Illustrating the SASI Framework Video
Transcript with Chinese Translation

In this video, I'm gonna describe and illustrate a framework that we developed that can help us in supporting students' approaches to statistical investigations, or the SASI framework for short.
在这个视频中，我将描述和说明一个我们开发的框架，这个框架可以帮助我们在学生参与统计调查活动时提供支持，我们把它建成为SASI（Supporting students' Approaches to Statistical Investigations）框架.

We developed this framework from the Guidelines for Assessment and Instruction in Statistics Education at the K-12 Level that was developed in 2007.
根据2007年出版的K-12阶段《统计教育测评和教学指南》，我们开发了这个框架。

We further adapted this framework and incorporated more recent research on students' reasoning and the ways that students can be productive in the thinking and the habits of mind that they develop.
我们进一步调整了这个框架，并且结合了最新的有关学生推理与学生思维习惯形成方式的研究。

In this framework, you'll notice that the four phases of a statistical investigation are at the center.
在该框架中，你将注意到处于中心位置的统计调查活动的四个阶段。

You pose a question, you collect your data, you analyze your data and you interpret results.
提出问题，收集数据，分析数据以及解释结果。

We also have the statistical habits of mind located at the outside in those little mind bubbles that you see.
在你没有看到的地方，我们也提供了一些统计思维习惯。

Here, we're trying to signify the kind of productive statistical habits that can be useful at each phase in an investigation.
在这里，我们尝试阐明在调查活动各阶段都有用的富有成效的统计习惯。

Then you see these gradations in color, these help us organize and think about three levels of statistical sophistication, Level A, B, and C.
然后会出现一些渐变的颜色，这有助于我们划分和思考统计思维的三个水平。

Now, these levels don't necessarily relate to elementary, middle and high school or grades or ages.
这三个水平的划分不一定和学段、年级或年龄相关。

A student working at any age, if it's their first introduction to statistics is going to need those experiences at Level A.
任何年龄的学生，如果是第一次接触统计学，他们所需的是水平1的经验。

Now, of course, ideally we're gonna have those experiences begin in the primary or elementary grades and allow many extended opportunities for students to develop their sophistication as they get older and as they have more opportunities to engage in statistics.

当然，理想情况下我们希望学生在小学获取这些经验，这样他们就可以随着年龄的增长得到养成思维的拓展的机会，并且有更多的机会参与统计。

To illustrate this framework, we're gonna use an example of a context that I have seen used in middle school through college introductory statistics.
为了说明这个框架，我们将使用是一个从初中到大学统计入门中使用的一个情景的例子。

Consider typical body measurements that we may be interested in, such as, "How tall are we?", "How wide is our arm span?", "What is our shoe size?"
考虑一下我们感兴趣的有关身体数据的问题，例如"我们多高？"，"我们的臂展是多少？"，"我们的鞋码是多少？"

If we wanted students to think about characterizing students at their school in relationship to these measurements, how might they approach it?
如果我们希望学生思考在他们的学校中描述与这些相关的学生的特征的测量，他们如何完成它？

Well, in the posing question stage, some students may start by asking questions such as, who has the biggest arm span, does the tallest student have the biggest shoe size?
好吧，在提出问题阶段，一些学生可能开始提出"谁的臂展最大""最高的学生是否鞋码也是最大的"之类的问题。

Well, these questions are not statistical in nature, they look for special cases.
然而这些问题本质上并不是统计问题，他们只是找一些特殊的例子。

These questions don't think about the group of students, and they aren't really anticipating variability.
这些问题没有把学生当作一个整体去思考，也没有真正地考虑到变化性。

So we need to be able to help students move beyond looking for special cases that have a single answer, and think about ways of asking questions to characterize groups, and to anticipate the variability.
所以我们应当能够帮助学生不只是回答单一答案的特殊案例，也要思考能够界定整体的问题，并且预测出其中的变化性。

For example, I could ask the students to consider the student with the largest arm span and how they compare with the rest of the students in the school.
例如，我可以要求学生思考具有最大臂展的学生并且和学校中其余学生做比较。

Do they anticipate that there are many students that have the same arm span?
他们能否预测出最大臂展的学生有多少？

Or do they think that our arm spans may be different, or maybe we all have similar arm spans.
他们是否想到我们的臂展有可能不同，或者我们有相同的臂展。

These types of questions could help a student start to consider group attributes and anticipate variability.
这些类型的问题可以帮助学生开始思考整体的属性并且预测变化性。

Let's consider a few other questions that students might pose.
让我们来想一些其它学生可能提出的问题。

What is a typical height of students in our class?
我们班学生身高的典型值是多少？

Do boys and girls have similar arm spans?
男女生是否具有相似的臂展？

Do middle school boys tend to be taller than girls?
中学男生是否趋向于比女生高？

These are good statistical questions that consider all the data values.
这些事考虑到所有数据值的好问题。

Two of these questions are explicitly seeking to compare measures from the boys to the measures from the girls.
其中两个问题明确要求将男孩的测量值和女孩的测量值进行比较。

Let's consider how these questions could be categorized into the different levels.
让我们考虑一下这些问题的不同水平划分。

For the first one, by asking the question to make a claim about height of students in a single class without seeking any generalization to a larger group, we would consider this a Level A.
对于第一个问题，通过提出问题对一个班级学生身高做出断言，没有对推广到更大的集体中，我们认为这处于水平1。

The second question is seeking to compare and account for variability that may exist between two groups, the boys and the girls.
第二个问题群体寻求比较和解释可能存在于男生和女生两个群体间的变化性。

Thus, this question is a Level B that can engage students in beginning to think about how generalization beyond the data may be possible.
因此，这个问题处于水平2，它使学生开始思考所给数据之外的一般化成为可能。

In the third question, the use of the word tend indicates that there's a goal to seek generalization beyond the sample data and it hints of the uncertainty in any claim that may be made.
第三个问题中，使用"趋向"一词表明，问题的目标是在样本数据之外寻求一般化，并暗示了可能出现的任何推断中的不确定性。

This question we would consider as a Level C question.
我们将认为这是水平3 的问题。

So, let's go on to discuss different levels of sophistication for collecting data.
现在，让我们继续讨论收据数据的不同水平。

If the original aim of the task is to characterize middle school students in your school with regards to their height, shoe size, and arm span, students may suggest different types of data collection.

如果该任务的最初目的是描述学校中的学生的身高，鞋子尺码和臂展，学生可能会建议不同类型的数据收集。

Let's consider a conversation among a group of students.
让我们思考一组学生间的对话。

The first girl says, "Well, maybe we can survey our entire school, it is the best way".
第一个女生说："好吧，也许我们可以调查整个学校，这个最好的方法。"

And then a boy responds, "No way, that would take too much effort to survey all our school. Maybe we can randomly select students from each grade!"
然后男孩回应到："不可能，调查我们学校太费力了。也许我们应该从每个年级随机抽取学生。"

And then, the girl says, "Yeah, that is a good idea, but maybe we can just collect data in our class. Maybe our class is not that representative of our whole school".
之后这个女生又说到："是的，这是个好主意，但是也许我们可以仅仅收集我们班级的数据。也许我们班不是我们整个学校的代表。"

This short conversation illustrates the different levels of sophistication regarding the collecting data phase.
这个简短的对话说明了收集数据阶段学生呈现出的不同水平。

For the first girl, she's gonna discuss the census of the whole school.
第一个女孩，她讨论了对于整个学校的普查。

So, there's no need to generalize.
所以，没有概括的必要。

All the data would be available to her.
所有的数据对她来说是可得的。

This plan can be considered as a Level A in the SASI framework.
这个计划可以认为是SASI框架中的水平1.

Collecting data from everyone that you are interested in.
收集你感兴趣的每个人的数据。

In contrast, for the boy, his idea about randomly selecting and making generalizations from the sample to the population makes this plan to be a Level C, but then in the second girl, starts reasoning in a level that we would consider a Level B.
相反的，男孩考虑到随机选择和从样本中得出结论并推断到整体，这个是水平2的计划，然而在第二个女孩那里，我们可以认为她是在水平2进行推理的。

When she starts to acknowledge the possibility of generalization from the class to the school, she's also recognizing though that there could be a danger in generalizing to the whole school based on a single class of students that are all in the same grade.
当她开始承认从班级到学校一般化的可能性后，她也认识到基于同一个年级选取的班级所得到的结论推广到整个学校是可能是有缺陷的。

We're gonna narrow in on a particular example of students ages 13 to 14 who are working on this task.
我们将针对13至14岁正在从事这项任务的的学生的特定范例进行研究

The teacher made a purposeful decision that they were only gonna collect data from their class.
教师做出了学生只能从班里收集数据的决定。

So, they came up with a measurement system and directions and method that they would be using so that everyone was going to measure the height in centimeters, using meter sticks, and a method of being able to do that.
因此，他们提出了一个将要运用测量系统以及方向和方法，每个人将用米尺测量高度，单位是厘米。

They were going to measure arm span, again, using meter sticks.
然后继续用米尺测量臂展。

And then, everyone was going to record their shoe size by taking off their shoe and recording the size of the shoe that they had on that day.
之后，每个人将记录他们当天所穿鞋的鞋码。

So, let's take a look at the ways that some students might be analyzing their data that they have collected.
让我们来看一下一些学生如何分析他们所收集到的数据。
Students working at Level A may make a dot plot of the heights of students in the class.
水平1的学生可能会根据班级学生的高度做一个点状图。

When looking at this graph they observe that there's one person that seems shorter than all the rest, but that many seem to be around 160-166 centimeters.
当看到点图的时候他们观察到有一个人似乎比其余的人矮，很多人似乎处于160到166厘米间。

In this way, these students used an appropriate graph for the data, and they compared a special case to the group, and they were informally describing a cluster of cases.
这样，这些学生对于数据使用了合适的图，他们比较和整体和特例，非正式地描述了例子的集合。

They may go on to compute a mean or median to help in their description of typical heights.
他们也许会继续计算平均数或中数来帮助他们描述典型身高。

Students working at Level B could start to compare the distributions of male and females in the dot plots and box plots and make some observations.
位于水平2的学生可能开始比较点图中男性和女性的分布并进行观察。

They may say things such as, "I think boys tend to be taller than girls, though many boys and girls are similar heights, and there's one boy much shorter than all the rest, I'm not sure that measurement is real".
他们可能会说这些事情："我认为男孩趋向于比女孩高，即使很多男孩和女孩有着相似的身高，这里有一个男孩的身高明显比其余人的更矮，我不认为测量是真实的。"

Well, moving on to Level C, students should be using both graphical and numerical statistics

to compare two groups.
当达到水平3的时候，学生应当运用统计图和数值统计量比较两个群体。

Students at this level would be coordinating these two types of representations and doing this in a more sophisticated way.
这个水平的学生将协调两种表示形式，并进行更为复杂的操作。

For example, they could observe that the mean and the median of the boys are smaller than that of the girls.
例如，他们将观察男生的平均数和中数比女生的更小。

The range of the boys' heights seems to be bigger than that of the girls.
男生身高的范围比女生的大。

Well, some of these students may wonder if there was something wrong with the outlier that they observe.
也有些学生认为他们观察到的异常值是有错误的。

They realized that there are not any boys in their class who are 121 centimeters or less than 4 ft tall.
他们认识到班级里没有一个男生的身高是121厘米或低于4英尺。

Some students decide to remove that outlier for further analysis.
一些学生为了更进一步的分析决定移除异常值。

Other students working on this task seriously consider the extreme case of 121.
其它进行这项任务的学生认真考虑121的极端情况。

As the students look around their room and ask, "Who was that case?", they decide to explore the arm span data, as well.
当学生环视教室问到"这个例子是谁？"，他们也决定找出臂展的数据。

And when they do this they noticed that when they click on the case of 121 centimeters for the height, that that student has an arm span of 178.
做这件事的时候，他们注意到当点击身高121厘米的例子时，这个学生拥有178的臂展。

When they go back to the data recording sheet, they realize that someone had recorded 121 when they, in fact, were one of the tallest students in the class.
他们回顾记录数据的表时，他们认识到121厘米的学生时全班最高的学生之一。

As the students discuss this they realize there was a measurement error in using the meter sticks and that that entry should have been 179.
当学生讨论时，他们认识到运用米尺时的错误并且应该输入179.

They update the data to reflect this new value, and they make different box plots for the height.
他们升级了数据，反映了新的数值，重新制作了身高的箱形图。

They observe that the height of boys varies more than the heights of the girls, and they use the shading feature to illustrate that about a third of the boys are taller than all the girls.

他们观察到男孩的身高比女孩的高，并且他们使用阴影覆盖来说明大约三分之一的男孩比所有女孩都高。

So, in the final interpreting results phase, students may also be making claims at various levels.
在最终的解释结果阶段，学生做出的判断也会有不同的水平划分。

So, let's start with our students working at Level A. They make the claim that most students in their class had heights that are about 160 – 166 centimeters.
所以，让我们从在水平1的学生开始。他们认为班上的大多数学生的身高都在160-166厘米左右。

They thought this was pretty typical and made sense for students of their age group, and that they still wondered about the realistic value that was so low that they saw in the graph.
他们认为这是非常典型的，对于他们年龄组的学生来说是有意义的，并且他们仍然想知道为什么在图表中看到的如此低的实际值。

The students working at Level B were making claims that boys tended to be taller than girls.
水平2的学生认为男孩趋向于比女孩高。

They, again, thought that this was appropriate for their age range, but that they might expect some little higher variability and that since their class was in the middle of a three grade school, that they might anticipate the same variability that their class might be representative of the entire school.
他们再次认为这适合他们的年龄范围，但他们可能期望一些更高的变化性，并且因为他们的班级是在学校三个年级的中间，他们可能预期他们的班级可能会有代表整个学校的相同变化。

For the students working at Level C, their sophistication was a little bit higher.
水平3的学生他们的思维层次是高一些的。

They were able to very much pay attention to the variability in those boys' heights, and they saw that 35% of the boys were taller than 100% of the girls.
他们能够注意到男生身高的变化性，也注意到了35%的男生高于全体女生。

This made them really wonder about how representative that classes' data could be to represent the entire school, and they were suggesting that in order to really make a claim about the whole school, that perhaps they should select a random sample of students from each grade level and analyze that data further.
这让他们真的很想知道这些班级的数据能够多大程度上代表整个学校，并且他们建议为了真正对整个学校提出推断，也许他们应该从每个年级中随机选择学生样本级别并进一步分析数据。

The SASI framework is intended to help you structure the ways that you can support the development of your students' statistical sophistication.
SASI框架意图帮助你建立能够培养学生统计思维水平的方式。

I hope that this video and the examples that I used in it can help you make sense of the framework, the ways that students might reason at different levels, and help you imagine the kinds of tasks and assessments that you could use in your classroom.
我希望这个视频和所用到的例子能够帮助你理解这个框架，学生会在不同的水平进行推理，也能帮助你想象可以用到课堂中的任务和评价。

I hope you find this framework useful in your own practice.
希望你能够在教学实践中发现这个框架的用处。