Aggregating data to draw meaningful conclusions

GACIP EGAD Workshop December 2018

Brian Frank, Jake Kaupp, Nerissa Mulligan

Session focus

Comparing merits of different approaches to aggregating data, going:

From: Task-level scores assigned to individual students (which most of us are doing)

To: Demonstrating that a student cohort possesses a graduate attribute

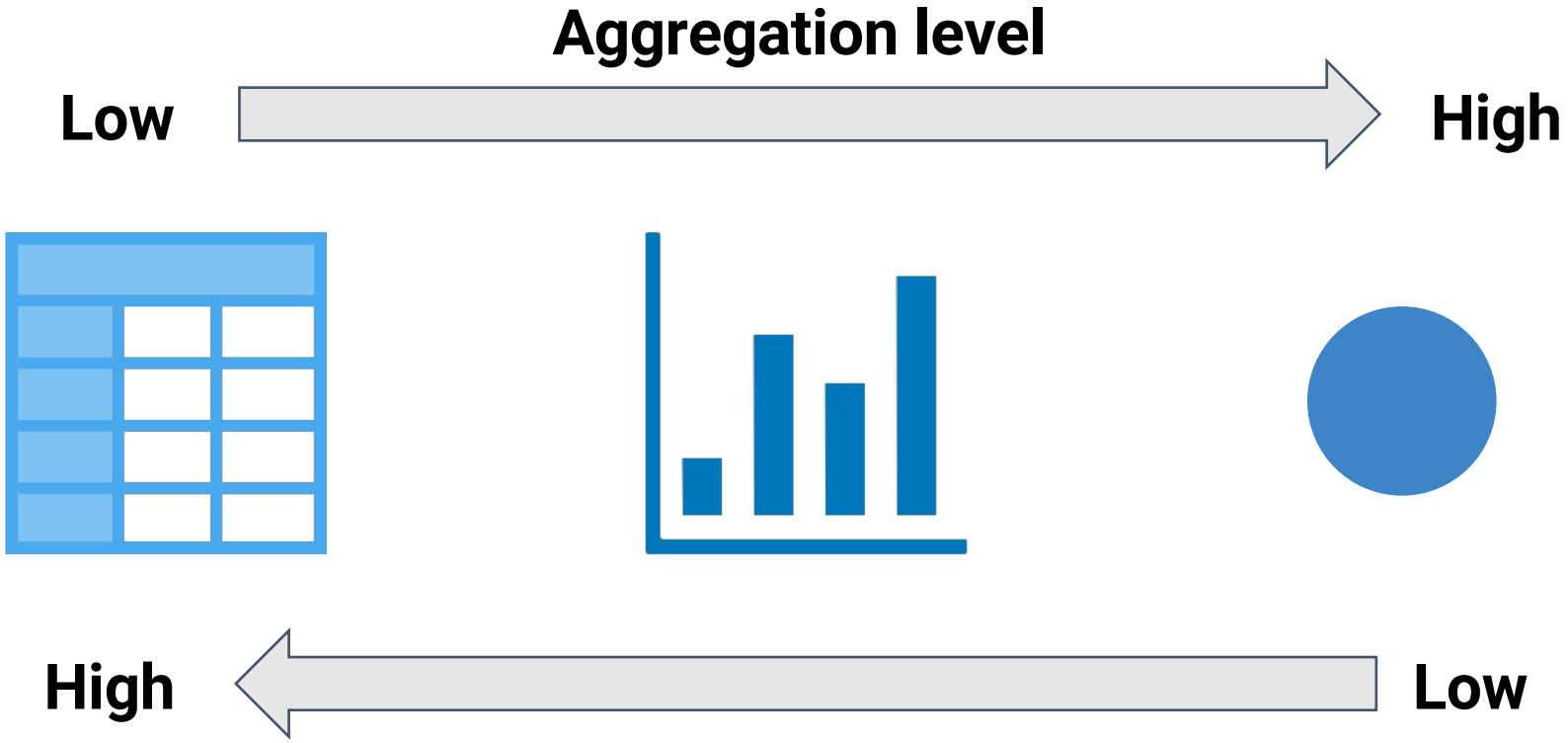
Slides will be available afterwards
We will circulate key elements arising from the discussion.

As required by:

3.1.5 Assessment results: At least one set of assessment results must be obtained for all twelve attributes over a cycle of six years of less. The results should provide clear evidence that graduates of a program possess the above list of attributes.

3.2.1 Improvement process: There must be processes in place that demonstrate that program outcomes are being assessed in the context of the graduate attributes, and that the results are validated, analyzed and applied to the further development of the program.

When you **aggregate data**, you replace **groups of observations** with **summary statistics** based on those observations.


Aggregation level
Low
High
High
Low
Amount of information conveyed

```
group_by(Student ID, Course, Attribute, Indicator)
```

Student ID	Program	Year of Study	Course	Attribute	Indicator	Assessment	Score
A	ENGR	1	ENGR 101	KB	ENGR-KB-1	Midterm	3
A	ENGR	1	ENGR 101	KB	ENGR-KB-1	Final	5
B	ENGR	1	ENGR 101	KB	ENGR-KB-1	Midterm	4
B	ENGR	1	ENGR 101	KB	ENGR-KB-1	Final	4
C	ENGR	1	ENGR 101	KB	ENGR-KB-1	Midterm	4
C	ENGR	1	ENGR 101	KB	ENGR-KB-1	Final	1

Raw Data

```
summarize(Score = mean(Score)
```

Student ID	Course	Attribute	Indicator	Score
A	ENGR 101	KB	ENGR-KB-1	4
B	ENGR 101	KB	ENGR-KB-1	4
C	ENGR 101	KB	ENGR-KB-1	2.5

Distribution

```
group_by(Course, Attribute, Indicator)
```

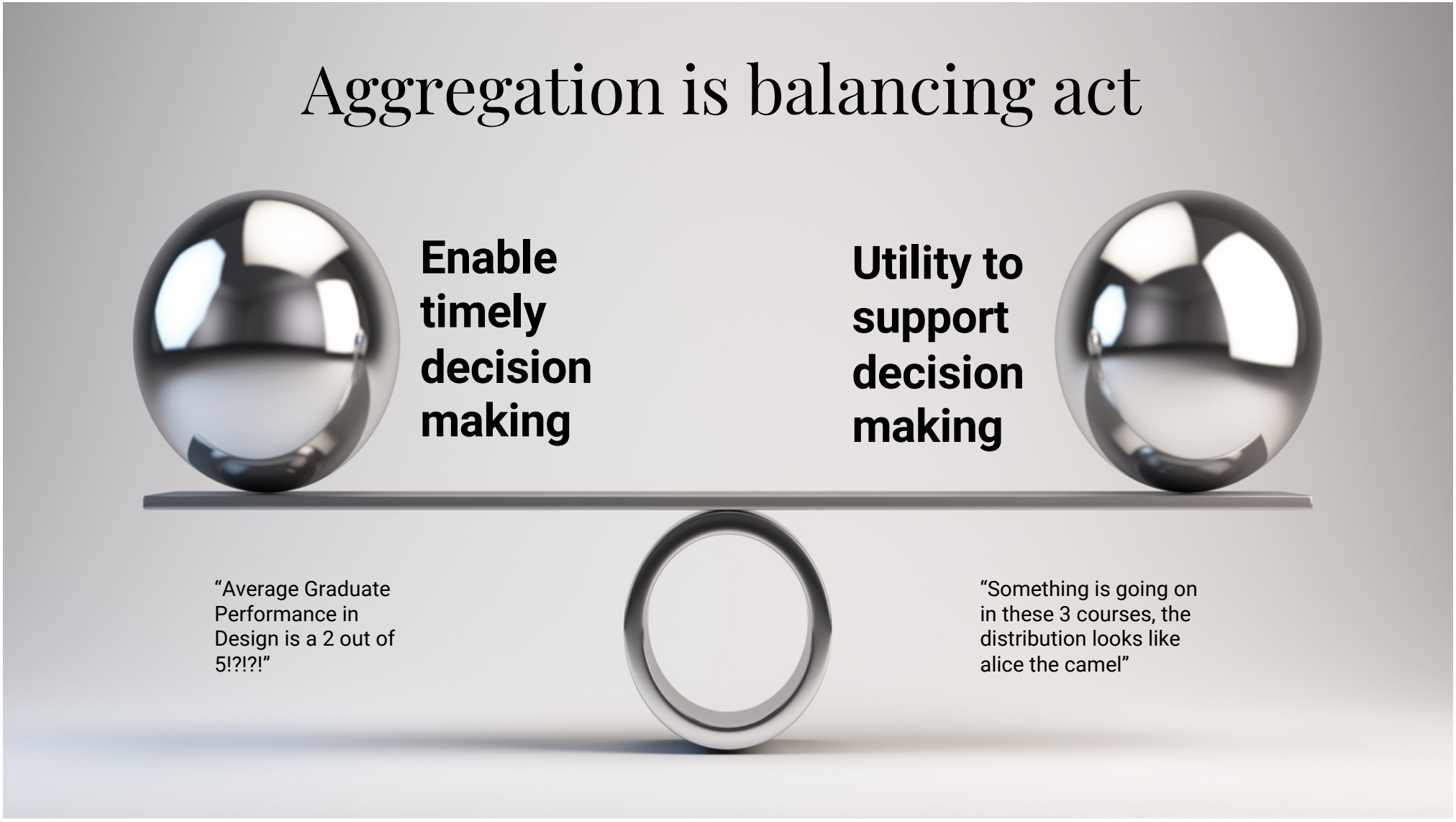
Student ID	Course	Attribute	Indicator	Score
A	ENGR 101	KB	ENGR-KB-1	4
B	ENGR 101	KB	ENGR-KB-1	4
C	ENGR 101	KB	ENGR-KB-1	2.5

```
summarize(Score = mean(Score)
```

Course	Attribute	Indicator	Score
ENGR 101	KB	ENGR-KB-1	3.5

Aggregate

Aggregation is balancing act



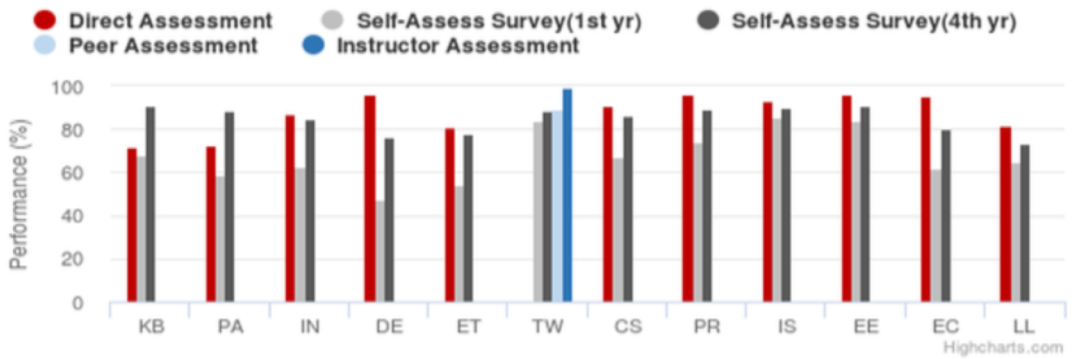
Let's start with assuming that:

1. The data collection process is encompassing all attributes and engaging instructors
2. There is a way to gauge how well instructors trust the underlying data
3. Indicators are accepted and taken to represent the range of key aspects of each attribute
4. The data is relatively stable over time
5. There are multiple measures of each indicator at roughly comparable times, allowing triangulation

Basically, assume the GA/CI process is working!

Compare aggregation approaches by factors:

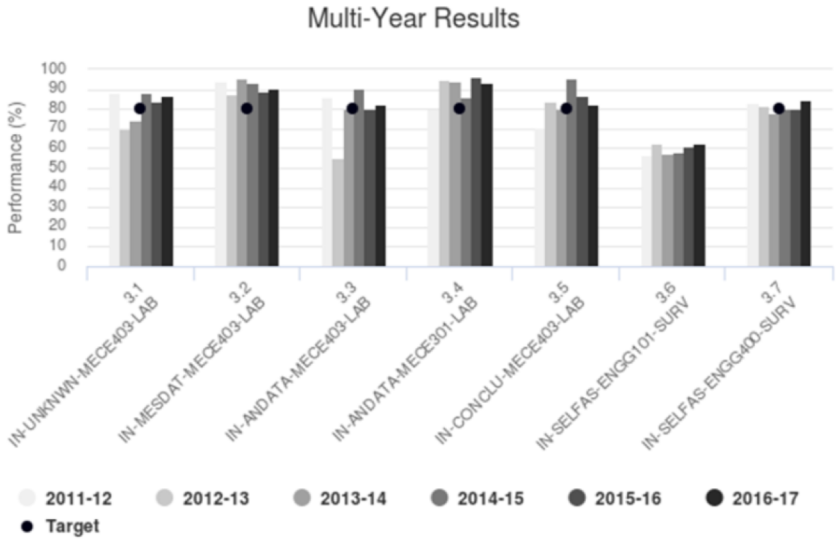
Factor	Possible options
Aggregation target	• single value (e.g. Design = 3.6/5) • distribution of performance, (e.g. histogram of student performance) • qualitative description (textual based analysis of results)
Aggregation level	• up to attribute (e.g. Design) • up to indicator within each attribute (e.g. “Problem definition”) • up to task within indicator within attribute (e.g. “Capstone design report”)
Differentiation factors	• Year of Program (Year 1 to 4) • IDA level (Introduce, Developed, Applied) • Program option (e.g. biomechanics vs. materials) • Summative vs. Formative • Assessment type (e.g. final report, exam, lab simulation, portfolio) • Student groups (first in family, racialized, Indigenous)
Reliability measure	• Qualitative rating by instructors (e.g. text comment by each instructor) • Quantitative rating by instructors (e.g. graded as “highly trustworthy”=4/4) • Correlation between tasks (e.g. correlation between three measures of “problem definition”) • Correlation between years (e.g. correlation between scores in 2016, 2017, and 2018)



#	KB		PA		IN		DE		ET		TW		CS		PR		IS		EE		EC		LL	
	Val	Targ	Val	Targ	Val	Targ	Val	Targ	Val	Targ	Val	Targ	Val	Targ	Val	Targ	Val	Targ	Val	Targ	Val	Targ	Val	Targ
1	87.3	80	82.6	80	96.7	80	100.0	80	88.2	80	97.2	-	91.6	80	95.7	80	89.9	80	95.0	80	92.8	80	73.7	-
2	75.9	-	67.0	-	90.5	80	87.2	80	74.8	80	96.4	80	98.5	80	96.7	80	99.8	80	84.7	-	57.5	-	76.1	80
3	93.9	80	93.5	80	82.1	80	97.4	80	68.6	80	70.8	-	82.8	80	95.5	80	89.2	80	93.9	80	76.5	80	70.0	-
4	82.9	80	63.8	80	93.1	80	97.4	80	75.9	80	78.8	80	84.7	80	74.2	-	100.0	80	93.7	80	94.4	80	73.5	80
5	80.5	80	52.6	-	92.1	80	100.0	80	97.1	80	88.7	80	87.4	80	89.0	80	86.6	80		80	100.0	80	84.4	80
6	73.7	-	85.1	80	62.5	-	96.9	80	54.3	-	99.1	80	100.0	80			85.5	-	99.6	80	66.8	-	43.3	-
7	87.5	80	64.8	80	94.7	80	100.0	80	78.2	80	94.3	80	66.8	-			90.4	80	83.1	-	83.5	80	55.4	80
8	80.7	80	60.2	-			88.8	80			99.1	80	86.4	80					88.2	80			70.9	80
9	73.9	-	89.6	80			100.0	80			82.0	80											89.3	80
10	96.4	80	77.9	80			47.6	-			99.1	80											73.9	-
11	48.0	-	54.1	-			76.4	80			91.8	80											88.1	80
12	86.7	80	87.0	80							93.5	-												
13	56.4	80									96.8	80												
14	68.3	80									75.5	-												
15	74.6	80									82.5	80												
16	65.8	80																						
17	80.7	80																						
18	58.1	80																						

Target met or exceeded
Performance below target
Missing data
No target

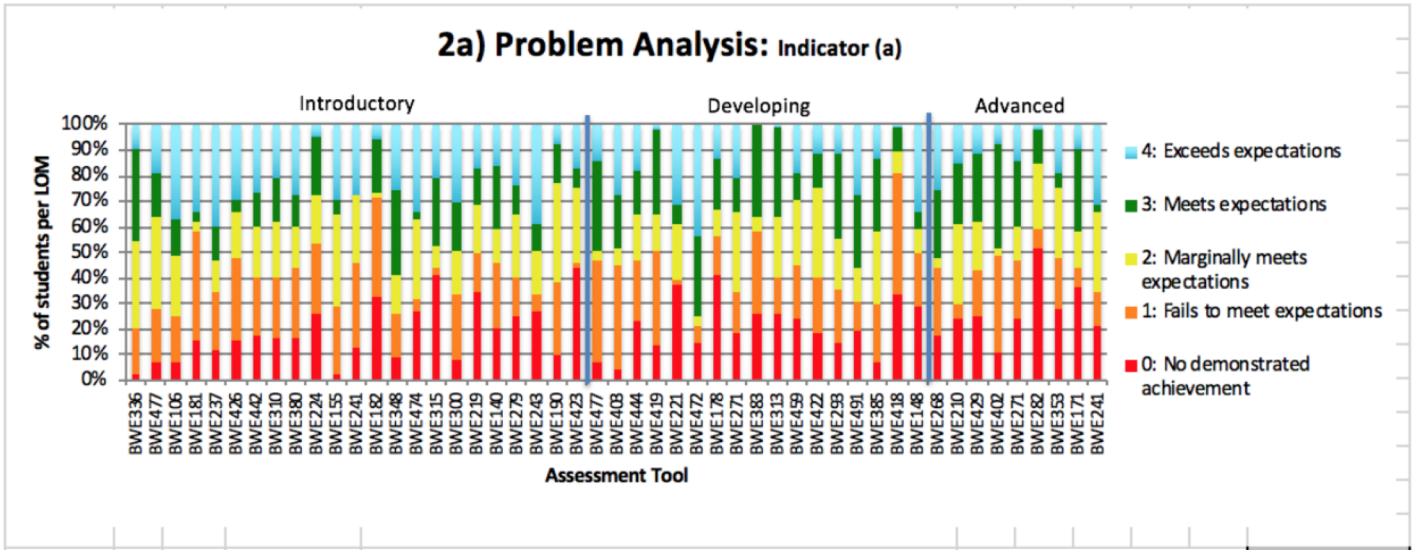
Factor	Approach
Aggregation target	Single values: Mean and % meeting target
Aggregation level	Attribute
Differentiation factors	Direct/peer/self assessment
Reliability measure	



NO.	MEASURE	CURRENT YEAR				PERFORMANCE (%)						Targ
		Uns.	Dev.	Sat.	Exc.	2016-17	2015-16	2014-15	2013-14	2012-13	2011-12	
3.1	IN-UNKNWN-MECE403-LAB	9	17	39	130	86.7	83.8	88.0	73.9	70.1	88.2	80
3.2	IN-MESDAT-MECE403-LAB	7	12	43	134	90.3	89.0	93.5	95.8	87.4	93.9	80
3.3	IN-ANDATA-MECE403-LAB	5	30	107	54	82.1	80.4	90.2	80.0	55.2	86.3	80
3.4	IN-ANDATA-MECE301-LAB	5	14	60	195	93.1	96.6	86.1	94.2	95.2	80.0	80
3.5	IN-CONCLU-MECE403-LAB	5	30	72	89	82.1	86.9	95.7	80.0	83.9	70.6	80
3.6	IN-SELFAS-ENGG101-SURV	30	247	315	147	62.5	60.9	57.9	57.7	62.5	56.3	
3.7	IN-SELFAS-ENGG400-SURV	8	118	412	287	84.7	80.4	80.2	78.1	81.7	83.2	80

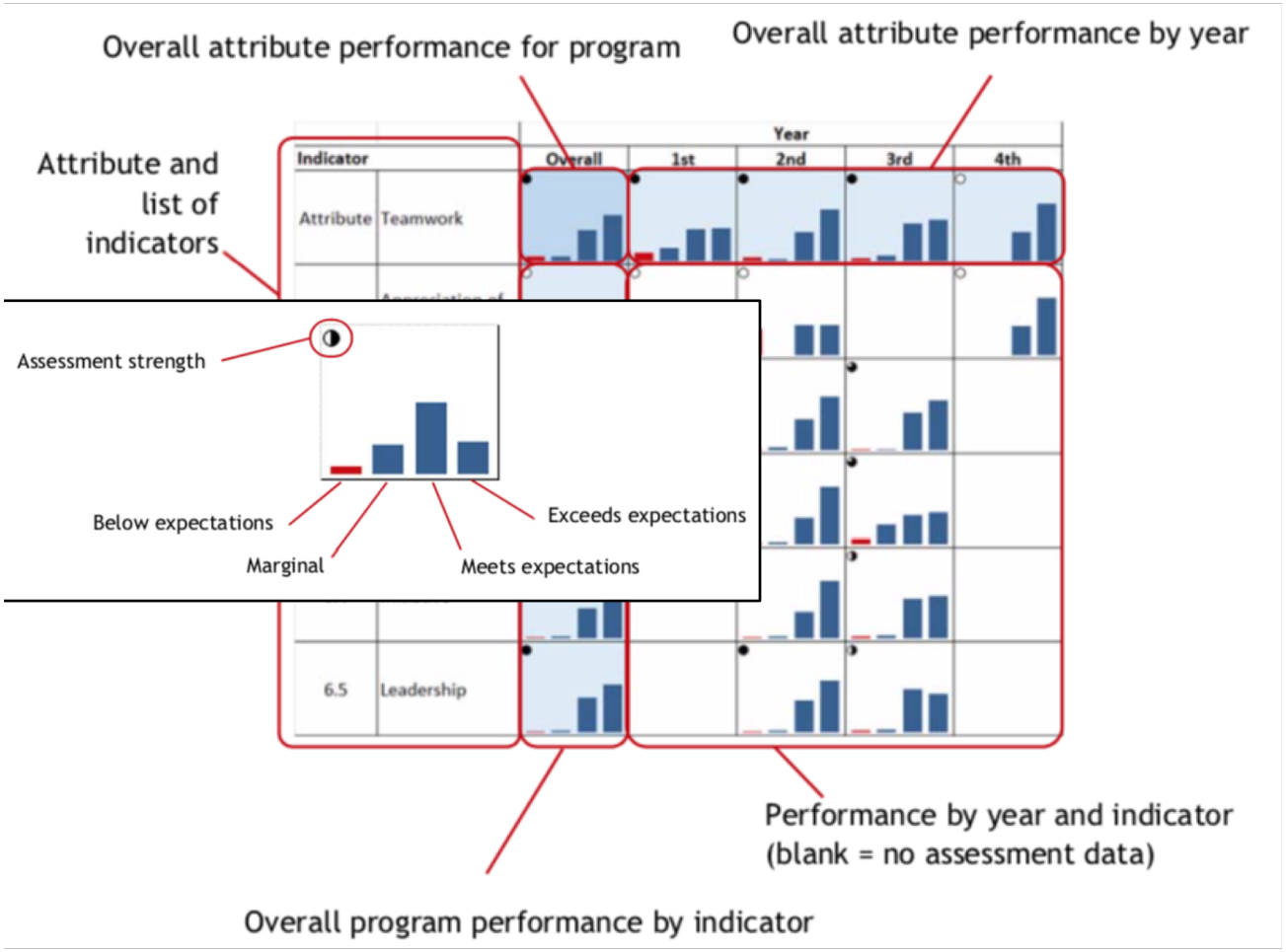
Target met or exceeded
Performance below target
Missing data
No target

Factor	Approach
Aggregation target	Single values: Mean and % meeting target
Aggregation level	Indicator
Differentiation factors	Year
Reliability measure	

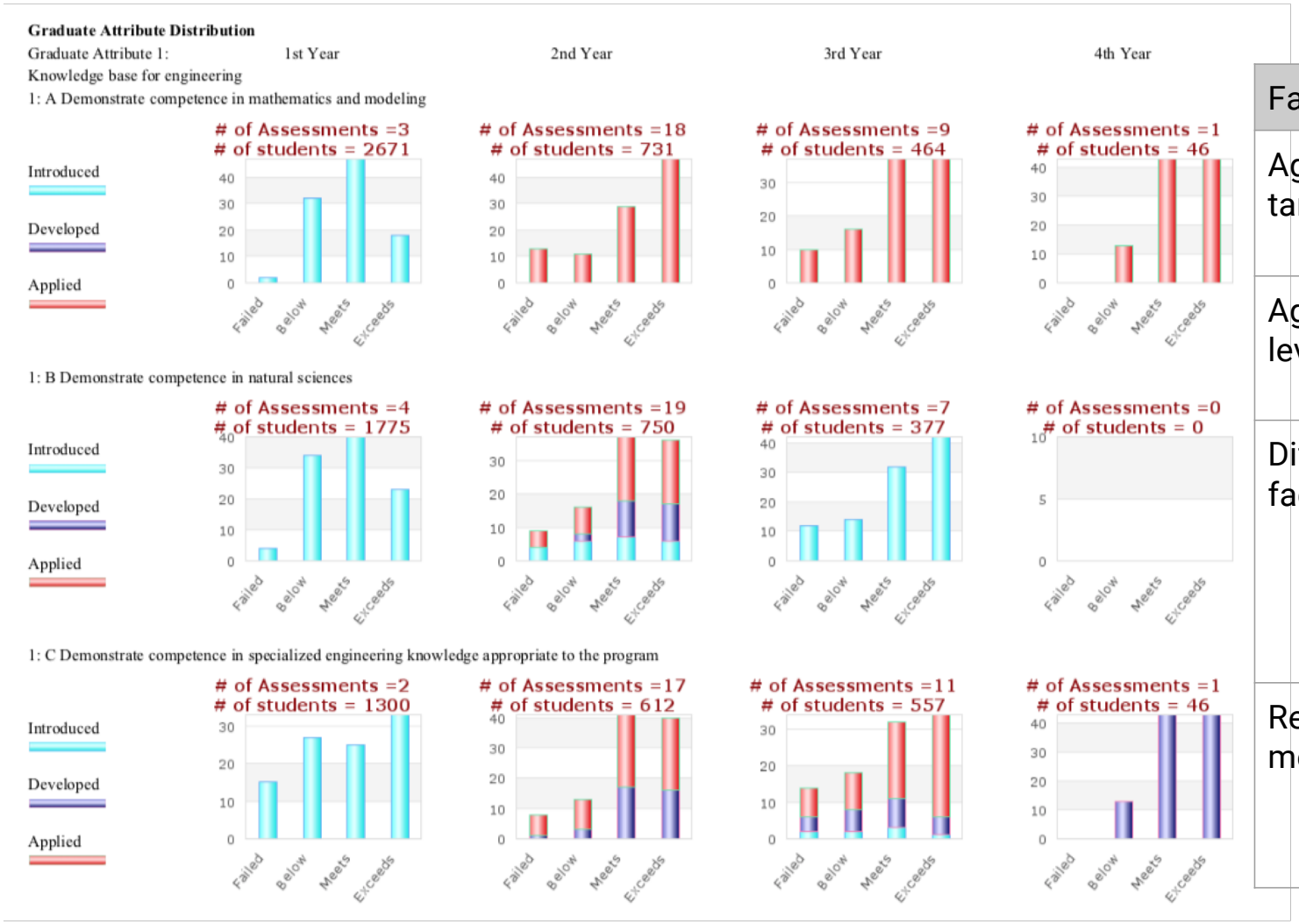


Tool	Ind	Level	Assessor	Question or course learning outcome	# of students at LOM...					% of students over threshold
					0	1	2	3	4	
BWE336	a	I	1. Instructor	CLO #5 (Awesome assessment method #5)	7	47	88	93	25	45%
BWE477	a	I	1. Instructor	CLO #4 (Awesome assessment method #5)	18	54	93	44	50	36%
BWE106	a	I	1. Instructor	CLO #11 (Awesome assessment method #7)	14	35	45	27	71	51%
BWE181	a	I	1. Instructor	CLO #11 (Awesome assessment method #6)	27	75	8	6	60	38%
BWE237	a	I	1. Instructor	CLO #11 (Awesome assessment method #1)	25	46	26	28	83	53%
BWE426	a	I	1. Instructor	CLO #7 (Awesome assessment method #8)	43	91	49	15	82	35%
BWE442	a	I	1. Instructor	CLO #6 (Awesome assessment method #1)	37	48	44	27	57	39%
BWE310	a	I	1. Instructor	CLO #3 (Awesome assessment method #2)	61	87	81	65	77	38%
BWE380	a	I	1. Instructor	CLO #8 (Awesome assessment method #1)	37	63	35	29	61	40%
BWE224	a	I	1. Instructor	CLO #2 (Awesome assessment method #6)	95	99	72	80	19	27%
BWE155	a	I	1. Instructor	CLO #11 (Awesome assessment method #6)	3	41	54	8	44	35%
BWE241	a	I	1. Instructor	CLO #6 (Awesome assessment method #3)	35	93	72	1	75	28%
BWE182	a	I	1. Instructor	CLO #10 (Awesome assessment method #3)	77	89	6	47	14	26%

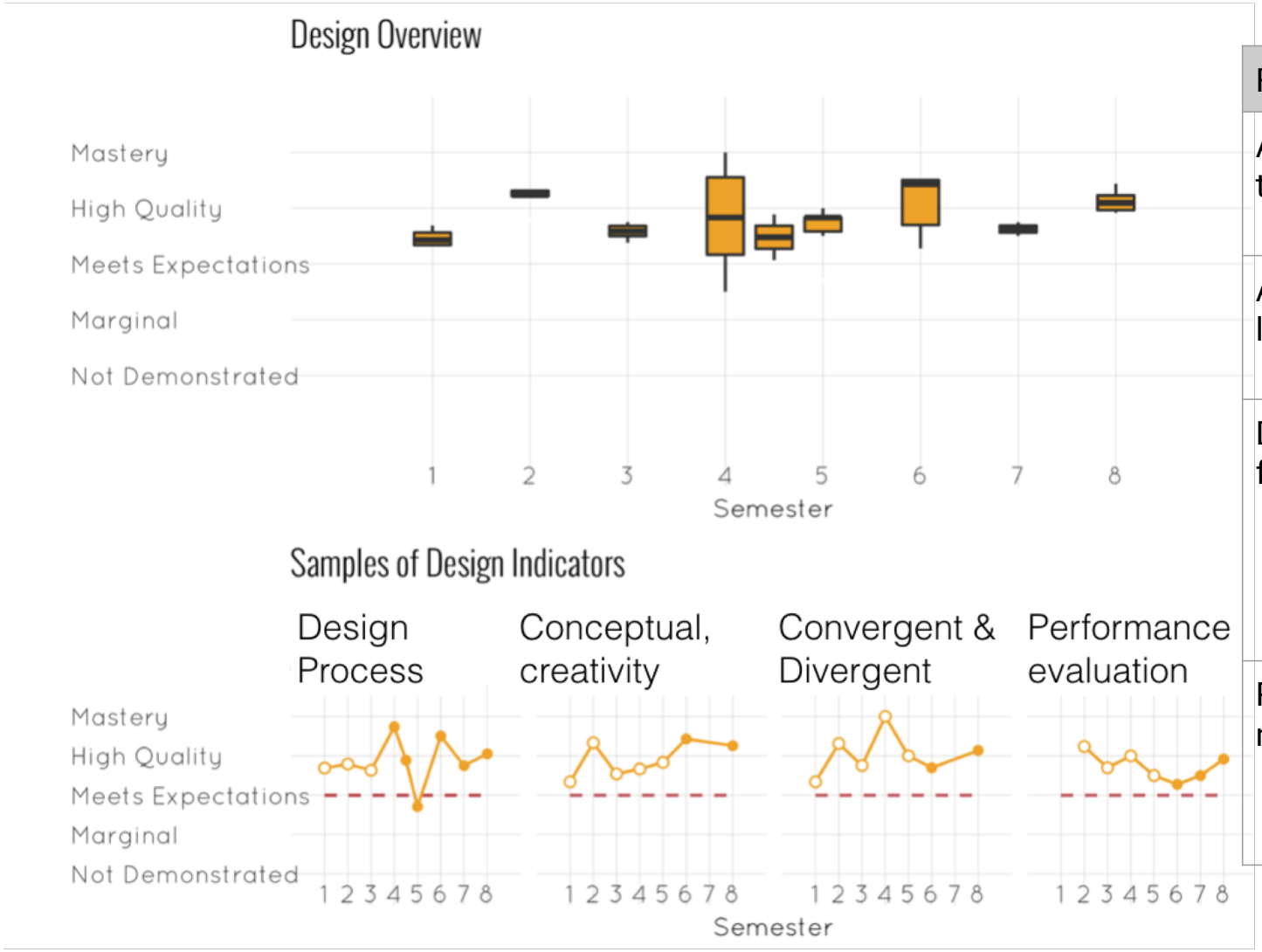
Factor	Approach
Aggregation target	Frequency distribution of performance
Aggregation level	Assessment measure
Differentiation factors	IDA
Reliability measure	



Factor	Approach
Aggregation target	Frequency distribution: of performance
Aggregation level	Indicator
Differentiation factors	Year
Reliability measure	Quantitative rating by instructor



Factor	Approach
Aggregation target	Frequency distribution: of performance
Aggregation level	Attribute
Differentiation factors	IDA, Year
Reliability measure	# students, # assessments



Factor	Approach
Aggregation target	Distribution: of performance
Aggregation level	Indicator, attribute
Differentiation factors	Semesters, time
Reliability measure	Qualitative by instructor

Task 1: Connect the factors to current institutional approaches

Within your table group:

1. Identify how well the four factors describe the institution's approaches represented within the group. Is there a key factor that is not captured by that list of four? If so add it to your list of factors to consider

2. Briefly describe each institution's approach using the factors

Task 1: Connect the factors to current institutional approaches

Report out:

1. Any key factors emerge that were not captured by the original list of four?
2. How well are institutional approaches captured by the factors? Are there two extremely different approaches between institutions represented at your table?

Task 2: Consider what aggregation means to key stakeholders

As a table group, identify what key stakeholders are looking for from aggregation:

- Course instructors
- Department administration
- Faculty administration
- CEAB visiting team

Task 2: Consider what aggregation means to key stakeholders

<u>Report out</u>: what are key stakeholders are looking for from aggregation?

- Course instructors
- Department committees, staff, and administration (which may also consult with broader stakeholders)
- Faculty committees, staff, and administration (which may also consult with broader stakeholders)
- CEAB visiting team

Task 3: What mix of aggregation approaches would meet the collective needs of stakeholders?

Consider needs of key stakeholders:

- Course instructors
- Department committees, staff, and administration (which may also consult with broader stakeholders)
- Faculty committees, staff, and administration (which may also consult with broader stakeholders)
- CEAB visiting team

Consider factors in aggregation:

1. Aggregation target: single value, distribution of performance, or qualitative description
2. Aggregated level: up to attribute, up to indicator within each attribute, up to task within indicator within attribute
3. Differentiation factor: differentiate by year, IDA level, program, student sub-group, student
4. Reliability measure: qualitative or quantitative instructor rating, correlation between tasks or years

Task 3: Report Out: What mix of aggregation approaches would meet the collective needs of stakeholders?

Consider needs of key stakeholders:

- Course instructors
- Department committees, staff, and administration (which may also consult with broader stakeholders)
- Faculty committees, staff, and administration (which may also consult with broader stakeholders)
- CEAB visiting team

Consider factors in aggregation

1. Aggregation target: single value, distribution of performance, or qualitative description
2. Aggregated level: up to attribute, up to indicator within each attribute, up to task within indicator within attribute
3. Differentiation factor: differentiate by year, IDA level, program, student sub-group, student
4. Reliability measure: qualitative or quantitative instructor rating, correlation between tasks or years