Education Accountability Act (SB141) Similar Districts Technical Manual

June 2026

Table of Contents

Introduction

This document provides a summary of the methodology used by the Oregon Department of Education (ODE) in the development of a similar school district model as required by the Education Accountability Act (SB 141).

The Education Accountability Act requires the State Board of Education to set metric targets for similar school districts and requires school districts and charter schools to use the metric targets when setting longitudinal performance growth targets for:

- K-2 Regular Attendance
- K-12 Regular Attendance
- 3rd Grade ELA
- 8th Grade Math
- 9th Grade On Track

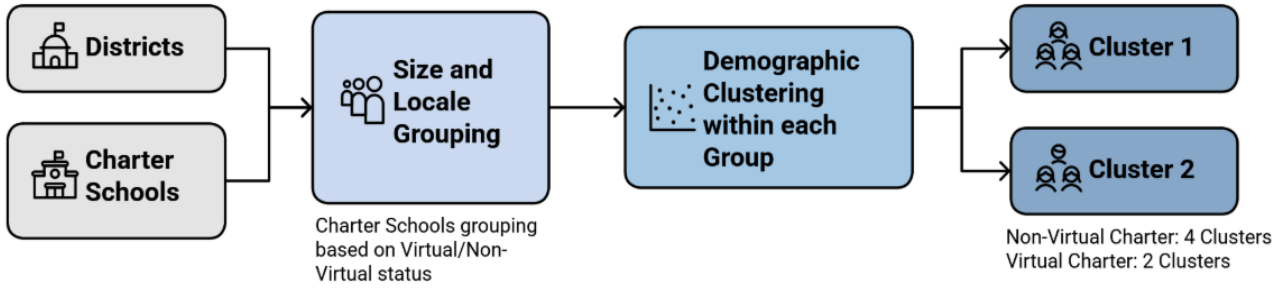
- Four-year Graduation
- Five-year Completion
- Local Metrics[1]

The aim of the similar districts grouping is to identify districts with similar local contexts by clustering districts and charter schools based on their student demographics and community characteristics (see Figure 1 for an overall process summary and the *Method* section for full details). These clusters of similar districts will then be used to establish metric targets for student outcomes over the next four years. Each district will use the targets for its similar district cluster to inform its own target setting in co-development with ODE.

The notion of similar districts acknowledges that districts differ both in their starting places and outcomes. As a result, districts can vary in what will be ambitious yet achievable targets and progress toward those targets.

In addition to target setting, these clusters can provide educators and administrators with a useful set of districts that are comparable in meaningful ways, offering additional context to inform decision-making, focus data analysis, foster learning, and support continuous improvement.

Figure 1. Diagram of Similar District Clustering Process



The diagram provides a broad summary for the Similar Districts clustering process.

1. Districts and charter schools were analyzed as separate groups for Similar Districts clustering.
2. Districts were first stratified into Groups based on their enrollment size and geographic locale, while charter schools were grouped based on their virtual or non-virtual status.
3. Within these initial Groups, a statistical clustering method based on student and community characteristics was applied.

[1] More information on the Local Metrics can be found on the ODE Education Accountability Act Website

4. This resulted in two clusters for each district Group (e.g., Group 1 Cluster 1 and Group 1 Cluster 2) and the Virtual Charter Group, while the Non-Virtual Charter group was divided into four clusters. See the *Method* section below for full details.

Similar District Method

Size and Locale Stratification

Before clustering, districts were first stratified into different Groups (e.g., Group 1, Group 2, etc. See Map 1 and Table 1) based on student enrollment[2] and NCES[3] geographic locale. Subsequent "nearest neighbors" statistical clustering was conducted only within each Group and not across Groups. For example, all districts in Group 1 were further classified into Cluster 1 or Cluster 2 within that same Group. No district will have similar districts from a different Group.

Charter schools were removed from their districts' data[4] and comprise two additional Groups: one for virtual charter schools and one for non-virtual charter schools.

Districts could appeal to be moved to a size and locale grouping that they felt offered better comparators (see *Appeal Process* for more details). For example, a larger district composed of many small towns may feel that they have closer comparators in a group containing small town districts. ODE also made some adjustments to the groupings based on the professional judgment of regional specialists at ODE, prior to the appeals process.

Student Enrollment Grouping:
Large: 7,500 or more students
Medium: 1,651 - 7,499 students
Small:[5] 80 - 1,650 students
Very small:[6] Fewer than 80 students

NCES Geographic Locale Grouping:
City - large densely populated principal cities

[2] Student enrollments as of the 1st school day in May 2025 with charter school enrollment separated out.
[3] National Center for Education Statistics Locale Classifications
[4] Except in cases where the district is composed entirely of charter schools.
[5] From HB3037, "Small ADM school district" means a school district that has an average daily membership, as defined in ORS 327.006, of less than 1,650.
[6] From Student Investment Account small and rural district determination.

Suburban - urban areas located outside principal cities
Town - near urban areas with smaller population
Rural - non urban territory

Map 1. School Districts by Final Similar District Size and Locale Grouping

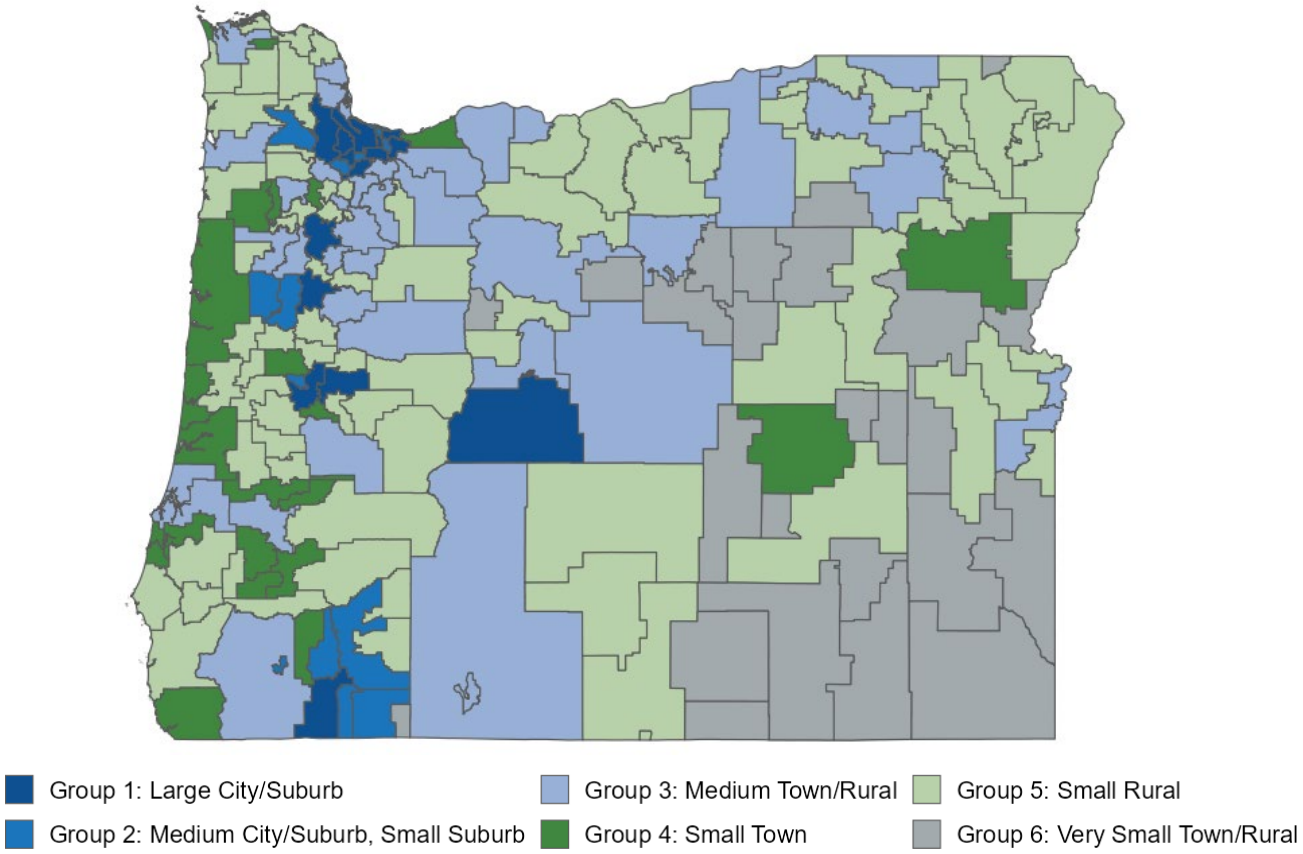


Table 1 below indicates the enrollment size and locale classification that corresponds to each final Similar District Group.

Table 1. Final Similar District Enrollment Size and Locale Classification

Group	Size and Locale
1	Large City/Suburb
2	Medium City/Suburb, Small Suburb
3	Medium Town/Rural
4	Small Town

5	Small Rural
6	Very Small Town/Rural

Charter schools were separated into Groups based on their virtual or non-virtual status.

For simplicity, the following sections describe the clustering process for districts. The same process was followed for charter schools except where differences are noted.

Student Demographic and District Variables

For the identification of similar districts, ODE utilized a hierarchical clustering method to calculate the statistical "nearest neighbors" of districts based on student demographic and community attributes.

The selection of the student demographic and community attribute variables was informed by past ODE accountability work on similar-school groupings, engagement with the Accountability Reporting Advisory Committee (ARAC), the Assessment & Accountability Technical Advisory Committee (TAC), engagement with schools and districts, including charter schools, and the performance of the variables in the model.

After the initial stratification, districts in each of the Groups were further separated into two Clusters (e.g, Group 1, Cluster 1 or Group 1, Cluster 2). The calculation and identification of similar district Clusters within each Group relied on the following student demographic and community attributes:

- The percent of students identified as experiencing poverty[7]
- The percent of students identified as ever English learners[8]
- The percent of students identified as belonging to a racial/ethnic group that has historically experienced academic disparities[9]
- The percent of students identified as mobile within the school year[10]

[7] These are students who participate in SNAP, participate in TANF, are foster students, houseless students, and/or migratory students.

[8] These are students who participated or are currently participating in a program to acquire academic English.

[9] These are students who are American Indian/Alaska Native, Asian, Black/African American, Hispanic/Latinx, Multiracial, or Native Hawaiian/Pacific Islander.

[10] These are students who experience one or more of the following: (a) attends more than one Oregon public school during the school year, (b) enters the Oregon public education system late (i.e., after the first school day in October), (c) exits the Oregon public education system early (i.e., on or before the first school day in May without

- The distance to the nearest 2- or 4-year public higher education institution,[11] as a proxy measure for service and community organization availability (This variable was not used in virtual charter schools clustering)

This variable selection and modeling process is intended to cluster districts by district and community traits – factors largely outside of district control – and not by outcomes. Outcomes should vary within clusters and change over time and the similar district identification should help districts that serve similar student groups in similar settings build from strength and learn from the successes and challenges of districts achieving different outcomes in similar contexts.

Calculation of Similarity and Clustering

Within each Group, the combination of the demographic and community variables is the basis for similarity. A statistical distance (Euclidean distance) was calculated between each district and every other district in the same Group using the selected variables above.

Prior to the statistical distance calculation, all variables were max-min normalized (subtract the minimum, then divide by the range) such that the minimum value becomes 0 and maximum value becomes 1. Before normalization, square root or log transformation was applied to non-normally distributed variables to redistribute them for improved model function. No weights were applied to the variables in the analysis, such that each variable contributed equally to the model.

The Euclidean distance is the distance between two sets of coordinates. In the example equation, letters A to E represent the 5 variables and the subscript number 1 represents District 1 and subscript number 2 represents District 2. The distance between the two districts is calculated as:

$$distance = \sqrt{\begin{matrix}(A_1 - A_2)^2 + \\ (B_1 - B_2)^2 + \\ (C_1 - C_2)^2 + \\ (D_1 - D_2)^2 + \\ (E_1 - E_2)^2\end{matrix}}$$

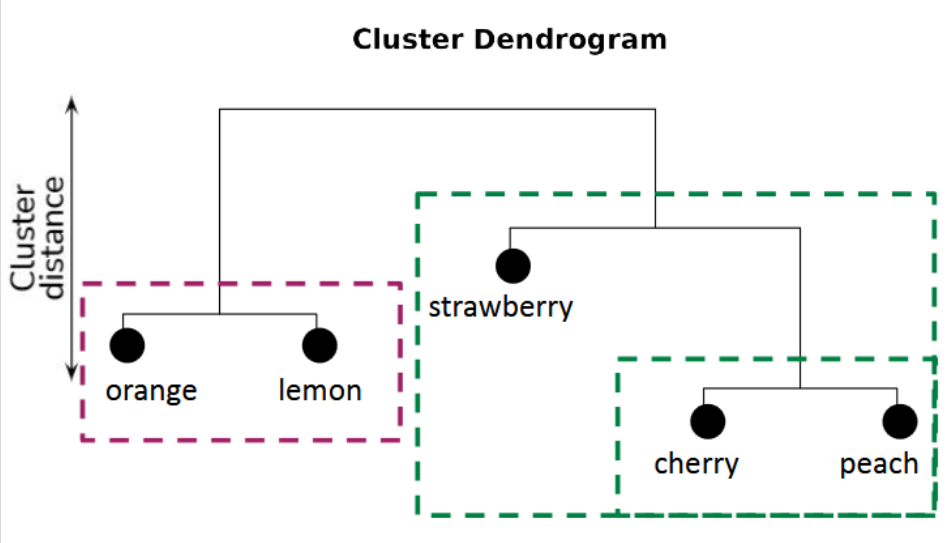
After the statistical distance calculation, hierarchical cluster analysis using Ward's linkage was conducted to identify clusters of similar districts within each Group. In hierarchical clustering, each observation or district starts out as its own individual cluster at the lowest level. At each step, the two closest clusters that are the most statistically similar are merged. This merging

earning a diploma, certificate, etc.), or (d) has significant gaps in enrollment during the school year totaling ten or more consecutive week days. See Student Mobility User Guide for details.

[11] The Euclidean distance (in miles) to the nearest 2- or 4-year public higher education institution, including institutions in Oregon, Washington, Idaho, California and Nevada.

process continues until all the observations in the Group are joined together (see Figure 2 for a detailed example).

Figure 2. Example of Hierarchical Clustering and Cluster Partitioning



The cluster dendrogram above illustrates hierarchical clustering by grouping fruits based on their phylogenetic traits. Initially, each fruit starts out as its own individual cluster at the lowest level (e.g., orange, lemon, strawberry, cherry, peach). At each subsequent step, the two most statistically similar clusters are merged, for example, cherry and peach are joined first, followed by orange and lemon. This merging process continues until all observations are joined together. In the fruit example, strawberry eventually joins the cherry and peach cluster as it is genetically more similar to them than to orange and lemon. The resulting tree-like dendrogram figure visually shows the sequential merging of the fruits. Finally, this dendrogram can be partitioned into broader clusters based on similarities, resulting in two clusters: Cluster 1 (orange and lemon) and Cluster 2 (strawberry, cherry and peach).

Through this clustering technique, districts with the least statistical variation (e.g., the most similar to one another) are grouped together. The number of districts in each cluster is allowed to vary, to maximize within cluster similarity rather than forcing equal sized clusters.

As a result, the number of districts in each cluster may vary, and districts may have similar districts in their cluster that have similar percentages to them on all variables and they may also have similar districts that are nearly identical to them in one variable and quite different from them in another variable (see

Ward's linkage method organizes districts into clusters by minimizing differences within each cluster and maximizing differences between clusters (minimizes the increase in within-cluster variance). Thus, the districts in Cluster 1 are statistically more similar to each other than they are to the districts in Cluster 2.

Table 2 and Table 3). The Appendix at the end of the document provides a summary of the student demographic and community characteristic variable for the final Similar District and Charter Clusters.

The final Similar District Groups and Clusters are published on the ODE Education Accountability Act website under Similar Districts and Charter Schools.

Table 2. Final Similar District Summary Table

Group	Size and Locale	Number of Districts Cluster 1	Number of Districts Cluster 2
1	Large City/Suburb	7	8
2	Medium City/Suburb, Small Suburb	7	7
3	Medium Town/Rural	35	6
4	Small Town	6	15
5	Small Rural	40	41
6	Very Small Town/Rural	16	7

Table 3. Final Summary Table for Charter Schools

Group	Number of Charter Schools Cluster 1	Number of Charter Schools Cluster 2	Number of Charter Schools Cluster 3	Number of Charter Schools Cluster 4
Non-Virtual Charters	26	14	28	12
Virtual Charters	7	11	–	–

Note. Non-Virtual charters were separated into four clusters.

Evaluation of Model Fit

A goodness-of-fit measure called the silhouette score was used to evaluate cluster model fit, with the goal to minimize differences within clusters and maximize differences between

clusters. Districts categorized into the same clusters should be the closest in similarity while districts in different clusters are distinct.

The silhouette score is a calculation that can be performed for each point (district) within the cluster, and measures how similar that point is to other points in its cluster, and how different it is from points not in its cluster. Higher values indicate a better fit within the Cluster for the point in question. The final silhouette score for the similar district clusters ranged from .25 to .67 with an overall average silhouette score of .37, which indicates adequate fit for this type of model.

Other Variables Considered

In the development of the Similar District model, the selection of the student demographic and community attribute variables was informed by past ODE accountability work on similar-school groupings, engagement with the Accountability Reporting Advisory Committee (ARAC), the Assessment & Accountability Technical Advisory Committee (TAC), engagement with schools and districts, and the performance of the variables in the model.

Variables that were suggested and considered during model development included: students with a home language that is not English; count of unique home languages; students with disabilities; students who experienced houselessness; students who are migrant; ever foster students; teacher retention/turnover; teacher experience; and change in student enrollment between 2022-23 to 2024-25.

The inclusion of variables in the final model was grounded in model fit and parsimony. Variables were added into the model that increased model fit and variables that did not increase model fit were left out. Some variables considered were duplicative or highly correlated with existing variables such that their inclusion did not contribute to additional model fit. For example, the home language variables were highly correlated with the percentage of ever English learners; the houseless, migrant and foster student variables were already included in the existing student experiencing poverty[12] variable. The percent of students with disabilities was not included as the rate did not vary substantially between districts. Variables such as teacher retention/turnover, and experience were considered and not included due to the level of influence that district policies may have on staffing. The goal of the clustering process was to identify districts that have similar contexts, while allowing district policy decisions to vary.

[12] These are students who participate in SNAP, participate in TANF, are foster students, homeless students, and/or migratory students.

Appeal Process

Since the initial Group determination was based on district size and geographic locale, some districts could find closer comparators in a different Group than their original assigned Group. For example, a district originally assigned to Group 1 may have closer comparators in Group 2. To account for districts' unique qualities and local context, ODE released a preliminary version of the similar district model results in early December 2025 and districts had the opportunity to appeal based on:

1. If the district and all of its sponsored charter schools wished to be considered together, rather than grouping the charter schools separately.
2. If the district determined that it has closer comparators in a different Group (i.e. a different size and locale classification).

While districts were able to appeal their Group assignment, the Clusters within Groups established through the statistical clustering process was not appealable. For example, a district in Group 3, Cluster 1 cannot appeal to move to Group 3, Cluster 2; however, that same district can appeal to move from Group 3 to Group 4 if desired.

The proposed Group moves were evaluated for statistical soundness using the candidate's proposed and the contextual information provided through the appeal review process to determine the final district classifications. When a Group move improved the model's silhouette score[13], the district was reassigned to the new Group and the clustering process was rerun to reflect the updated Group and Clusters.

The final Similar District Groups and Clusters are published on the ODE Education Accountability Act website under Similar Districts and Charter Schools.

[13] See the section *Evaluation of Model Fit* for details on the silhouette score.

Appendix. Similar District Group and Clusters Summary

The table summarizes the student demographic and community characteristic variables used in the Similar District and Charter Clusters. The values shown are the unweighted overall district averages (e.g., the average percent of Ever English Learners in 2024-25 was 13.5% for districts in Group 1, Cluster 1). The summary table also demonstrates how these demographic variables differ across the clusters as a result of the statistical clustering process, distinguishing districts based on common characteristics (e.g., the average percent of Ever English Learners was 33.5% for districts in Group 1, Cluster 2, which distinguishes them from the districts in Cluster 1). Data source: Oregon Department of Education student data, 2024–25.

Group 1: Large City/Suburb Districts

Student Demographic and District Variables	Group 1, Cluster 1	Group 1, Cluster 2
Average Size	15,655	18,135
Average Students Historically Experienced Disparity	36.6	59.7
Average Ever English Learners	13.5	33.3
Average Students Experiencing Poverty	27.2	38.2
Average Mobility	12.5	16.7
Average Distance to Higher Ed Institution (miles)	2.5	6.4

Group 2: Medium City/Suburb, Small Suburb Districts

Student Demographic and District Variables	Group 2, Cluster 1	Group 2, Cluster 2
Average Size	3,884	4,102
Average Students Historically Experienced Disparity	27.6	52.4
Average Ever English Learners	6.4	26.0
Average Students Experiencing Poverty	20.3	43.7
Average Mobility	10.3	16.4
Average Distance to Higher Ed Institution (miles)	3.4	11.8

Group 3: Medium Town/Rural Districts

Student Demographic and District Variables	Group 3, Cluster 1	Group 3, Cluster 2
Average Size	3,146	3,019
Average Students Historically Experienced Disparity	32.1	71.7
Average Ever English Learners	12.0	46.8
Average Students Experiencing Poverty	36.5	45.6
Average Mobility	14.6	16.0
Average Distance to Higher Ed Institution (miles)	10.5	24.6

Group 4: Small Town Districts

Student Demographic and District Variables	Group 4, Cluster 1	Group 4, Cluster 2
Average Size	1,604	946
Average Students Historically Experienced Disparity	32.1	24.0
Average Ever English Learners	12.3	< 5%
Average Students Experiencing Poverty	33.9	43.1
Average Mobility	12.2	17.6
Average Distance to Higher Ed Institution (miles)	8.4	28.2

Group 5: Small Rural Districts

Student Demographic and District Variables	Group 5, Cluster 1	Group 5, Cluster 2
Average Size	281	606
Average Students Historically Experienced Disparity	16.6	27.2
Average Ever English Learners	< 5%	10.4
Average Students Experiencing Poverty	36.9	31.5
Average Mobility	13.4	11.5
Average Distance to Higher Ed Institution (miles)	36.2	19.1

Group 6: Very Small Districts

Student Demographic and District Variables	Group 6, Cluster 1	Group 6, Cluster 2
Average Size	33	8
Average Students Historically Experienced Disparity	25.2	14.0
Average Ever English Learners	< 5%	< 5%
Average Students Experiencing Poverty	38.8	32.1
Average Mobility	25.6	< 5%
Average Distance to Higher Ed Institution (miles)	65.1	82.7

Non-Virtual Charter

Student Demographic and District Variables	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Average Size	297	275	210	121
Average Students Historically Experienced Disparity	17.9	59.2	26.8	45.1
Average Ever English Learners	< 5%	28.7	< 5%	< 5%
Average Students Experiencing Poverty	21.8	31.1	18.5	54.4
Average Mobility	10.8	9.1	14.0	30.0
Average Distance to Higher Ed Institution (miles)	14.6	6.0	2.4	3.5

Virtual Charter

Student Demographic and District Variables	Cluster 1	Cluster 2
Average Size	1,225	945
Average Students Historically Experienced Disparity	24.1	36.5
Average Ever English Learners	< 5%	10.6
Average Students Experiencing Poverty	25.3	46.5
Average Mobility	25.6	47.1

Note. Average Distance to Higher Ed Institution not used in Virtual Charter clustering