2026-27 School Year
Issue #3

Instructional Data Warehouse News

November 12, 2026

www.nassauboces.org/idw

"Comparing Apples to Apples" Is More Than Just a Phrase

Finding Legitimate Comparisons Takes Work

by Fred Cohen

"Disaggregating the data" seems like an oxymoron, a self-contradictory term. After all, "data" is a plural word and only gains legitimacy as we aggregate more and more data points. Adding data points to a sample size sharpens statistical estimates. After all, isn't that the point of analyzing data in the first place!

But sometimes a large sample size must be broken down into components when each component has distinct characteristics. The simplest example is a school district that has a number of schools with the same grade levels yet very different enrollment patterns. One school may have a unique self-contained special education population, while another school houses the district's bilingual program. In such cases, it is wise to **aggregate the entire district population for some data analyses** but **disaggregate the data for other purposes** because some schools are really not comparable to others and need separate and distinct analyses.



Though the "No Child Left Behind" initiative certainly had its weaknesses, it did make clear that not only did whole populations of students have to succeed, but distinct subgroups of populations also had to succeed for a district to meet federal standards.

Because of this mandate, the Instructional Data Warehouse (IDW) began including these disaggregations into many of our reports to make these comparisons easier. Such disaggregated reports allow districts to compare not only their whole populations, but their unique subgroups populations as well.

Most important is that these reports include county benchmarks so that each school and district can compare itself to similar subgroups countywide making for a fairer comparison to a large sample size (typically 10,000 to15,000 students per grade). The intent is to compare one district's subgroups to similar subgroups countywide.



Upcoming IDW Events

IDW Trainings:

Beyond the Basics: Advanced Reports for Power Users
11/12/26
9:00am to 11:00am

Teacher Interface:
11/18/26
9:00am to11:00am

Grade 3-8 Admin
11/18/26
1:00pm to 3:00pm

Achieving Equitable Outcomes 12/2/26
9:00am to 11:00am

Training time is included in each district's IDW subscription.

To book an in-district training, please call Stephanie Witt 516-608-6623.

NASSAURIC REGIONAL INFORMATION CENTER

Department of Curriculum, Instruction and Technology
Questions? Contact: Stephanie Witt, Supervisor II, Instructional Data Warehouse
Phone: 516-608-6623 • Email: switt1@nasboces.org

The IDW provides a report like the **sample below** for every state assessment from Grade 3-8 ELA and Math to all Regents exams. The reports can be found on the main "Reports" menu under the heading, "Subgroup Analysis by Performance Level." Note the characteristics of the report for this **Grade 4 ELA 2025-26 report**. The subgroup disaggregations are named in the first column. The numbers of students (and percentages) of the whole tested population are listed by subgroup divisions in the second and third columns. Most important is the level of proficiency achieved by each subgroup in the district (last column).

	Total Students	% of Total	Final Numeric Score	Level 1 Count	% at Level 1	Level 2 Count	% at Level 2	Level 3 Count	% at Level 3	Level 4 Count	% at Level 4	L3 + L4	% at L3 + L4
Female	202	46.4%	459.3	16	7.9%	37	18.3%	69	34.2%	80	39.6%	149	73.8%
Male	233	53.6%	458.7	27	11.6%	32	13.7%	79	33.9%	95	40.8%	174	74.7%
American Indian or Alaska Native	1	0.2%	488.0	0	0.0%	0	0.0%	0	0.0%	1	100.0%	1	100.0%
Asian	191	43.9%	461.3	13	6.8%	28	14.7%	65	34.0%	85	44.5%	150	78.5%
Black or African American	16	3.7%	453.5	4	25.0%	2	12.5%	4	25.0%	6	37.5%	10	62.5%
Hispanic or Latino	90	20.7%	452.7	18	20.0%	15	16.7%	27	30.0%	30	33.3%	57	63.3%
Multiracial	16	3.7%	457.9	2	12.5%	3	18.8%	3	18.8%	8	50.0%	11	68.8%
Native Hawaiian/Other Pacific Islander	3	0.7%	457.7	0	0.0%	0	0.0%	3	100.0%	0	0.0%	3	100.0%
White	118	27.1%	460.5	6	5.1%	21	17.8%	46	39.0%	45	38.1%	91	77.1%
Disabled	51	11.7%	433.6	28	54.9%	8	15.7%	6	11.8%	9	17.6%	15	29.4%
Non-Disabled	384	88.3%	462.3	15	3.9%	61	15.9%	142	37.0%	166	43.2%	308	80.2%
Not Low Income	305	70.1%	460.8	24	7.9%	40	13.1%	114	37.4%	127	41.6%	241	79.0%
Poverty - From Low Income Family	130	29.9%	454.6	19	14.6%	29	22.3%	34	26.2%	48	36.9%	82	63.1%
ELL Eligible	11	2.5%	423.5	8	72.7%	2	18.2%	1	9.1%	0	0.0%	1	9.1%
Not ELL Eligible	424	97.5%	459.9	35	8.3%	67	15.8%	147	34.7%	175	41.3%	322	75.9%
Total	435	100.0%	459.0	43	9.9%	69	15.9%	148	34.0%	175	40.2%	323	74.3%

One can see the total student proficiency, as listed at the bottom right of the table, is 74.3%. But also note the circled subgroup division for "Not Low Income" and "Poverty - From Low Income Family." Knowing the devastating learning effects associated with poverty, it is not surprising that the proficiency levels of the two subgroups are quite different. This "Subgroup by Performance Level" IDW report allows us to see the percentage of students in each subgroup (column 3) and the proficiency level also (final column, circled).

The question remains, however: Are the comparisons in proficiency levels fair? Yes, students from low-income families have numerous learning challenges not associated with middle income families, but how do we find appropriate benchmarks for each of these two subgroups as well as benchmarks for all the other subgroups on this report. It's clear that comparing students from different income groups to each other is NOT comparing apples to apples. So whom should we compare each subgroup to?

The IDW has a companion report that makes this possible. The "Subgroup Analysis Report by Performance Level - Regional" provides similar data for all Nassau County test takers making an apples-to-apples comparison legitimate. How do my students from Low-Income families compare to students from Low Income families countywide? That's a fair comparison which can inform our analysis. We can also see how each district's overall population compares to the county average population. For example the district in question has a low-income family population of 29.9%. Is that typical for Nassau County? By using the regional report to compare, each district gains important insights into how their own subgroup enrollments compare.

NASSAURIC REGIONAL INFORMATION CENTER

Department of Curriculum, Instruction and Technology
Questions? Contact: Stephanie Witt, Supervisor II, Instructional Data Warehouse
Phone: 516-608-6623 • Email: switt1@nasboces.org

The report below is Nassau County's Regional Subgroup Performance benchmark report for ELA 4 for 2025-26. The report lists the same subgroups as the district report and notes the percentage of test takers in each subgroup. The report also lists the performance level percentage achieved for each subgroup (last column).

Note that the district's overall performance (bottom right) was 74.3% proficient while the county earned 60.3% proficient Not surprising, the county had a higher percentage of Low-Income students than the district (33.8% compared to 29.9%), perhaps accounting for some of the difference in overall proficiency. Note as well the differences in ELL (or LEP) students. The district had 2.5% ELL; the county had 9.3%. These percentages are, of course, not the sole determining factors in proficiency, but one can imagine the impact special populations have on district resources.

	Total Students	% of Total	Level 1 Count	% at Level 1	Level 2 Count	% at Level 2	Level 3 Count	% at Level 3	Level 4 Count	% at Level 4	L3 + L4	% at L3 + L4
Female	*5,909*	*49.6%*	767	13.0%	1,419	24.0%	1,804	30.5%	1,919	32.5%	*3,723*	*63.0%*
Male	*5,996*	*50.4%*	962	16.0%	1,583	26.4%	1,774	29.6%	1,677	28.0%	*3,451*	*57.6%*
American Indian or Alaska Native	*45*	*0.4%*	7	15.6%	13	28.9%	16	35.6%	9	20.0%	*25*	*55.6%*
Asian	*2,883*	*24.2%*	221	7.7%	549	19.0%	896	31.1%	1,217	42.2%	*2,113*	*73.3%*
Black or African American	*1,141*	*9.6%*	212	18.6%	303	26.6%	351	30.8%	275	24.1%	*626*	*54.9%*
Hispanic or Latino	*3,527*	*29.6%*	911	25.8%	996	28.2%	929	26.3%	691	19.6%	*1,620*	*45.9%*
Multiracial	*350*	*2.9%*	30	8.6%	74	21.1%	116	33.1%	130	37.1%	*246*	*70.3%*
Native Hawaiian/Other Pacific Islander	*18*	*0.2%*	0	0.0%	8	44.4%	7	38.9%	3	16.7%	*10*	*55.6%*
White	*3,941*	*33.1%*	348	8.8%	1,059	26.9%	1,263	32.0%	1,271	32.3%	*2,534*	*64.3%*
Disabled	*1,455*	*12.2%*	630	43.3%	475	32.6%	243	16.7%	107	7.4%	*350*	*24.1%*
Non-Disabled	*10,450*	*87.8%*	1,099	10.5%	2,527	24.2%	3,335	31.9%	3,489	33.4%	*6,824*	*65.3%*
Not Low Income	*7,885*	*66.2%*	741	9.4%	1,836	23.3%	2,517	31.9%	2,791	35.4%	*5,308*	*67.3%*
Poverty - From Low Income Family	*4,020*	*33.8%*	988	24.6%	1,166	29.0%	1,061	26.4%	805	20.0%	*1,866*	*46.4%*
LEP Eligible	*1,112*	*9.3%*	649	58.4%	358	32.2%	87	7.8%	18	1.6%	*105*	*9.4%*
Not LEP Eligible	*10,793*	*90.7%*	1,080	10.0%	2,644	24.5%	3,491	32.3%	3,578	33.2%	*7,069*	*65.5%*
Total	*11,905*	*100.0%*	1,729	14.5%	3,002	25.2%	3,578	30.1%	3,596	30.2%	*7,174*	*60.3%*



As good as these subgroup reports are in making apples to apples comparisons, they are far from perfect and come with a key caveat. **Subgroup reports make it sound like students from identical subgroups, from one district to another, are indeed equal** and therefore allow the data analyst free reign in making comparisons. *Simply put, identically named subgroups may be as **dissimilar** as a granny smith apple is from a macintosh apple.* Imagine this: there are over 4,020 students from low-income families who took the Grade 4 ELA assessment last year. Were they all at the same level of "poverty?" Clearly some barely qualified for the designation while others might be homeless and living at a barely subsistence level. **We can't assume that gradations do not occur within subgroups.**

NASSAURIC REGIONAL INFORMATION CENTER

Department of Curriculum, Instruction and Technology
Questions? Contact: Stephanie Witt, Supervisor II, Instructional Data Warehouse
Phone: 516-608-6623 • Email: switt1@nasboces.org

One way to dramatize the vast differences within a single subgroup is to try to view the subgroup “Asian” as if it were a single homogenous subgroup. There are almost five billion Asians worldwide or almost 60% of the world’s population. Does it make sense to say that my neighbor from New Delhi, as well as my friend from Tokyo and his neighbor from Tibet belong to the same subgroup as another friend from Bangladesh? Subgroup categories make it convenient to narrow aggregations, but we should not feel that we really have met the burden of comparing like populations by simply assigning a subgroup category. We have merely scratched the surface. Let’s always remember the purpose of our analyses and honor obligations to only compare like things.

We must be more perceptive than the limited subgroup categories our IDW reports include. The insight and experience of knowledgeable data analysts are still needed. Artificial intelligence agents may someday replace the judgement of a wise data analyst, but, as of this newsletter’s publication date, that day has not yet arrived.

As always, IDW personnel are available to provide telephone assistance as well as virtual and in-district training for any district seeking to learn more about this newsletter, or about other best data practices. To schedule a training session or ask any IDW-related questions, please call Stephanie Witt at (516) 608-6623, Fred Cohen at (516) 608-6640, Barbara Dwyer at (516) 608-6695, Dr. Wanda Toledo at (516) 608-6648, Dr. David Erlanger at (516) 832-2786, or Joseph Wiener at (516) 832-2708.

What’s New in the IDW?

- **Check out the new “Favorites” tab as featured in the previous issue.**
- **This is the ideal time to use the IDW to review results from all Grade 3 through 8 assessments and the June Regents exams as well.**
- **The full suite of reports are available for all 2025-26 assessments**

NASSAURIC REGIONAL INFORMATION CENTER

Department of Curriculum, Instruction and Technology
Questions? Contact: Stephanie Witt, Supervisor II, Instructional Data Warehouse
Phone: 516-608-6623 • Email: switt1@nasboces.org