Module 5: Real World Data-Driven Decisions

Module Five: Real World Data-Driven Decisions

5.01 Learning Objectives

Learning Objectives

After completing this module, you should be able to:

1. Explain the management implications of the use of business intelligence and knowledge management systems
2. Define Big Data and describe its current uses for analysis and future potential and its implications
3. Explain common analytics for business and quality improvement
4. Recommend manufacturing business decisions based on data analytics
5. Explain common analytics used in healthcare
6. Recommend healthcare decisions based on data analytics
7. Explain common analytics used in education
8. Recommend educational decisions based in data analytics
9. Explain common analytics used in government
10. Recommend governmental decisions based on data analytics

5.03 Results-based Management

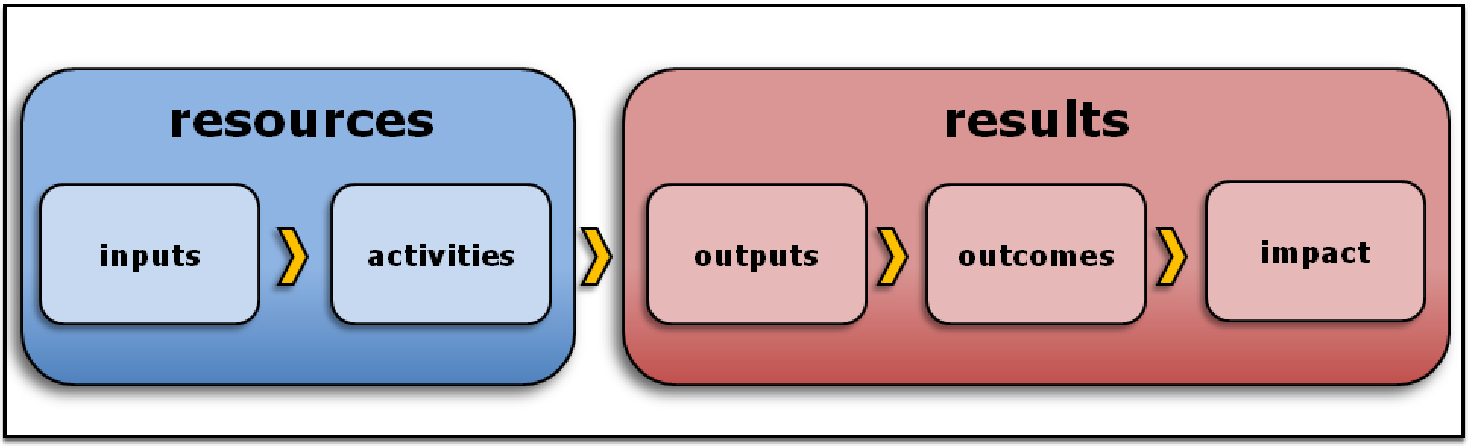
Results-based Management

Results-based management (RBM) is a management strategy that uses results as the central measurement of performance. It has been adopted by many nonprofit and governmental institutions.

The United Nations Development Group more formally defines RBM as "a management strategy by which all actors on the ground, contributing directly or indirectly to achieving a set of development results, ensure that their processes, products, and services contribute to the achievement of desired results (outputs, outcomes, and goals). RBM rests on clearly defined accountability for results and requires monitoring and self-assessment of progress towards results, including reporting on performance."

RBM takes a life-cycle approach in which the processes are continuous and cyclical. Results-based management uses clearly defined terms to continuously measure and evaluate performance, as it relates to results, and these results must be measurable using data.

The Result Chain



The Steps of RBM

Here are the basic steps in an RBM system.

RBM Stage	Explanation	Example
1. Input	The first step of RBM is to define the resources, human or financial, used by the RBM system. This could include the people who are doing the work, the funds used to finance the work, and the information being accessed.	A nonprofit identifies the people and funding needed for a special project to extend its counseling services to recent immigrants.
2. Activities	This second step involves the process that converts inputs to outputs. It includes the actions necessary to produce results. This could include training, evaluating, developing action plans, or working with media.	The nonprofit develops an action plan, recruits new counselors, and prepares a media campaign to reach recent immigrants.
3. Output	This next step is when the outputs have been created by the RBM activities. This could include goods and services, publications, systems, evaluations, or changes in skills.	The nonprofit's media campaign to reach recent immigrants launches on local radio and television stations.
4. Outcome	This is the short-term effect that the outputs will have. It could include greater efficiency, more viability, better decision making, social action, or changed public opinions.	A significant number of the nonprofit's new clients are recent immigrants seeking counseling.
5. Impact	This last step when applying results-based management is to study the long-term effects that the output will have. This could include economic, environmental, cultural, or political change.	The number of recent immigrants seeking help from the nonprofit grows as the community becomes more aware of the services being offered.

A results-based management system requires collaboration from all participants. RBM is built upon a system of partnership and inclusiveness. There must be a mutually beneficial partnership in which all participants have the common goal of improving results. Participants would also share expectations about how the organization would perform. Sharing these expectations leads everyone in the organization to manage and accept responsibility jointly. This includes reviewing one's own work and reporting on one's own progress toward results.

As it is a life-cycle approach, RBM requires constant planning, monitoring, evaluating, and executing the plan.

Results-based management systems channel the performance data from the monitoring and evaluation processes continuously. With organized and analyzed data as the input, the output is the ability to learn as an organization and work jointly towards the end goal.

Transparency, simplicity, and flexibility are also vital to an RBM system's success. All stakeholders need to have clear and well-defined roles, and they must also understand how they and the organization benefit from the realization of goals. Rather than having many complex indicators, RBM simplifies the data to better allow for a constant changing of individual's roles as their performance is evaluated. Flexibility allows for changes in data collection and evaluation techniques as it relates to the results-based management system.

5.04 Business Improvement Analytics

Business Improvement Analytics

A common analytic for business improvement, index numbers are employed to measure the change in quantity or price over time for a good or a number of goods and services. This is extremely important when looking at the direction that things are headed in an industry as it gives a clear comparison of what something is now to what something was a certain number of years ago.

One of the most commonly used index numbers throughout all of business is the consumer price index, or CPI. The CPI is a defined "basket" of assorted consumer goods and services that are purchased by a common household. In the United States the CPI is determined by the Bureau of Labor Statistics, and the CPI in other countries is generally determined by a national statistical source in that country. CPI is one of the most closely watched statistics across the world because it is the main measure of inflation for a given country. This number gives everyone an idea of how the economy, as a whole, has changed over time.

To determine how a price or quantity for a good or service compares over time by calculating an index number, it is important to choose a base period. A base period is the first time that a researcher compares to other times. For example, if a researcher was looking at the change of a certain bond price in 1995 compared to its price in 2010, the base period would be 1995. An index number of 100 indicates the same price or quantity as the base period. The base period is important because it is used to give an initial reference point to compare all other things to.

Simple Index Number

A simple index number shows the change in price or quantity of a single good or service over time. This can be determined in three steps.

1. Determine the price or quantity of a good at the base period (first time that the researcher is comparing).
2. Determine the price or quantity of a good at the time being compared to the base period.
3. Consider the following variables:

   $I_t =$ index number
   $Y_t =$ price at time being compared to base period
   $Y_0 =$ price at base period

   Use the following equation to calculate the index number.

   $$I_t = \left(\frac{Y_t}{Y_0}\right) \times 100$$

Example: Transporting soda

Lucas' company transports soda throughout the country. Last year, Lucas' company delivered 2,000,000 bottles of soda and paid $500,000 for gasoline. This year Lucas' company delivered 3,000,000 bottles of soda to the same area and paid $1,000,000 for gasoline. Lucas is wondering if the increase in gasoline costs means that he should increase the price of his soda or if the increase in gasoline costs is just because of the increase in bottles delivered.

To help Lucas, let's compare two simple index numbers: the cost of gasoline, and the quantity of bottles delivered.

	Cost of Gasoline	**Bottles Delivered**
Base values	Y_0 = \$500,000	Y_0 = 2,000,000
Second measure	Y_t = \$1,000,000	Y_t = \$3,000,000
Ratio of now to then	$\left(\frac{1,000,000}{500,000}\right)$	$\left(\frac{3,000,000}{2,000,000}\right)$
Ratio written as index number	$I_t = 2 \times 100 = 200$	$I_t = 1.5 \times 100 = 150$

The index number on the cost of gas is greater than the index number for bottles delivered meaning that the cost of gas grew faster than the bottles delivered over the same time period. Based on these two variables, we can recommend that Lucas increase the cost of the soda he's delivering to help pay for the increasing cost of gas.

A composite index compares the prices or quantities of a number of goods or services over time. There are two main types of composite indexes: simple and weighted.

Simple Composite Index

A simple composite index is created when a researcher gathers data from many different sources without weighing any data more significantly than any other data. Usually, this means that there is only one good or service being measured so that all of the information is weighted equally.

For example, Rebecca is the manager of the local car dealership, and she is measuring the change in cars sold over the past month. Here is a summary of her data.

	March	April
Salesperson A	15	11
Salesperson B	11	13
Salesperson C	18	15

Once Rebecca gathered this information, she weighed each salesperson's sales equally and determined the sales index number over the last month:

$$\frac{(11+13+15)}{(15+11+18)} \times 100 = 88.6$$

From this simple composite index, Rebecca determined that April's sales were only 89 percent of March's sales.

Weighted Composite Index

A weighted composite index is created when a researcher applies more weight to certain goods or services than others as they are calculating the index number. More weight gets applied to certain goods or services based on quantity sizes or prices. This index gives an understanding that is more proportionate to actual changes over time.

Optional Enrichment: Weighted Composite Indexes

There are different ways to weigh the different goods or services. Let's look at two of these weighted composite indexes: the Laspeyres index and the Paasche index.

Laspeyres index

A Laspeyres index is a comparison of the same quantity of goods with the same weight over a period of time. The CPI is an example of a Laspeyres index. This is determined by measuring the same quantity of goods and prices over time. The following is the Laspeyres index equation:

$$I_t = \frac{\sum_{i=1}^{k} Q_{it_0} P_{i_t}}{\sum_{i=1}^{k} Q_{it_0} P_{i_0}} \times 100$$

Where:

k = number of different goods and services

Q_{it_0} = the quantity for the certain good/service at the base period

P_{i_0} = the price for the certain good/service at the base period

P_{it} = the price for the certain good/service at the final studied time

This equation has a lot in it, so let's use an example to get a better understanding of how to apply this index.

George is working for Seattle Rain, a coat company. He is studying the sales of trench coats and raincoats.

These are the sales figures from 2000 to 2010:

	Quantity sold in 2000	2000 Price	2010 Price
Rain coats	5000	50	60
Trench coats	1000	60	100

The weighted total for 2000 is:

(5000 × 50) + (1000 × 60) = 310,000

The weighted total for 2010 is:

(5000 * 60) + (1000 * 100) = 400,000

Therefore, the Laspeyres index is:

$$\frac{400{,}000}{310{,}000} * 100 = 129$$

George determines that considering the Laspeyres index is 129, the price of coats has increased close to 29 percent from 2000 to 2010. This increase is not exactly 29 percent, because it is unknown if the same number of rain and trench coats were sold in 2000 as in 2010.

The Laspeyres index often slightly exaggerates inflation as it does not consider a consumer's reaction to price changes.

Paasche Index

A Paasche index calculates the difference over time between the weighted totals of the quantities purchased at each time. This has the benefit of getting the real information of the actual spending habits but loses the controls of observing a certain set over time. The following is the equation for the Paasche index:

$$I_t = \frac{\sum_{i=1}^{k} Q_{it} P_{i_t}}{\sum_{i=1}^{k} Q_{it} P_{i_0}} \times 100$$

Where:

k = number of different goods and services

Q_{it_0} = the quantity for the certain good/service at the base period

Q_{it} = the quantity for the certain good/service at the final studied time

P_{i_0} = the price for the certain good/service at the base period

P_{i_t} = the price for the certain good/service at the final studied time

This equation is very similar to the Laspeyres index. The only difference is that the quantity in the Laspeyres index does not change while the Paasche does.

Getting back to George at Seattle Rain, he still wants to know more about the changing sales of coats. Let's calculate the Paasche index for coats.

	Quantity sold in 2000	Price 2000	Quantity sold in 2010	Price 2010
Rain coats	5000	50	5500	60
Trench coats	1000	60	500	100

The weighted total for 2000 is:

(5000 * 50) + (1000 * 60) = 310,000

The weighted total for 2010 is:

(5500 * 60) + (500 * 100) = 380,000

Therefore, the Paasche index is:

$$\left(\frac{380{,}000}{310{,}000}\right) * 100 = 123$$

George now determines that considering the Paasche index is 123, the price of coats has increased close to 23 percent from 2000 to 2010. This increase is not exactly 23 percent, because the number of each type of coat sold in 2000 and 2010 was not constant. The Paasche index's consideration of quantity sold in 2010 already reflects the consumer behavior's reaction to the price increase, and therefore the measured increase is diminished.

The Laspeyres index often overstates inflation, whereas the Paasche index often understates inflation. From this knowledge, George can conclude that the real increase in the price of coats is likely between 23 and 29 percent.

5.05 Case Study: Business Improvement

Tomato Farming: Analyzing Operating Costs and Decisions about New Practices

Coleen Wentzel is a third generation tomato farmer in central California. The oldest of three children, she started helping out on the farm when she was very young, learning the family business. She is a member of the California Tomato Farmers, a cooperative of multi-generational farming families who have made commitments to follow sustainable farming practices that aim to preserve the land while growing and harvesting tomatoes responsibly.

In the past decade, Wentzel and other farmers in the cooperative have been grappling with rising input prices on crop production, while commodity prices have remained stagnant. While Wentzel knows that her overall production costs have been rising, she is not clear about which inputs are having the most dramatic increase on overall costs and what measures she should take to make her business more efficient.

Wentzel's operating costs for tomato production consist of labor, materials, fuel, repairs on equipment, and interest on operating costs. Labor for workers is paid on an hourly, monthly or "per job" basis, and includes wages plus Wentzel's share of federal and state payroll taxes, worker's compensation insurance and other benefits. Materials include inputs such as seed, fertilizer, pesticides (as needed to prevent crop failure) and water. Her field operations use lots of large-scale equipment, including tractors, trucks and other specialized harvesting equipment that require periodic maintenance and repairs. Following are some comparisons of what Wentzel spent on various operating inputs in 2007 (base comparison year) vs. her most recent year (2013).

	Base Cost/Acre (2007)	Current Cost/Acre (2013)	% of Total Operating Costs	Cost Index
Diesel Fuel	$108	$282	16%	261
Fertilizer (per acre)	$67	$131	8%	196
Labor (per acre)	$379	$344	20%	91
Pesticides	$106	$94	5%	89
Water	$92	$125	7%	136
Equipment maintenance/repair	$162	$301	17%	186
Interest	$50	$66	4%	132
Seeds, other materials	$378	$398	23%	105
Overall				132

One thing that Wentzel is unclear about is whether the changes in her operating costs have been caused by market price increases or by greater usage of certain items. In an attempt to better understand the impact of market prices, she consulted the National Agricultural Statistics Service that provides an Index of Prices Paid by Farmers, calculating the relative change in prices over a base year for agricultural products. She learned the following:

- Diesel fuel prices in California increased from just under $1.00 a gallon in 2007 to $3.99 per gallon in 2013
- Fertilizer prices tripled between 2007 and 2013, due largely to energy price increases as well as increases in prices of nitrogen, phosphorous and potassium
- Pesticides include a wide range of materials that, on an individual basis, have fluctuated wildly in price but on the whole have remained flat.
- The hourly labor rates increased from 2007 to 2013 at a steady rate, while the agricultural workers' compensation rates decreased.

A Shift in Farming Practices

Like other farmers in the cooperative, Wentzel is looking at conservation tillage as a way to decrease fuel and labor consumption. Traditionally, this practice involves minimal or no disruption to the soil during seedbed preparation, which results in crop residues remaining on the surface to retain soil moisture and help reduce soil erosion. The result is the increased availability of water for crop production and improved soil quality that retains organic matter such as carbon and nitrogen (reducing the need for fertilizer). It also has the ancillary benefit of requiring less machine tilling, which results in decreased fuel and labor usage and improved air quality.

These practices are relatively new to tomato farming, compared to other crops such as corn, beans, and wheat, and there are potential drawbacks. For example, reducing or eliminating tillage can result in more weeds, increasing reliance on herbicides for weed control. In addition, there is a risk to learning new methods and incorporating them properly, while not adversely impacting tomato production. In general, in areas with a greater drought potential, such as California, conservation tillage systems have been shown to have a yield advantage over conventional systems, but this data is variable and no guarantee that the new system would work on Wentzel's farm.

With rising fuel and fertilizer costs and stagnant commodity prices, however, she is ready to do something aggressive to reduce her costs, insulate her business from steep rises in input prices, and remain true to the sustainability mission of the cooperative.

Question 1. Calculate the indexes of prices paid for fuel in the California in 2013, where 2007 is the base year.

Question 2. Compare the answer you got in Question 1 (price index for fuel) with Wentzel's cost index for fuel on her farm specifically. Which is higher? What does it tell you about how Wentzel likely used fuel during this period of rising fuel costs?

Question 3. In the table of operating costs per acre that Wentzel compiled, are her cost index figures weighted? Why or why not?

Question 4. Do you think it makes sense for Wentzel to move toward conservation tillage and crop rotation practices? Why or why not?

Suggested/Sample Responses

Question 1:

$I_t = \left(\frac{Y_t}{Y_0}\right) \times 100$, where Y_t is \$3.99/gallon, and Y_0 is roughly \$1.00/gallon.

$I_t = 399$

Question 2:

The US price index for fuel (399) is much higher than the cost index for Wentzel's fuel cost per acre (261), calculated for the year 2013, with a base year of 2007. This tells us that, although the market rate for fuel increased substantially during this period, Wentzel likely used less fuel per acre in 2013 than she did in 2007. This may have been due to technological innovations or different farming practices that reduced the amount of fuel needed to farm an acre.

Question 3:

Yes. The cost per acre comparisons are a weighted index because they take into consideration the price increases per unit (e.g., a gallon of fuel) as well as the amounts she used. For example, for some inputs, the market price decreased, but she used more of the input so her cost index might be the same or lower than the market price index for the same time period comparison.

Question 4:

While there are risks to changing any significant business practice, conservation tillage seems like a good idea on many levels. Since Wentzel is worried about rising operating costs, the new practices have strong potential to reduce many operating inputs, including fuel, fertilizer, water, labor, and equipment maintenance. Several of these inputs have seen steady to dramatic increases in market rates (fuel, fertilizer, labor), and any efforts to curb usage of these costs makes sense.

There are also significant environmental benefits, including air and soil quality and water conservation, that are in keeping with the sustainability commitment of the California Tomato Growers cooperative.

5.06 Healthcare Analytics

Healthcare Analytics

Nearly every industry in the 21st century is data driven, and the healthcare industry is no exception. We have alluded to there being many applications of the concepts we've learned within the field of medicine, but in addition to these statistical concepts, there are a variety of analytics designed specifically for the healthcare industry. Throughout this assignment, we will examine in detail some of the common analytics used in the healthcare industry.

Epidemiology

Epidemiology is the study of the incidence, distribution, and possible control of diseases and other factors relating to health. Public health and medical professionals rely on these analytics to help shape the decisions they make every day. Rate, ratio, and proportion are fundamental elements of statistics that are essential to epidemiology. In epidemiology, it is imperative to understand how to use these metrics to study raw data.

A rate is the measure of an event occurring over a period of time. For example, the birth rate measures the number of births per 1000 people in a given year. A ratio measures one quantity in relation to another quantity. For example, the gender ratio measures the number of males in a population to the number of females in the same population. Lastly, a proportion is a ratio in which a part of a group is compared to the whole group. For example, a proportion can be used to measure the number of people with diabetes in relation to the total population.

In epidemiology, it is important to convert this raw data into meaningful statistics. For example, rates, ratios, and proportions can be used to study healthcare statistics such as disease frequency. In doing so, it is crucial that we use the most appropriate units of analysis. Consider the table below:

Country	People living with HIV/AIDS (estimate)	Population	Proportion of population living with AIDS
Swaziland	200,000	1,200,000	0.167
United States	1,100,000	317,100,000	0.003
Source: World Health Organization			

The table above exemplifies how important it is to use the proper measures of analysis when studying disease frequency in epidemiology. If we simply compare the number of people living with HIV/AIDS in Swaziland to the number of people living with HIV/AIDS in the United States, we see that the following ratio:

$\mathbf{200{,}000 : 1{,}100{,}000 \text{ or } 0.18 : 1}$

According to this measure, there are 18% percent as many people living with HIV/AIDS in Swaziland as in the United States. This figure makes it seem like HIV/AIDS is more prevalent in the United States than in Swaziland. But is that really the case?

A better method for measuring the prevalence of HIV/AIDS in a country is to measure the proportion of that country's population that is living with HIV/AIDS. The ratio of the proportion of people living with HIV/AIDS in Swaziland to the proportion of people living with HIV/AIDS in the United States is as follows:

$\mathbf{0.167 : 0.003 \text{ or } 55.67 : 1}$

Therefore, the proportion of people living with HIV/AIDS is nearly 56 times greater in Swaziland than in the United States! This statistic tells us that HIV/AIDS is far more rampant in Swaziland than the United States. We can see from this example that public health decisions can be drastically influenced by the analytics used in epidemiology.

Prevalence and Incidence

Prevalence and incidence are two similar measures used when studying disease frequency, but they do have differences, and it is important to understand when to use each one. Prevalence counts all of the existing cases of a disease, while incidence only counts *new* cases. As you can imagine, from a public health and healthcare perspective, differentiating between these two measures is crucial.

Prevalence measures the number of cases of a particular disease that exist in a population. This measure does not differentiate between cases — it counts both existing cases and new cases. Because it measures the number of individuals that have a specific disease compared to the population, prevalence is a type of proportion. Prevalence, P, can be calculated as follows:

$$P = \frac{\textbf{number of cases}}{\textbf{total population}}$$

Incidence, on the other hand, measures the number of new cases that arise in a population over the course of a designated time period. There are two types of incidence that are commonly used; cumulative incidence (CI) and incidence rate (IR). Cumulative incidence is a proportion that measures the number of new cases that arise in a particular time period compared to the population. Cumulative incidence, CI, can be calculated as follows:

$$CI = \frac{\textbf{number of new cases in a particular time period}}{\textbf{population}}$$

The other commonly-used type of incidence is incidence rate. A rate is a proportion that includes an amount of time in the denominator. When studying disease frequency, incidence rate uses person-time units, which is a cumulative measure of the amount of time each person was studied. Incidence rate, IR, can be calculated as follows:

$$IR = \frac{\textbf{number of new cases}}{\textbf{person-time units}}$$

It is important to recognize when to use each of the three measures mentioned above. Since prevalence measures the number of cases that exist in a population at a given point in time, it can be helpful in determining where relief is needed most. For example, one might measure the prevalence of HIV/AIDS in different regions when determining where relief efforts should be focused.

Incidence, while more complicated to calculate than prevalence, is very helpful in certain situations. Cumulative incidence, which measures the number of new cases that arise in a period of time, can be helpful in determining which populations are at the highest risk of contracting a disease. If we were interested in disease prevention, we could use cumulative incidence to measure the proportion of new cases based on a variety of factors, such as age or gender. Incidence rate, which uses person-time units, can be used to measure changing levels of risk over a period of time. For example, we could use incidence rate to determine if a treatment or preventative measure is effective in slowing the spread of a disease.

These epidemiological measures of disease frequency are only a few examples of how analytics can be used in healthcare. Understanding data is of utmost importance for professionals within the medical and healthcare industries. With the proper use of data and analytics, public policy can be influenced, effective medicine can be identified, and ineffective medicine can be eliminated.

5.07 Case Study: Health Care Analytics

The Case to Require Bacterial Meningitis Vaccinations for College Students

Omar Saddiq is an epidemiologist, conducting research on trends in bacterial meningitis cases among college students living on campuses in the United States. While bacterial meningitis can occur in anyone, college students have a slightly elevated risk of contracting the disease. They are more likely to engage in high-risk behaviors that can result in exposure, such as sharing drinks or cigarettes and living in close contact.

Saddiq is involved in discussions with university officials and lawmakers in the state of Illinois about whether or not to require students to receive a bacterial meningitis vaccination prior to enrollment. There is strong support for implementing the vaccine requirement to protect college students from the devastating illness, including an updated recommendation from the CDC Advisory Committee. Saddiq knows, however, that there will be opposition to this requirement from some lawmakers who argue that it represents too much government intrusion into students' personal lives and decision making.

Saddiq had collected data from five states' university health departments over a five-year time period to compare incidence rates among different population sub-groups: those who received the vaccine prior to enrollment and those who didn't. He also broke these groups down by residential status (whether they lived in dorms on campus or off campus). A control population of all people (college and non-college) between the ages of 18 and 23 had an incidence rate of 1.4 bacterial meningitis cases per 100,000.

Vaccine received?	Residential Status	BM incidence rate per 100,000 per year
Yes	On Campus Dorm	0.57
No	On Campus Dorm	4.6
Yes	Off Campus	0.34
No	Off Campus	1.6

Based on the different incidence rates above, Saddiq drew the following conclusions:

- the rate ratio of cases among non-immunized dorm dwellers is 8:1 or eight times higher than the rate of cases among those vaccinated
- the risk among non-vaccinated students who lived in on-campus housing was about three times higher than non-vaccinated students who lived off campus
- the risk among non-vaccinated students who lived in on-campus housing was some three times as high as the risk for the general population of the same age.

The available meningitis vaccines are about 85% - 90% effective against certain strains of bacteria, according to recent studies. Since the vaccines do not protect against all causes of meningitis, it is possible for someone to receive the vaccine and still get meningitis, but the risk of contracting it is significantly lower after the vaccine.

Opponents have raised issues about the cost of requiring all students to be immunized. The vaccine, which ranges from $50-$90 per dose, will likely be covered by insurance or students would pay out-of-pocket. But the increased administrative record-keeping, follow-up and enforcement for students who don't comply will likely add expense and administrative burden for universities staff members. Some state lawmakers have questioned whether the cost-benefit analysis of requiring universal bacterial meningitis vaccination makes sense on economic terms, especially since the infection rate is low, compared to other infectious diseases.

Saddiq believes that, from an economic cost-benefit perspective, requiring vaccinations for all students will prevent enough cases of bacterial meningitis to save on hospitalization costs, including ongoing medical costs for those cases resulting in ongoing debilitation (amputation, brain trauma) or "lost value" from premature death. Any discussion of costs imposed by requiring the vaccinations, he believes, must be weighed against the costs avoided (financial and emotional) by preventing cases of the disease.

Exercise

A. You would be ____ times more likely to get bacterial meningitis if you lived on campus without getting vaccinated than if you lived off campus having been vaccinated. The answer should be rounded to the nearest tenth (for example, 00.0)

B. If there were 500,000 students living on campuses throughout the country, and the incidence rates were to stay the same as in Saddiq's sample, about how many more students per year would get bacterial meningitis if they were not vaccinated than if they were? The answer should be an integer, so round to the nearest integer.

C. If there were 100,000 students living on campus, 70 percent of whom had taken the vaccine, about how many would get bacterial meningitis in a year? The answer should be an integer, so round to the nearest integer.

Exercise Answers

A: 13.5. You would be 13.5 times more likely to get bacterial meningitis. 4.6/.34 = 13.53

B: 20. There would be 20 more students per year infected. 5*(4.6 -.57) = 20.15

C: 2. There would be 2 students getting infected each year. 0.30*4.6 + 0.70*.57 = 1.779

Review Checkpoint

Question 1. Based on the data and information presented above, how would you construct an argument in favor of universities and colleges requiring bacterial meningitis vaccines prior to enrollment? How would you argue against the likely opponents?

Question 2. If the effort to require bacterial meningitis immunizations for ALL incoming students fails on the basis that it is not economically feasible to require everyone to be vaccinated, could you suggest a modified requirement, based on the data above?

Suggested/Sample Responses

Question 1:

Every year on college campuses, there are a number (albeit small) of students that contract bacterial meningitis with potentially devastating consequences. Students that live on campus in dorms are at elevated risk of infection due to living in close quarters and certain behaviors, as well as stress and lack of sleep. Many of these cases are preventable with vaccines that are 85-90% effective in preventing bacterial meningitis. While each vaccination dose has a cost that ranges from $50 to $100, the benefits of saving even a percentage of students from the effects of meningitis are worth it. Hospitalization and follow-on medical costs are staggering; and the emotional toll of premature death or lifelong disability resulting from a percentage of bacterial meningitis cases are huge costs. There is a positive network effect as well: the more students are vaccinated, the more it reduces the risk of contagion.

Yes, the aggregate cost to vaccinate every incoming student is significant, and the administrative burden of enforcing the requirement will increase. However, studies show that the incidence rate of the disease is much lower among students that have been vaccination. IF everyone is vaccinated, there will be far fewer students that have to experience the devastating infection, and ultimately lives will be saved.

Question 2:

In addition to providing clear information and recommendations that everyone gets vaccinated, colleges could require all students living in dorms to receive vaccinations. The data suggests that this is the most at-risk sub-group, with students living in close contact. Requiring immunization for this sub-group would be less costly than requiring that all students get vaccinated, but it would target the most vulnerable population.

5.08 Education Analytics: Data Driven Decision Making in Education

Education Analytics: Data Driven Decision Making in Education

Ganesan (Ravi) Ravishanker discusses "the value of academic analytics in bringing efficiencies, easy access to data, and decision making" in *Doing Academic Analytics Right: Intelligent Answers to Simple Questions* Ravishanker maintains that data-driven decision making has the potential for dramatic improvement in higher education, but that these analytics have been underutilized.

REFERENCE: Click to review normal distribution and ANOVA content

The Normal Distribution

When data tends to occur around a central value with no bias right or left, it gets close to a normal distribution. All normal distributions look like a symmetric, bell-shaped curve.

Please click the slideshow's "Next" button to advance.

Applications

The normal distribution exists in many different fields. Click on each field below to see an example.

ANOVA

Analysis of Variance (ANOVA) is a technique used to determine if there is sufficient evidence from sample data of three or more populations to conclude that the means of the population are not all equal. Ordinary hypothesis testing can be used to determine whether there is a basis to conclude that two populations have different means.

The null hypothesis states a claim that all population means are equal so, for example, if three populations are considered the null hypothesis would be Ho: $\mu_1 = \mu_2 = u_3$. The alternate hypothesis would represent that not all of the population means are equal, that is, at least one of the population means are considered significantly different statistically.

The test-statistic is a ratio where the numerator is more generally affected (overestimating the variance) by the fact that the null hypothesis is not true, while the denominator is generally less affected by the same fact. Thus, the ratio would be relatively close to 1 if the null hypothesis is true as both the numerator and denominator would be reasonable estimates of the variance. However, if the null hypothesis is false, the numerator tends to overestimate the variance so the ratio increases and strays from the value 1.

The F-distribution is utilized to set the rejection region for analysis of variance. If the null hypothesis is false, the test statistic is likely to lie in the rejection region, so the conclusion that at least one of the population means is significantly different from others would be valid, subject to the stated level of significance.

A sample with a smaller variance will have more data points clustered around the mean and give us more confidence in a small range of potential means. A sample with a large variance will require a larger range of potential means to be considered possible. Therefore, by analyzing the variances within each sample, we can determine whether the means between the samples are significantly different. The result of our ANOVA will inform us as to whether we can conclude from the samples that the populations we are sampling have different means.

A one-way ANOVA tests the difference between populations' means using a single factor. For example, let's say you are studying the effect of different pest management techniques on the yield of apples. You use three pest management systems on three sections of the orchard and record the weekly yield from each orchard section. In this experiment, the pest management technique is the single factor.

Consider the case of examining the average size of household for various ethnicity groups. By collecting a sample for each group, we may test the hypothesis that the average size-of-household is equal for each of the groups. Sample data for household size within groups is provided below:

Caucasian	**African-American**	**Latino**	**Asian-American**
5	3	5	4
2	4	3	3
6	4	6	3
3	6	5	2
4	5	4	3
Sample Mean			
4.0	4.4	4.6	3.0
Sample Variance			
2.50	1.30	1.30	0.50
Sample Standard Deviation			
1.58	1.14	1.14	0.71

The calculation of the numerator of the test-statistic yields 2.53 while the denominator is 1.4. Thus, the ratio equals 1.81. Referring to the F-distribution table using a five percent level of significance, a degree of freedom for the numerator of 3 (number of groups minus one) and a degree of freedom of 16 for the denominator (total number of observations minus the number of groups) we find an F-ratio of 3.24.

The rejection region is therefore $F > 3.24$ and since our test-statistic is 1.81 we find the test-statistic does not lie in the rejection region. Thus, we fail to reject the null hypothesis. This means, for example, that although the sample mean of the Latino household is more than fifty percent greater than that of the Asian American household, the data is not sufficient evidence to conclude the population means are statistically significantly different.

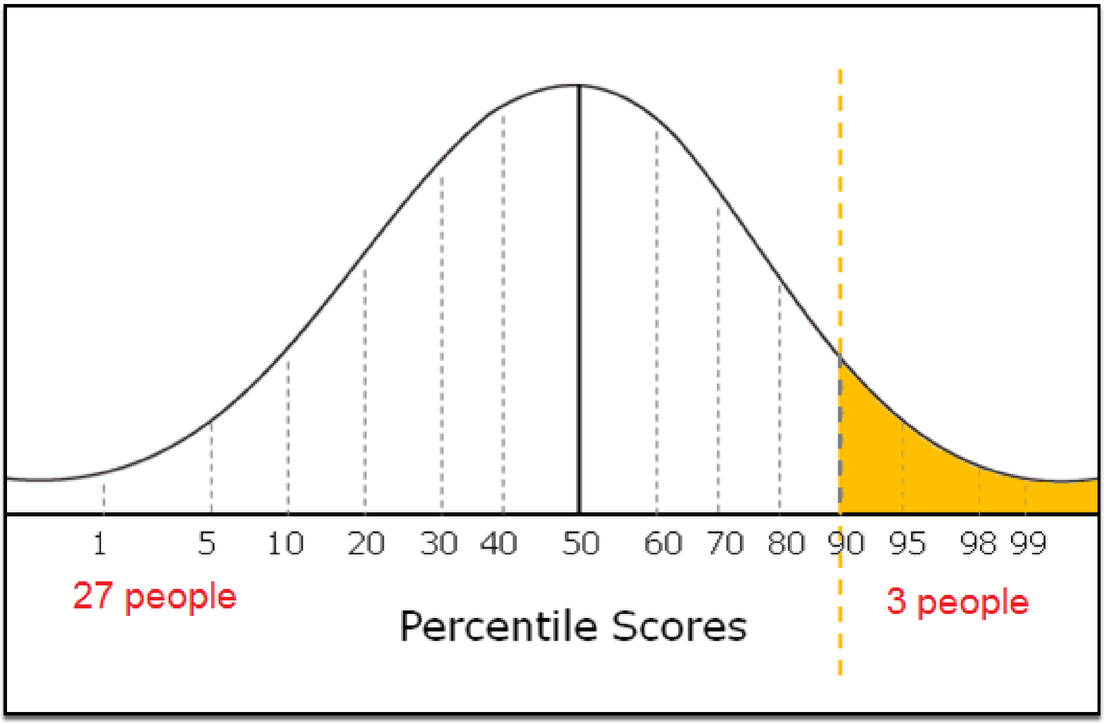
Had the test-statistic calculation been 3.81 instead of 1.81 we would observe that it does exceed the 3.24 value from the table so the test-statistic would lie in the rejection region. In this case, we would reject the null hypothesis and accept the alternative hypothesis that not all the group population means are equal, that is, at least one of the group population means is not equal to others, at the five percent level of significance.

Key Statistical Tools

Statistics can help education leaders to a better understanding of student progress, the effectiveness of different questions, and the construction of tests.

Percentiles

A percentile signifies the percent of the population that falls below a certain value. If someone was in a class of 30 and he or she was in the 90th percentile for a test, there would be three people who had better scores.



Percentiles are frequently used in education because many employers and admissions offices look at students' relationships to their peers. Studying a student's relationship to his or her peers is important because average grades vary from institution to institution, and so a student's score in one place might not be equal to the same student's score somewhere else.

Standard Scores

Standard scores, or z-scores, measure the distance of a piece of data from the mean compared to the entire population. The unit of measure for a standard score is a standard deviation. Therefore, if someone's test result has a standard score of one, then their score was one standard deviation above the mean. Standard scores are used in education because they can put a result in context to other scores that are not necessarily from the same source.

REFERENCE: Click to review standard scores content

Standard Scores (z-scores)

Often, if we are comparing two or more distributions to one another, the scales will be different. Consider the book sales data and the restaurant spending data:

Student	$ Spent
1	14.00
2	38.00
3	15.00
4	29.00
5	10.00
6	11.00
7	45.00
8	4.00
9	50.00
10	15.00
11	39.00
12	27.00
13	10.00
14	20.00
15	30.00
16	9.00
17	16.00
18	24.00
19	28.00
20	7.00

Book Sales	f	Book Sales	f
38	1	52	2
39	0	53	2
40	0	54	2
41	2	55	3
42	4	56	0
43	2	57	1
44	4	58	1
45	6	59	2
46	2	60	1
47	2	61	2
48	5	62	1
49	2	63	0
50	3	64	0
51	3	65	1

It is very difficult to see how these distributions are similar or different from one another because of the differences in scale. Z-scores will transform different datasets to the same scale. In order to do this, we figure out how far away each individual data point is from its respective mean. In other words, a z-score tells us the number of standard deviations a data point is from its mean. If a data point from the book sale data has the same z-score as a data point from the restaurant spending data, the two data points have the same relative location in their respective datasets.

The formula for z-score is:

$$z = \frac{x - \bar{x}}{s}$$

where:
z = the number of standard deviation units a raw score is from its mean
x = the raw score
$\bar{x}$ = the sample mean
s = the sample standard deviation (also denoted by σ for a population)

Because the mean is at the center of the distribution if the score falls below the mean, the z-score is negative. If the score falls above the mean, the z-score is positive. Once transformed to z-scores, all distributions have a mean equal to zero and a standard deviation of 1.

Example

Consider the book sales data. The mean of this dataset is 49.481. The standard deviation is 6.345. How many standard deviations away from the 49.481 is the data point of 44?

$$Z = \frac{X - \bar{X}}{S}$$

$$Z = \frac{(44 - 49.481)}{6.345} = -0.864$$

With a z-score of -0.864, the data point of 44 books is 0.864 standard deviations *below* the mean for this sample of book sales. We know this data point is below the mean because the z-score is a negative number.

Further, if you wanted to find a comparable data point among the restaurant spending data, you can use the same formula, solving for a different variable. If the standard deviation of the restaurant spending data is 13.31 and the mean is 22.05:

$$Z = \frac{X - \bar{X}}{S}$$

$$-0.864 = \frac{(x - 22.05)}{13.31}$$

$x = 10.55$

Since both data points, 44 from the book sale and 10.55 from the restaurant spending, have the same z-score, we know that they fall at the same relative point in their respective datasets even though one set of data is listed in quantity of books and the other in dollars.

Test Construction

The goal of most tests in educational settings is to measure the ability levels of an individual in a certain subject area. When a researcher is studying an area that has not been studied before, a test has to be created and validated to determine if it is measuring what it says it is measuring. There are two types of tests: norm-referenced tests and criterion-referenced tests. Norm-referenced tests compare an individual to other individuals. A standard score is a type of norm-referenced measurement. Criterion-referenced tests compare an individual to certain defined standards. A driver's permit test where someone has to answer a certain number of questions correctly in order to pass is an example of a criterion-referenced test. The difference between these tests is very important when considering if a test result should be compared to other test results or if there is a defined amount of knowledge that should have been learned.

Most tests are comprised of questions that are all worth the same amount of points (are equally-weighted), which are then combined to create a composite test score. When combining these equally-weighted questions, the difference between the difficult and easy questions gets lost. This is important because there is a lot of information about a student's range of abilities that do not exist in a composite score.

True Score Model – Classical Test Theory

In education, someone might have a bad day and not do as well on a test as they probably should have done. Also, someone might have all of his or her guesses fall his or her way, and he or she might do better than they probably should have. An observed score is the score that is actually achieved by an individual on a test. A true score is the average score an individual would achieve if he or she were to take the test infinite times. In an ideal world, someone's observed score would be the same as that person's true score. Also, in an ideal world, the test would measure what it said it was measuring.

Because this uncertainty exists, it is important to consider measurement error. There are two types of measurement error: random measurement error and systematic measurement error. Random error is because of chance coincidental situations. Random error can eventually be eliminated over time as the number of tests increases. Systematic errors occur when something that is entirely unrelated to the test is affecting the results. An example could be a student that continues to do poorly on his or her Chemistry tests because the tests are in Chinese and the student is French (and does not know Chinese). Systematic error can be eliminated by adjusting the test so that the test is unaffected by unrelated factors. The True Score Theory states that, in a test without systematic error, the observed score is the true score plus the random error. Observed or true scores and types of error are incredibly important when creating and analyzing a test because they help determine the validity of those tests.

Item Response Theory

Item Response Theory (IRT), also known as latent trait theory, is a model of designing, analyzing, and scoring tests. It is often considered more useful than classical test theory because it tries to pull as much information out of each answer as possible. The GRE and GMAT are two examples of tests that apply IRT to their tests. The IRT does not assume that each question is of equal difficulty and focuses on each question's implications of a correct answer on different scales. Classical test theory looks at the implications of the overall test on each individual question, whereas IRT looks at each individual question and tries to determine the meaning of a correct answer. There are a number of different IRT models, including the three-parameter logistic (3PL) model, the logistic and normal IRT model, and the Rasch model. IRT can be more difficult to apply than classical test theory but generally gives stronger findings and does a better job of scaling people.

For additional reading on the reliability of testing and item analysis, please click to expand the topics below.

Reliability of a Test

There are a number of different ways that a person can measure the reliability of a test.

A reliability index is taken to determine how well a person's observed score on a test represents that person's true score. This is the ratio of the standard deviation of the true scores and the standard deviation of the observed scores. Because it's not possible to get true scores, and therefore standard deviations for true scores, one of the following models has to be used to represent true scores.

In the test-retest method, the same test is given to the same sample on two different occasions. Correlation is lost as the time between the tests increases.

In the parallel-forms method, two tests are made in the exact same way and given to the same sample. The parallel-forms method and the test-retest methods compare the two samples to each other.

The split-half method and Cronbach's Alpha measure internal consistency, and therefore do not take any information outside of the one sample test's results. The split-half method splits a sample into halves and measures the correlation between the two halves. Cronbach's Alpha takes the average of all of the possible variations of the split-half method within a sample. All of these techniques may be used when determining how reliable a test is.

Item Analysis

An item is a single account from a test, often a written question. Item analysis is the study of the results for each individual item, or question, on a test. This is very important in education as it can show how a test should be adjusted and gives insight into an individual's test results. A greater variance between a certain question's results means the question was successful in discriminating between individuals' abilities. The greatest variance is when the observed score has a 50 percent chance of being correct. When there is a question with given possible answers, the average observed score (p) will be higher than the true score (p_0). This is because someone without any of the related knowledge or ability has a chance to get an answer correct. The following equation determines the likely observed score to account for this:

$$p = p_0 + \frac{1-p_0}{m}$$

Where:

p = item's observed score
p_o = item's difficulty (true score)
m = item's number of possible answers

This takes into account the difficulty of an item as well as the chances that someone randomly picks the correct answer from the options available. An example could be a multiple choice question on a test with four possible answers. If 60 percent of the class knows the correct answer, then 40 percent of the class will guess between the four answers. The question's observed score is 70 percent because 60 percent of the class knows the correct answer and 10 percent (one-fourth of 40 percent) got the correct answer by luck.

Often, the top 20 percent of the sample's test results are compared to the bottom 20 percent of the sample's test results. The rule of thumb for any item is that the upper 20 percent's item x observed score should be at least 40 percent higher than the bottom 20 percent's item x observed score. If 85 percent of the top 20 percent of students get Question 9 correct, then, at most, 45 percent of the bottom 20 percent should get Question 9 correct. If the difference is less than 40 percent, then the question is not consistent with the rest of the test and does not discriminate enough between individuals' abilities.

Item Analysis Example

Professor Andrews receives tests from her 50 students at the end of class. She wants to determine if she should keep any of the questions for next year's test.

She decides to use item analysis to determine if the true score is what she wants; she also wants to see if each question separates the class enough. Andrews wants each question's true score to be between 50 and 80. All of the questions are multiple choice questions with four possible answers.

Andrews sees that 82 percent of her students answered Question 5 correctly. Using Item analysis:

$$p = p_o + \frac{(1-p_o)}{m} = \frac{1+p_o(m-1)}{m}$$

$$0.82 = p_0 + \frac{1-p_0}{4}$$

$$0.82 = 0.75 \times p_0 + 0.25$$

$$p_0 = \frac{0.82-0.25}{0.75} = \frac{57}{75} = 0.76$$

These results for Question 5 question fall within the desired true score. She does this with all of the questions on the test and determines which questions she will keep.

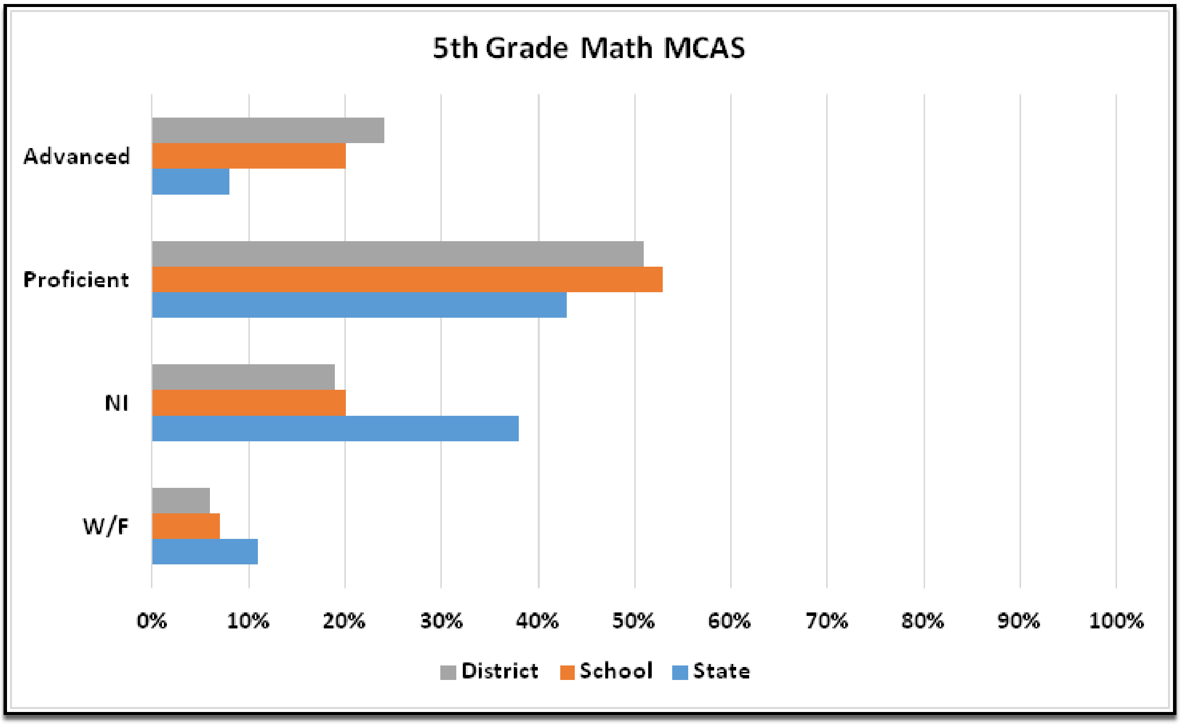
Andrews then separates the top ten tests and the bottom ten tests. For each question, she then compares the number of correct answers from the bottom ten and top ten. If the difference between the groups is 40 percent or greater, in favor of the top ten students, then the question discriminated enough between the students' abilities. For example, if eight of the top ten students and seven of the bottom ten students were correct on Question 7, then the question does not discriminate enough for it to be useful in differentiating between the students' abilities.

5.09 Case Study: Education Analytics

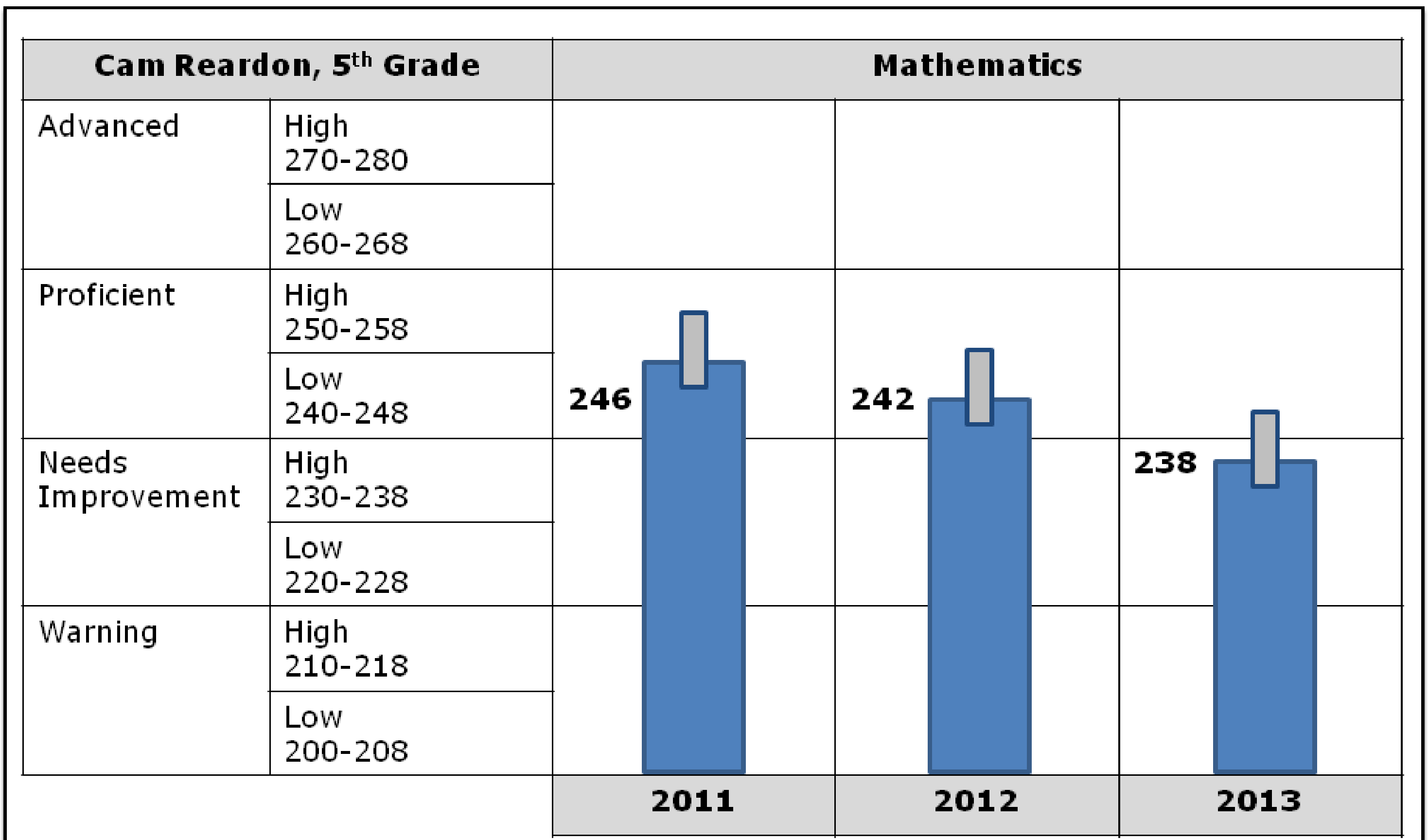
Education Mini Case Module 5: Reliability and Measurement Error in 5^{th} Grade MCAS Scores

Cam Reardon is a 5^{th} grade student at North Shore Elementary School. He is also a hockey fanatic. On the night before his school's Math state achievement (MCAS) testing date, Cam had the opportunity to attend a Bruins game that happened to go into double overtime and a shootout before ending close to midnight. By the time Cam and his mom managed to get out of the arena parking garage and drive home, it was close to 1 a.m.

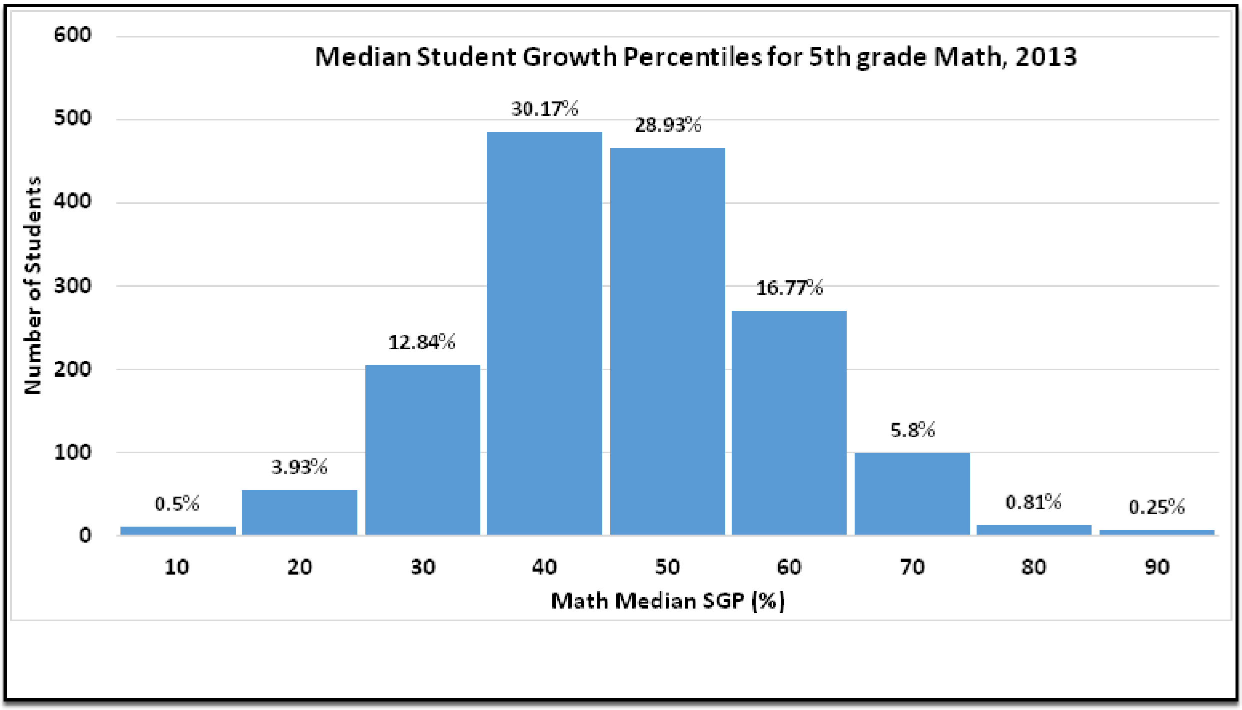
When Cam took the MCAS test the next day in school, he felt tired and had trouble concentrating. Not surprisingly, he scored in the "Needs Improvement" category, even though he was reasonably strong in math and had previously scored in the "Proficient" category in prior years. When compared with several different peer groups (state, district, and school) his math MCAS score of 238 placed him in the 38th percentile in the state and the 20th percentile in his school (see the chart below).



The following part of Cam's MCAS report showed his performance on math testing for the latest three years, with the blue bars indicating how well Cam actually did on the tests, and the small gray bar showing the range of scores Cam would likely get if he took the test many times.



Interestingly, although Cam's scores dropped from 2012 to 2013, his Student Growth Percentile (SGP) was actually at a 72, meaning that he performed at the 72nd percentile when compared to "academic peers" statewide. The SGP is calculated by comparing one student's progress to the progress of all other students across the state who took the same test and have a similar history of MCAS scores for two or more years. An SGP of 72 means that Cam actually did better in 2013 than 72% of the students across the state who scored about the same as he did in 2012 and 2011. The following chart shows the percentage of students from Cam's peer group that was in each student growth percentile:



As it turned out, the 5th grade teachers (and a handful of teachers from other districts) filed complaints with the State Department of Education about the reliability of this particular test. They claimed it was much more difficult than previous fifth-grade math MCAS tests and that it rendered comparisons such as growth percentiles and other year-over-year performance improvement metrics to be somewhat questionable. In the context of the District's new Teacher Performance Evaluations, it also made it appear as though they had not done a good job preparing their students for the test.

Based on a number of these complaints, and the fact that many students saw significant performance declines, the State DOE responded by preparing a new fifth-grade math MCAS test that would be administered to all of the state's 5th graders the following month. On this new test, Cam scored much higher, as did many others in his school and across the state.

REFERENCE:

Reliability and Validity of Data

In statistics (as well as science) measurements need to be both reliable and valid. Reliable data is both consistent and repeatable. If you were to administer the same test to the same person three times and the scores were similar each time, the test could be categorized as reliable. If the results varied greatly, the test would be unreliable. Similarly, valid data is data resulting from a test that accurately measures what it is intended to measure. For instance, if a test reflects an accurate measurement of a student's abilities, it is said to be valid.

Errors: Random versus systematic

All measurements contain some degree of error. This error may be random or systematic. Random errors should cancel themselves out over a large number of measurements if they are NOT related to the true score and if there is no correlation between the errors. If the errors do not meet these criteria, they are not random. Systematic errors are not due to chance, and although they can be corrected, correcting them takes time and attention to detail. Systematic errors are more difficult to find because the data is likely consistently askew in the same direction. Skewness is a measure of the degree to which data "leans" toward one side. If data is collected using an instrument that needs calibration, this error will repeat itself consistently across the data set.

Measurement bias

Measurement bias can invalidate the results of any study, so it is important not to let bias creep into your experiment.

Because it is usually impossible to measure results for an entire population, a sample of the population is measured. To produce unbiased results the sample tested must be sufficiently random. If, for example, you ask 100 people in Annapolis, MD whether they prefer the Army or Navy, your sample is likely to be biased in favor of the Navy because the U.S. Naval Academy is located there. A better methodology would be to randomize your sample and get opinions from people in many locations throughout the country.

Information bias

Assuming your sample is properly randomized, the second way bias can enter your model is when data is collected. This is called information bias and may occur for a variety of reasons. Your interviewer might be biased and may ask questions in a way designed to elicit a certain response. The questions themselves may be biased in favor of a certain outcome, or a responder may lie in their responses. The best way to combat information bias is to keep both interviewers and responders in the dark about the purpose of the survey, inasmuch as this is possible.

For example, consider a transit authority trying to determine whether a new passenger rail line from Big City to Springfield would draw a sufficient ridership. As a way of assessing demand, it interviews people in Springfield employing the following interview questions:

- "Would you be willing to ride a new, affordable train on your next trip to Big City if the service existed?"
- "Are you in favor of limited train service into Big City if it involves lengthy disruptive and costly construction?"

Notice the likely responses to each question. While many people might answer "yes" to the first question, many of those same people might answer "no" to the second question. The results of either survey would include informational bias.

Questions

Question 1. How do you explain the variance in Cam's performance on the 2013 math test, compared to his typical ability and previous scores? Is his 2013 score an "observed score" or a "true score?" Explain why.

To the extent that Cam's late night at the hockey game the night before made him tired and may have impacted his score, would this be an example of random measurement error or systematic measurement error? Why?

Question 2. How do you interpret the fact that Cam's math score went down in 2013, but his growth percentile was 72%? What does this likely say about the 2013 test?

Question 3. When the DOE decides to issue a new 5th grade math test in response to teachers' complaints, is this "test-retest" or "parallel forms" method that is used to get something closer to true scores? Why would the DOE agree to prepare a new test? Do you think the teachers should be evaluated in part based on the 5th grade math scores on the first test?

Suggested/Sample Responses

Question 1:

This is an example of random measurement error. The fact that Cam stayed out late the night before this MCAS testing session probably resulted in him being tired and not testing well. If there were many testing sessions, it is very unlikely that Cam would stay up late at a double-overtime game before each one. Therefore, he might test better after a good night's sleep on other occasions. Systematic measurement error, on the other hand, is a condition that would impact his result every time — not due to his ability but due to some other consistent barrier.

Question 2:

The fact that Cam's score declined, yet he was in a high growth percentile relative to his peers means that many of the students in Cam's academic peer group (kids who scored similarly in previous years) saw their scores go down in 2013. And 72% of them had their scores go down more than Cam's did. This is likely an indicator that the test was harder this year — perhaps unrealistically so, as the teachers are contending to the DOE.

Question 3:

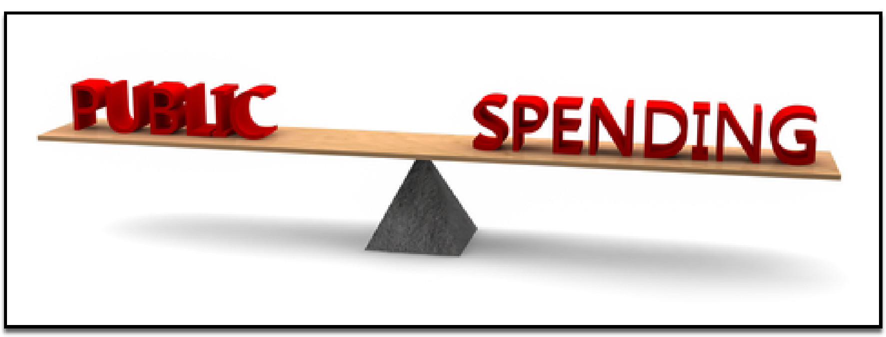
Parallel forms test since it gives two tests measuring the same content to the same sample.

In response to complaints about the reliability of the first math test as a measure of students' true math abilities, the DOE is issuing a second test to help gauge the reliability of the initial scores. If the scores on the new test diverge significantly, it will indicate that the first scores were not likely a very accurate measure of students' true scores.

It is probably not a great idea to evaluate teacher performance based on MCAS scores, especially one that turns out to have reliability issues and might not be a great indicator of students' true abilities. It does, however, make more sense to use Student Growth Percentiles (SGPs) for a teacher's students to help measure teaching performance. This measure controls for the variability in test difficulty by measuring a student's improvement, relative to other students that have performed similarly in the past. It does not, however, control for the random measurement error caused by Cam's late night at the hockey game.

5.10 Public Sector Analytics

Public Sector Analytics



Analytics is of increasing importance in the public sector. In the private sector, companies use analytics to maximize their profits and cut costs. Similarly, analytics are used in government to understand past performance and deliver public services at a lower cost.

Many countries and states are required to have a balanced budget each year. While the United States federal government does not have a balanced budget, nearly every state and many countries have stipulations that require a balanced budget. Like a private corporation, these government agencies must use financial and statistical analytics to ensure spending does not exceed income. Recently, there has been added pressure on many public entities, including the United States federal government, to analyze and reform their spending habits and maximize the allocation of public benefits.

To achieve their goals and better allocate their resources, government agencies often use cost-benefit analysis. Cost-benefit analysis is seen as a staple of analytics within the private sector, but it is a technique that actually began in the public sector. In the 1930s, the United States used cost-benefit analysis to determine which water resource projects would be prioritized and funded.

A cost-benefit analysis performed in the public sector can be quite different from cost-benefit analysis in the private sector. Many public sector projects can cost much more than the amount of money they generate. Some projects don't generate any obvious money at all, but they can still be prioritized. For example, when prioritizing water resource projects that could help control flooding, there was no clear economic benefit for the government. The benefit that the government needed to measure was the benefit to the general welfare of the population. Therefore, governments need to evaluate which projects benefit the public enough to justify their costs, even when they will earn no obvious revenue.

Example

In 2004, the state of Illinois launched a program to keep juvenile offenders in their home communities during treatment, rather than being sent to a state facility, based on a cost-benefit analysis. In 2005, it cost the state $70,827 to house each youth offender in a state detention center. Through a cost-benefit analysis, it was determined that keeping these offenders in their home communities was less expensive and decreased crime rates.

By 2010, the original cost-effectiveness analysis proved correct as every $1 million spent on the program led to a calculated cost avoidance of $3.55 million that the state would have had to pay on juvenile incarceration.

Source: Click to access "National Conference of State Legislature's Cost-Benefit Analysis of Juvenile Justice Programs."

In the military, a cost-benefit may not be easily evaluated, so a cost-effectiveness analysis is often used. In a cost-effective analysis, a quantifiable goal is determined, such as recognition of an enemy assault. To conduct the cost-effective analysis, the cost of achieving the predetermined goal is analyzed.

In using analytics within the public sectors, many different interests need to be taken into account. While corporations can have a somewhat unified goal of delivering the best goods and services and maximizing revenues, governments often have less clear or competing goals within itself. The process of prioritizing some people's best interests over others is often controversial. For example, when constructing interstate highways in the late twentieth century, the United States government often prioritized the speed and ease of automobile travel over the impacts to the environment and the residents of the neighborhoods which it passed through. With changing attitudes, the public sector may now use different statistics to analyze whether a public project is deemed worthy.

Analytics can be vital to a public agency's success. While the data used and analysis performed may be different from the private sector, the public sector needs analytics to succeed in today's world. While a new set of challenges can arise when using data in government, those equipped with analytical skills can use data to help address many of the world's dilemmas.

5.11 Case Study: Public Sector Analytics

Costs, Benefits and Decision Making in a City's Efforts to Improve Environmental Sustainability

Rich Bragan is the Director of the City of Northboro's Department of Public Works (DPW). Among his many responsibilities, he has been charged with implementing the City's ambitious new Sustainability Plan that aims to incorporate environmental sustainability into the design construction, operation, maintenance and renovation of the town's facilities and buildings. The Mayor and City Council established a Sustainability Task Force that has worked with Bragan to evaluate and develop cost estimates for potential projects that could reduce greenhouse gas emissions at facilities. One recent project the team implemented involved investment in lighting improvements at various town offices that resulted in an annual savings of 1.3

million kWh and $171,000.

Currently the task force is focusing on Northboro's massive wastewater treatment facility (WWTF) that is the city's largest energy consumer. To support the task force's mission to use more cost-efficient renewable energies throughout the city, Bragan proposed the installation of a rooftop solar array that would provide 860 kW of electricity for the facility. The WWTF has a large rooftop that is ideal for solar panel placement, and a new roof installation on the building is scheduled within the next two years. This presents an ideal opportunity to coordinate the installation of a solar array with the necessary roofing project.

Bragan summarized the benefits of installing a solar energy system at the WWTF:

1. A solar installation establishes a secondary source of power, projected to generate more than 1,000,000 kWh per year.
2. Solar generated power will provide over 1,000 tons of CO_2 offset per year.
3. The WWTF is in a highly visible location adjacent to a major highway, with over 300,000 people passing the facility on a daily basis. The project will create positive public perception and demonstrate the City administration's commitment to environmental and fiscal responsibility.
4. Use of a renewable energy source will avoid the price volatility of fossil fuel markets.
5. Beyond initial installation and maintenance, solar energy requires no expensive raw materials or transportation and requires lower operational costs, contributing to reduced overall utility costs.

Recommended Solution

Despite the many compelling benefits of pursuing a rooftop solar installation at the WWTF, the task force is concerned that the substantial up-front investment will be a tough sell for the City Council and ultimately the taxpayers. To address this concern, Bragan prepared a cost analysis to present to the task force for consideration.

He also pointed out that the cost of solar panels has decreased significantly (approximately 60%) in the last five years due to advancements in photovoltaic (PV) cell technology and the increase in the number of solar providers. Based on initial scope estimates, the WWTF requires approximately 3,500 240-watt PV cells to provide approximately 860 kW of electricity. The expected cost of this installation is $2-3 million, with a payback period of approximately 5.4 years.

The payback on the system would come from two sources:

1. Savings in traditional energy consumption amounting to roughly $150,000 per year
2. Solar Renewable Energy Certificates (SRECs)

In states that have a Renewables Portfolio Standard (RPS), regulated electricity suppliers must include 15% renewable energy in the electricity they sell. Each time a solar energy system (such as the one proposed for the WWTF) produces 1,000 kWh of electricity, one SREC is awarded to the facility. Those SRECs can then be sold on a market to allow electric suppliers to buy these certificates in order to meet their solar RPS requirements. In other words, SRECs operate as a kind of currency that translates into revenue that offsets the solar investment costs.

Bragan estimates that the proposed WWTF solar energy system could generate 1,055 SRECs per year, at a value of approximately $130 per SREC, for a total of about $140,000 per year.

In summary, Bragan believes there is significant opportunity to experience an attractive return on a solar installation investment, both in financial terms and in terms of shifting to cleaner, more environmentally sustainable technologies. He also feels this investment will generate positive PR among the public that is generally interested in cleaner, more sustainable energy usage.

Like many municipalities, Northboro's budget is constrained and must be balanced. Getting approval for major capital expenditures is challenging, involving competition with other projects and extensive scrutiny of financial details. Often the Capital Committee members make little effort to measure the non-financial benefits of proposed projects, weighing them strictly with ROI analysis or payback periods. Bragan needs to break down all of the many costs and benefits involved in the project and convince them that this is an investment worth considering on many levels.

Review Checkpoint

Question 1. List out the financial and non-financial benefits and costs associated with the proposed solar panel installation at Northboro's Wastewater Treatment Facility.

Question 2. Bragan recognizes that the project has significant merit from both a sustainability perspective and a long-term financial perspective. The challenge is clearly the short term investment needed to fund the installation. Considering that the City's Capital Planning and Investment Committee (a subcommittee of the Mayor's Budgetary Office) is ultimately responsible for making decisions about large capital investments, how do you think Bragan can influence their decision in favor of the project?

Suggested/Sample Responses

Question 1:

Financial Costs

- $2-3 million investment in solar panel system
- Operating costs

Financial Benefits

- Reduced fuel energy consumption/costs
- Avoid price volatility of fossil fuels
- Ability to sell SRECs to electricity providers, source of revenue
- Solar technology prices have come down significantly

Non-Financial Benefits

- Cleaner energy source better for environment, significant reduction in CO_2 emissions
- Positive public perception
- Reduce reliance on non-renewable energy sources

Question 2:

It is important for him to understand how the committee bases its decisions and who or what influences the committee members. If they are appointed by the Mayor, it makes sense to work to get the Mayor's support for the project so that he/she can influence their thinking. If they are elected by taxpayers, then Bragan might consider launching a public information campaign through local media and/or advertising to convince residents of the financial and environmental benefits of the solar project so that they will weigh in with their opinions, their approval ratings and voting decisions.

Another important initiative would be to look for alternative financing options for the solar energy system investment. Bragan should look at all possible grant options and investigate third-party financing arrangements that would fund some or all of the up-front costs in exchange for the future annual savings realized.

He should repeatedly remind City officials that they created and committed to the Sustainability Plan. He should emphasize that this project is consistent with the Plan's objectives and will go a long way to establishing credibility and results.

5.12 Flashcards

Flashcards

Term	Definition

5.13 Module 5 Printable PDF

This assignment does not contain any printable content.

5.15 Module Feedback

This assignment does not contain any printable content.