

The following content is provided under a Creative Commons license. Your support will help MIT OpenCourseWare continue to offer high quality educational resources for free. To make a donation, or to view additional materials from hundreds of MIT courses, visit MIT OpenCourseWare at [ocw.mit.edu](http://ocw.mit.edu).

**JOHN GUTTAG:** Hello, everybody. Some announcements. The last reading assignment of the semester, at least from us. Course evaluations are still available through this Friday. But only till noon. Again, I urge you all to do it.

And then finally, for the final exam, we're going to be giving you some code to study in advance of the exam. And then we will ask questions about that code on the exam itself. This was described in the announcement for the exam.

And we will be making this code available later today. Now, I would suggest that you try and get your heads around it. If you are confused, that's a good thing to talk about in office hours, to get some help with it, as opposed to waiting till 20 minutes before the exam and realizing you're confused.

All right. I want to pick up where we left off on Monday. So you may recall that we were comparing results of KNN and logistic regression on our Titanic data. And we have this up using 10 80/20 splits for KNN equals 3 and logistic regression with  $p$  equals 0.5.

And what I observed is that logistic regression happened to perform slightly better, but certainly nothing that you would choose to write home about. It's a little bit better. That isn't to say it will always be better. It happens to be here.

But the point I closed with is one of the things we care about when we use machine learning is not only our ability to make predictions with the model. But what can we learn by studying the model itself? Remember, the idea is that the model is somehow capturing the system or the process that generated the data. And by studying the model we can learn something useful.

So to do that for logistic regression, we begin by looking at the weights of the different variables. And we had this up in the last slide. The model classes are "Died" and "Survived."

For the label Survived, we said that if you were in a first-class cabin, that had a positive impact on your survival, a pretty strong positive impact. You can't interpret these weights in and of

themselves. If I said it's 1.6, that really doesn't mean anything.

So what you have to look at is the relative weights, not the absolute weights. And we see that it's a pretty strong relative weight. A second-class cabin also has a positive weight, in this case, of 0.46.

So it was indicating you better had a better-than-average chance of surviving, but much less strong than a first class. And if you are one of those poor people in a third-class cabin, well, that had a negative weight on survival. You were less likely to survive.

Age had a very small effect here, slightly negative. What that meant is the older you were, the less likely you were to have survived. But it's a very small negative value. The male gender had a relatively large negative gender, suggesting that if you were a male you were more likely to die than if you were a female. This might be true in the general population, but it was especially true on the Titanic.

Finally, I warned you that while what I just went through is something you will read in lots of papers that use machine learning, you will hear in lots of talks about people who have used machine learning. But you should be very wary when people speak that way. It's not nonsense, but some cautionary notes.

In particular, there's a big issue because the features are often correlated with one another. And so you can't interpret the weights one feature at a time. To get a little bit technical, there are two major ways people use logistic regression.

They're called L1 and L2. We used an L2. I'll come back to that in a minute. Because that's the default in Python, or in [INAUDIBLE].

You can set that parameter at L2 and do that to L1 if you want. I experimented with it. It didn't change the results that much. But what an L1 regression is designed to do is to find some weights and drive them to 0.

This is particularly useful when you have a very high-dimensional problem relative to the number of examples. And this gets back to that question we've talked about many times, of overfitting. If you've got 1,000 variables and 1,000 examples, you're very likely to overfit.

L1 is designed to avoid overfitting by taking many of those 1,000 variables and just giving them 0 weight. And it does typically generalize better. But if you have two variables that are

correlated, L1 will drive 1 to 0, and it will look like it's unimportant.

But in fact, it might be important. It's just correlated with another, which has gotten all the credit. L2, which is what we did, does the opposite. It spreads the weight across the variables. So have a bunch of correlated variables, it might look like none of them are very important. Because each of them gets a small amount of the weight.

Again, not so important when you have four or five variables, is what I'm showing you. But it matters when you have 100 or 1,000 variables. Let's look at an example.

So the cabin classes, the way we set it up,  $c_1 + c_2 + c_3$ -- whoops-- is not equal to 0. What is it equal to? I'll fix this right now. What should that have said? What's the invariant here?

Well, a person is in exactly one class. I guess if you're really rich, maybe you rented two cabins, one in first and one in second. But probably not. Or if you did, you put your servants in second or third. But what does this add up to? Yeah?

**AUDIENCE:** 1.

**JOHN GUTTAG:** Has to add up to 1. Thank you. So it adds up to 1. Whoa. Got his attention, at least.

So what this tells us is the values are not independent. Because if  $c_1$  is 1, then  $c_2$  and  $c_3$  must be 0. Right? And so now we could go back to the previous slide and ask the question well, is it that being in first class is protective? Or is it that being in second or third class is risky?

And there's no simple answer to that. So let's do an experiment. We have these correlated variables.

Suppose we eliminate  $c_1$  altogether. So I did that by changing the init method of class passenger. Takes the same arguments, but we'll look at the code. Because it's a little bit clearer there.

So there was the original one. And I'm going to replace that by this. combine that with the original one.

So what you see is that instead of having five features, I now have four. I've eliminated the  $c_1$  binary feature. And then the code is straightforward, that I've just come through here, and I've just enumerated the possibilities.

So if you're in first class, then second and third are both 0. Otherwise, one of them is a 1. So my invariant is gone now, right? It's not the case that we know that these two things have to add up to 1, because maybe I'm in the third case. OK, let's go run that code and see what happens.

Well, if you remember, we see that our accuracy has not really declined much. Pretty much the same results we got before. But our weights are really quite different. Now, suddenly,  $c_2$  and  $c_3$  have large negative weights. We can look at them side by side here.

So you see, not much difference. It actually performs maybe-- well, really no real difference in performance. But you'll notice that the weights are really quite different. That now, what had been a strong positive weight and relatively weak negative weights is now replaced by two strong negative weights.

And age and gender change just a little bit. So the whole point here is that we have to be very careful, when you have correlated features, about over-interpreting the weights. It is generally pretty safe to rely on the sign, whether it's negative or positive.

All right, changing the topic but sticking with logistic regression, there is this parameter you may recall,  $p$ , which is the probability. And that was the cut-off. And we set it to 0.5, saying if it estimates the probability of survival to be 0.5 or higher, then we're going to guess survived, predict survived. Otherwise, deceased.

You can change that. And so I'm going to try two extreme values, setting  $p$  to 0.1 and  $p$  to 0.9. Now, what do we think that's likely to change?

Remember, we looked at a bunch of different attributes. In particular, what attributes do we think are most likely to change? Anyone who has not answered a question want to volunteer?

I have nothing against you, it's just I'm trying to spread the wealth. And I don't want to give you diabetes, with all the candy. All right, you get to go again.

**AUDIENCE:** Sensitivity.

**JOHN GUTTAG:** Pardon?

**AUDIENCE:** The sensitivity and specificity.

**JOHN GUTTAG:** Sensitivity and specificity, positive predictive value. Because we're shifting. And we're saying, well, by changing the probability, we're making a decision that it's more important to not miss survivors than it is to, say, ask gets too high.

So let's look at what happens when we run that. I won't run it for you. But these are the results we got. So as it happens, 0.9 gave me higher accuracy. But the key thing is, notice the big difference here.

So what is that telling me? Well, it's telling me that if I predict you're going to survive you probably did. But look what it did to the sensitivity. It means that most of the survivors, I'm predicting they died.

Why is the accuracy still OK? Well, because most people died on the boat, on the ship, right? So we would have done pretty well, you recall, if we just guessed died for everybody.

So it's important to understand these things. I once did some work using machine learning for an insurance company who was trying to set rates. And I asked them what they wanted to do.

And they said they didn't want to lose money. They didn't want to insure people who were going to get in accidents. So I was able to change this  $p$  parameter so that it did a great job.

The problem was they got to write almost no policies. Because I could pretty much guarantee the people I said wouldn't get in an accident wouldn't. But there were a whole bunch of people who didn't, who they wouldn't write policies for. So they ended up not making any money. It was a bad decision.

So we can change the cutoff. That leads to a really important concept of something called the Receiver Operating Characteristic. And it's a funny name, having to do with it originally going back to radio receivers. But we can ignore that.

The goal here is to say, suppose I don't want to make a decision about where the cutoff is, but I want to look at, in some sense, all possible cutoffs and look at the shape of it. And that's what this code is designed to do. So the way it works is I'll take a training set and a test set, usual thing.

I'll build one model. And that's an important thing, that there's only one model getting built. And then I'm going to vary  $p$ .

And I'm going to call apply model with the same model and the same test set, but different p's and keep track of all of those results. I'm then going to plot a two-dimensional plot. The y-axis will have sensitivity. And the x-axis will have one minus specificity.

So I am accumulating a bunch of results. And then I'm going to produce this curve calling `sklearn.metrics.auc`, that's not the curve. AUC stands for Area Under the Curve. And we'll see why we want to get that area under the curve.

When I run that, it produces this. So here's the curve, the blue line. And there's some things to note about it.

Way down at this end I can have 0, right? I can set it so that I don't make any predictions. And this is interesting. So at this end it is saying what?

Remember that my x-axis is not specificity, but 1 minus specificity. So what we see is this corner is highly sensitive and very unspecific. So I'll get a lot of false positives. This corner is very specific, because 1 minus specificity is 0, and very insensitive.

So way down at the bottom, I'm declaring nobody to be positive. And way up here, everybody. Clearly, I don't want to be at either of these places on the curve, right?

Typically I want to be somewhere in the middle. And here, we can see, there's a nice knee in the curve here. We can choose a place. What does this green line represent, do you think?

The green line represents a random classifier. I flip a coin and I just classify something positive or negative, depending on the heads or tails, in this case. So now we can look at an interesting region, which is this region, the area between the curve and a random classifier. And that sort of tells me how much better I am than random.

I can look at the whole area, the area under the curve. And that's this, the area under the Receiver Operating Curve. In the best of all worlds, the curve would be 1. That would be a perfect classifier.

In the worst of all worlds, it would be 0. But it's never 0 because we don't do worse than 0.5. We hope not to do worse than random. If so, we just reverse our predictions. And then we're better than random. So random is as bad as you can do, really.

And so this is a very important concept. And it lets us evaluate how good a classifier is

independently of what we choose to be the cutoff. So when you read the literature and people say, I have this wonderful method of making predictions, you'll almost always see them cite the AUROC.

Any questions about this or about machine learning in general? If so, this would be a good time to ask them, since I'm about to totally change the topic. Yes?

**AUDIENCE:** At what level does AUROC start to be statistically significant? And how many data points do you need to also prove that [INAUDIBLE]?

**JOHN GUTTAG:** Right. So the question is, at what point does the AUROC become statistically significant? And that is, essentially, an unanswerable question. Whoops, relay it back. Needed to put more air under the throw. I look like the quarterback for the Rams, if you saw them play lately.

So if you ask this question about significance, it will depend upon a number of things. So you're always asking, is it significantly better than  $x$ ? And so the question is, is it significantly better than random?

And you can't just say, for example, that 0.6 isn't and 0.7 is. Because it depends how many points you have. If you have a lot of points, it could be only a tiny bit better than 0.5 and still be statistically significant.

It may be uninterestingly better. It may not be significant in the English sense, but you still get statistical significance. So that's a problem when studies have lots of points.

In general, it depends upon the application. For a lot of applications, you'll see things in the 0.7's being considered pretty useful. And the real question shouldn't be whether it's significant, but whether it's useful. Can you make useful decisions based upon it?

And the other thing is, typically, when you're talking about that, you're selecting some point and really talking about a region relative to that point. We usually don't really care what it does out here. Because we hardly ever operate out there anyway. We're usually somewhere in the middle. But good question. Yeah?

**AUDIENCE:** Why are we doing 1 minus specificity?

**JOHN GUTTAG:** Why are we doing 1 minus specificity instead of specificity? Is that the question? And the answer is, essentially, so we can do this trick of computing the area.

It gives us this nice curve. This nice, if you will, concave curve which lets us compute this area under here nicely if you were to take specificity and just draw it, it would look different. Obviously, mathematically, they're, in some sense, the same right. If you have  $1 - x$  and  $x$ , you can get either from the other. So it really just has to do with the way people want to draw this picture.

**AUDIENCE:** [INAUDIBLE]?

**JOHN GUTTAG:** Pardon?

**AUDIENCE:** Does that not change [INAUDIBLE]?

**JOHN GUTTAG:** Does it not--

**AUDIENCE:** Doesn't it change the meaning of what you're [INAUDIBLE]?

**JOHN GUTTAG:** Well, you'd have to use a different statistic. You couldn't cite the AUROC if you did specificity directly. Which is why they do  $1 -$  minus.

The goal is you want to have this point at 0 and this 0.00 and 1.1. And playing  $1 -$  minus gives you this trick, of anchoring those two points. And so then you get a curve connecting them, which you can then easily compare to the random curve. It's just one of these little tricks that statisticians like to play to make things easy to visualize and easy to compute statistics about. It's not a fundamentally important issue. Anything else?

All right, so I told you I was going to change topics-- finally got one completed-- and I am. And this is a topic I approach with some reluctance. So you have probably all heard this expression, that there are three kinds of lies, lies, damn lies, and statistics.

And we've been talking a lot about statistics. And now I want to spend the rest of today's lecture and the start of Wednesday's lecture talking about how to lie with statistics. So at this point, I usually put on my "Numbers Never Lie" hat.

But do say that numbers never lie, but liars use numbers. And I hope none of you will ever go work for a politician and put this knowledge to bad use. This quote is well known. It's variously attributed, often, to Mark Twain, the fellow on the left.

He claimed not to have invented it, but said it was invented by Benjamin Disraeli. And I prefer to believe that, since it does seem like something a Prime Minister would invent. So let's think



about this. The issue here is the way the human mind works and statistics.

Darrell Huff, a well-known statistician who did write a book called *How to Lie with Statistics*, says, "if you can't prove what you want to prove, demonstrate something else and pretend they are the same thing. In the daze that follows the collision of statistics with the human mind, hardly anyone will notice the difference." And indeed, empirically, he seems to be right.

So let's look at some examples. Here's one I like. This is from another famous statistician called Anscombe. And he invented this thing called Anscombe's Quartet. I take my hat off now. It's too hot in here.

A bunch of numbers, 11 x, y pairs. I know you don't want to look at the numbers, so here are some statistics about them. Each of those pairs has the same mean value for x, the same mean for y, the same variance for x, the same variance for y.

And then I went and I fit a linear regression model to it. And lo and behold, I got the same equation for everyone,  $y$  equals  $0.5x$  plus 3. So that raises the question, if we go back, is there really much difference between these pairs of x and y?

Are they really similar? And the answer is, that's what they look like if you plot them. So even though statistically they appear to be kind of the same, they could hardly be more different, right? Those are not the same distributions.

So there's an important moral here, which is that statistics about data is not the same thing as the data itself. And this seems obvious, but it's amazing how easy it is to forget it. The number of papers I've read where I see a bunch of statistics about the data but don't see the data is enormous. And it's easy to lose track of the fact that the statistics don't tell the whole story.

So the answer is the old Chinese proverb, a picture is worth a thousand words, I urge you, the first thing you should do when you get a data set, is plot it. If it's got too many points to plot all the points, subsample it and plot of subsample. Use some visualization tool to look at the data itself.

Now, that said, pictures are wonderful. But you can lie with pictures. So here's an interesting chart.

These are grades in 6.0001 by gender. So the males are blue and the females are pink. Sorry for being such a traditionalist. And as you can see, the women did way better than the men.

Now, I know for some of you this is confirmation bias. You say, of course. Others say, impossible, But in fact, if you look carefully, you'll see that's not what this chart says at all.

Because if you look at the axis here, you'll see that actually there's not much difference. Here's what I get if I plot it from 0 to 5. Yeah, the women did a little bit better. But that's not a statistically-significant difference. And by the way, when I plotted it last year for 6.0002, the blue was about that much higher than the pink. Don't read much into either of them.

But the trick was here, I took the y-axis and ran it from 3.9 to 4.05. I cleverly chose my baseline in such a way to make the difference look much bigger than it is. Here I did the honest thing of put the baseline at 0 and run it to 5. Because that's the range of grades at MIT.

And so when you look at a chart, it's important to keep in mind that you need to look at the axis labels and the scales. Let's look at another chart, just in case you think I'm the only one who likes to play with graphics.

This is a chart from Fox News. And they're arguing here. It's the shocking statistics that there are 108.6 million people on welfare, and 101.7 with a full-time job. And you can imagine the rhetoric that accompanies this chart.

This is actually correct. It is true from the Census Bureau data. Sort of. But notice that I said you should read the labels on the axes.

There is no label here. But you can bet that the y-intercept is not 0 on this. Because you can see how small 101.7 looks like. So it makes the difference look bigger than it is.

Now, that's not the only funny thing about it. I said you should look at the labels on the x-axis. Well, they've labeled them. But what do these things mean?

Well, I looked it up, and I'll tell you what they actually mean. People on welfare counts the number of people in a household in which at least one person is on welfare. So if there is say, two parents, one is working and one is collecting welfare and there are four kids, that counts as six people on welfare.

People with a full-time job, is actually does not count households. So in the same family, you would have six on the bar on the left, and one on the bar on the right. Clearly giving a very different impression.

And so again, pictures can be good. But if you don't dive deep into them, they really can fool you. Now, before I should leave this slide, I should say that it's not the case that you can't believe anything you read on Fox News. Because in fact, the Red Sox did beat the St. Louis Cardinals 4 to 2 that day.

So the moral here is to ask whether the things being compared are actually comparable. Or you're really comparing apples and oranges, as they say. OK, this is probably the most common statistical sin. It's called GIGO. And perhaps this picture can make you guess what the G's stand for GIGO is Garbage In, Garbage Out.

So here's a great, again, quote about it. So Charles Babbage designed the first digital computer, the first actual computation engine. He was unable to build it.

But hundreds of years after he died one was built according to his design, and it actually worked. No electronics, really. So he was a famous person. And he was asked by Parliament about his machine, which he was asking them to fund.

Well, if you put wrong numbers into the machine, will the machine have right numbers come out the other end? And of course, he was a very smart guy. And he was totally baffled.

This question seems so stupid, he couldn't believe anyone would even ask it. That it was just computation. And the answers you get are based on the data you put in. If you put in garbage, you get out garbage.

So here is an example from the 1840s. They did a census in the 1840s. And for those of you who are not familiar with American history, it was a very contentious time in the US. The country was divided between states that had slavery and states that didn't. And that was the dominant political issue of the day.

John Calhoun, who was Secretary of State and a leader in the Senate, was from South Carolina and probably the strongest proponent of slavery. And he used the census data to say that slavery was actually good for the slaves. Kind of an amazing thought. Basically saying that this data claimed that freed slaves were more likely to be insane than enslaved slaves.

He was rebutted in the House by John Quincy Adams, who had formerly been President of the United States. After he stopped being President, he ran for Congress. From Braintree, Massachusetts. Actually now called Quincy, the part he's from, after his family.

And he claimed that atrocious misrepresentations had been made on a subject of deep importance. He was an abolitionist. So you don't even have to look at that statistics to know who to believe. Just look at these pictures.

Are you going to believe this nice gentleman from Braintree or this scary guy from South Carolina? But setting looks aside, Calhoun eventually admitted that the census was indeed full of errors. But he said that was fine. Because there were so many of them that they would balance each other out and lead to the same conclusion, as if they were all correct.

So he didn't believe in garbage in, garbage out. He said yeah, it is garbage. But it'll all come out in the end OK.

Well, now we know enough to ask the question. This isn't totally brain dead, in that we've already looked at experiments and said we get experimental error. And under some circumstances, you can manage the error.

The data isn't garbage. It just has errors. But it's true if the measurement errors are unbiased and independent of each other.

And almost identically distributed on either side of the mean, right? That's why we spend so much time looking at the normal distribution, and why it's called Gaussian. Because Gauss said, yes, I know I have errors in my astronomical measurements.

But I believe my errors are distributed in what we now call a Gaussian curve. And therefore, I can still work with them and get an accurate estimate of the values. Now, of course, that wasn't true here.

The errors were not random. They were, in fact, quite systematic, designed to produce a certain thing. And the last word was from another abolitionist who claimed it was the census that was insane. All right, that's Garbage In, Garbage Out.

The moral here is that analysis of bad data is worse than no analysis at all, really. Time and again we see people doing, actually often, correct statistical analysis of incorrect data and reaching conclusions. And that's really risky. So before one goes off and starts using statistical techniques of the sort we've been discussing, the first question you have to ask is, is the data itself worth analyzing? And it often isn't.

Now, you could argue that this is a thing of the past, and no modern politician would make

these kinds of mistakes. I'm not going to insert a photo here. But I leave it to you to think which politician's photo you might paste in this frame.

All right, onto another statistical sin. This is a picture of a World War II fighter plane. I don't know enough about planes to know what kind of plane it is.

Anyone here? There must be an Aero student who will be able to tell me what plane this is. Don't they teach you guys anything in Aero these days? Shame on them. All right.

Anyway, it's a plane. That much I know. And it has a propeller. And that's all I can tell you about the airplane.

So this was a photo taken at a airfield in Britain. And the Allies would send planes over Germany for bombing runs and fighters to protect the bombers. And when they came back, the planes were often damaged.

And they would inspect the damage and say look, there's a lot of flak there. The Germans shot flak at the planes. And that would be a part of the plane that maybe we should reinforce in the future.

So when it gets hit by flak it survives it. It does less damage. So you can analyze where the Germans were hitting the planes, and you would add a little extra armor to that part of the plane. What's the flaw in that? Yeah?

**AUDIENCE:** They didn't look at the planes that actually got shot down.

**JOHN GUTTAG:** Yeah. This is what's called, in the jargon, survivor bias. S-U-R-V-I-V-O-R.

The planes they really should have been analyzing were the ones that got shot down. But those were hard to analyze. So they analyzed the ones they had and drew conclusions, and perhaps totally the wrong conclusion.

Maybe the conclusion they should have drawn is well, it's OK if you get hit here. Let's reinforce the other places. I don't know enough to know what the right answer was. I do know that this was statistically the wrong thing to be thinking about doing.

And this is an issue we have whenever we do sampling. All statistical techniques are based upon the assumption that by sampling a subset of the population we can infer things about the population as a whole. Everything we've done this term has been based on that. When we

were fitting curves we were doing that. When we were talking about the empirical rule and Monte Carlo Simulation, we were doing that, when we were building models, with machine learning, we were doing that.

And if random sampling is used, you can make meaningful mathematical statements about the relation of the sample to the entire population. And that's why so much of what we did works. And when we're doing simulations, that's really easy. When we were choosing random values of the needles for trying to find pi, or random value if the roulette wheel spins. We could be pretty sure our samples were, indeed, random.

In the field, it's not so easy. Right? Because some samples are much more convenient to acquire than others. It's much easier to acquire a plane on the field in Britain than a plane on the ground in France.

Convenient sampling, as it's often called, is not usually random. So you have survivor bias. So I asked you to do course evaluations. Well, there's survivor bias there. The people who really hated this course have already dropped it.

And so we won't sample them. That's good for me, at least. But we see that. We see that with grades. The people who are really struggling, who were most likely to fail, have probably dropped the course too.

That's one of the reasons I don't think it's fair to say, we're going to have a curve. And we're going to always fail this fraction, and give A's to this fraction. Because by the end of the term, we have a lot of survivor bias. The students who are left are, on average, better than the students who started the semester. So you need to take that into account.

Another kind of non-representative sampling or convenience sampling is opinion polls, in that you have something there called non-response bias. So I don't know about you, but I get phone calls asking my opinion about things. Surveys about products, whatever. I never answer. I just hang up the phone.

I get a zillion emails. Every time I stay in a hotel, I get an email asking me to rate the hotel. When I fly I get e-mails from the airline. I don't answer any of those surveys. But some people do, presumably, or they wouldn't send them out.

But why should they think that the people who answer the survey are representative of all the people who stay in the hotel or all the people who fly on the plane? They're not. They're the

kind of people who maybe have time to answer surveys. And so you get a non-response bias. And that tends to distort your results.

When samples are not random and independent, we can still run statistics on them. We can compute means and standard deviations. And that's fine. But we can't draw conclusions using things like the Empirical Rule or the Central Limit Theorem, Standard Error. Because the basic assumption underlying all of that is that the samples are random and independent.

This is one of the reasons why political polls are so unreliable. They compute statistics using Standard Error, assuming that the samples are random and independent. But they, for example, get them mostly by calling landlines. And so they get a bias towards people who actually answer the phone on a landline. How many of you have a land line where you live?

Not many, right? Mostly you rely on your cell phones. And so any survey that depends on landlines is going to leave a lot of the population out. They'll get a lot of people of my vintage, not of your vintage. And that gets you in trouble.

So the moral here is always understand how the data was collected, what the assumptions in the analysis were, and whether they're satisfied. If these things are not true, you need to be very wary of the results. All right, I think I'll stop here.

We'll finish up our panoply of statistical sins on Wednesday, in the first half. Then we'll do a course wrap-up. Then I'll wish you all godspeed and a good final. See you Wednesday.