN	School Name	Grade	Year	Category	Number Tested	Mean Scale Score	# Level 1	% Level 1	# Level 2	% Level 2	# Level 3	% Level 3	# Level 4	% Level 4	# Level 3+4	% Level 3+4	District	Borough	Number
1	01M015	P.S. 015 ROBERTO CLEMENTE	3	2023	Black	6	434	1	16.7	4	66.7	1	16.7	0	0.0	1	16.7	1	Manhattan	015
2	01M015	P.S. 015 ROBERTO CLEMENTE	3	2023	Hispanic	14	437	5	35.7	5	35.7	3	21.4	1	7.1	4	28.6	1	Manhattan	015
7	01M015	P.S. 015 ROBERTO CLEMENTE	4	2023	Hispanic	14	438	5	35.7	3	21.4	6	42.9	0	0.0	6	42.9	1	Manhattan	015
10	01M015	P.S. 015 ROBERTO CLEMENTE	5	2023	Black	13	431	6	46.2	5	38.5	2	15.4	0	0.0	2	15.4	1	Manhattan	015
11	01M015	P.S. 015 ROBERTO CLEMENTE	5	2023	Hispanic	13	431	7	53.8	4	30.8	2	15.4	0	0.0	2	15.4	1	Manhattan	015

[14]:
```
alt.Chart(data[data['Grade']=='All Grades'].groupby('Borough',␣
 ↪as_index=False)['Number Tested'].sum()
).mark_arc().encode(
    theta=alt.Theta('Number Tested:Q', title='Number Tested'),
    color=alt.Color('Borough:N', title='Borough', scale=alt.
 ↪Scale(scheme='spectral')),
    tooltip=['Borough', 'Number Tested']
).properties(
    title='Number Tested by Borough'
)
```

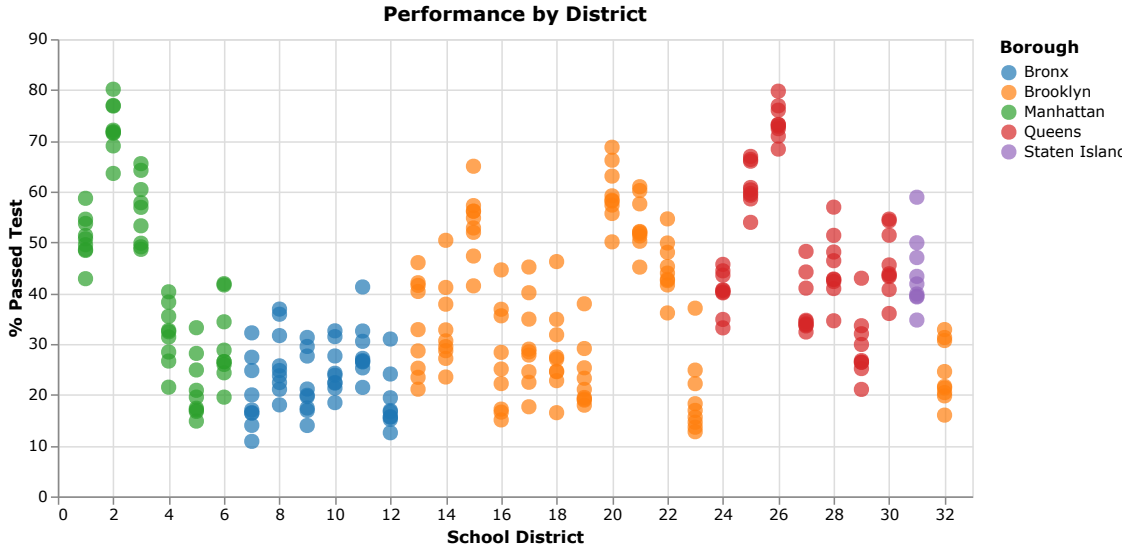
[14]:



The pie chart shows the distribution of test-takers among the various boroughs of New York City. The largest number of students take the tests are in Brooklyn and Queens, consistent with their high demographic density. These boroughs provide the bulk of the data, which can distort interpretation if the analysis does not take into account the proportionality of the data. The chart illustrates the ratio of test-takers among the boroughs, highlighting the disparities in the distribution.

[15]:
```
grouped_data = data[data['Grade']=='All Grades'].groupby(['Borough',␣
↪'District','Year'], as_index=False).agg(
    Tested=('Number Tested', 'sum'),
    Passed=('# Level 3+4', 'sum')
)
grouped_data['Passed'] = grouped_data['Passed']/grouped_data['Tested']*100

alt.Chart(grouped_data).mark_circle(size=100).encode(
   x=alt.X('District:Q', title='School District', scale=alt.Scale(domain=[0,␣
↪33])),
   y=alt.Y('Passed:Q', title='% Passed Test'),
   color=alt.Color('Borough:N', scale=alt.Scale(scheme='category10'))
).properties(
   title='Performance by District',
   width=600,
   height=300
)
```

[15]:

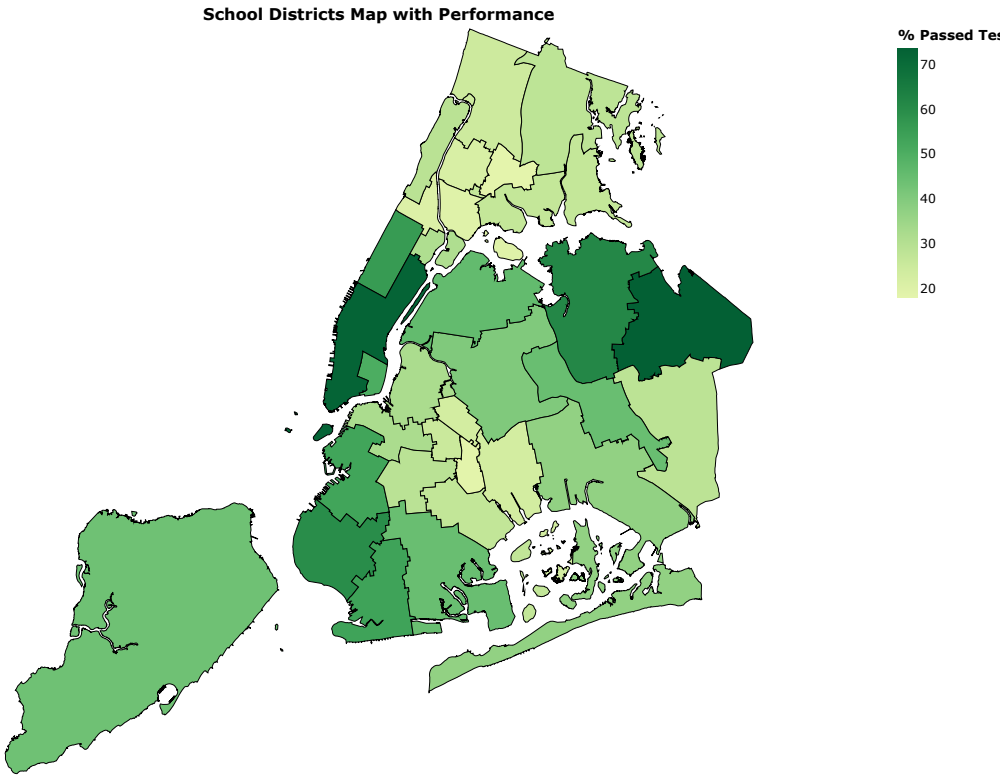


The graph shows the range of performance across school districts. The greater the range, the less consistent the district's performance is from year to year. Districts in different boroughs are colored differently.

[16]:
```
grouped_data = data[data['Grade']=='All Grades'].groupby(['Borough',␣
 ↪'District'], as_index=False).agg(
    Tested=('Number Tested', 'sum'),
    Passed=('# Level 3+4', 'sum')
)
grouped_data['Passed'] = grouped_data['Passed']/grouped_data['Tested']*100
grouped_data['school_dist'] = grouped_data['District']

borough_selector = alt.selection_point(
    fields=['Borough'],
    bind=alt.binding_select(options=['All', 'Bronx', 'Brooklyn', 'Manhattan',␣
 ↪'Queens',
                                     'Staten Island'], name='Borough'),
    value='All',
    name = "borough_selector",
)

alt.Chart(alt.Data(values=geojson_data['features'])).mark_geoshape(
    stroke='black',
    strokeWidth=0.5
).encode(
    color=alt.condition(
        'borough_selector.Borough == "All" || datum.Borough == borough_selector.
 ↪Borough',
```

```
        alt.Color('Passed:Q', title='% Passed Test', scale=alt.
↪Scale(scheme='yellowgreen')),
        alt.value('lightgray')
    ),
    tooltip=[
        alt.Tooltip('Borough:N', title='Borough'),
        alt.Tooltip('properties.school_dist:N', title='District'),
        alt.Tooltip('Passed:Q', title='% Passed Test', format=".2f")
    ]
).transform_lookup(
    lookup='properties.school_dist',
    from_=alt.LookupData(grouped_data, key='school_dist',␣
↪fields=['Borough','Passed'])
).add_params(
    borough_selector
).properties(
    width=800,
    height=600,
    title='School Districts Map with Performance'
).project('mercator')
```

[16]:

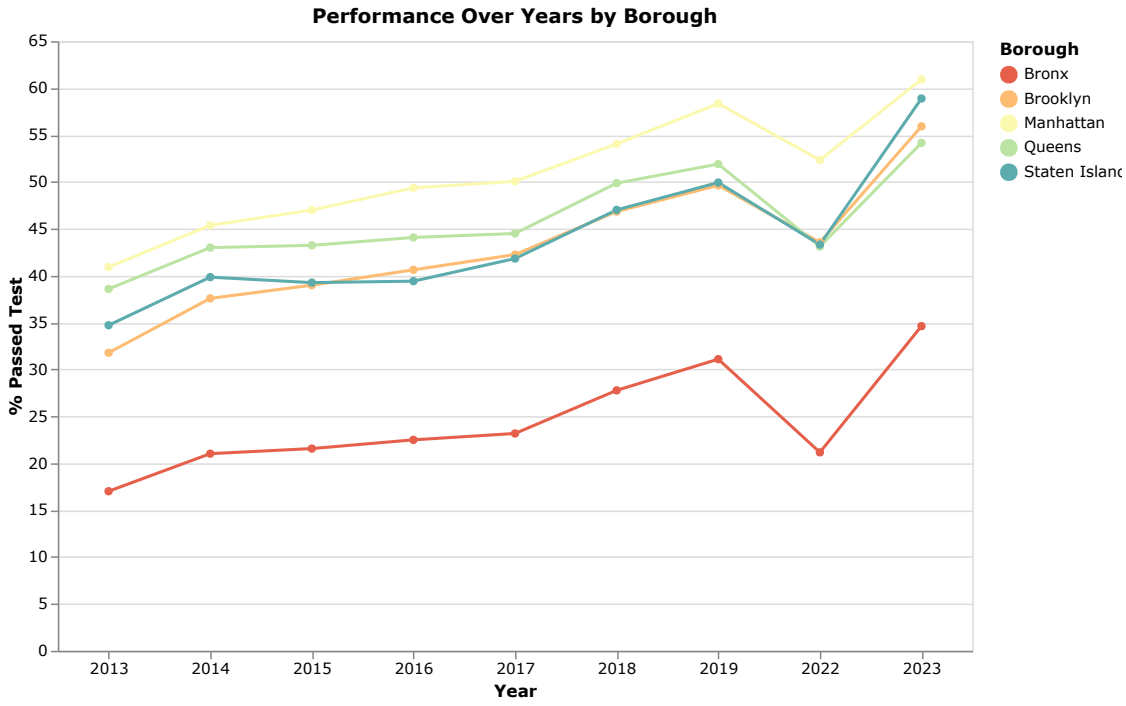


The map visually shows the differences in performance between New York City school districts. Lighter areas indicate low pass rates, darker areas indicate high pass rates. A specific borough can

be selected using the selector. From the graph, we can see high pass rates in Manhattan schools, as well as certain districts in Brooklyn and Queens. However, Brooklyn also has districts with extremely low pass rate, as does the Bronx, which has the lowest pass rate overall. The map visually shows the geographic distribution of performance, highlighting areas where educational interventions are needed.

[17]:
```
grouped_data = data.groupby(['Borough','Year','Grade'], as_index=False).agg(
    Tested=('Number Tested', 'sum'),
    Passed=('# Level 3+4', 'sum')
)
grouped_data['Passed'] = grouped_data['Passed']/grouped_data['Tested']*100

grade_selector = alt.selection_point(
    fields=['Grade'],
    bind=alt.binding_select(options=grouped_data['Grade'].unique()),
    value='All Grades',
    name='Select'
)

alt.Chart(grouped_data).mark_line(point=True).encode(
    x=alt.X('Year:O', title='Year', axis=alt.Axis(labelAngle=0)),
    y=alt.Y('Passed:Q', title='% Passed Test'),
    color=alt.Color('Borough:N', title='Borough', scale=alt.
  ↪Scale(scheme='spectral')),
    tooltip=['Year', 'Borough', 'Passed', 'Grade']
).add_params(
    grade_selector
).transform_filter(
    grade_selector
).properties(
    title='Performance Over Years by Borough',
    width=600,
    height=400
)
```

[17]:



This graph shows year-over-year performance trends across NYC boroughs from 2013 to 2023. The overall trend is very similar across all boroughs, with gradual improvements in test scores year over year and sharp drops in performance during the COVID years (2020-2022), some of which are missing from the data, likely due to remote learning and no testing.

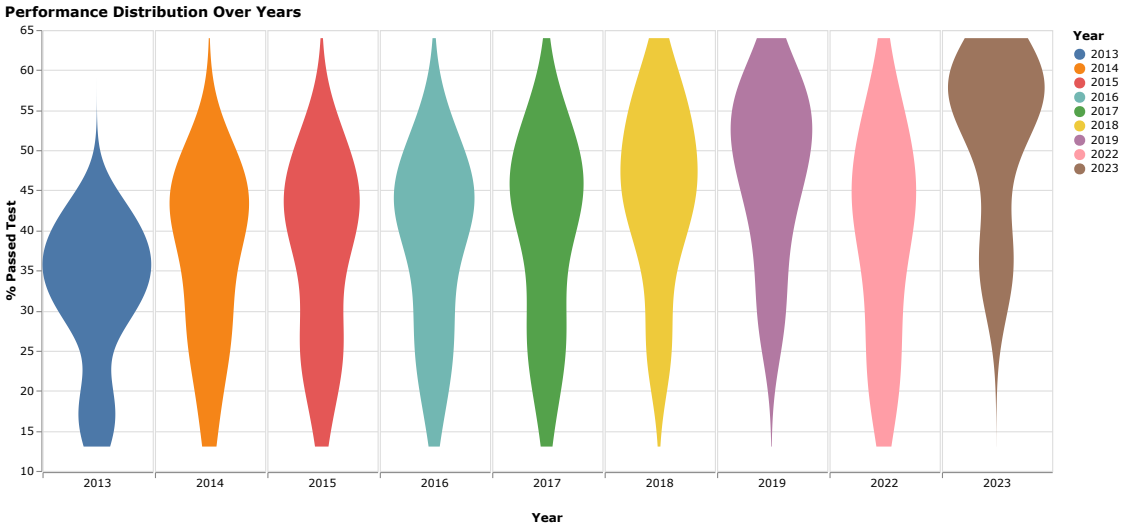
The Bronx, while showing similar yearly trends to other boroughs, is significantly worse in academic performance, as we've already seen on the map. And Brooklyn, which also has districts with education problems, averaged its scores due to its high-performing districts.

Using the selector, you can view graphs for individual grades.

[18]:
```
violin_data = grouped_data[grouped_data['Grade']!='All Grades'].copy()

alt.Chart(violin_data, width=100).transform_density(
    'Passed',
    as_=['Passed', 'density'],
    groupby=['Year']
).mark_area(orient='horizontal').encode(
    alt.X('density:Q')
        .stack('center')
        .impute(None)
        .title(None)
        .axis(labels=False, values=[0], grid=False, ticks=True),
    alt.Y('Passed:Q', title='% Passed Test'),
    alt.Color('Year:N'),
    alt.Column('Year:N')
```

```
        .spacing(0)
        .header(titleOrient='bottom', labelOrient='bottom', labelPadding=0)
).properties(
    title='Performance Distribution Over Years',
    height=400
)
```

[18]:



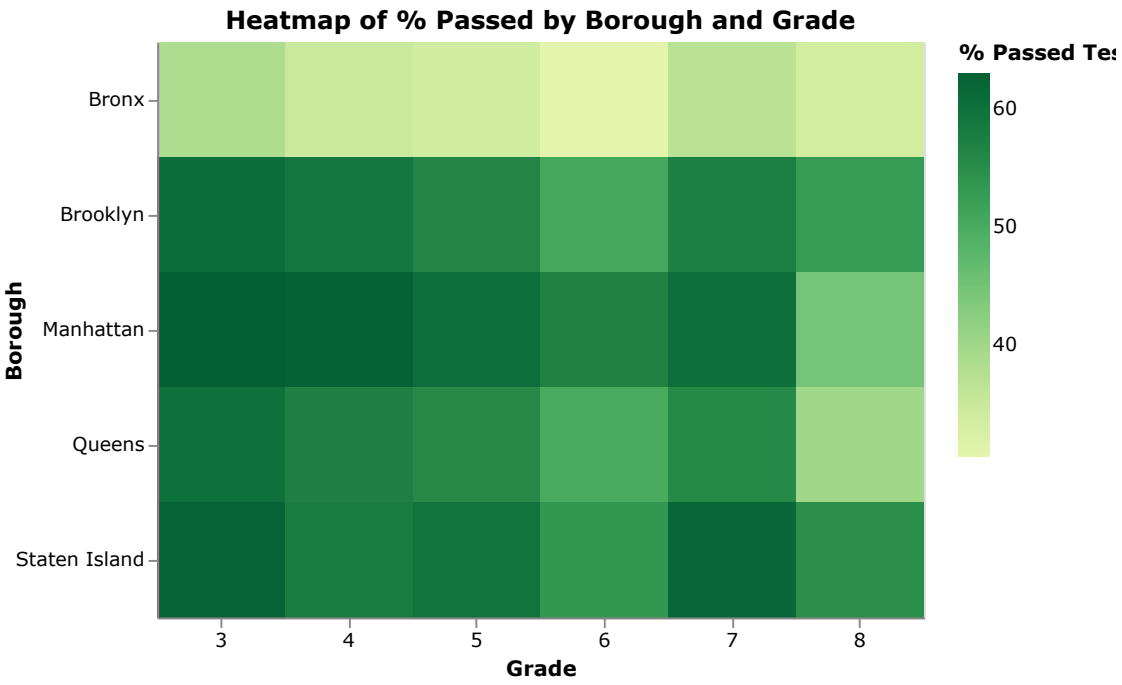
Violin plot illustrates the distribution of student performance over the years. This graph is similar to a histogram, with the width of the violin plot showing the frequency of values in different performance ranges. Wide plots mean that many students are in a given performance range, while narrow plots mean that values are less frequent.

[19]:

```
year_selector = alt.selection_point(
    fields=['Year'],
    bind=alt.binding_select(options=data['Year'].unique()),
    value=2023,
    name='Select'
)

alt.Chart(grouped_data).mark_rect().transform_filter(
    alt.datum.Grade != 'All Grades'
).encode(
    x=alt.X('Grade:O', title='Grade', axis=alt.Axis(labelAngle=0)),
    y=alt.Y('Borough:N', title='Borough'),
    color=alt.Color('Passed:Q', title='% Passed Test', scale=alt.
 ↪Scale(scheme='yellowgreen')),
    tooltip=['Year', 'Borough', 'Grade', 'Passed']
).add_params(
    year_selector
).transform_filter(
```

```
    year_selector
).properties(
    title='Heatmap of % Passed by Borough and Grade',
    width=400,
    height=300
)
```

[19]:



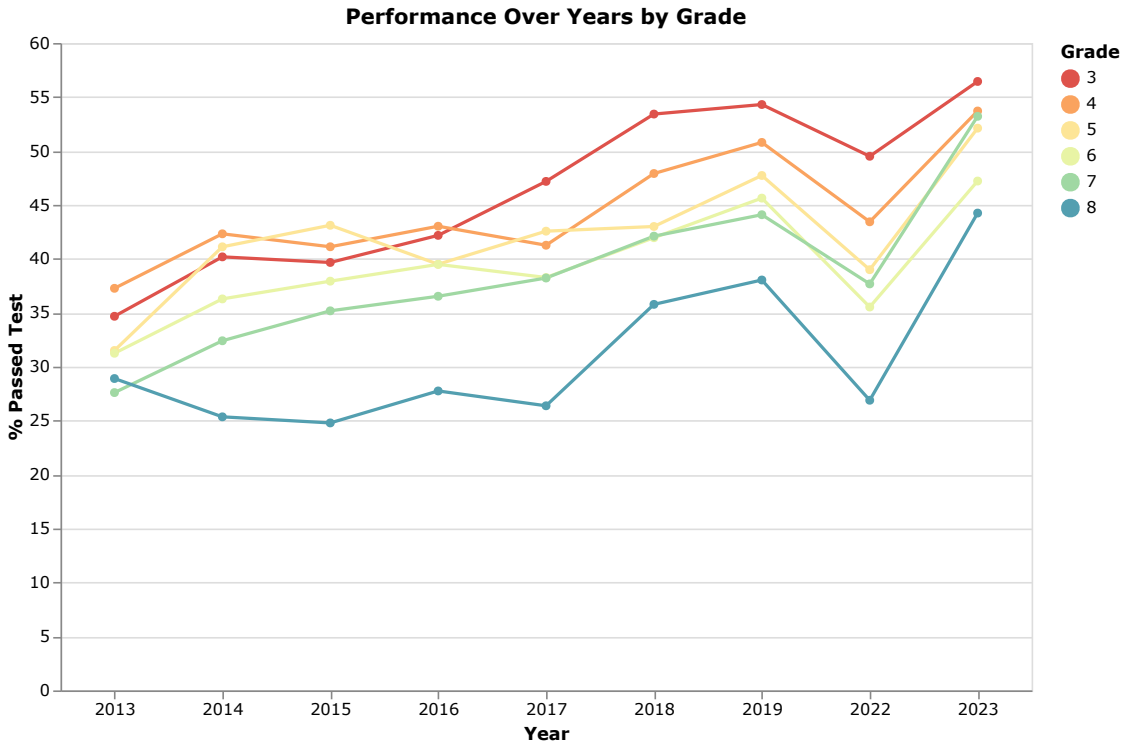
The heat map shows the pass rates for different grades in each borough. Lighter areas indicate low pass rates, darker areas indicate high pass rates. This graph also shows the worst pass rates in the Bronx, regardless of grade. We can also see that in all boroughs, the pass rates in the higher grades (6-8) are lower than in the lower grades (3-5). Brooklyn and Staten Island show more uniform results, even in the higher grades, but we must not forget about the variability of the districts in Brooklyn.

Using the selector, you can view the graph for different years.

[20]:
```
grouped_data = data.groupby(['Year','Grade'], as_index=False).agg(
    Tested=('Number Tested', 'sum'),
    Passed=('# Level 3+4', 'sum')
)
grouped_data['Passed'] = grouped_data['Passed']/grouped_data['Tested']*100

alt.Chart(grouped_data).mark_line(point=True).transform_filter(
    alt.datum.Grade != 'All Grades'
).encode(
    x=alt.X('Year:O', title='Year', axis=alt.Axis(labelAngle=0)),
```

```
    y=alt.Y('Passed:Q', title='% Passed Test'),
    color=alt.Color('Grade:N', title='Grade', scale=alt.
 ↪Scale(scheme='spectral')),
    tooltip=['Year', 'Grade', 'Passed']
).properties(
    title='Performance Over Years by Grade',
    width=600,
    height=400
)
```

[20]:



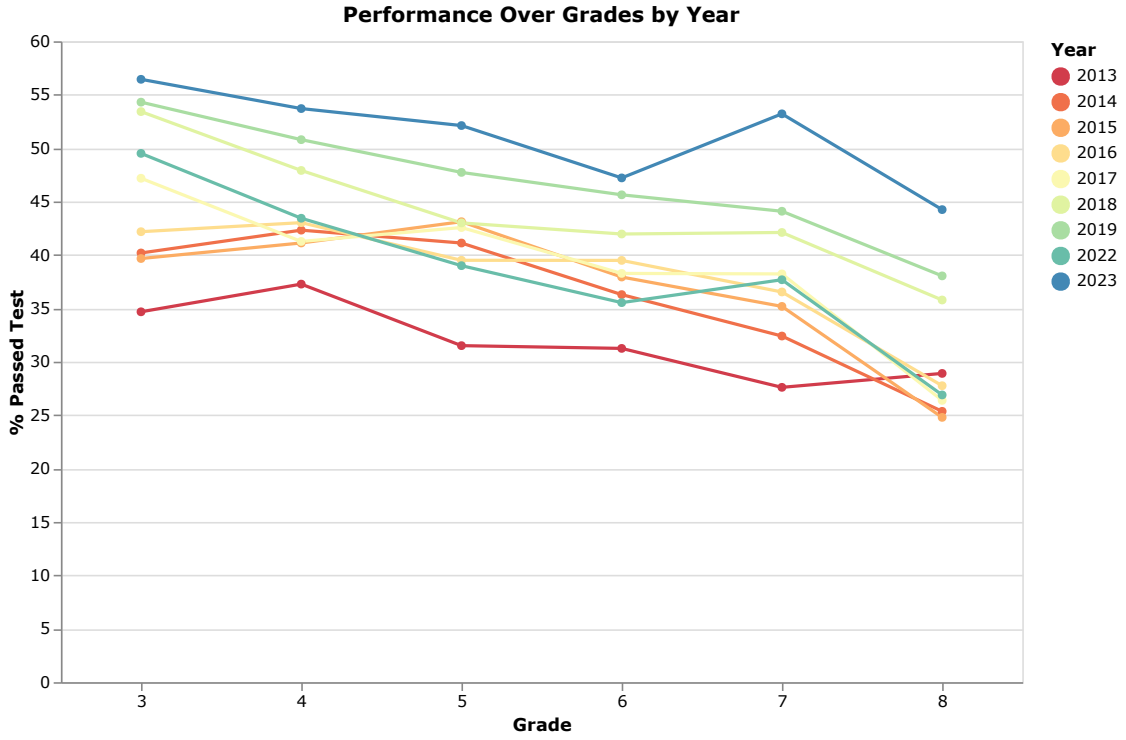
The lines on this graph show how the results have changed over the years for each grade. The graph shows that younger grades generally have better academic performance, regardless of the year. They are also less susceptible to the negative impact of external factors, such as distance learning during the Covid years.

[21]:

```
alt.Chart(grouped_data).mark_line(point=True).transform_filter(
    alt.datum.Grade != 'All Grades'
).encode(
    x=alt.X('Grade:O', title='Grade', axis=alt.Axis(labelAngle=0)),
    y=alt.Y('Passed:Q', title='% Passed Test'),
    color=alt.Color('Year:N', title='Year', scale=alt.Scale(scheme='spectral')),
    tooltip=['Year', 'Grade', 'Passed']
).properties(
```

```
    title='Performance Over Grades by Year',
    width=600,
    height=400
)
```

[21]:



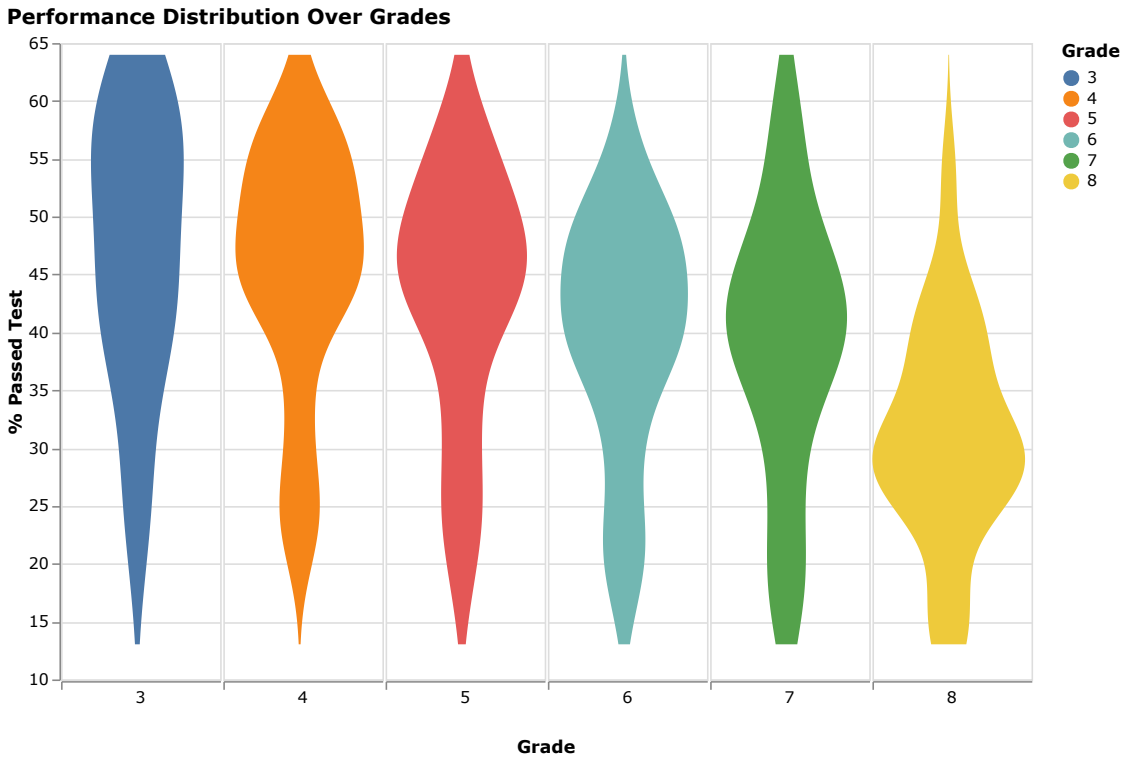
This graph is similar to the previous one, but looks at the same data from a different angle. Its lines show how the results changed by grade for each year. The graph shows that although academic performance improves year over year, but for each year higher grades have worse test scores than younger grades.

[22]:

```
alt.Chart(violin_data, width=100).transform_density(
    'Passed',
    as_=['Passed', 'density'],
    groupby=['Grade']
).mark_area(orient='horizontal').encode(
    alt.X('density:Q')
        .stack('center')
        .impute(None)
        .title(None)
        .axis(labels=False, values=[0], grid=False, ticks=True),
    alt.Y('Passed:Q', title='% Passed Test'),
    alt.Color('Grade:N'),
    alt.Column('Grade:N')
        .spacing(0)
```

```
        .header(titleOrient='bottom', labelOrient='bottom', labelPadding=0)
).properties(
    title='Performance Distribution Over Grades',
    height=400
)
```

[22]:



Violin Plot illustrates the distribution of student performance across grades. Like the previous Violin Plot, this graph is similar to a histogram, with the width of the violin plot showing the frequency of values in different performance ranges. Wide plots mean that many students are in a given performance range, while narrow plots mean that values are less frequent.

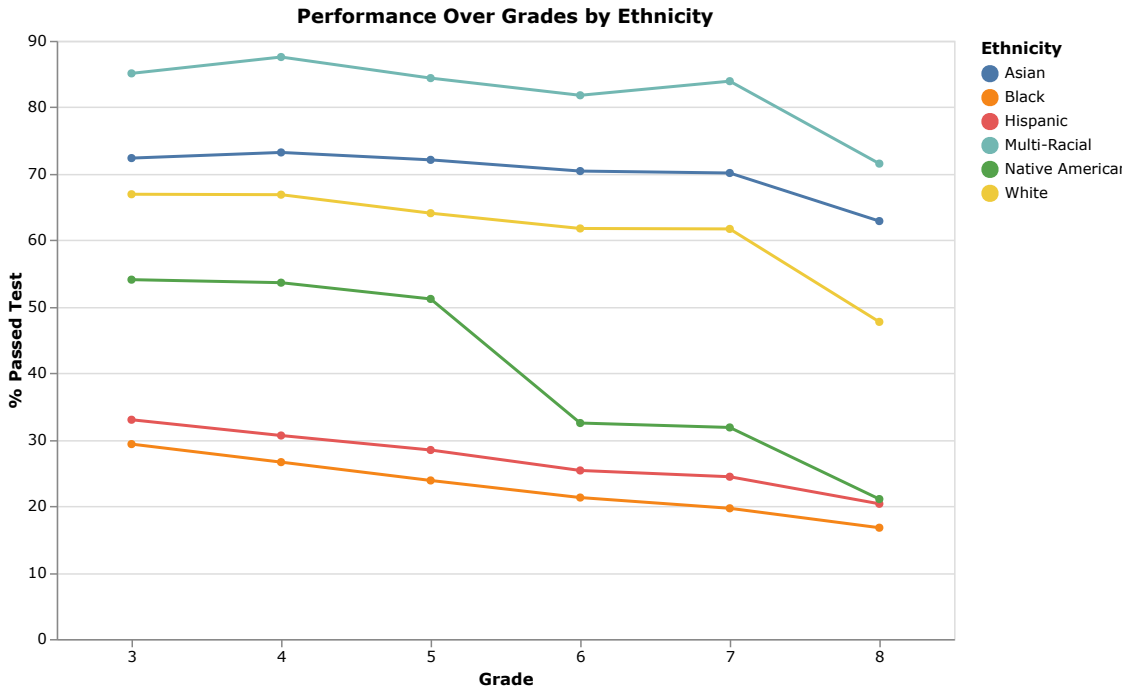
On the graph we can see not only a decline in academic performance from grade to grade, but also a change in the distribution of test scores, also in the direction of worsening academic performance.

[23]:

```
grouped_data = data[data['Grade']!='All Grades'].groupby(['Grade','Category'],␣
 ↪as_index=False).agg(
    Tested=('Number Tested', 'sum'),
    Passed=('# Level 3+4', 'sum')
)
grouped_data['Passed'] = grouped_data['Passed']/grouped_data['Tested']*100

alt.Chart(grouped_data).mark_line(point=True).encode(
   x=alt.X('Grade:O', title='Grade', axis=alt.Axis(labelAngle=0)),
   y=alt.Y('Passed:Q', title='% Passed Test'),
```

```
    color=alt.Color('Category:N', title='Ethnicity'),
    tooltip=['Grade', 'Category', 'Passed']
).properties(
    title='Performance Over Grades by Ethnicity',
    width=600,
    height=400
)
```

[23]:



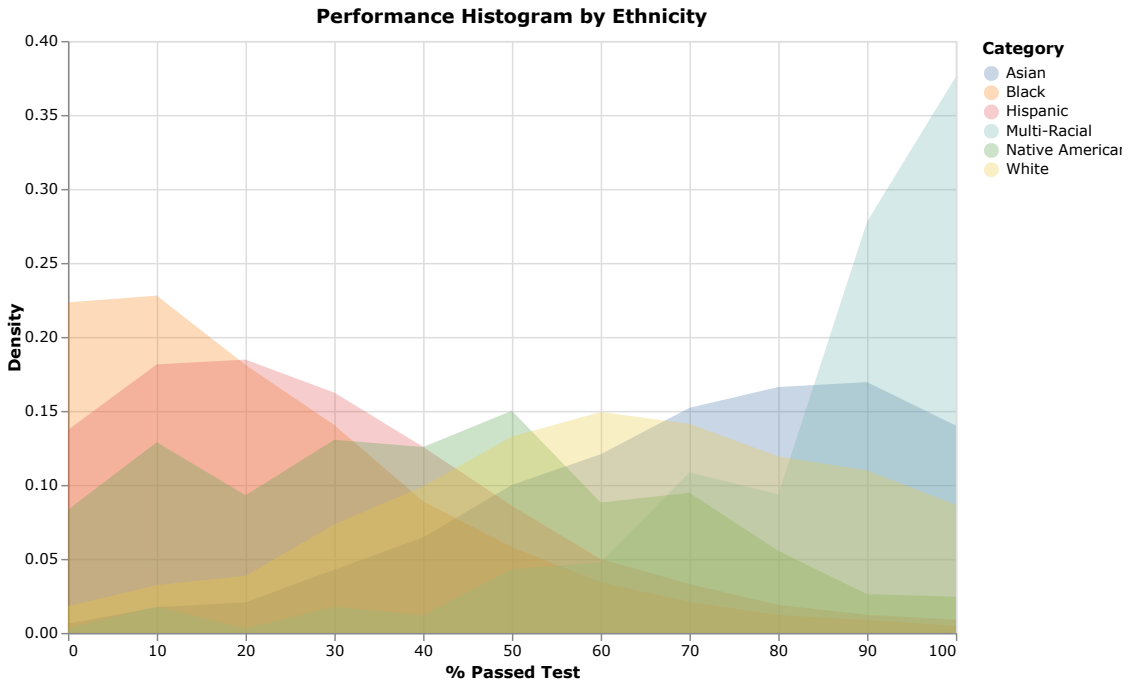
The graph shows the changes in academic performance among different ethnic groups depending on grades. All ethnic groups show a similar trend of declining academic performance towards high school, but we can also see significant differences in academic performance among different ethnic groups. Multi-racial and Asian students show significantly better academic performance than Hispanics and Blacks, although the latter do not show such a significant drop in academic performance at grade 8.

[24]:

```
filtered_data = data[data['Grade'] != 'All Grades'].copy()
filtered_data['Bins'] = pd.cut(filtered_data['% Level 3+4'], bins=11,␣
 ↪include_lowest=True)
binned_data = filtered_data.groupby(['Bins', 'Category']).size().
 ↪reset_index(name='Count')
binned_data['Total'] = binned_data.groupby('Category')['Count'].transform('sum')
binned_data['Density'] = binned_data['Count'] / binned_data['Total']
binned_data['Bins'] = binned_data['Bins'].map({edge: idx*10 for idx, edge in␣
 ↪enumerate(binned_data['Bins'].unique())})
```

```
category_selector = alt.selection_point(
    fields=['Category'],
    bind=alt.binding_select(options=['All', 'Asian', 'Black', 'Hispanic',␣
 ↪'Multi-Racial',
                                     'Native American', 'White'],␣
 ↪name='Category'),
    value='All',
    name = "category_selector",
)

alt.Chart(binned_data).mark_area(opacity=0.3).encode(
    x=alt.X("Bins:Q", title="% Passed Test", scale=alt.Scale(domain=(0, 100)),
            axis=alt.Axis(values=list(range(0, 101, 10)))),
    y=alt.Y("Density:Q").stack(None),
    color=alt.condition(
        'category_selector.Category == "All" || datum.Category ==␣
 ↪category_selector.Category',
        alt.Color('Category:N', legend=alt.Legend(title='Category')),
        alt.value('lightgray')
    )
).add_params(
    category_selector
).properties(
    title="Performance Histogram by Ethnicity",
    width=600,
    height=400
)
```

[24]:



The smoothed density histogram illustrates the distribution of academic performance across ethnic categories. The semi-transparent color regions are superimposed on each other, allowing comparison of the density distributions. The graph illustrates how academic performance is distributed among students of each ethnic group. The Multi-Racial and Asian groups are concentrated closer to the high scores (70–100%), while the Black and Hispanic groups show a significant number of students in the 20–50% range, indicating difficulties in academic performance.

Conclusion

Data analysis based on the math tests results among school students of New York in 2013-2023 showed significant and constant differences in academic performance between boroughs, ethnic groups and classes. Bronx demonstrates consistently the lowest degrees compared to other boroughs, especially in high school (grades 6-8). The academic performance here falls behind even in better years, and during the COVID-19 pandemic the situation got even worse, which reveals the presence of big structural problems in the education system of this borough. Manhattan and Queens, by contrast, show good results, including the crisis years. Brooklyn occupies the middle position and demonstrates more significant dispersion reflecting its variety in demographic composition and educational conditions. Despite the differences, all boroughs similarly tend to improve their academic performance through the years.

The COVID pandemic obviously had great influence on academic performance. During the years of the epidemic there were no tests arranged, as there is no data for this period. However, in the first following year the dramatic decrease of average results and bigger diversity in academic performance distribution are observed in all boroughs, regardless of ethnicity and grade. This can be related to the introduction of remote education and its uneven availability. Nevertheless, starting from 2023 the results are restored, which is seen clearly in Staten Island and Manhattan, possibly being the consequence of faster returning to offline education or the result of more effective adaptation of the educational system in these boroughs. In the Bronx and Brooklyn this restoration is slower, which shows that additional work is required there.

The demographic analysis revealed permanent differences in results of various ethnic groups. Multi-Racial and Asian students consistently demonstrate the highest results, moreover, their performance remains relatively stable even in the last grades. This makes significant contrast with the results of Black and Hispanic groups, where the results are much lower, and the drop of the performance in the last grades is the most obvious. The distribution of the academic performance among Black and Hispanic students is also more diverse, which indicates a significant number of students with low results in these groups. White students are in the middle position and demonstrate average results with less dramatic fluctuations in comparison with Black and Hispanic groups. These differences highlight the influence of both social and economic factors, as well as cultural peculiarities on the educational results.

Trends in class levels also indicate important patterns. Academic performance in all ethnic groups decreases from first to last grades, but this pattern is more dramatic among Black and Hispanic students, which can be related to the increase of academic loading and insufficient support during the studying process. Asian students keep high results in all grades, which indicates the presence of such factors ensuring their success as family support, cultural values or quality of educational programs. These differences require additional studies to understand the mechanisms boosting the

results of successful groups and the barriers suppressing the progress of more vulnerable categories. The general analysis proves the existence of significant educational inequality in New York. Bronx, Black and Hispanic groups, as well as last grades require special attention, since they demonstrate the lowest results. The found differences indicate the necessity of target educational intervention directed at supporting vulnerable groups and elimination of educational gaps between boroughs. Furthermore, restoration of academic performance after the pandemic shows that timely measures, such as improvement of educational resources availability and school programs adaptation can boost the results significantly. These data can be used as a base for development of strategies directed to elimination of educational barriers and creation of equal opportunities for all students of New York.

Final description and justification of design

Design of visualization was created for data analysis of math tests results in New York, including geographical, temporal and demographic aspects. The key elements include:

- **Geographical maps**: thermal maps using color scale made it possible to visualize differences in academic performance between boroughs. This made it easier to understand spatial distribution of data, particularly to reveal weaker areas, such as the Bronx.
- **Linear graphs**: they were used to reflect temporal trends of academic performance according to years and grades. It helped to highlight long-term trends, including the decrease of performance during the pandemic and its restoration afterwards.
- **Violin Plot and histograms**: these graphs illustrated distribution of academic performance among ethnic groups and years, which allowed to explore diversity and find groups with higher and lower results.
- **Interactivity**: filtration elements (for instance, the possibility to select an ethnic group or a borough) improved the convenience of analysis and allowed users to focus on certain aspects of their interest.

Justification of the design was based on the principles of efficiency and intuitivity, so that every visualization solved a certain analytical problem. Thermal maps were chosen to reveal geographical differences, linear graphs - for trends and histograms - for distribution analysis.

Final evaluation

To evaluate the final design I used Journaling Study, for which I engaged three colleagues. All participants work in the IT sphere, have different experience in data analysis and therefore they gave complex feedback.

Positive aspects, highlighted by the participants

- **Clarity of trends and differences**: Thermal maps were highly evaluated for their illustrative input. The participants quickly determined the borough with the lowest academic performance (Bronx) and stated that the color scale allowed to understand the difference between the areas intuitively. Linear graphs were found convenient to track temporal alterations. The participants could notice the decrease of academic performance in 2020 and observe its restoration after the pandemic.
- **Interactivity**: The possibility to sort the data according to the year borough or ethnic group improved the analysis experience significantly. One of the participants mentioned that

the selectors helped to focus on certain aspects without being distracted by additional data. Interactive cues with details about the data (for instance, the exact percentage value for students who passed their tests) allowed deeper analysis.

- **Histograms and Violin Plot**: Although interpretation of Violin Plot was questioned, the participants mentioned that it demonstrates data distribution effectively. For instance, they stated that in 2020 the range of academic performance was the widest, which indicates significant differences within groups.

Difficulties and concerns

- **Filtration interface**: In spite of their benefits, the selections didn't appear to be intuitive. The participants stated that at first they didn't notice the possibility to select the parameters (like the year or ethnic group) from the drop-down menu. This required additional explanation. One of the participants suggested adding visual hints, such as marked buttons or interactive indicators, to make the usage easier.
- **Violin Plot**: Less experienced participants found it difficult to interpret the graph. They doubted that the wide audience will be able to understand quickly that the width of the "violin" reflects the data density. Pop-up cues or introductory text could eliminate this obstacle.
- **Perception difficulties in the district map**: One of the participants mentioned that the range of values on the thermal map was not completely obvious without the legend. This was particularly true for the participants who are not used to working with color codes.

Interesting observations

- All participants stated that the linear graphs for the year trends were the most understandable intuitively. They could see straight away how the pandemic affected academic performance and which boroughs restored it quicker.
- The participants with experience in programming suggested using more interactive dashboards, which allow to compare temporal, geographical and demographic aspects at the same time.

Conclusions of the evaluation

The evaluation proved that visualizations solve the main problem - the demonstration of key differences and trends in data. However, there are a few aspects, which require refinement. Interactivity and clarity of trends were highly evaluated, but some elements like Violin Plot and filtration interface require upgrading to improve intuitive comprehensiveness and accessibility. Adding supporting text explanations and visual hints could improve the usage experience. In general, the project has demonstrated successful implementation of visualization to analyze complex educational data, but it also revealed some areas for further development.

Thank you for reviewing!