Chi-square Test of the Educational Opportunity Gap in Washington State Schools

Research Question

Statewide, Washington State Public School students are a majority White, with 61.7% of learners identifying as White. The next three most prominent racial groups in public schools are Hispanic or Latinx (21.1%), Black or African American (2.8%) and Asian (2.4%) (OSPI, 2019). Given the various observed discrepancies in median family wealth (Bhutta, Neil, Chang, Dettling, & Hsu, 2020), incarceration rates (Gramlich, 2020), and college attendance (NCES, 2019) between these races later in life, much focus has been placed on the way these groups thrive earlier on in public school systems.

In 2008, in light of a growing body of data demonstrating that certain racial groups were not performing equally well on standardized tests, the Washington State legislature mandated the formation of a committee to study the observed achievement gap in the Washington student population across racial groups. As a result, the Achievement Gap Oversight and Accountability Committee was formed and later renamed to the Educational Opportunity Gap Oversight and Accountability Committee or EOGOAC (OSPI, 2009). Their findings and subsequent policy recommendations have aimed to make the state's school system more equitable, with a key metric being student performance on standardized tests. The rate at which students pass is a critical measurement of overall societal opportunity as well, as the outcomes from these tests can dramatically affect students' future academic and career prospects (SBE, 2019). The EOGOAC's own findings in their annual publications demonstrate that the observed gap along racial lines has not significantly diminished since the committee's formation in 2008 (Flores & Rees, 2020). These findings correspond to findings from the Organization for Economic Cooperation and Development (OECD) in their test of 15-year-olds worldwide. Their 2018 Program for International Student Assessment (PISA) test results indicate that there is a high level of achievement disparity among various racial groups in the United States (PISA 2018).

The aim of this study is to search for statistical significance in the so-called "Opportunity Gap" from Washington State's most recent available testing data. This data from the 2018-2019 school year allows insight into the performance of various racial groups 10 years after the legislature formed the EOGOAC and enacted ongoing system-wide reforms. Naturally, one hopes to see better achievement rates in light of these statewide efforts to address these observed societal inequalities. The research question of this study is as follows: is there a significant difference between the rates at which High School Students of different races pass high-stakes tests in 2019, given the relatively small sizes of minority populations in public schools?

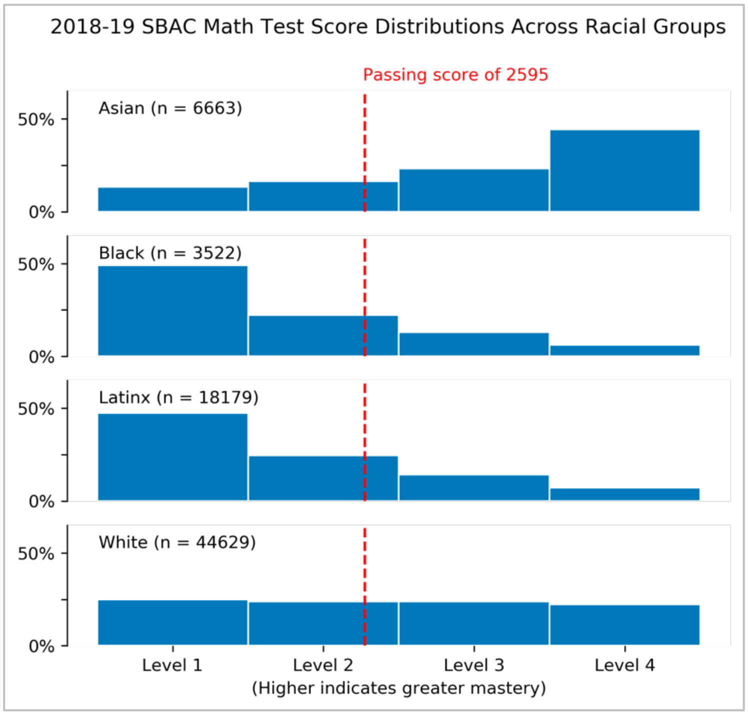

FIGURE 1

This study assumes the null hypothesis

$\underline{H_0}$: The proportion of students passing standardized math tests is the same among the four largest racial groups in Washington State,

with an alternative hypothesis of

$\underline{H_1}$: Test score data show a significant difference in the proportions of passing students among these racial groups.

In order to demonstrate significance from testing data, this study focuses on one major barrier to High School graduation, the Mathematics SBAC (Smarter Balanced Assessment Consortium) test given to all 10th graders in Washington State, and tabulates statewide passing rates across racial groups. A scaled score of 2595 is required for the traditional graduation pathway (SBE, 2019). A Chi-square test of independence may then be conducted in order to conclude if one may reject the null hypothesis or not, and if so, pairwise post-hoc analysis may be used to find which groups pass at significantly different rates.

Furthermore, this study aims to explore the available data to find a statistical basis for some of the various policy recommendations of the EOGOAC, including increasing school funding for hiring more qualified teachers and instruction aids, enacting programs aimed at increasing minority representation among the teacher population, and calling for both highly qualified and experienced teachers in areas with higher concentrations of underperforming learners.

DATA COLLECTION

One result of the 2008 legislature mentioned previously is the call for disaggregated data for the purpose of studying the relative performances of various subgroups in the public school system. According to the Revised Code of Washington, § 28A.300.042:

"All student data-related reports required of the superintendent of public instruction in this title must be disaggregated by at least the following subgroups of students: White, Black, Hispanic, American Indian/Alaskan Native, Asian, Pacific Islander/Hawaiian Native…"

This legislature provides the impetus for why this kind of data is collected and distributed by the Washington State Office of Superintendent of Public Instruction (OSPI). Large volumes of data are published on OSPI's data portal each year and this data is publicly available for analysis. However, because of the Family Educational Rights and Privacy Act (FERPA) laws, there is a limit to what information can be publicly disclosed. FERPA laws require that data records be suppressed in any case where a record could be used to identify a student at the individual level.

The "**Report Card Assessment Data 2018-19**" dataset, available from OSPI's data portal (dataset fully qualified and available for download at https://data.wa.gov/Education/Report-Card-Assessment-Data-2018-19-School-Year/5y3z-mgxd) provides testing counts, scoring level distributions, and passing outcomes disaggregated by race at the school level, the higher up school district level, as well as state totals for each test administered during the 2018-2019 school year. The counts from this dataset allow for the main hypothesis testing for the purposes of this study.

To investigate the statistical basis for school system reforms called for by the EOGOAC, 4 other datasets available from the OSPI's Data portal are obtained and mined:

(1) The "**Report Card Enrollment 2018-19 School Year**" dataset, used to obtain student demographic information at the school district level	More info and download data here: https://data.wa.gov/Education/Report-Card-Enrollment-2018-19-School-Year/u4gd-6wxx
(2) The "**Report Card Per Pupil Finance Data 2018-19**" dataset, used to extract district-wide spending data and exact student enrollment counts (FTE)	Download here: https://www.k12.wa.us/sites/default/files/public/dataadmin/dataportal/Report%20Card%20Per%20Pupil%20Finance%20Data%202018-19%20%281%29.xlsx Linked from OSPI's data portal here: https://www.k12.wa.us/data-reporting/data-portal
(3) The "**Report Card Teacher Qualification Summary**" dataset, used to find what percentage of teachers are employed at various districts who are considered "limited certification status," "inexperienced" and "out-of-field"	More info and download data here: https://data.wa.gov/Education/Report-Card-Teacher-Qualification-Summary/kyqv-48pz
(4) The "**Report Card Teacher Demographics 2018-19 School Year**" dataset, used to demonstrate the racial make-up of teachers as well as the total number of teachers employed within a district	More info and download data here: https://data.wa.gov/Education/Report-Card-Teacher-Demographics-2018-19-School-Ye/7eq8-772m

DATA EXTRACTION AND PREPARATION

In order to address the research question concerning statewide performance among these four racial groups, the "**Report Card Assessment Data 2018-19**" dataset must be queried for state totals for the 10th grade math SBAC disaggregated by race. From this query the 4x2 contingency table is formed showing Asian, Black, Latinx and White on the rows and Total Tested and Total Passed on the columns (see Table 1).

The follow-up investigation studying various interactions between Math SBAC achievement, student demographic distributions, district spending per student FTE, teacher experience and qualifications, and teacher demographic distributions requires a great deal of data querying, reshaping and merging of the datasets described above. The desired data format for such exploration is a tidy dataset that contains a snapshot of each district. The completed dataset features these variables along the columns and the school districts of Washington State along the rows.

The tool chosen to carry out the data extraction, preparation and analysis in this study is the Python language with the Pandas data analysis package. As discussed by Lasser, Python together with the portable Jupyter Notebook format is ideal for communicating and reproducing the findings of an analysis in a transparent way as it is both open-source and widely adopted (Lasser, 2020). According to a recent Kaggle survey of thousands of data scientists and an analysis by Dilmegani (2020) demonstrates that Python is recommended by a large margin as the tool of choice for data science work, which reinforces both its utility for analysis and wide adoption base.

The following Python code and annotations document the process taken to prepare, clean, and merge the 5 datasets. The full code is available in the accompanying Jupyter Notebook: "prepare_data.ipynb"

First, the datasets were imported into multiple Python Pandas DataFrames:

```
In [1]: import pandas as pd
        import numpy as np
        import matplotlib.pyplot as plt
```

```
In [2]: pd.set_option('display.max_columns', 500)
        pd.set_option('display.max_rows', 2000)
```

```
In [3]: files=['Report_Card_Teacher_Qualification_Summary.csv',
        'Report_Card_Assessment_Data_2018-19_School_Year.csv',
        'Report_Card_Enrollment_2018-19_School_Year.csv',
        'Report_Card_Teacher_Demographics_2018-19_School_Year.csv']

        df1=pd.read_csv(files[0], low_memory=False)
        df2=pd.read_csv(files[1], low_memory=False)
        df3=pd.read_csv(files[2], low_memory=False)
        df4=pd.read_csv(files[3], low_memory=False)
        df5=pd.read_excel('Report Card Per Pupil Finance Data 2018-19.xlsx')
```

An inconsistency in spelling and an extra space character were cleaned from the Finance data set:

```
In [4]: df5.DistrictName=df5.DistrictName.str.replace('Seattle School District #1','Seattle School District No. 1')\
                        .str.replace('Green Dot Destiny ','Green Dot Destiny')
```

Certain locales defined as districts in some of the datasets were not included in each of the five datasets. This was either due to slightly different spellings (such as “Seattle School District No. 1” vs “Seattle School District #1”) or locations that may not be widely accepted as a School District, such as various Technical Colleges where certain High School students may attend.

In the process of merging the extracted columns from each of the 5 datasets, careful investigation of which rows were unique to each dataset allowed for renaming certain rows uniformly across all 5 datasets before running the join operations. Only districts that were named across all 5 datasets were kept for the sake of data completeness. “Impact Public Schools” was identified from this code snippet for not being in any of the other 4 datasets.

```
In [5]: districts=df5.DistrictName.value_counts().sort_index()

In [6]: master=[]
        outs4=[]
        for district in list(df4.DistrictName.value_counts().sort_index().index):
            if district not in districts:
                outs4.append(district)
            if district in districts:
                master.append(district)

In [7]: outs3=[]
        for district in list(df3.DistrictName.value_counts().sort_index().index):
            if district not in districts:
                outs3.append(district)
            if district in districts:
                master.append(district)

In [8]: outs2=[]
        for district in list(df3.DistrictName.value_counts().sort_index().index):
            if district not in districts:
                outs2.append(district)
            if district in districts:
                master.append(district)

In [9]: outs1=[]
        for district in list(df3.DistrictName.value_counts().sort_index().index):
            if district not in districts:
                outs1.append(district)
            if district in districts:
                master.append(district)

In [10]: master=pd.Series(master).value_counts()

In [11]: [i for i in list(districts.index) if i not in master.index]
Out[11]: ['Impact Public Schools']
```

Next, the Test Scores data was queried for the desired test, grade level, and student racial group at the District Level. Some small effort shown at `In [14]:` was required for certain columns containing percent values stored as text instead of numeric data. These columns were parsed and recalculated as floating-point values representing the percentages indicated. Pivot tables turned repeated row entries into tidy columns. Columns were renamed according to the snake case convention adopted for this project. Finally, an inconsistent spelling for Seattle School District was corrected.

Test Scores

```
In [12]: sbac=\
df2.query("GradeLevel=='10th Grade' and\
           TestAdministration=='SBAC' and\
           TestSubject=='Math' and\
           StudentGroup in ('All Students',\
                            'Asian','Black/ African American',\
                            'Hispanic/ Latino of any race(s)','White') and\
           SchoolName=='District Total'")
```

```
In [13]: notnull_sbac=sbac.dropna(subset=['CountMetStandard']).copy()
```

```
In [14]: notnull_sbac.PercentMetStandard=notnull_sbac.PercentMetStandard.str.strip('%').astype(float)/100
```

```
In [15]: percent_met=\
notnull_sbac[['DistrictName','StudentGroup','PercentMetStandard']]\
.pivot(index='DistrictName', columns='StudentGroup').droplevel(0,axis=1)
```

```
In [16]: percent_met['students_testing_n']=\
sbac[sbac.StudentGroup=='All Students'].groupby('DistrictName')\
                                       .mean()['Count of Students Expected to Test']
```

```
In [17]: percent_met.columns=['sbac_pass_all',
                      'sbac_pass_asian',
                      'sbac_pass_black',
                      'sbac_pass_latinx',
                      'sbac_pass_white',
                      'students_testing_n']
```

```
In [18]: percent_met=\
percent_met[['students_testing_n',
             'sbac_pass_all',
             'sbac_pass_black',
             'sbac_pass_latinx',
             'sbac_pass_white',
             'sbac_pass_asian']]
```

```
In [19]: percent_met.rename({'Seattle School District #1':'Seattle School District No. 1'},inplace=True)
```

Next, the Teacher Qualifications dataset was prepared for merging. Rows describing overall qualifications from the 2018-19 school year were selected, and the desired qualification categories were selected. These rows were also reshaped into corresponding columns using the pivot method, and the columns were renamed to snake case:

Qualifications

```
In [20]: df1=df1[(df1.ContentAreaName=='All')&(df1.SchoolName.isnull())&(df1.SchoolYear=='2018-19')]
```

```
In [21]: qualified=df1[df1.TeacherQualification\
                               .isin(['Inexperienced status',
                                      'Out-of-Field status',
                                      'Limited Certificated status'])][['DistrictName',
                                                                         'TeacherQualification',
                                                                         'TeacherPercent']]
```

```
In [22]: qualified2=qualified.pivot(index='DistrictName',columns='TeacherQualification').droplevel(0,axis=1)
```

```
In [23]: qualified2.columns=['teacher_percent_inexperienced',
                     'teacher_percent_limited_cert',
                     'teacher_percent_out_of_field']
```

Next, the student enrollment demographic data was prepared for merging. District totals were selected and the racial counts for “Black/African American” (referred to as Black for the remainder of this report), “Hispanic/Latino of any races(s)" (referred to as the generally accepted inclusive term "Latinx" for the remainder of this report, as advised by the APA (2019)), White, and Asian demographic, as well as total counts called "All." Demographic percentages were calculated, and columns were again renamed to snake case:

Enrollment

```
In [24]: demographics=\
df3[(df3.SchoolName=='District Total')&\
    (df3.GradeLevel=='AllGrades')][['DistrictName',
                                    'All Students',
                                    'Black/ African American',
                                    'Hispanic/ Latino of any race(s)',
                                    'White',
                                    'Asian']].set_index('DistrictName')
```

```
In [25]: demographics['percent_black']=demographics['Black/ African American']/demographics['All Students']
demographics['percent_latinx']=demographics['Hispanic/ Latino of any race(s)']/demographics['All Students']
demographics['percent_white']=demographics['White']/demographics['All Students']
demographics['percent_asian']=demographics['Asian']/demographics['All Students']
```

```
In [26]: demographics=demographics.rename({'All Students':'student_count'},axis=1)
```

```
In [27]: demographics=demographics[['student_count',
                            'percent_black',
                            'percent_latinx',
                            'percent_white',
                            'percent_asian']]
```

Next, the Teacher Demographics dataset was queried for the various teacher demographic categories (including "Not Provided"). The data was reshaped using the pivot method and the columns were renamed to snake case. The teacher demographic dataset had some missing data in certain districts for certain demographics. This was interpreted to mean no teachers were employed from that racial demographic, thus 0 was imputed for these percentages.

Teacher Demographics

```
In [28]: teacher_demographics=\
df4[df4.SchoolName.isnull()&\
    df4.DemographicCategory.isin(['All',
                                  'Not Provided',
                                  'Asian',
                                  'White',
                                  'Black/African American',
                                  'Hispanic/Latino of any race(s)'])]\
                           .groupby(['DistrictName','DemographicCategory'])\
                           .sum()[['TeacherCount','TeacherPercent']]\
                           .reset_index()
```

```
In [29]: teacher_demographics2=teacher_demographics.pivot(index='DistrictName',columns='DemographicCategory')
```

```
In [30]: teacher_demographics3=teacher_demographics2.TeacherPercent.loc[:,'Asian':'White']
```

```
In [31]: teacher_demographics3['num_teachers']=teacher_demographics2['TeacherCount','All']
```

```
In [32]: teacher_demographics3.columns=['teacher_percent_asian',
                               'teacher_percent_black',
                               'teacher_percent_latinx',
                               'teacher_percent_unknown_demographic',
                               'teacher_percent_white',
                               'num_teachers']
```

```
In [33]: teacher_demographics3=teacher_demographics3.fillna(0)[['num_teachers',
                                                          'teacher_percent_black',
                                                          'teacher_percent_latinx',
                                                          'teacher_percent_white',
                                                          'teacher_percent_asian',
                                                          'teacher_percent_unknown_demographic',
                                                         ]]
```

Lastly, the funding data was grouped by district and summed across the various categories, including federal and state. The results from this method were carefully compared against numbers from OPSI's data portal found at https://washingtonstatereportcard.ospi.k12.wa.us/ to ensure accuracy of these numbers across districts. The total student FTE was also extracted as a column. Finally, a column called "`spending_per_student`" was calculated and the other columns were renamed in snake case.

Funding

```
In [34]: expenditure=\
         df5.query("SchoolName=='District Total'")\
            .groupby('DistrictName').sum()
```

```
In [35]: expenditure['student_FTE']=df5.query("SchoolName=='District Total'")\
                                        .groupby('DistrictName')\
                                        .mean().EnrollmentTotal
```

```
In [36]: expenditure=expenditure[['student_FTE','Expenditure']].copy()
```

```
In [37]: expenditure['spending_per_student']=expenditure.Expenditure/expenditure.student_FTE
```

```
In [38]: expenditure.columns=['student_FTE','total_spending','spending_per_student']
```

The full list was joined on the expenditure dataset, and the extraneous row noted earlier "Impact Public Schools" was dropped. Some districts that did not appear in the Teacher Demographic and the Teacher Qualifications dataset were dropped in `In [41]`. These districts were locations that may not be widely accepted as a School District, such as various Technical Colleges where some testing students may attend. Nulls in the Assessment dataset were left in place. This is because the Assessment dataset contains missing data due to record suppression that protects the privacy of individual students. Thus, imputation is not helpful or warranted for this dataset, and missing values are left as null values. To drop all rows where these values were null would severely limit the number of districts that could be analyzed and quality analysis across districts can still be carried out with these missing values.

Merging Full List

```
In [39]: df_wa=expenditure.join(teacher_demographics3,how='left')\
                          .join(qualified2,how='left')\
                          .join(demographics,how='left')\
                          .join(percent_met,how='outer')
```

```
In [40]: df_wa.drop('Impact Public Schools',inplace=True)
```

```
In [41]: df_wa.dropna(subset=['teacher_percent_black',
                              'teacher_percent_inexperienced'],inplace=True)
```

```
In [42]: df_wa.to_csv('wa_school_district_data.csv')
```

Lastly, the following are the steps taken to prepare the state totals dataset for the Chi-square analysis aimed at answering the main research question. As before, the Assessment dataset was queried according to the 10th grade math SBAC results, and this time only state totals were kept. The percent level 1-4 variables were kept for visualization purposes (See Figure 1). Percentages were calculated, and integer columns read in by pandas as floats were cast as integer columns. Both final DataFrames were saved as .csv files for later analysis, as shown in steps `In [42]` and `In [48]`.

State Totals Dataset

```
In [43]: state_totals=\
df2.query("GradeLevel=='10th Grade' and\
            TestAdministration=='SBAC' and\
            TestSubject=='Math' and\
            StudentGroup in ('All Students',\
                             'Asian','Black/ African American',\
                             'Hispanic/ Latino of any race(s)','White') and\
            SchoolName=='State Total'")[['StudentGroup',
                                          'Count of students expected to test including previously passed',
                                          'CountMetStandard']\
                                        + ['PercentLevel'+str(i) for i in range(1,5)]]
```

```
In [44]: state_totals.columns=['demographic',
                           'count_total_tested',
                           'count_met_standard',
                           'percent_level_1',
                           'percent_level_2',
                           'percent_level_3',
                           'percent_level_4']
```

```
In [45]: state_totals=state_totals.reset_index(drop=True)
```

```
In [46]: state_totals.percent_level_1=state_totals.percent_level_1.str.strip('%').astype(float)/100
state_totals.percent_level_2=state_totals.percent_level_2.str.strip('%').astype(float)/100
state_totals.percent_level_3=state_totals.percent_level_3.str.strip('%').astype(float)/100
state_totals.percent_level_4=state_totals.percent_level_4.str.strip('%').astype(float)/100
```

```
In [47]: state_total=\
state_totals.astype({'count_total_tested':int,
                     'count_met_standard':int})
```

```
In [48]: state_total.to_csv('state_totals.csv',index=False)
```

The final district level dataset contains 307 school districts and 22 variables. The state-wide dataset for the Chi-square analysis and visualization contains 5 rows and 7 columns and was used to create both Figure 1 and Table 1. Both datasets are summarized below:

```
Index: 307 entries, Aberdeen School District to Zillah School District
Data columns (total 23 columns):
 #   Column                            Non-Null Count  Dtype
---  ------                            --------------  -----
 0   student_FTE                       307 non-null    float64
 1   total_spending                    307 non-null    float64
 2   spending_per_student              307 non-null    float64
 3   num_teachers                      307 non-null    float64
 4   teacher_percent_black             307 non-null    float64
 5   teacher_percent_latinx            307 non-null    float64
 6   teacher_percent_white             307 non-null    float64
 7   teacher_percent_asian             307 non-null    float64
 8   teacher_percent_unknown_demographic 307 non-null  float64
 9   teacher_percent_inexperienced     307 non-null    float64
 10  teacher_percent_limited_cert      307 non-null    float64
 11  teacher_percent_out_of_field      307 non-null    float64
 12  student_count                     307 non-null    float64
 13  percent_black                     307 non-null    float64
 14  percent_latinx                    307 non-null    float64
 15  percent_white                     307 non-null    float64
 16  percent_asian                     307 non-null    float64
 17  students_testing_n                226 non-null    float64
 18  sbac_pass_all                     226 non-null    float64
 19  sbac_pass_black                   28 non-null     float64
 20  sbac_pass_latinx                  116 non-null    float64
 21  sbac_pass_white                   161 non-null    float64
 22  sbac_pass_asian                   46 non-null     float64
```

```
 #   Column              Non-Null Count  Dtype
---  ------              --------------  -----
 0   demographic         5 non-null      object
 1   count_total_tested  5 non-null      int64
 2   count_met_standard  5 non-null      int64
 3   percent_level_1     5 non-null      float64
 4   percent_level_2     5 non-null      float64
 5   percent_level_3     5 non-null      float64
 6   percent_level_4     5 non-null      float64
```

ANALYSIS

In order for this study to reject the null hypothesis, the result of the Chi-square Test for Independence must show that the ratio of test takers to test passers among these four races is highly unlikely to occur by chance. This study adopts the common metric for statistical significance of $p < 0.05$. In this context this would mean the passing rates observed in this data would only be seen in 5% of an infinite number of repeated experimental trials if the students from these races all had an equal likelihood of passing.

	count_total_tested	count_met_standard
asian	6663	4497
black	3522	671
latinx	18179	3913
white	44629	20673

TABLE 1

The result of the Chi-square test of the data shown in Table 1 shows a high level of significance ($p < 0.0001$), indicating these populations are very unlikely to have come from a single population with equally likely passing rates. Based on this value, we reject the null hypothesis in favor of the alternative: there is a significant difference in the achievement rates among test takers of these different racial backgrounds. We can conclude from this outcome that as of the 2018-2019 school year, the Washington State public school system has not achieved its goal of equipping all learners to an equivalent level of mathematics mastery regardless of racial background. The following code was run using the state totals dataset described earlier:

```
In [1]: import pandas as pd
        df=pd.read_csv('state_totals.csv')

In [2]: df.set_index('demographic',inplace=True)

In [3]: df.index=['all','asian','black','latinx','white']

In [4]: from scipy.stats import chi2_contingency

        chi2,p = chi2_contingency(df[['count_total_tested','count_met_standard']].drop('all'),
                                  correction=False)[:2] #Turns off Yate's correction
        table = pd.DataFrame(index=['Value'], columns=['Chi^2', 'p'])
        table.loc[:,:] = [chi2, p]

        table
```

χ^2	p
2628.77	< 0.0001

To follow up this statistical discovery, post hoc analysis comparing each racial pair can help understand which racial populations are significantly different from one another. A pairwise Fisher's Exact test showing unconditional Maximum Likelihood Estimates was carried out and the output is given in Table 2 below. Since we are performing 6 individual pairwise comparisons, it is important to lower the significance threshold according to the Bonferroni correction, dividing our .05 significance threshold value by n = 6 tests to ensure we are not stumbling upon significance simply from performing many repeated tests. This particular method for a Chi-square post-hoc analysis is described at length by Shan and Gerstenberger (2017). Here is the code for the pairwise post-hoc tests:

```
In [18]: demographics=['asian','black','latinx','white']
         p_vals2=pd.DataFrame(index=demographics, columns=pd.Index(demographics,name='(Odds, p)'))

         for group1 in demographics:
             for group2 in demographics:
                 contingency=df.loc[[group1,group2],'count_total_tested':'count_met_standard']
                 odds,p_val=fisher_exact(contingency)
                 p_vals2.loc[group1,group2]=(np.round(odds,2),np.round(p_val,4))

         p_vals2.T.style.set_caption("Fisher's Exact Post Hoc Tests with Odds Ratios")
```

Even with our new p threshold of 0.0083, we find high levels of significance in each pairwise comparison. As shown in Table 3, every group showed significantly different rates of passing the SBAC from every other group, demonstrating a high-level of group independence. The least significantly different pair was the comparison between Latinx students and Black students (p = 0.0072), but it still meets the threshold for significance after the Bonferroni correction. According to these tests, it is not the case that all races succeed at the same levels in the Washington State school system, and in fact, each racial group in this study has a significantly different achievement rate from every other racial group.

Fisher's Exact Post Hoc Comparison Tests

(Odds, p)	asian	black	latinx	white
asian		(3.54, < 0.0001)	(3.14, < 0.0001)	(1.46, < 0.0001)
black	(0.28, < 0.0001)		(0.89, 0.0072)	(0.41, < 0.0001)
latinx	(0.32, < 0.0001)	(1.13, 0.0072)		(0.46, < 0.0001)
white	(0.69, < 0.0001)	(2.43, < 0.0001)	(2.15, < 0.0001)	

TABLE 2

The Maximum Likelihood Estimates given in Table 2 show the odds of meeting standard to a rather discouraging effect. Black students are 0.89 times as likely to pass the 10th grade math SBAC as Latinx students, 0.41 times as likely as White students, and 0.28 times as likely as Asian students. White students, on the other hand, are 2.15 times as likely to pass as Latinx students, and 0.69 times as likely to pass as Asian students.

As shown in Figure 2, this approximate order of success rates persists in nearly every school district where data was available. There are a small number of districts that do not follow this pattern, which could make for an interesting follow-up case study.

Another striking quality of Figure 2 is the level at which district average passing rates vary from each other. Some districts have a greater than 80% passing rate, while elsewhere only around 10% of students pass on average. Furthermore, districts that have greater overall success rate do not always demonstrate higher success rates for all races. Shoreline School district, for example, is fairly successful on average, but exhibits a great deal of inequity of achievement among races. Meanwhile, Stanwood-Camano School District students do not pass as much on average, but the district shows greater success among Latinx students who meet standard at near the district average rate, and even more frequently than the Latinx students in Shoreline.

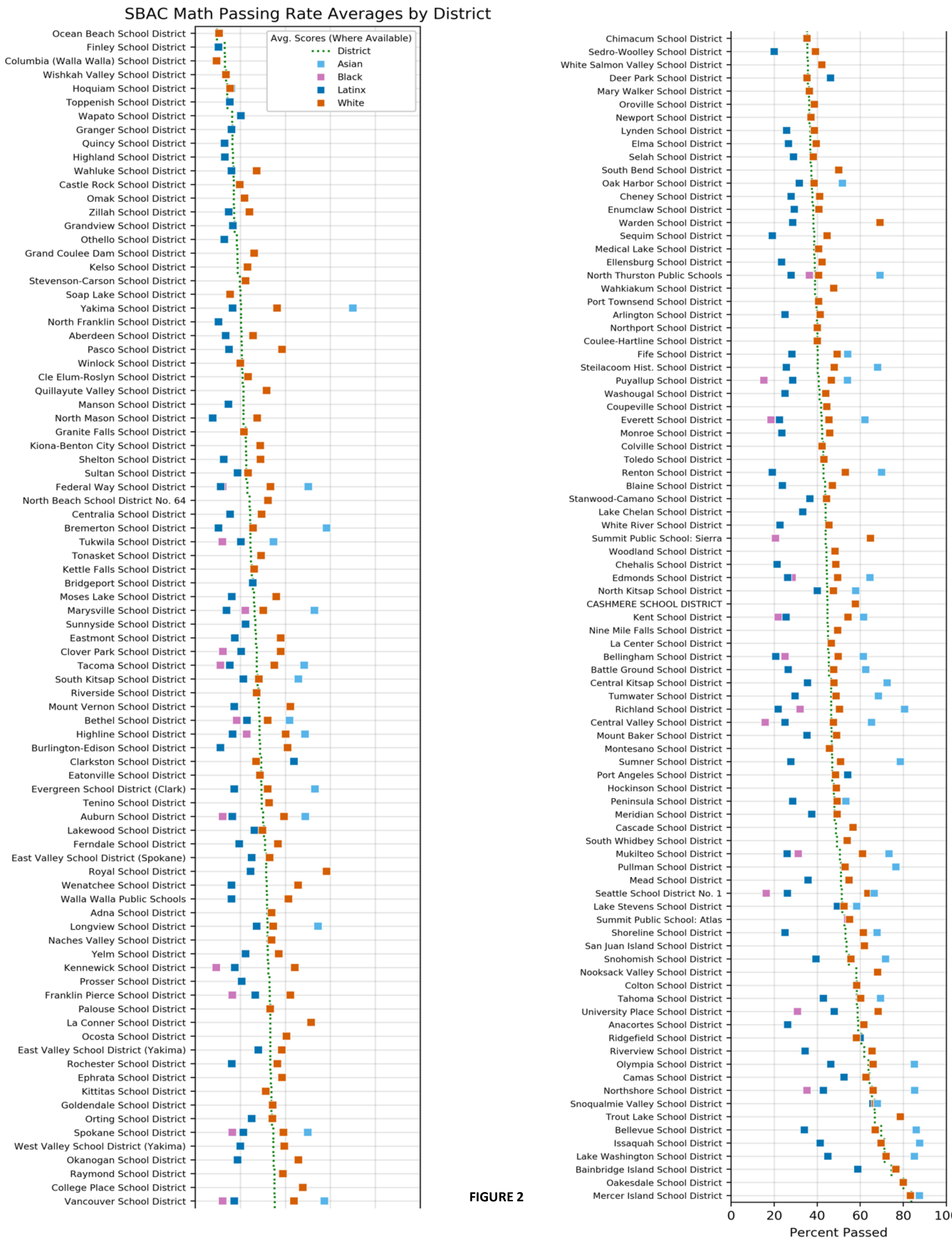
SBAC Math Passing Rate Averages by District
Avg. Scores (Where Available)
District
Asian
Black
Latinx
White
Ocean Beach School District
Finley School District
Columbia (Walla Walla) School District
Wishkah Valley School District
Hoquiam School District
Toppenish School District
Wapato School District
Granger School District
Quincy School District
Highland School District
Wahluke School District
Castle Rock School District
Omak School District
Zillah School District
Grandview School District
Othello School District
Grand Coulee Dam School District
Kelso School District
Stevenson-Carson School District
Soap Lake School District
Yakima School District
North Franklin School District
Aberdeen School District
Pasco School District
Winlock School District
Cle Elum-Roslyn School District
Quillayute Valley School District
Manson School District
North Mason School District
Granite Falls School District
Kiona-Benton City School District
Shelton School District
Sultan School District
Federal Way School District
North Beach School District No. 64
Centralia School District
Bremerton School District
Tukwila School District
Tonasket School District
Kettle Falls School District
Bridgeport School District
Moses Lake School District
Marysville School District
Sunnyside School District
Eastmont School District
Clover Park School District
Tacoma School District
South Kitsap School District
Riverside School District
Mount Vernon School District
Bethel School District
Highline School District
Burlington-Edison School District
Clarkston School District
Eatonville School District
Evergreen School District (Clark)
Tenino School District
Auburn School District
Lakewood School District
Ferndale School District
East Valley School District (Spokane)
Royal School District
Wenatchee School District
Walla Walla Public Schools
Adna School District
Longview School District
Naches Valley School District
Yelm School District
Kennewick School District
Prosser School District
Franklin Pierce School District
Palouse School District
La Conner School District
Ocosta School District
East Valley School District (Yakima)
Rochester School District
Ephrata School District
Kittitas School District
Goldendale School District
Orting School District
Spokane School District
West Valley School District (Yakima)
Okanogan School District
Raymond School District
College Place School District
Vancouver School District
Chimacum School District
Sedro-Woolley School District
White Salmon Valley School District
Deer Park School District
Mary Walker School District
Oroville School District
Newport School District
Lynden School District
Elma School District
Selah School District
South Bend School District
Oak Harbor School District
Cheney School District
Enumclaw School District
Warden School District
Sequim School District
Medical Lake School District
Ellensburg School District
North Thurston Public Schools
Wahkiakum School District
Port Townsend School District
Arlington School District
Northport School District
Coulee-Hartline School District
Fife School District
Steilacoom Hist. School District
Puyallup School District
Washougal School District
Coupeville School District
Everett School District
Monroe School District
Colville School District
Toledo School District
Renton School District
Blaine School District
Stanwood-Camano School District
Lake Chelan School District
White River School District
Summit Public School: Sierra
Woodland School District
Chehalis School District
Edmonds School District
North Kitsap School District
CASHMERE SCHOOL DISTRICT
Kent School District
Nine Mile Falls School District
La Center School District
Bellingham School District
Battle Ground School District
Central Kitsap School District
Tumwater School District
Richland School District
Central Valley School District
Mount Baker School District
Montesano School District
Sumner School District
Port Angeles School District
Hockinson School District
Peninsula School District
Meridian School District
Cascade School District
South Whidbey School District
Mukilteo School District
Pullman School District
Mead School District
Seattle School District No. 1
Lake Stevens School District
Summit Public School: Atlas
Shoreline School District
San Juan Island School District
Snohomish School District
Nooksack Valley School District
Colton School District
Tahoma School District
University Place School District
Anacortes School District
Ridgefield School District
Riverview School District
Olympia School District
Camas School District
Northshore School District
Snoqualmie Valley School District
Trout Lake School District
Bellevue School District
Issaquah School District
Lake Washington School District
Bainbridge Island School District
Oakesdale School District
Mercer Island School District
0
20
40
60
80
100
Percent Passed

FIGURE 2

The rest of the analysis was focused on parsing out any trends or meaningful insight inherent in the school district level dataset, attempting to better understand why various races perform so differently from one another. A few guiding questions for the research come from the EOGOAC's reports and policy recommendations to address various inequities in the education system that they claim disproportionately affect these lower achieving populations (Flores & Rees, 2020).

The first such claim is the fact that student demographics are not represented proportionately within the teacher demographic for most districts, which may contribute to a lacking level of "cultural competence" among teaching staff. The dataset shows that schools with fewer than 50% White students, the bottom quartile of districts in this metric, still have an average of 81% White teachers. Furthermore, as shown in Figure 3, only White students experience 1:1 or greater representation among their teaching staff in nearly all districts.

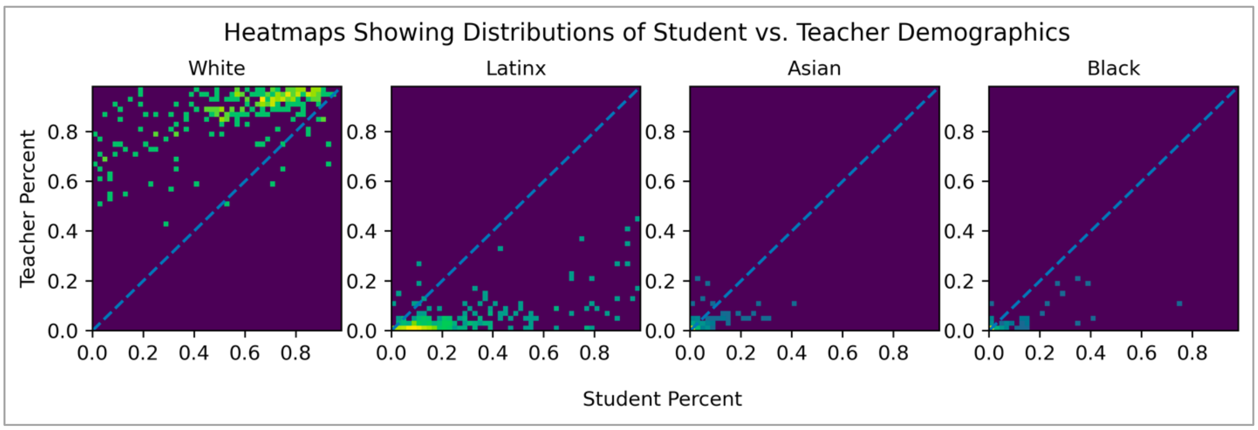

FIGURE 3

Does this mean that districts with better teacher representation for these minority demographics see better achievement? The data is inconclusive regarding this claim. Very few districts currently have high enough representation levels to even study this effect. A simple regression on these variables in Figure 4 suggests that the

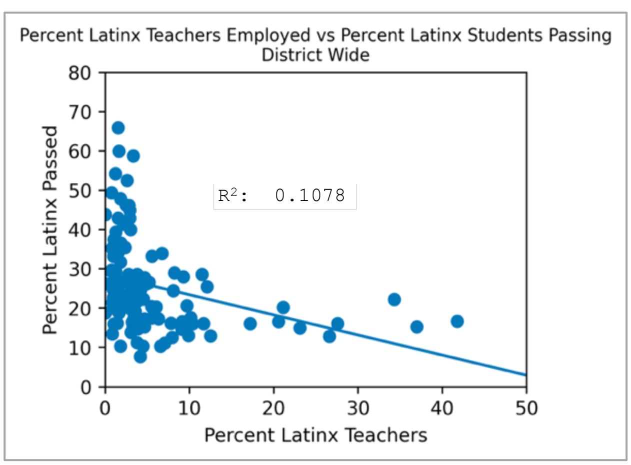

FIGURE 4

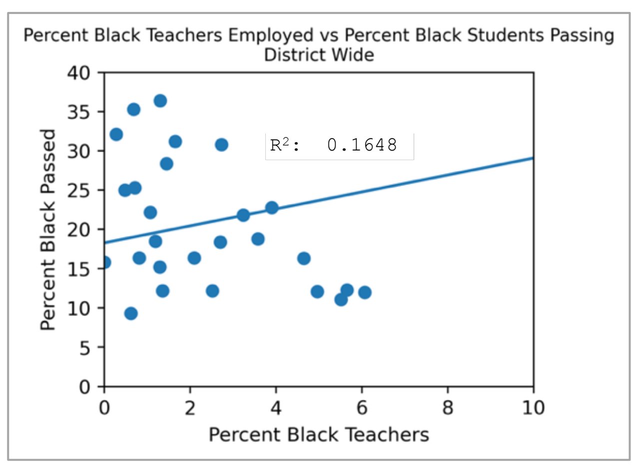

FIGURE 5

percentage of Latinx teachers employed in a district does not have a positive association with the SBAC passing rate of Latinx students in that district, but the R^2 value is just 0.107, suggesting no meaningful conclusions can be drawn.

Similarly, the data cannot demonstrate a meaningful positive effect for the Black teacher employment rate on the Black student passing rate, though it should be noted that no disaggregated SBAC passing rate data is recorded for any district employing greater than 7% Black teachers.

Regression model for predicting percent inexperienced teachers employed at a given Washington State School District

	coef	P>|t|	[0.025	0.975]
const	34.0372	0.000	25.527	42.548
percent_latinx	-0.0230	0.664	-0.127	0.081
percent_black	1.0016	0.000	0.771	1.233
percent_asian	-0.2512	0.110	-0.560	0.057
percent_white	-0.1390	0.005	-0.235	-0.043

TABLE 3

The next of the EOGOAC's claims investigated in this follow-up exploratory analysis was that Black and Latinx students are disproportionately taught by less qualified, less experienced teachers. A multiple linear regression with each of the student demographic percentages set as independent variables and the percentage of inexperienced teachers employed at the district being the dependent variable shows some significance for this claim. According to the regression model shown in Table 3, with everything else held constant, an increase in the White student demographic makeup will decrease the rate of inexperienced teachers in that district ($p = 0.005$), whereas the opposite is true for an increase in the Black demographic makeup ($p < 0.001$). The p values for the coefficients on the Latinx and Asian demographic levels do not suggest statistical significance for these variables. The overall R^2 value for this model is 0.314, suggesting that about 31 percent of the variance across school districts of the rate of inexperienced teachers employed can be explained by this demographic distribution. The definition of an "inexperienced" teacher in this dataset is a teacher having 5 or fewer years of teaching experience. The code for this model output is the following:

```
import statsmodels.api as sms

x=df[['percent_latinx', 'percent_black', 'percent_asian', 'percent_white']]
y=df['teacher_percent_inexperienced']
x=sms.add_constant(x)

model = sms.OLS(y,x).fit()

model.summary()
```

To a weaker, but still significant degree, the above independent variables can also predict the percent of teachers teaching outside their field of expertise and the percent of teachers teaching with a limited teachers' certification. These multiple regression models showed a positive correlation coefficient for `percent_black` with a p value of 0.002 and 0.022 respectively, although with lower R^2 values and coefficient values, meaning these models do not account for as much of the variance and the predictive effect is weaker.

Another issue discussed by the EOGOAC concerns funding allocations for various school districts that are in need of more experienced teachers and higher quality teaching materials. The question arises then, is there any relationship between the lower achieving demographics and state funding allocation? To test this, first a multiple regression model with demographics as the independent and per-student spending as the dependent variable. At first, no interactions between the student demographic percentage rates were included in the model, and this led to a model that was difficult to interpret. Though the p values were significant, the fact that each individual demographic increase predicted a decrease in funding (albeit at different rates) made the model in Table 4 lack much meaning. However, when the model included all interactions between the variables representing the various demographic percentages as shown in Table 5, the model found that, with everything else held constant, an increase in Asian demographic alone predicts a large increase (between \$391 and \$3329) in per-student spending district-wide. However, when an increase in Latinx alone or White alone demographic occurs, a decrease in per-student funding is predicted by the model. The Black alone demographic prediction has a very high p value, indicating the model sees a great amount of variation in this characteristic among districts.

As a final piece of exploratory data analysis, clustering of the districts by certain characteristics was carried out in order to detect any tangible common characteristics across districts in the state. The clustering method used was the HDBSCAN clustering python package, a density-based clustering model that looks for observations that exist in groups that are close together across the data dimensions. According to McInnis and Healy (2016), HDBSCAN is a particularly useful clustering

Model predicting per-student funding district wide from demographic characteristics (interactions excluded)

	coef	P>|t|	[0.025	0.975]
Intercept	2.692e+04	0.000	2.24e+04	3.14e+04
percent_latinx	-131.0995	0.000	-186.449	-75.750
percent_white	-95.4533	0.000	-146.218	-44.688
percent_black	-62.1448	0.319	-184.613	60.324
percent_asian	-359.0094	0.000	-522.640	-195.378

TABLE 4

Model predicting per-student funding district wide with demographic characteristics (interactions included)

	coef	P>|t|	[0.025	0.975]
Intercept	2.671e+04	0.000	2.22e+04	3.12e+04
percent_latinx	**-102.1092**	**0.004**	**-171.710**	**-32.508**
percent_white	**-76.2449**	**0.007**	**-131.474**	**-21.016**
percent_latinx:percent_white	-0.8040	0.387	-2.630	1.022
percent_black	-139.6575	0.737	-956.851	677.536
percent_latinx:percent_black	1.9420	0.926	-39.415	43.299
percent_white:percent_black	6.3495	0.474	-11.065	23.764
percent_latinx:percent_white:percent_black	-0.2047	0.709	-1.285	0.875
percent_asian	**1860.0337**	**0.013**	**391.071**	**3328.997**
percent_latinx:percent_asian	-56.7116	0.189	-141.474	28.051
percent_white:percent_asian	-35.9425	0.009	-62.732	-9.153
percent_latinx:percent_white:percent_asian	0.4130	0.611	-1.182	2.008
percent_black:percent_asian	0.2507	0.998	-251.350	251.852
percent_latinx:percent_black:percent_asian	-0.7532	0.873	-10.021	8.514
percent_white:percent_black:percent_asian	-1.9996	0.530	-8.256	4.257
percent_latinx:percent_white:percent_black:percent_asian	0.1125	0.393	-0.146	0.371

TABLE 5

algorithm for exploratory data analysis as it makes very few assumptions about the data other than the fact that some noise likely exists. Since SBAC testing data was not available for all districts, clustering only took place on the 225 districts for which this data was available. After some trial and error, the model excluded variables that contained information about school district size. When district size information was included among the parameters, the clustering algorithm was fixated on the various district sizes, and yet this is among the least useful or interesting features for the purposes of this study. (One cannot simply create larger or smaller school districts in order to affect education outcomes! District size is very much geographically dependent.) Additionally, one variable was calculated for the purposes of clustering, namely students per teacher, reflecting the educational metric of "class size."

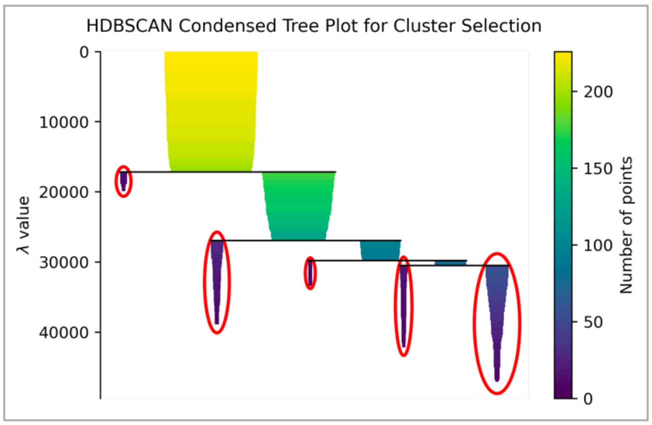

FIGURE 6

The HDBSCAN algorithm allows for some districts to be classified as "noise" among clustered districts, and this "noise" group became the largest cluster group (n = 91). The algorithm found 5 distinct groups in the 225 districts, in addition to what will be referred to as the "noise" group, or "Cluster 0," for those districts classified that did not fit into any densely clustered pattern. (Note: in order to select the smaller terminal clusters shown in the tree plot shown in Figure 6, the "leaf" clustering method was selected. To see a full list of Districts that comprise each of the clusters, see Appendix A.) To investigate the qualities where clusters differed,

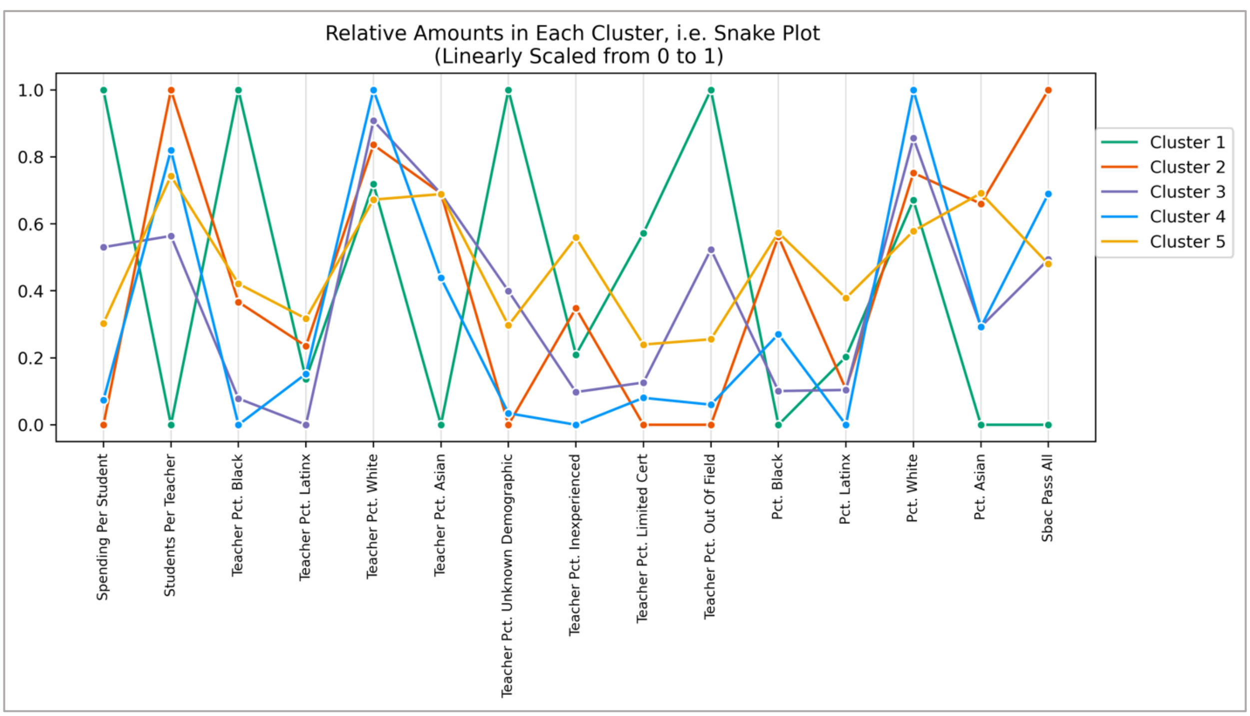

FIGURE 7

a snake plot was used (Figure 7) and to see the relative distributions across all variables by cluster, a faceted violin plot was used (Figure 8). This combination of the mean-points from the snake plot in addition to the overview of the full distribution on the violin plot allows for a more precise and careful characterization of the clusters: while the snake plot emphasizes the differences between the clusters, the violin plot helps to put these differences into perspective. (Note: In the violin plot, Cluster 0 represents the set of school districts classified as "noise" by the HDBSCAN algorithm, and thus can naturally be seen to have the greatest amount of variation across each of the variables.)

Cluster 2 appears to have the highest performance on the math SBAC, whereas Cluster 1 appears to have the lowest average performance, so these clusters are of particular interest to this study. Cluster 2 has the lowest rate of teachers teaching under a limited certification and teachers teaching out of their field of expertise. They also have a decisively lower spending per student and higher number of students per teacher. They have a prominent bulge in

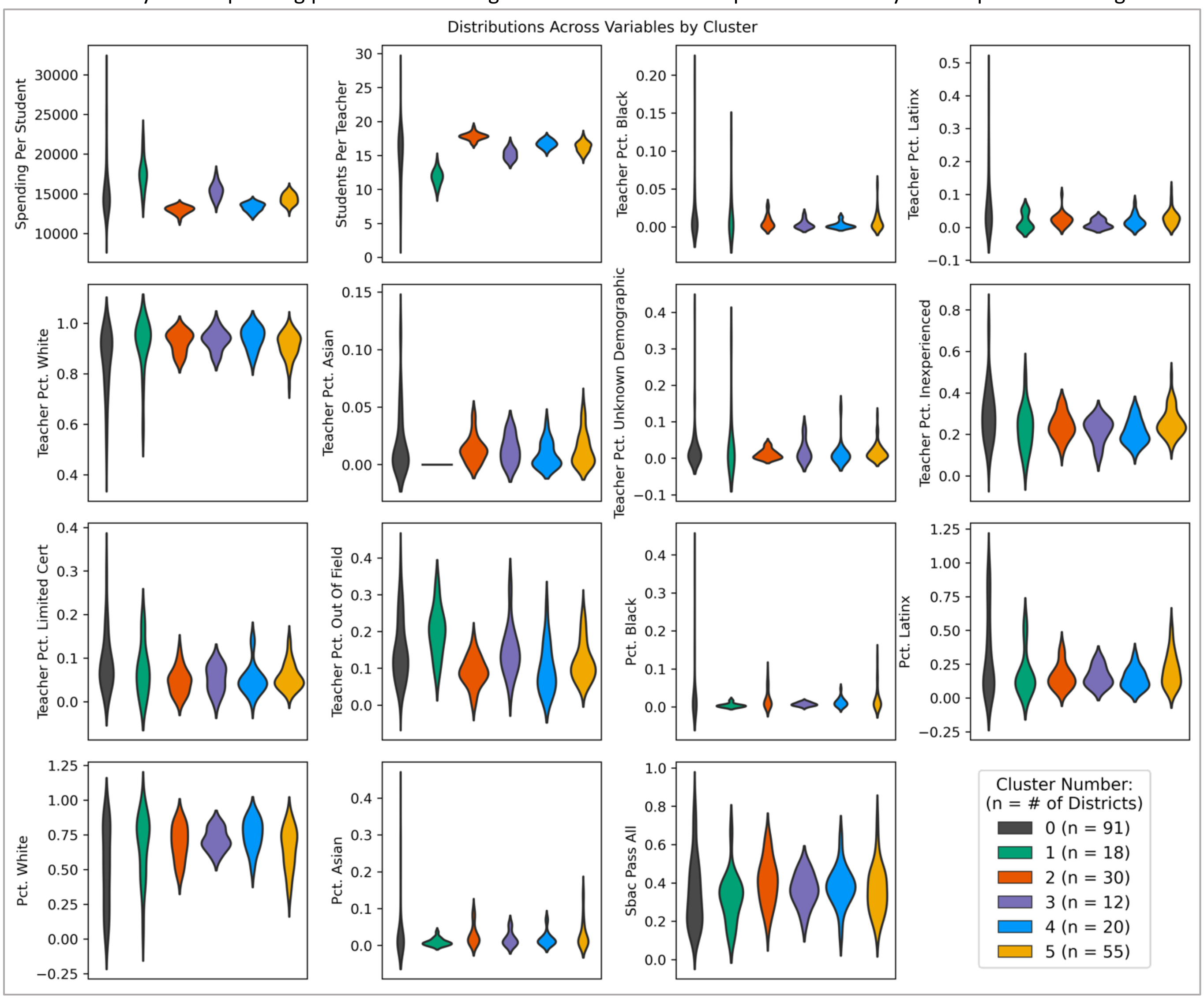

FIGURE 6

Asian teachers with the distribution centered above all the other clusters for this variable, though there are districts in Cluster 0 with more Asian teachers. There is some right skew to the Black student demographic in Cluster 2.

Cluster 1 has the most amount of variation, or in other words, it's the least dense cluster across many variables (barring Cluster 0). However, it is marked by its smaller number of students per teacher, total lack of Asian teachers, and low percent Black and Asian student demographic. There also appears to be some bi-modality to the Latinx population, both student and teacher – indicating some districts with few, and some districts with comparatively many Latinx persons, with a gap in-between. Cluster 1 also has many teachers teaching of their field, indicating these districts may be experiencing teacher shortage.

In addition to SBAC performance clues, the clusters also reiterate and illuminate some of the findings from the prior regression analyses. Cluster 5's prominent features include a heavy right skew in Latinx and Black student demographics, as well as the highest average ratio of inexperienced teachers. These districts have the lowest average ratio of White students (though Cluster 1 encompasses some districts with even fewer White students). Overall, Cluster 5 fares 2nd to worst on the SBAC passing rate, albeit with a lot of variation shown in the violin plot (there is some very high right skew). Cluster 5 also has a somewhat low district spending and somewhat high class-size on average. It is also notable that this is the largest cluster with n = 55 districts.

This analysis will conclude with some general observations on the cluster analysis. All clustered districts have a high concentration of White teachers, and districts not employing predominantly White teachers are considered "noise" by the clustering algorithm, meaning they are not part of any sort of recognizable norm state-wide. Just clusters 1 and 5 have some left skew on this variable. It is also notable that while certain clusters did better on average on the SBAC, there is no clear association from the HDBSCAN clustering in terms of SBAC performance, as all 5 clusters have a great deal of variation in this variable. The cluster with the least variance on the SBAC test appears to be made up of mostly White and Asian student demographics with a somewhat high spending per student and somewhat low class-size (students per teacher). However, this cluster only performs right in the middle in terms of SBAC achievement, and it is also the smallest cluster. Finally, it is interesting to note that the 91 districts categorized as "noise" by the algorithm seem to fair in general worse than any of the 5 clusters in terms of their distribution, although Cluster 1 still has the lowest average, due to the very long tails on "Cluster 0."

Data Summary and Implications

Ever since standardized student achievement data disaggregated by race has been widely analyzed in the US, there has been some significant discrepancy in terms of how students from various racial backgrounds performed (NCES, 2013). Although this nation-wide longitudinal data demonstrates that Black and Hispanic/Latinx students have made relative gains, it is discouraging to see this new evidence of the same achievement gap persisting amongst students in Washington State more than four decades later. Various attempts to explain or quantify the root causes of this gap have led to policy proposals, both in Washington State and across the US, but neither government nor society have managed to enact the necessary change to overcome this persistent metric of racial inequality.

Educational research conducted worldwide demonstrates that it is possible for school systems to enact change and observe significant increases in their student performance. The PISA international study, which tests 15-year-olds from countries all around the world, has prompted countries to address their own achievement based on an international standard. Some countries have responded to performance gaps with great success based on these measures. In the early 2000's, in response to receiving news of very low scores in Germany, especially for immigrant and impoverished populations, Germany nearly doubled its spending on public education. They expanded early childhood education and implemented a national set of standards to monitor their own progress towards equity (OECD, *Germany's PISA Shock*). Germany saw academic achievement change dramatically - students went from performing below average in every metric (reading, mathematics, and science), with the greatest rates of inequality between student groups observed in any of the countries tested by PISA, to performing above average nationwide. This improvement was largely due to immigrant populations making significant gains and narrowing Germany's achievement gap to a level that is about average worldwide in 2012 (Berwick, 2015). Since 2012, Germany has seen some decline from its students, but remains at or above average performance levels overall.

By comparison, students in the US perform near worldwide average levels in reading and science but below average in mathematics. Nationwide, according to the OECD (2019, p. 4), the US has seen no significant gains since the year 2000 other than some small gains in science performance.

Traditionally, education in the US has been left up to state and local governments to decide on many aspects of how the system functions. States also have significantly varying demographics both in terms of socio-economic and racial status. Due to these regional differences, one may expect that there is potential for a great deal of variation to occur from state to state. This study was focused on examining the most recent Washington State testing data in light of the legislation's recognition and prioritization of the achievement gap. This study finds that significant issues of unequal math achievement among the races remain. The reasons for these gaps in achievement are likely extremely complex and depend on many factors both in and outside the school system. However, this study also finds some statistical basis for a number of feasible claims that there are certain school-system-wide factors that disproportionately affect these lower achieving student groups. This is one reason that the EOGOAC replaced the term “Achievement Gap” with the term “Opportunity Gap” - this shifts the blame away from individual low performing students and onto policymakers for failing to enact the reforms needed to correct these institutional patterns. This study identifies a number of such patterns that could be contributing to the lower levels of subject mastery observed from the groups affected.

The regression model shown in Table 3 uses the 2018-2019 school year data to demonstrate that districts with higher concentrations of Black students employ more inexperienced teachers, while those with increased White and Asian demographics employ fewer inexperienced teachers. These finding are similar for districts employing teachers that are not fully certified and/or teaching out of their field of expertise. While it is necessary to provide new teachers with opportunities to practice their craft so that they can gain effective teaching skills, a large ratio of inexperienced teachers can be detrimental to a student population in a number of ways. Primarily, inexperienced teachers often lack the complex skills necessary to engage all learners at a high level. Furthermore, students may be less likely to open up and trust a new teacher while that teacher is just getting their footing in a school. If districts are comprised of high levels of new teachers, students may see many teachers leaving their schools throughout their academic careers instead of investing in that district long enough to make a difference in the lives of students. Finally, teachers who are teaching with limited certification and/or teaching outside of their field may be a marker for a district facing teacher shortages, and these districts may be hiring less qualified candidates from a lower-talent hiring pool in order to simply fill positions. When the demographic makeup of a school district can explain nearly 32 percent of the variation in the experience level of its teaching staff, to the benefit of higher performing demographics and to the detriment of lower performing demographics, this is an issue that demands a response from policy makers. The system must find ways to incentivize highly experienced and skilled teachers to stay in districts that traditionally face staff turn-over and/or a significant population of less-experienced teachers.

Furthermore, this 2018-2019 school year data demonstrates that Black and Latinx students seldom see their own demographics represented in the teaching staff where they attend school. The relationship between students and teachers is a delicate power dynamic that requires a great deal of mutual respect and understanding. Due to the long history of racial tensions in the United States, there is a real potential for cultural misunderstandings leading to a strained or diminished student-teacher relationship, ultimately resulting in less optimal educational outcomes. Another very practical and tangible concern for the Latinx population in this regard is the fact of language barriers. According to the Pew Research Center, just under two-thirds of the adult Hispanic population in the US either speak English or are bilingual (Krogstad & Gonzalez-Barrera 2015). This means many Hispanic children are raised in a Spanish speaking household and could benefit from school staff speaking their preferred language. Employing more Latinx teachers helps to remove both the cultural and the language barriers that may interfere with a Latinx learner’s academic work. While the Revised Code of Washington, § 28A.600.015 (2016) notes that “the legislature intends to adopt policies and programs to implement” the EOGOAC’s recommendations to “Enhance the cultural competence of current and future educators and classified staff,” and “Invest in the recruitment, hiring, and retention of educators of color,” the current data still shows an overwhelmingly White teacher population state-wide. We must find additional ways to avoid disenfranchising students of color and train teachers that can effectively identify with and reach these student populations.

This summary will conclude by naming few factors not taken into account in this study. Socioeconomic status data has not been considered in this analysis and many studies have demonstrated that wealth levels have an overwhelming impact on student achievement. It is also the case that Black and Latinx families have significantly lower median wealth levels than those of White and Asian families (Bhutta, Neil, Chang, Dettling, & Hsu, 2020). However, as PISA and the EOGOAC studies demonstrate, even after accounting for wealth, achievement gaps persist along racial boundaries (Florence & Rees, 2020). They show that on the 8th grade SBAC, Black students who are not low income perform just above white students who *are* classified as low income, meeting standard a fraction of the time as non-low-income White and Asian students. So, while income can explain some of the achievement gap seen between students of different races, it is far from the only contributing factor. As this study shows, race alone is a highly significant variable contributing to a student's odds of meeting mathematics learning standards.

Another such factor that this study neglects to consider is the disciplinary rates seen for students from various demographics. In response to the school punitive measure rates observed to be unequal across different states, Revised Code of Washington, § 28A.600.015 also includes the recommendation to "Reduce the length of time students of color are excluded from school due to suspension and expulsion and provide students support for reengagement plans." This factor likely plays some part in the disenfranchisement of learners of color who experience more frequent suspension and expulsion, but this study does not take this variable into account. The author suggests this additional variable as important to include in further clustering-based exploration to identify if there is some underlying pattern defining districts where punishment is especially biased against students of color.

One final variable that could yield fruitful results in future analysis of this data includes factoring in trends across geographic regions. Since schools only draw from students and teachers living nearby, it is likely that a great deal about the characteristics of school districts stem from underlying geographic relationships. Analyses should be carried out to see if students are performing better in certain locations around the state so that underperforming locales can be targeted for various kinds of intervention and aid.

These unequal test scores across racial lines dubbed the "Opportunity Gap" are signals from these students of color that something in the system is not working. Washington State must find ways to engage these learners; the well-being of its future citizens, the success of its local businesses, and the flourishing of scientific research within its state borders all depend on it. This paper concludes with a quote from the Washington State Constitution - Article IX, Section 1 is a reminder that it is of fundamental importance to the identity of this State to pursue all avenues available to close the highly significant learning gap that divides students across racial lines to this day:

"It is the paramount duty of the state to make ample provision for the education of all children residing within its borders, without distinction or preference on account of race, color, cast or sex."

SOURCES

American Psychological Association. (2019). 5.7 Racial and Ethnic Identity. In *Publication manual of the American Psychological Association*. Washington, DC: Americam Psychological Association.

Berwick, C. (2015, November 3). The Great German School Turnaround. *The Atlantic*. Retrieved December 29, 2020, from https://www.theatlantic.com/education/archive/2015/11/great-german-scool-turnaround/413806/

Bhutta, Neil, Chang, A. C., Dettling, L. J., & Hsu, J. W. (2020, September 28). Disparities in Wealth by Race and Ethnicity in the 2019 Survey of Consumer Finances. Retrieved December 29, 2020, from https://doi.org/10.17016/2380-7172.2797

Dilmegani (2020). *60 Top Data Science Tools: The Ultimate Guide [2020 update].* Retrieved December 15, 2020 from https://research.aimultiple.com/data-science-tools/

Flores, M., & Rees, H. (2020). *Closing the Opportunity Gap in Washington's Public Education System* (Rep. No. 2020 Annual). Retrieved December 15, 2020, from The Educational Opportunity Gap Oversight and Accountability Committee (EOGOAC) website: https://www.k12.wa.us/sites/default/files/public/workgroups/eogoac/pubdocs/2020%20EOGOAC%20Report.pdf

Gramlich, J. (2020, August 06). Black imprisonment rate in the U.S. has fallen by a third since 2006. Retrieved December 29, 2020, from https://www.pewresearch.org/fact-tank/2020/05/06/share-of-black-white-hispanic-americans-in-prison-2018-vs-2006/

Highlights of U.S. PISA 2018 Results Web Report (NCES 2020-166 and NCES 2020-072). Retrieved December 15, 2020, from U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics. Retrieved December 15, 2020, from: https://nces.ed.gov/surveys/pisa/pisa2018/index.asp.

Krogstad, J. M., & Gonzalez-Barrera, A. (2015, March 24). A majority of English-speaking Hispanics in the U.S. are bilingual. *Pew Research Center*. Retrieved December 29, 2020, from https://www.pewresearch.org/fact-tank/2015/03/24/a-majority-of-english-speaking-hispanics-in-the-u-s-are-bilingual/

Lasser, J. (2020). Creating an executable paper is a journey through Open Science. *Communications Physics, 3*(1). doi:10.1038/s42005-020-00403-4

McInnes, L., & Healy, J. (2016). Comparing Python Clustering Algorithms (1142673943 860130485 S. Astels, Ed.). Retrieved December 29, 2020, from https://hdbscan.readthedocs.io/en/latest/comparing_clustering_algorithms.html

National Center for Education Statistics. (2013). *The Nation's Report Card: Trends in Academic Progress* (The Nation's Report Card, pp. 38-40, Rep.). Washington, DC: US Department of Education.

National Center for Education Statistics. (2019, February). Indicator 19: College Participation Rates. Retrieved December 29, 2020, from https://nces.ed.gov/programs/raceindicators/indicator_REA.asp

OECD. (2019). *United States - Country Note - PISA 2018 Results* (p. 4, Rep.). Paris: OECD Publishing.

OECD. (n.d.). Germany's PISA Shock. Retrieved December 29, 2020, from http://www.oecd.org/about/impact/germany-pisa-shock.htm

OSPI. (2009). EOGAOC Reports and Resources. Retrieved December 29, 2020, from https://www.k12.wa.us/about-ospi/workgroups-committees/currently-meeting-workgroups/educational-opportunity-gap-oversight-and-accountability-committee/eogaoc-reports-and-resources

SBE. (n.d.). Graduation Pathway Options. Retrieved December 29, 2020, from https://www.sbe.wa.gov/our-work/graduation-pathway-options

Shan, G., & Gerstenberger, S. (2017). Fisher's exact approach for post hoc analysis of a chi-squared test. *Plos One, 12*(12). doi:10.1371/journal.pone.0188709

Wash. Rev. Code § 28A.300.042 (2016)

Wash. Rev. Code § 28A.600.015 (2016)

Wash. Const. art. IX, sect. I.

Appendix A

Table of Districts by Cluster Grouping

According to the HDBSCAN algorithm with a min distance of 7 and after row vector normalization using scikit-learn's normalize method, the following clusters were identified, as illustrated in Figures 7 and 8:

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Coulee-Hartline School District	Arlington School District	Bellingham School District	Adna School District	Anacortes School District
Dayton School District	Bethel School District	Mary Walker School District	Asotin-Anatone School District	Auburn School District
La Conner School District	CASHMERE SCHOOL DISTRICT	Mossyrock School District	Chehalis School District	Bainbridge Island School District
Lyle School District	Camas School District	Mount Baker School District	Cheney School District	Battle Ground School District
Northport School District	Castle Rock School District	Ocosta School District	Clarkston School District	Blaine School District
Oakville School District	Colville School District	Port Townsend School District	Ellensburg School District	Cascade School District
Odessa School District	Ephrata School District	Raymond School District	Enumclaw School District	Central Kitsap School District
Pateros School District	Fife School District	Sedro-Woolley School District	Freeman School District	Centralia School District
Pe Ell School District	Hockinson School District	South Kitsap School District	Kalama School District	Chewelah School District
Republic School District	Kennewick School District	South Whidbey School District	Longview School District	Clover Park School District
Ritzville School District	Kettle Falls School District	Stevenson-Carson School District	Napavine School District	Columbia (Walla Walla) School District
South Bend School District	Lake Stevens School District	Willapa Valley School District	North Thurston Public Schools	Coupeville School District
Tekoa School District	Lakewood School District		Olympia School District	East Valley School District (Spokane)
Touchet School District	Lynden School District		Onalaska School District	Edmonds School District
Trout Lake School District	Mead School District		Rochester School District	Elma School District
Wahkiakum School District	Monroe School District		Stanwood-Camano School District	Everett School District
Waterville School District	Montesano School District		Toutle Lake School District	Evergreen School District (Clark)
Wellpinit School District #49	Pullman School District		Tumwater School District	Ferndale School District
	Puyallup School District		Vancouver School District	Finley School District
	Quillayute Valley School District		White Salmon Valley School District	Franklin Pierce School District
	Rainier School District			Goldendale School District
	Reardan-Edwall School District			Granite Falls School District
	Ridgefield School District			Kelso School District
	Riverview School District			Kiona-Benton City School District
	Steilacoom Hist. School District			Kittitas School District
	Sumner School District			Lake Chelan School District
	Tenino School District			Marysville School District
	University Place School District			Medical Lake School District
	West Valley School District (Yakima)			Mukilteo School District
	Yelm School District			Naches Valley School District
				Newport School District
				Nine Mile Falls School District
				Nooksack Valley School District
				North Kitsap School District
				North Mason School District
				Oak Harbor School District
				Peninsula School District
				Pomeroy School District
				Port Angeles School District
				Quilcene School District
				Riverside School District
				Selah School District
				Sequim School District
				Shoreline School District
				Snoqualmie Valley School District
				Soap Lake School District
				Spokane School District
				Sultan School District
				Tacoma School District
				Toledo School District
				Tonasket School District
				Washougal School District
				Wenatchee School District
				Winlock School District
				Woodland School District

The following School Districts were not found to share any strong pattern with any cluster above:

Aberdeen School District
Bellevue School District
Bremerton School District
Brewster School District
Bridgeport School District
Burlington-Edison School District
Cape Flattery School District
Central Valley School District
Chimacum School District
Cle Elum-Roslyn School District
Colfax School District
College Place School District
Colton School District
Columbia (Stevens) School District
Concrete School District
Crescent School District
Creston School District
Curlew School District
Cusick School District
Davenport School District
Deer Park School District
East Valley School District (Yakima)
Eastmont School District
Eatonville School District
Entiat School District
Excel Public Charter School
Federal Way School District
Grand Coulee Dam School District
Grandview School District
Granger School District Highland School District
Highline School District
Hoquiam School District
Inchelium School District
Issaquah School District
Kent School District
La Center School District
Lake Washington School District
Liberty School District
Mabton School District
Manson School District
Mary M Knight School District
Mercer Island School District
Meridian School District
Methow Valley School District
Morton School District
Moses Lake School District
Mount Vernon School District
Naselle-Grays River Valley School District
North Beach School District No. 64
North Franklin School District
Northshore School District
Oakesdale School District
Ocean Beach School District
Okanogan School District
Omak School District
Orcas Island School District
Oroville School District
Orting School District
Othello School District
Palouse School District
Pasco School District
Prosser School District
Quincy School District
Renton School District
Richland School District
Rosalia School District
Royal School District
San Juan Island School District
Seattle School District No. 1
Selkirk School District
Shelton School District
Snohomish School District
St. John School District
Summit Public School: Atlas
Summit Public School: Olympus
Summit Public School: Sierra
Sunnyside School District
Tahoma School District
Toppenish School District
Tukwila School District
Wahluke School District
Walla Walla Public Schools
Wapato School District
Warden School District
White Pass School District
White River School District
Wilbur School District
Wishkah Valley School District
Yakima School District
Zillah School District

Finally, the following columns were omitted from clustering due to missing testing data. This is an indicator of having a small testing population where results have been suppressed:

Almira School District
Benge School District
Bickleton School District
Boistfort School District
Brinnon School District
Carbonado School District
Centerville School District
Conway School District
Cosmopolis School District
Damman School District
Darrington School District
Dieringer School District
Dixie School District
Easton School District
Endicott School District
Evaline School District
Evergreen School District (Stevens)
Garfield School District
Glenwood School District
Grapeview School District
Great Northern School District
Green Dot Destiny
Green Mountain School District
Griffin School District
Harrington School District
Hood Canal School District
Rainier Prep Charter School District
Rainier Valley Leadership Academy
Roosevelt School District
Index Elementary School District 63
Innovation Schools
Kahlotus School District
Keller School District
Klickitat School District
LaCrosse School District
Lake Quinault School District
Lamont School District
Lind School District
Loon Lake School District
Lopez School District
Mansfield School District
McCleary School District
Mill A School District
Mount Adams School District
Mount Pleasant School District
Nespelem School District #14
North River School District
Onion Creek School District
Orchard Prairie School District
Orient School District
Orondo School District
PRIDE Prep Charter School District
Palisades School District
Paterson School District
Pioneer School District
Prescott School District
Queets-Clearwater School District
SOAR Academy Charter District
Satsop School District
Shaw Island School District
Skamania School District
Skykomish School District
Southside School District
Spokane International Academy
Sprague School District
Star School District No. 054
Starbuck School District
Stehekin School District
Steptoe School District
Summit Valley School District
Suquamish Tribal Education Department
Taholah School District
Thorp School District
Union Gap School District
Valley School District
Vashon Island School District
Waitsburg School District
Washtucna School District
West Valley School District (Spokane)
Wilson Creek School District
Wishram School District