GABRIEL SANCHEZ-MARTINEZ: So today, we'll talk about cost estimation. It's the first of the modeling lectures. We'll talk very briefly about the role-- the different roles that cost models have in transit agencies. And then we'll discuss three types of-- three categories of cost models.

So all right, let's get into it. So what are the roles? How could cost models be used? And to start, a cost model is a model, a mathematical model that takes some variables and predicts some costs for the agency. It could be operational cost. It could be fixed cost.

So how could these be used? The first row is predicting the cost change associated with a service change. So this is a very traditional typical example. So you might extend the route as we are thinking about in this assignment, Route 1770A, extending it. And you might wonder, how much will that cost to operate? So you could use a cost model in that situation.

Or maybe there's a wider increase in service. Maybe service is being expanded from ending at 2 AM to running 24 hours. How much would that cost? So these are the kinds of questions, service changes that you might want to have an idea of how much will it cost me.

So they are often concerned with marginal or incremental costs, especially in this application. So if you expand service to cover all night, for example, you may not need that-- that change may not trigger the need for a new bus maintenance facility, or some large capital expenditure. So we're talking here about per-- incremental cost. As I add buses, or as I add bus hours, or as I add driver hours, how much extra do I pay?

You have different results over different time periods. This is important. It doesn't cost the same to run service off-peak or at night as it does on-peak. I will talk more about that later in this lecture. But having that in mind, and having a cost model that is sensitive to time periods might be very important to get accurate estimates.

OK, so these models can also be used for routine performance monitoring and service policy triggers. You might remember from the short-term planning lectures that there are some

processes where every bus route might be reviewed. And you might decide that you need to add service based on its performance.

Or you might decide that a bus route isn't meeting the requirements to be financially sustainable. So if you are considering removing, dropping service [INAUDIBLE], how much will you save by doing that? All of those are applications of bus mode.

OK, the second point here is that you can predict the cost change associated with the change in a production process. Now, we're not talking so much about service, but we're thinking of, what happens if I'm in a labor agreement that does not allow part timers, and now has negotiated the ability to hire part timers? What will that do to cost, sort of cost structure?

What about if I want to contract out maintenance work? I'm doing that right now in-house. And now maybe for a portion or for all of it, we are looking at public-private partnerships to contract out maintenance work.

Or contracting out suburban bus routes-- for instance, I have low ridership. And they're kind of on the outskirts. You don't want to drop service, but maybe you don't need big 40-foot buses and you can run them with smaller buses. Could you contract those out? What will that do to cost?

New fare technology, the MBTA-- I don't know who was here a long time ago. We had tokens here. And we moved to the CharlieCard. And now the MBTA is thinking about the next generation of fare collection for Boston.

AUDIENCE:	When was the CharlieCard invented?
GABRIEL SANCHEZ-MARTINEZ:	I don't remember the year.
AUDIENCE:	2005.
GABRIEL SANCHEZ-MARTINEZ:	But it was-- yeah, it was the mid-2000s--
AUDIENCE:	2005.

GABRIEL SANCHEZ-MARTINEZ: --because I remember in 2004 there being tokens still.

AUDIENCE: I think the last station with tokens was Government Center in 2006.

GABRIEL SANCHEZ-MARTINEZ: So anytime something like this happens, there is a change in costs, administration costs. And the recovery of the fare recovery ratio might change. So can we model that? And can we use cost models to predict those changes in cost or in revenue?

OK, then another sort of third and very important application, not so much in Boston right now, but certainly in other jurisdictions, is the allocation of subsidy required by jurisdiction. So if you look at, for example, the WMATA in Washington, DC, there are different districts. And they all have to pay for service.

So there are many examples of this over the states and also internationally, where different jurisdictions are essentially partnering to subsidize public transportation service. And we know what the total cost of operation is. So how much is fair that each restriction pay?

According to size, according to how many people-- often, it's according to how many people ride from those jurisdictions. So you might have a data collection to figure out how many people are taking buses from each of the jurisdictions. And then you need to allocate a cost allocation model to take all of the costs, including the things that are sort of more obvious, like the bus hours that operate in those jurisdictions, but also things that are not directly connected to the jurisdictions, such as administration costs, and facility maintenance, and those things.

So all of that needs to be allocated to each jurisdiction so that the cost can be evenly or fairly split. And doing that right is critical, often, even for the decision to participate or not in this program. Question?

AUDIENCE: For jurisdictions, you mentioned Washington. Would that be like Virginia and Washington DC?

GABRIEL SANCHEZ-MARTINEZ: Yeah.

AUDIENCE: Or in New York, it would be New York-- Port Authority of New York and New Jersey?

GABRIEL SANCHEZ-MARTINEZ:	Yeah.
AUDIENCE:	Or would you also mean, like, counties?
GABRIEL SANCHEZ-MARTINEZ:	There are many examples and different structures of how different jurisdictions come together to provide service. So in some cases, you have many different operators, and more commonly in the West of the US, many different operators all providing service somehow. In other cases, you have one major operator funded by different jurisdictions.

All right, so you might remember, we've only talked about modal capacities and costs that, especially in the US, in North America, costs are divided into capital expenditures and operations expenditures. So capital costs have to do with the purchase of vehicles, or the [INAUDIBLE] of vehicles, heavy maintenance, the fixed facility construction, so construction or major repairs of trucks, garage, stations, bus stops, things like that, and then other long-term physical assets, so administration buildings, and things like that.

And then there's operating costs. So those have to do with-- they're more proportional to how much service you output, so labor wages and benefits, materials and supplies. Materials includes fuel, and tires, and electricity to run the-- power the metro. Then there's agency administration and other kinds of expenses of this nature. So it's important to keep those two in mind, because the cost models that we specify should take this into consideration and be able to allocate the capital and operating expenses separately.

All right, so there are three types of cost models that we're going to discuss. The first is a fully allocated causal factor model. It's the simplest of ones that we'll discuss. One of the problems with that model is that it allocates all of the costs to some categories. And it's not sensitive to time periods. As I said earlier, some costs vary a lot by time period.

So the second model essentially is a variation or an extension of the first one, where we do take some specific costs that are very sensitive to time and capture that, those differences, over time. And then we'll talk about incremental fixed/variable cost models. We're going to give examples of all of these using MBTA bus from the '90s as an example.

Let's start with the first one, the fully allocated causal factor model. So here's what you do. You select some causal factors. So these are not necessarily-- maybe causal factor is one of the

best term, but that's what this model is-- class is called. I would say, these are explanatory variables more than causal factors. They don't have to cause the cost directly, but they are used as explanatory variables in a model of cost.

So then you assign each expense type to each of the factors. So for example, operator wages and benefits, where would that go? If you were thinking of three kinds of explanatory variables, vehicle hours being one, vehicle miles being the second, and peak vehicles being the third, and you're going to assign every cost you incur to one of those three, then operator wages and benefits are more appropriately assigned to vehicle hours, because the pay drivers by the hour, so this is the one that it relates to most closely.

Fuel-- so fuel may be a combination of vehicle hours and vehicle miles. Or if you have to assign them to one, then vehicle miles-- you think of the fuel mileage of vehicles. So that's the one that most closely relates vehicle miles to fuel.

And then administration-- well, administration, you might want to assign it to peak vehicles, because peak vehicles is a proxy for how big is this agency. And the bigger that agency, the bigger the administration costs. So it's one way of assigning these things. So this is an example.

After you do all those assignments, you calculate the average costs for each of the factors. So you have, for example, the cost assigned to vehicle hours divided by the total number of vehicle hours. And you have those unit costs.

And then the cost model is simply each of those unit costs divided-- or multiplied by the three explanatory variables that you included. So you multiply your unit vehicle hour cost by the number of vehicle hours, your unit vehicle miles cost by vehicle miles, and your unit peak vehicles cost by peak vehicles. And that is the total cost of the agency. Questions?

AUDIENCE: This assumes that you're somehow able to split the data into those categories, to say the costs assigned to vehicle hours are x.

GABRIEL SANCHEZ-MARTINEZ: Yes, so that's one of the characteristics and perhaps weaknesses of this kind of model, that it requires the judgment of the person to assign-- to first, to generate these classes and these explanatory variables, and second, to assign them to each cost of this class or each account in a ledger to each of these explanatory variables. So But these models are used in the industry.

So if you plug in the number of vehicle hours, and vehicle miles, and peak vehicles that you

used to estimate this model, then you're just going to get the total cost of operations or of-- or total combined cost on the same period that this data came from. But more interestingly, if you are considering expanding service, and that will increase vehicle hours, vehicle miles, and peak vehicles, then this model will work in some ways to forecast what the total cost will be after you grow your agency. So this is one application of this simple model.

AUDIENCE: Is there a time unit that you're supposed to use, like a full year to capture [INAUDIBLE]?

GABRIEL SANCHEZ-MARTINEZ: This is up to you as well, another judgment call. It shouldn't be one day, but-- because you want to see all the kinds of expenses. So I think a year makes sense, a fiscal year, because you will have to put out a statement for [INAUDIBLE]. You do accounting on that fiscal year basis.

OK, so here's an example from 1996 of a bus on the MBTA. So we have those three variables that we talked about-- vehicle hours, vehicle miles, and peak vehicles. We divide each of those into variable or fixed.

So some of the expenses are considered variable, because they sort of scale up more linearly as you increase the number of vehicle hours. And others are more fixed expenses. You wouldn't expect to get extra fixed costs by adding a few buses, right?

So one example of that is, if you get 10 more buses for a new bus route, you don't need a new maintenance facility. So your maintenance facility might be considered a fixed expense. And you might want a variable cost model to predict the changes of that small change or that marginal change. So that's why we divide things into fixed and variable.

The total cost to be allocated is $173.6 million. And we divide it over these six categories, so the three variables and their fixed and variable variance. Here we have the percentages of each and the unit costs.

So now, what are possible cost models following this approach? A first and simple one, a very traditional model, is the first one listed here. So 39.82, which is the total unit cost of revenue vehicle hours-- it's the sum of 37.13 and 2.69-- times the number of revenue vehicle hours, and then 2.41, which is the sum of 2.270 and 0.14 times revenue vehicle miles times some factor. And that factor, essentially here, we have-- we're not capturing the difference between fixed and variable costs. So that's why we add them up here, 39.82 and 2.41.

The other characteristic of this model is that it does not have peak vehicles as an explanatory variable. So we address that with this adjustment factor. If you take the total cost, which is 173.6, and you divide it by the sum of the unit costs for revenue vehicle hours and revenue vehicle miles, you get 1.26. And that's the adjustment factor. So any questions--

AUDIENCE: Sorry--

GABRIEL SANCHEZ-MARTINEZ: --with that approach?

AUDIENCE: You said the 1.26 is calculated by doing what?

GABRIEL SANCHEZ-MARTINEZ: Let me write it down for you. So you have a total of 173.6. And that's a total. And you divide it by the total of the total number of unit costs that you are capturing with explanatory variables, which is 29 plus 5.7 plus 50--

AUDIENCE: Plus 3.

GABRIEL SANCHEZ-MARTINEZ: --plus 3.

AUDIENCE: [INAUDIBLE]

AUDIENCE: And last [INAUDIBLE]--

GABRIEL SANCHEZ-MARTINEZ: --equals 1.261.

AUDIENCE: [INAUDIBLE]

GABRIEL SANCHEZ-MARTINEZ: So we're dividing the total by the amount that we did capture. And then that factors out of the model. So this is a very simple model of a full annual cost for the agency. Here we're looking at the fiscal year. And it's sensitive to two variables. It's not sensitive to time period. And it's mixing fixed and variable costs. So this model may not be the best one for some applications.

AUDIENCE:	But just, what's the intuition behind using vehicle hours and vehicle miles?
GABRIEL SANCHEZ-MARTINEZ:	It's something that you-- that agencies always measure and report, so it's something that there's already a data collection effort for it. If you think about especially the traditional applications without automatic data collection, there is an effort to measure anything you deliver to collect how much output you produce. And so we already have that.
	It's something that then you can use as an explanatory variable. And it relates to service. So it's your service output. These are variables that are also going to be found in any service funding proposal in some way or another. If you extend routes, then that increases both of these.
AUDIENCE:	I guess my other question was, like, what's the intuition behind vehicle hours being relatively more expensive than vehicle miles?
GABRIEL SANCHEZ-MARTINEZ:	That depends on how you assign each of these cost categories to it.
AUDIENCE:	It's not that, like, operating a bus implies--
GABRIEL SANCHEZ-MARTINEZ:	Well, I mean, in this case, wages are-- wages and benefits are a principle or one of the key--
AUDIENCE:	Got it, OK.
GABRIEL SANCHEZ-MARTINEZ:	--one of the key costs for the agency. And that's being reflected entirely in vehicle hours. And potentially fuel is a second one. And that is at least partially covered by running vehicle hours.
	It could be. I'm not sure in this model if it is. I suspect that everything was assigned to vehicle miles, but it could be. So yeah, vehicle hours is a proxy for more things. More costs are being assigned to that variable. Eli?
AUDIENCE:	Can you give some examples of variable and fixed costs?
GABRIEL SANCHEZ-	Yeah, we'll actually see more of those later in the lecture. Fuel is variable, because as you drive more and you consume more-- and if you all of a sudden operate less service, then you'll

MARTINEZ: see that drop immediately. A fixed cost could be the administration costs for, let's say, your service planner team.

So you've hired people. And if you drop service a little bit, you're not necessarily going to fire people. Or if you expand service, you're not necessarily going to hire a bunch of people. Or you may not need new offices for them. So all of those things are more fixed.

AUDIENCE: Thanks.

GABRIEL SANCHEZ-MARTINEZ: OK, second model, so here we have the same first part of the model, but now, instead of using the adjustment factor, we directly take the unit cost of peak vehicles and add it in. So now we have a model that is sensitive to peak vehicles as well. And we don't need the adjustment, because all of the things that have been assigned to peak vehicles are being directly related to peak vehicles. So it's a simple variation.

And here, we have a variable annual cost model. So these first two have all of the costs, fixed and variable. But for some applications, you're only interested in variable annual cost. And I've given the example of minor expansion in service, where you add a little bit of service or remove a little bit of service.

And you wouldn't expect to incur fixed costs as a result. So here, we then only consider the unit costs of the variable portion of revenue vehicle hours, which is 37.13, and 2.57 for revenue vehicle miles. Any questions on these simple models?

OK, as I mentioned before, one of the key weaknesses of these models is that they're not sensitive to time periods. So it's much more expensive to operate sort of on the peak than it is off-peak. So if you were to use, say, variable annual cost model for bus routes 70 and 78 and now you want to expand it to go to Kendall, or Lechmere, or somewhere else, and you have a greater number of revenue vehicle hours than your revenue vehicle miles, if you-- maybe if you expand service proportionally throughout the peak and off-peak, you're fine. But if you only expand service in the peak or do it disproportionately peak/off-peak, then this model will not capture those differences very well. And for that, we need temporal variational, which is our next--

AUDIENCE: Just a second, but why wouldn't you capture it? Because suppose I had service. Suppose I had a bus in the peak. So now peak vehicles, that gets impacted.

But suppose I take fuel and I say, OK, there is fuel. Fuel affects revenue miles, but also it

affects revenue hours. And I put fuel in twice, both in the hours and the miles. Now I can also take into account the fact that the bus is sitting in traffic and burning fuel and also-- but I'm taking that into account. So why is this model necessarily [INAUDIBLE].

GABRIEL SANCHEZ-MARTINEZ: So the questions why does this model not adequately capture peak versus off-peak costs? So it goes back to this calculating average costs by factor. So these are averages. You're taking averages of all the vehicles at all times of day and dividing by-- explaining them with an explanatory variable that is not necessarily was directly causing it.

And so especially driver costs are going to vary a lot by sort of peak/off-peak. And we're taking an average. So if you increase service proportionately, you're fine. But the unit cost is not reflecting the difference between peak and off-peak.

So it's the unit cost itself more than the independent variable that you put in, which is what you're saying. You're saying, I can measure the independent variable precisely, so I can plug it in here. Well, but your factor is average.

AUDIENCE: I see.

GABRIEL SANCHEZ-MARTINEZ: So the approach of this model, which is an extension of the first one, is to do everything that we did except for driver costs, because these costs are very sensitive to time of day. And for those costs, per crew costs, we take the day. We divide it into 30-minute periods. And then we look at all runs.

Now, a run is what a driver does in a day. That's their shift. That's a sequence of trips that are being operated by a single driver. And we look at all runs and all periods. And for each period, any run that is at least this-- has at least 15 minutes in that period gets included in that period.

And then we compute the average paid per vehicle hour by dividing the daily pay by the number of vehicle hours over all runs in that period. So each of these runs is going to have a different cost. Some runs are going to operate more in the peak. And they're going to have more overtime, for example, or more spread penalties. And we'll talk about those now.

And so those runs will be more expensive. Other runs that are straight are going to be less expensive. And you can calculate an average for each 30- or 50-minute period.

So now we have an average again, but it's by time of day for a small bucket of time. And then

for each of those, we find the minimum, the average, and the maximum. So we're not only going to look at averages. We're going to also look at the distribution of costs along each period.

And this is just an equation for average. So what we have here is the sum over all runs in a specific period. This is the wage, the total paid for that run, divided by the number of hours for that run. So we compute the average.

OK, if we look at the driver requirements for the MBTA back in 1983, the early 1980s, this is what we get. So many more drivers are required in the AM peak and the PM peak than are required off-peak, or early in the morning, or late in the afternoon and evening. Why is that? We've talked a little bit about it already.

AUDIENCE: There's more service.

GABRIEL SANCHEZ-MARTINEZ: So you provide more service. Why do you need to provide more service?

AUDIENCE: Because that's when the--

AUDIENCE: More demand.

AUDIENCE: --commuters are moving.

GABRIEL SANCHEZ-MARTINEZ: Because there's more demand for it, right? So if provided the same level of service, then your buses would be too crowded and you would not deliver enough capacity. What else?

AUDIENCE: They're also in traffic more. So buses are out--

GABRIEL SANCHEZ-MARTINEZ: So every trip that you run takes longer, and stuck in traffic. And the dwell times are longer. OK, so if you need more, if you need to provide more frequent service and each run that you do, each trip that you run is slower, than that means that you have a greater cycle time and a greater frequency, and therefore, a greater vehicle requirement from your first assignment. And of course you need drivers the drive those vehicles, so your driver requirement also goes up.

Great, so what's wrong with the time? What's inconvenient about the timing of the peaks?

AUDIENCE:	They're eight hours apart.
GABRIEL SANCHEZ-MARTINEZ:	They are eight hours apart. And how long are shifts typically?
AUDIENCE:	Eight hours.
GABRIEL SANCHEZ-MARTINEZ:	Eight hours long, so if we hire someone to start their shift at 6:00 AM, they can't cover the PM peaks. They only cover the AM peak. OK, so what do we do? One option is to hire part timers.

So use full-time employees for the base. And then hire part timers for the peaks. That's the least expensive and most efficient option. But unions don't like it, because if you have a high percentage of part-time drivers, then it's not to the benefit of their union employees.

If you don't allow part timers or you cap the allowance on part timers, then the agency is required to hire full-time employees to cover the AM and the PM peak. And if you only do that, then that's terribly inefficient, because you're hiring for eight hours. Then you can only operate two or three. That's all that you need that driver for.

But there is some arrangement. You can actually have shifts that have a break in the middle of the day. And these are called spread shifts. So some of these shifts operate for two or three hours in the morning, have a long break during the day, and operate for two or three hours in the afternoon. And often, these shifts are paid a spread penalty.

So in this example, what we have in white here are shifts that last between 10 hours and 11 hours. So the spread is the difference between the check-out time and the check-in time. And shifts in white here have a spread between 10 and 11 hours. Those drivers are paid a 50% premium for any hours above 8 hours.

And then there are these other shifts that are between 11 and 13 hours. So 11, 12 in dark gray, and 12 and 13 shaded diagonally, these shifts are paid 100%. So they're paid double for any hours in excess of the-- I guess in excess of 11, in this case.

So that's going to drive costs up, because if you look at the total pay for that shift over the day, then the unit pay per hour for each of those drivers is higher. And if you look at any time slice

of the day, peaks are going to have many more of those drivers than off-peak. So the average across the drivers on any bucket of time that is 30 minutes wide is going to be higher.

AUDIENCE: Sorry, could you explain maybe for like, let's say I'm a driver that runs the shifts with 12 hours less than spread, 13 hours, the bottom one.

GABRIEL SANCHEZ-MARTINEZ: This one, yeah.

AUDIENCE: [INAUDIBLE] they're driving, the peak ones-- like, when would I work and what hours do I get paid for?

GABRIEL SANCHEZ-MARTINEZ: Excellent segue for the next slide. So here we have essentially the runs by a driver. And you can see the-- if you draw a line here, this is 8:00 AM, so AM peak. And then this is 5:00 PM right around here.

So you see that, if you draw a line between 8:00 and 5:00 PM, these are the shifts that start during the AM peak. And those shifts have the biggest spreads. So any one of these rows is an example of one run.

And if you could be one of these drivers-- so if you have this day, you check in at 6:30. You have a pretty long break during the day. And you check out at 8:00-- at 6:00 PM. So you do an hour in the PM. And you can have longer ones too. There will be some rules that prevent excessively long ones.

AUDIENCE: And when you were saying I get paid 50% more as that bus driver, do I get paid 50% more for the-- only the hours that I'm working?

GABRIEL SANCHEZ-MARTINEZ: Yes.

AUDIENCE: OK, got it.

GABRIEL SANCHEZ-MARTINEZ: And that will vary by agreement. So that's specific to this labor agreement.

AUDIENCE: And during in between the shifts?

GABRIEL SANCHEZ-MARTINEZ: So you are paid the spread penalty. And you are paid-- there are a bunch of rules. like, No matter how many hours you work, you're paid a minimum of eight. There's a swing pay bonus. So if you start your day somewhere different than you end the day, then you have to pay the person, I think, 30 minutes extra for their driving time, or whatever, for getting to the other location where you potentially left your car.

So there are compensations for all of these things. And the point here in this cost modeling lecture is that these drivers with spread, with large spreads, cost more. So we want to build a cost model that recognizes those differences, especially with regards to driver hours. So we extend service and require more driver hours more in the peak than in the off-peak.

Then we need to account for those differences. Here is a graph by time of day of the wage cost per platform hour, just by hour, divided by the base pay rate. So off-peak, you see that the average is variables to one.

But if you go to the peak, then you're paying an average of, say, 25% higher, because some of the drivers are straight shifts, but there's a mix of drivers that are spread, large spread, and being paid some spread penalties. So that drives the costs up. OK, here's an exercise to solidify this concept. Oh, before that we have a question.

AUDIENCE: Sorry, is it uncommon that the employees, they night [INAUDIBLE] price. I mean that someone starts working at night and drives some buses or trains. And they go to sleep, and after the day, restart their work in the morning?

GABRIEL SANCHEZ-MARTINEZ: So for safety reasons, a lot of labor agreements require-- and any work rules require a minimum amount of break between the end of a run and the start of the next one. And that could be 12 hours. So you know, where that person sleeps is up to them. I suppose most of them go home.

Normally the rules won't allow you to have a run that, if you end service at 5:00, you won't be able to check in at 8:00 again. If you did a full day that end at 5:00 AM, you're probably not going to be allowed to start the next full day at 8:00 AM, because that's not a long enough break. And that would be unsafe.

AUDIENCE: Yeah, I said this because in Japan, it is very common to have some place to stay and provide them. And also it enable us to run the very late night [INAUDIBLE] services. So yeah, I'm just

curious about this.

GABRIEL SANCHEZ-MARTINEZ: I didn't know about that. I'm interested to learn more about it.

AUDIENCE: Would people want to do that every day? Or would they, like, maybe expect to do that once a week?

AUDIENCE: It depends on the shift, but no one continue that kind of--

AUDIENCE: Yeah you'd--

AUDIENCE: --shift.

AUDIENCE: --live at work.

AUDIENCE: Sorry, what?

AUDIENCE: You would just live at work and do your whole life just not doing anything else.

AUDIENCE: [INAUDIBLE]

GABRIEL SANCHEZ-MARTINEZ: I mean, these are maybe cultural differences for work rules differences across cultures. I'd love to learn more about it. If you have any reference that you can send me, that would be great.

AUDIENCE: Sure.

GABRIEL SANCHEZ-MARTINEZ: Any other questions? Ari, I think you had a question.

AUDIENCE: Oh, I was going to say that based on Chicago, which runs overnight service, there is so few enough shifts that, basically, everything is [INAUDIBLE] as a straight shift. But I think when you bid for an overnight shift, you basically get all overnight shifts. You keep on that rhythm so you wouldn't have an overnight and a day. And there are some operators that really like the overnight shift, because it provides-- they can care for kids during the day, or just they like that-- like working outside the peaks.

GABRIEL SANCHEZ-MARTINEZ: I don't know what the labor rules are there, but some, personally, they also get a premium for taking the--

AUDIENCE: [INAUDIBLE]

GABRIEL SANCHEZ-MARTINEZ: There are usually premiums for running on holidays or running late at night. So the least desirable runs might have an extra premium associated to compensate for the inconvenience or the undesirability thereof.

Great, so we have an example here where we've done this exercise and we've looked at all periods of the day. We've looked at the minimum, the average, and the maximum unit costs for drivers, so screw, for peak, off-peak, and the combination of peak and off-peak. And we have six different scenarios.

And we want to choose the unit cost for each of these. So let's go one by one. So what happens if we want to increase the peak and off-peak service proportionally? Which unit cost from this matrix would you use?

AUDIENCE: Combined.

GABRIEL SANCHEZ-MARTINEZ: So we're hearing combined-- agreement, yeah? And from combining, would you take minimum, average, or maximum?

AUDIENCE: Average.

GABRIEL SANCHEZ-MARTINEZ: Average combined? Yeah, OK, and why is that? Because it's proportional on both ends. And we don't actually need the temporal variation model for this. So we can go without. What if it's a proportional decrease in both peak and off-peak? Would we use the same one?

AUDIENCE: Minimum combined?

GABRIEL SANCHEZ-MARTINEZ: Some people are saying the same one.

AUDIENCE: [INAUDIBLE]

GABRIEL SANCHEZ- MARTINEZ: You're saying minimum combined. We have one minimum combined. What's your idea for minimum combined?

AUDIENCE: [INAUDIBLE]

GABRIEL SANCHEZ- MARTINEZ: Retracting, OK.

AUDIENCE: When you say minimum cost or average cost, is that for the peak period or is that for [INAUDIBLE]?

GABRIEL SANCHEZ- MARTINEZ: So if it's peak, it will be for the peak hours. If it's off-peak, it's for off-pea, hours.

AUDIENCE: No, I mean the--

GABRIEL SANCHEZ- MARTINEZ: But minimum maximum hours, like, within the peak hours, then there are some drivers that are paid more on average and others that are paid less, because some of them are-- have spread penalties, and others don't. So even in the peaks, are straight shifts. And those are on the low end. Question?

AUDIENCE: Well, I guess, wouldn't you want to use-- like, if were going to add to the peak, is it a bad assumption to assume that it's going to be more expensive than your most [INAUDIBLE]?

GABRIEL SANCHEZ- MARTINEZ: We're getting to that. So the third one is, increases in peak period services only, so--

AUDIENCE: But I guess, even combined, like, wouldn't you want to use the max or the min because it's incremental?

GABRIEL SANCHEZ- MARTINEZ: The idea is that, if you increase things proportionally, then you're going to have the same mix that you have now but a little more each. If you add service both peak and off-peak in proportion, which means that, if there is twice as much peak service, then you're going to have twice much more peak service, then you're going to have the same distribution of spread

penalties, and therefore, the same average costs. But now we go to the next question, which is, increases in peak period services only, which of these would you use?

AUDIENCE: The max.

GABRIEL SANCHEZ-MARTINEZ: We're hearing max. And I assume for peak, max for peak?

AUDIENCE: Yeah.

GABRIEL SANCHEZ-MARTINEZ: Any other ideas?

AUDIENCE: Peak average.

GABRIEL SANCHEZ-MARTINEZ: Peak average, two ideas-- OK, let's debate.

AUDIENCE: All right, so basically, if we're increasing in the peak period only, we're going to have to do more of these spread shifts to do that.

GABRIEL SANCHEZ-MARTINEZ: Yes.

AUDIENCE: And those are our most expensive shifts.

GABRIEL SANCHEZ-MARTINEZ: Yes, are you convinced?

AUDIENCE: I'm convinced.

GABRIEL SANCHEZ-MARTINEZ: OK, good so great, so $45 right, the maximum of the peak-- what if it's a decrease in peak period service? So now we're coming back here and we're only taking the top of the top off and we're not giving anything to the off-peak [INAUDIBLE].

AUDIENCE: The inverse.

AUDIENCE: [INAUDIBLE]

GABRIEL SANCHEZ-MARTINEZ: Same, 45--

AUDIENCE: Maximum

GABRIEL SANCHEZ-MARTINEZ: --going down now, OK, great-- OK, increases in off-peak period services only, so now we're leaving the peak where it is. And we're decreasing service off-peak-- so here and maybe at night.

AUDIENCE: Increasing, you said?

GABRIEL SANCHEZ-MARTINEZ: Decreasing-- is it decreasing or increasing?

AUDIENCE: Increase.

AUDIENCE: Increase.

GABRIEL SANCHEZ-MARTINEZ: Let's start with increasing.

AUDIENCE: Then maybe minimum in off-peak?

GABRIEL SANCHEZ-MARTINEZ: OK, that's actually the high end of what it could be. So of the cells here, yes, that's the one that you should pick. You should use the minimum off-peak. But it could actually be closer to 0.

If you're depending on the cost structure and the spread penalties, you will be able to convert some of the drivers that are peak to spread-- to a non-spread shift. So you'll get straight shifts from these drivers. And that may actually compensate so much that it's a free conversion.

So what about decreases in off-peak period service? So now we are actually decreasing the off-peak and leaving everything the same. So what two things happen?

Presumably, off-peak services are straight runs. They're not spread. So if you remove those but keep the peak level operation the same, then for any of those that you remove, you have to--

AUDIENCE: Add in the most expensive--

GABRIEL SANCHEZ-MARTINEZ: --add in the most expensive ones. So what is the net effect?

AUDIENCE: Adding $15.

GABRIEL SANCHEZ-MARTINEZ: It's probably going to be close to 0 again because of this effect.

AUDIENCE: But can you explain--

GABRIEL SANCHEZ-MARTINEZ: And again, this depends on the specific costs for the labor agreement and the spread penalties and everything. But it gives you an idea. So if you only decrease off-peak service--

AUDIENCE: You're getting rid of those.

GABRIEL SANCHEZ-MARTINEZ: --then you have to remove a few of the straight runs and not pay them the full day, but you have to then get-- you have to convert some of the existing spread runs to--

AUDIENCE: [INAUDIBLE] straight.

GABRIEL SANCHEZ-MARTINEZ: --from straight to spread.

AUDIENCE: Well wait, why is it at a cost of 0 then?

GABRIEL SANCHEZ- Because you-- even though you are reducing the number of driver hours, you will increase the proportion of the most expensive ones only in the peak.

MARTINEZ:

AUDIENCE: So you might not wind up saving any money.

GABRIEL SANCHEZ-MARTINEZ: Exactly.

AUDIENCE: Where would you put it in the matrix?

GABRIEL SANCHEZ-MARTINEZ: Well, it's not in the matrix. But if anything, you would look at off-peak minimum. But it's going to be in the range of 0 to 30. And again, depending on the specific numbers, it could even be that you have to pay more. Or many things can happen.

[INTERPOSING VOICES]

AUDIENCE: --thinking it was going to be like the occupant minimum and the peak maximum.

GABRIEL SANCHEZ-MARTINEZ: Yeah, it could be. That's a good idea. But again, it depends on the specifics. And this is a cost model that isn't so precise.

AUDIENCE: You're saying, because you're not actually increasing peak service, but you're proportionally increasing it.

GABRIEL SANCHEZ-MARTINEZ: Not proportionally, we're just decreasing the off-peak and not touching the peak.

AUDIENCE: OK, but then, I mean, proportionally, more of your hours are on the peak.

GABRIEL SANCHEZ-MARTINEZ: Yes, exactly.

AUDIENCE: [INAUDIBLE]

GABRIEL SANCHEZ-MARTINEZ: Exactly, yeah. All right--

AUDIENCE:	[INAUDIBLE]
GABRIEL SANCHEZ-MARTINEZ:	So that was a good. There's a mistake in this slide. Please fix it. It says here, "see page 8." It should be, "see page 6."

OK, so from page six, the total fixed cost to be allocated is $44.6 million. And that is the sum of-- we want to allocate fixed costs. So that is the sum of all the fixed costs that appear on page six. There's 5.7, 3, and 35.9. That adds up to 44.6. And now we have here number of buses operating on the peak, on the base, on the evening, Saturday and Sunday, the number of hours-- so peak average, 4.5, 6 base hours, 4 evening hours for a total of 14.5 hours per week day, 12 hours operating Saturday, 12 hours operating Sunday.

And the objective of this model is to take the fixed costs incurred by the agency and allocate them to vehicle hours. So let's do that. We have to do this allocation.

And we can start with-- one way of looking at this is, these evening buses, 250 evening buses are also operating in the base and in the peak. Then there's an additional 125 buses that are operating only in the base and at Saturdays, during Saturdays. And then there's an additional 400 buses that are operating only in the peak during the weekday. So let's allocate costs to the 250, the additional 125, and the additional 400 according to how long they operate and how much it costs based on the unit costs that we calculated in this example.

So we start with 250 buses across all time periods. What's the share of fixed cost to be allocated to them? It's 32%. It's 250 buses divided by the total number of buses, which is 775. That's 32%. 32% of the total cost is $14.4 million. That's how much money we want to allocate to these buses.

And how many bus hours are operated in the whole year? Well, there's 250 buses. And we want to add up the hours operated during weekdays, during Saturdays, and during Sundays for the whole week.

So there are roughly 250 weekdays in a year, 58 Saturdays, and 57 Sundays. Those add up to 365 days. And there are 14.5 hours during the weekday and 12 on the weekends.

So we add those up. We multiply by the number of buses. And we get 1.25 million bus hours

for the whole year. And if we divide the $14.4 million by 1.25 million bus hours, we get an average cost per bus hour of $11.52.

And that's the base, which is the only thing that applies in the evening and Sundays. And it is a portion of what applies in Saturday, weekday days, and peak, so only a portion. So now let's move to the next 125 buses operating on all time periods except Sundays and weekday evenings.

So how much money do we have to allocate? The $44.6 million total divided by the additional 125 buses divided by the total number of buses. That's $7.2 million. How many total bus hours are operated by those 125 buses?

Well, this is 10.5 hours excluding the evening during the weekdays times 250 weekdays plus the 12 hours on Saturdays times the 58 Saturdays. So that's 0.42 million bus hours. And 7.2 divided by 0.42 is $17.14 average cost per bus hour. Do you follow? Any questions on this?

OK, now let's do the peak. That's the most expensive part. So how much do we have to allocate? It's $23 million. How do we get that? 44.6 total times 400 divided by 725.

And how many annual bus hours do we need to-- do we operate during the peak? It's 400 additional buses only in the peak times 4.5 hours of the peak times 250 weekdays. And that's 0.45 million hours. And 23 million by 0.45 is $51.11 average cost per bus hour, or per peak bus hour, in this case.

OK, so now we have unit costs for peak, for the base, and then the off-peak, for the evenings' and Sundays' operation, which is great. So the variable vehicle cost was $37.13. We had calculated that from the previous model that you can find on page six.

And now we can adjust these costs by time period, which is what we want to do. We want to have a model that is sensitive to time periods. So for Sunday evening service, we will increase the $37.13 by $11.52. For Saturday and weekday service, we want to increase the cost, the base cost of $37.13, by $13.97. And for weekday peak, we want to increase it by $32.86.

And notice that when we look at Saturday and weekday, we have to combine the two unit costs, because we have a portion of Saturday and weekday-based service is from the first 375 buses. And the portion of it is from the additional more expensive 125 buses. And when we do peak, the same thing applies.

So we start with the 250 base buses at $11.52. Then we have 125 at $17.14. And then we have the peak of the peak buses, which are 400 times the unit cost for those buses, which is much higher. And that's the total cost, the total unit cost for a weekday peak service.

Questions on this? You have a question, Eli? This is all sort of written here. And all the numbers are copied from previous slides, so you should be able to follow this on your own if you're a little lost. But if not, come see me after class, then we can talk about it.

OK, so now let's compare the models. We started out with a very traditional simple model, and the first one we gave, which was full annual cost. And it had this factor which accounted for the fact that we weren't including the peak vehicle, the costs associated with peak vehicles. So that was the first example.

We also had in that page a variable cost model, which only accounts for variable costs, so it excludes fixed costs. And now we can look at the peak period model and off-peak period model. So we make an adjustment to the first model, which-- to this variable cost model with the unit cost that we just calculated. So if we take the $37.13 of revenue vehicle hours that was in our initial variable cost model and we add the unit cost for peak revenue hours, which is $32.86 from right here-- so that's the weekday peak service unit cost-- then we get $69.99. Now that becomes our unit cost in the peak.

And we have the 2.41. Give me a second here. That comes from page six directly, I believe. Yes, so that's the combination of 2.27 and 0.14.

Because we're not counting the temporal difference on peak vehicle or revenue vehicle miles. We're only doing it for revenue vehicle hours. It's for driver cost. And driver costs were associated with hours, not miles. So that stays the same.

So this is our full annual peak cost model. And if we want to increase peak-only service, we use this model, not the other one. And it will reflect more faithfully what the costs will be.

We can do the same with off-peak service. So we take the $37.13. And we can add the combination of the portion of it that is $11.52 and the portion of it that is $13.97. Let's do that. So let me have-- let's do that on the board so that we have the example.

So for off-peak period model, we start out with the base cost of $37.13. And then we add $11.52. And that's going to apply only for the fraction of off-peak hours that are sort of Sunday and weekday evening.

So if we go back to $11.52, this is Sunday and evening service only. So the question is, how many Sunday and evening service hours are there in a whole year? The proportion is this one that I'm writing here, 1684 divided by 3,880. And I'll calculate that too in a second.

We also want to add the other cost, which is $13.97. And the proportion of that is 2,196 divide by 3,880. So that's 37.13 plus about 5 plus about 7.91. And that comes out to 50.04.

OK, so where do we get 1,684 and 2,196 from? It's similar to before, right? We just count how many hours there are. So let's just do it quickly here for the record. So let's just compute the bottom one, actually. And you'll get the idea for the others, because it doesn't make sense to do all of them.

So off peak hours-- so we have 10 off-peak hours in the weekday times 250 weekdays per year plus-- so this is weekdays-- plus 12 hours on Saturdays. And we said there are 7-- or 58 Saturdays per year-- plus 12 hours on Sundays, and we said 57 Sundays per year. And that comes out to 3,880 hours. So that's combinations of these to get the 1,684 and the 2,196. Yeah?

AUDIENCE: So for the 1,684, that's Sunday evening service, right?

GABRIEL SANCHEZ-MARTINEZ: Right, so yes, exactly. So that's Sunday and evenings, right.

AUDIENCE: And then the other is Saturday and weekend base?

GABRIEL SANCHEZ-MARTINEZ: Yes, exactly. So did people follow this allocation example? So what we're doing is taking-- dividing the day into periods. Some periods, we recognize are more expensive than others throughout.

Now, we have fixed cost. And we want to allocate them across periods. And we want to allocate more cost to the peak, because that sort of costs more. So we then use these unit costs to-- and we count how many hours we provide off-peak and during the peak. And we do things in proportions to that, essentially.

These are not the only cost models that you can come up with. These are not the only ways of allocating cost. But they provide an example of how you can go about allocating costs.

Allocating cost is a science. It's also an art. And you encounter it in agencies, because you have these huge budgets, and huge expenditures. And many projects require you to allocate to a portion or allocate costs. And if you want to have a cost model, you have to do it, because your cost model should reflect these different variables, explanatory variables.

The third and last type of model that we are going to discuss is the incremental fixed variable model. It's somewhat similar to the ones that we saw before, but now we classify costs as being variable, semi-variable, and fixed. So we have a little more

AUDIENCE: [INAUDIBLE]

GABRIEL SANCHEZ- MARTINEZ: --precision--

AUDIENCE: --[INAUDIBLE] marked by--

GABRIEL SANCHEZ- MARTINEZ: --a little more precision in how-- in what we consider fixed or variable. We also have the same explanatory variables that we had before, so vehicle hours, vehicle miles, and peak vehicles. So now, the combination of these are nine different explanatory variables.

And we do the same thing as we have done before. We have a schedule of expenses. These are our accounts in your ledger, so crew wages, vehicle servicing, fuel, tires, insurance, traffic stuff, things like publicity, and rent, and building maintenance. Everything is included here. So all of your budget is going to fall into some category. It's divided among these categories.

So now you decide, for each of these, which of these three resources apply and which cost type you consider this to be. So for example, crew wages, let's associate that with bus hours, as we did before. And that's fully variable. So that, we should feel comfortable with.

Things like building utilities, that shouldn't be variable. So that's a fixed cost. And we'll associate that with peak buses, because peak buses is the one that we say is a good proxy for how big the agency is. So peak buses is, again, like agency size, if you will.

So then there are things in the middle. So publicity, well, semi-variable, because if you-- so the extent that you increase service, there is going to be some amount of additional publicity. It's not going to be necessarily a linear proportion increase with as you-- you know, we add one more bus, then you have that much more marketing or publicity expense. Some of those

expenses are having an office and staff, so some salaries, so that portion is fixed. But then there is extra space that you have to cover on each additional bus, so that part is variable.

So now we have a semi-variable cost. And you can associate that, in this case, with peak buses, because for each peak bus you have, then that's how many buses you have to plaster with ads. That gives you an example.

So you can go through each of these. Again, this requires judgment, and therefore, each person may do it slightly differently. But you use a schedule. And you then build your-- you calculate your unit costs, much like we did with the earlier examples. And you then apply the model.

But now we have more precision. So if you have some kind of new expense, or you decide you're going to save some money somehow, then you decide, what-- how am I going to spend less or spend more? Would that be a variable, semi-variable, or fixed cost? Would I associate that with vehicle hours, vehicle miles, or peak vehicles?

And then you would use that unit cost to bring the total costs up or down from the agency. Even more detailed models, you could just go by expense category and do your engineering-style sort of budget if you want to really calculate exactly how much cost you'll save, or exactly how much extra cost you'll incur.

But these cost models are applied more-- a little more bluntly in projects where you are evaluating, especially at an early stage, what it would cost. And you have different scenarios. And you want to quickly know how much it would cost. OK, so questions on cost modeling? No questions on cost-- yeah, one question?

AUDIENCE: So in this case, you would have more reliability on the proportion of each expense assigned to each category?

GABRIEL SANCHEZ-MARTINEZ: Yeah, so you-- essentially, you're lumping things a little less. You have more variables. So your unit costs are more precise. And I think the ability to move to more sophisticated models is increasing now that everything is computer-based. [INAUDIBLE], you had a question?

AUDIENCE: [INAUDIBLE]

AUDIENCE: [INAUDIBLE]

GABRIEL SANCHEZ-MARTINEZ: Oh, the three of you had a question. No, not you? OK.

AUDIENCE: I had a question.

GABRIEL SANCHEZ-MARTINEZ: OK, Ethan.

AUDIENCE: Back on Slide 11, you said it's possible that reducing driver hours in off-peak periods might have no effect or could even potentially actually increase costs.

AUDIENCE: It could-- well, yeah, it shouldn't increase. It shouldn't increase costs.

AUDIENCE: Well, I was going to ask, is there-- is it ever the case that off-peak service, when there is really low demand, exists simply for the convenience of scheduling shifts rather than meeting maximum headways, or anything like that?

GABRIEL SANCHEZ-MARTINEZ: Potentially, but because driver costs are not the only cost of operating service, you would then also agree to spend extra fuel and extra maintenance costs. And that is also a significant portion of a cost. It's not just free. Or it's only-- only this part may be free, not the whole thing.

But yeah, I think, if you find yourself in a situation where your labor agreement requires you to have-- maybe there is a cap on how much spread you can have. So you have all these straight runs. And they're only working for four hours. And then you say, well, you know, let's add service. And let's make the off-peak service very good.

So yeah, that could happen, but it's a management decision. That's not something with the science. It's a policy managerial decision. Yeah?

AUDIENCE: On the last slide when you were talking about peak buses, if we're decreasing [INAUDIBLE], is it indicative of agency size?

GABRIEL SANCHEZ-MARTINEZ: Yes. So in this example, we only have three-- well, we have nine exploratory variables, which are the combination of the three resource variables and the three cost types. But if you go back to the first example in this lecture, we had three explanatory variables-- bus hours, bus miles, and peak buses. And we have to assign everything to one of these three.

So of those three, peak buses is something that most closely relates to how large your agency is, much more so than bus hours or bus miles. And if you had other variables, then maybe those would be even closer to agency size. But in this example, we only have these three-- yeah.

And why peak buses? Well, because if you have more buses, you need a bigger parking garage for them. You need a bigger refueling and maintenance facility. You might need more mechanics and more-- right? So the many costs that are mostly fixed scale up with how many vehicles you have on your garage, and how many garages you have, and all those things.

Another detail which I want to emphasize is that, with regards to the variable and fixed, so again, if you have to incur an expense that is more marginal, then you might want to use a variable cost model, because you may not need another garage, for example, for a few extra buses. But at some point, you do incur on-- you reach an investment threshold, where, well, now you've crossed the line. And you do need an extra facility.

And you have to hire people to run the facility. And that's going to be-- that's going to increase your fixed cost. So that happens in lumpy fashion rather than variable and scaling up [INAUDIBLE].

We have an additional 10 minutes. And maybe it would be a good opportunity to discuss the first assignment, which I have not finished grading. I'm sorry about that. I hope to have that back to you soon. But we might as well go over it and just discuss the solution. So let me switch to that.

OK, so does everybody remember the first assignment? Or is that sort of back in your [INAUDIBLE]-- you had some sample data in a spreadsheet. And you had to calculate-- the first question was, calculate statistics by direction.

So this is the spreadsheet, and time of day, and running time. This data was given to you. And I like to use the indirect function. I don't know if people will know it or not, but it's convenient.

So if you type the range name in a cell, then you can compute things as a function of that range name. And that's very convenient. So I selected the range B2 to B751. And that corresponds to all the data on this, all the data on this first sheet.

And then I did the same for direction, too. So therefore, when I go to mean, I say, equals

average indirect of B4. And I'm pointing the mean to the range, which is a variable in itself. That's very neat.

So you can do that for average mean, again, min, max, and percentile, and then fill to the right, because you have a different range here. So it'll reflect that. So it's a little trick. So these are the statistics. This was very straightforward.

Then Question 4, it starts getting more at transportation itself. This requires some judgment. So we're saying, what's the running time? What's the recovery time? What's the half cycle time? And what's the cycle time?

So running time should be something that happens typically, because this is what you publish on your schedule. This is what's going to go on your journey planner. So people need to know what to expect. And therefore, running time should be based on the average or the median.

Of these statistics, that you computed before, these should be the two that you use. I usually prefer the median, because the average is affected by outliers and by heavy tails on the right quite a bit. So I usually go for the median, or the 50th percentile.

Recovery time-- so what is recovery time for?

AUDIENCE:	For making up-- well, first of all, so the drivers can go to the bathroom and things like that [INAUDIBLE]--
GABRIEL SANCHEZ-MARTINEZ:	Sure, yeah.
AUDIENCE:	--and then second of all, to make up for problems in the schedule. So say I'm driving out on this bus route. And it takes me eight minutes longer than I thought it would.
GABRIEL SANCHEZ-MARTINEZ:	I wouldn't characterize it as a problem with the schedule per se, but-- so the schedule is a deterministic decision about how long it will take you. And the reality is that the running time is a stochastic variable. It varies from run to run.

So you want to say-- you want to make a claim about what the typical running time is. But then you have to have some recovery time to account for anything later than that. And you recognize-- it's a way of recognizing that these running times are stochastic.

So the half cycle time is how much time you budget for the run, including those late ones, right? So I set that to a high percentile, the 95th, in this case. The reading was suggesting that you could use 15% times the average, or things like that. When you have data, then you should use the data.

So the technique of taking an average and multiplying it by a percentage is what the industry typically does when you don't have automatically collected data, when you essentially send someone out to drive behind the bus or to do-- or a ride checker to collect data on the bus. And you do five, or six, maybe 10 trips. And then you get an average.

And that's not enough to estimate your high percentiles, because you probably didn't see the high percentile in those only 10 trips. So therefore, you have to rely on a different method. And then you could multiply by a percentage. But if you have data, then you should use the data.

And 10%, 15%, 20%, whatever you use might not be adequate for a specific bus route. So because you have data, you can actually measure how variable it is. And therefore, high percentile is a better choice. 95% is 1 in 20, right? So if you think of there being 20 weekdays in a month, then you're saying, well, about-- for the 8:00 trip, I expect it to be late about once in a month.

AUDIENCE: [INAUDIBLE] it start out late.

GABRIEL SANCHEZ-MARTINEZ: To start out late, because it was-- yeah, for there to be a knock on effect on my next trip, essentially. And for every other time of the month, I expect it to run smoothly. So that's why 95 is a typical choice here. Lower numbers of percentiles are common, especially at high frequency service-- 90, 85. Question in the back, yeah?

AUDIENCE: So how do you use the 95th percentile for the recovery? You said for the traffic--

GABRIEL SANCHEZ-MARTINEZ: Exactly, the half cycle is the 95. And the recovery time is the difference between the half cycle and the typical. So on your journey planner, you only put the running time, but then between trips, you add the recovery time. And then the cycle time is the sum of both half cycle times. So that's pretty straightforward.

Question 5 was, how many vehicles do you need to operate this? And there is an equation, which is maybe the most important equation in the whole course-- number of vehicles equals cycle time divided by headway. And you need to round up. So cycle time is 84 minutes. The headway we said was 10 minutes. So that's 84 divided by 10. It's 8.4. Round up, that's 9

vehicles. OK, you need nine vehicles to operate the bus route.

Question 6 asks you to plot the times. Hopefully you labeled your axes and did everything like you should. And so here's time of day. And here is running times in direction one, running times in direction two.

And we identify that not all of these running times come from the same operating environment. We can see that there are some peaks and some off-peaks. And uh oh, we just calculated everything wrong, because we mixed up all the data. And we shouldn't have.

So the question was, what are these periods? So AM peak going from 7:00 to 9:00, midday from 9:00 to 16:00, and PM from 16:00 to 18:00. And we assume a 24-hour format. And now we have to redo all the work. We have to compute the same statistics.

I've just changed the ranges and copied the formulas from below, and the fill to the right. And we have all the statistics. So that's great. We repeat the calculation. And we get the vehicle requirement by period now. We still need nine vehicles in the AM and PM peaks, but it turns out that we only need seven in the base, between peaks. So that's great.

What's the cycle time? The cycle time, once you've calculated, you can flip this equation. I could have an entire lecture on this equation. Believe, it gets-- you can flip it around. And there is many ways of interpreting this equation.

But if you solve for c, cycle time equals number of vehicles times headway. Now you don't round any direction. So once you've decided what n is, then you solve this equation for c and revise your cycle time.

Because you rounded up, your cycle time will have gone up. And therefore, your recovery time will have gone up. The revised cycle time is 90 minutes, AM and PM peaks, 70 minutes off-peak. And the recovery time is the difference between the typical running time and the cycle time.

So then I asked, what if you had done the mistake of using all the data combined? That means that you're fine on AM and PM peaks, but you would be excessively provisioning resources in the off-peak. So your cycle time was 90, because you said it was nine vehicles. And that recovery time would be 30 minutes and a half, which would be excessive.

AUDIENCE: It's fine if we used the average instead of the median, right?

GABRIEL SANCHEZ-MARTINEZ:	Yeah. I prefer-- I think median is a better choice, but both are fine, especially if you remove outliers from the average. And this data didn't have outliers. As you can see on the graphs, it's very suspiciously clean data, and very precise, places where the period begins and ends. The data you will get for Assignment 3 does not look like that.
	So all right, Question 9 said, now imagine that this bus runs in a loop. It only has one turn. And I provide data. We have data here for that combined operation of running inbound, then outbound, then coming back.
	We said we have a headway of 7 minutes. And we are running with a fleet size of 10 vehicles. How reliable is this service? So we can compute from the headway the fleet size and cycle time using this equation right here. And we have 70 minutes.
	And what we want to do then is, from all these observations, you can think of this as some sort of probability distribution. We have some distribution of running time. This is probability density. And this is running time.
	And we said the cycle time is 70. So 70 will fall somewhere around here it turns out, if you look it up. That's very close to 50%. The way I computed it used the spreadsheets solver, you will see. There are different names for it.
	So I said, calculate the percentile of whatever percentile I give you here. So initially, I set it to-- I don't know, maybe I set it to 95. And it gave me the percentile. This percentile is running on the AM peak data, or whatever it is on spreadsheet 9.
	And then I said, change that percentile such that the cell equals 70. And it solved it. And it gave 0.476 probability. Is that reliable, yes or no? No.
AUDIENCE:	So what's that probability refer to?
GABRIEL SANCHEZ-MARTINEZ:	That's the probability that you can run in 70 minutes or less. And therefore, you don't cause delays on the next trip. Your next trip can begin on time. Sorry?
AUDIENCE:	Not reliable.
GABRIEL	Not reliable, thank you-- OK, so Question 10 is the challenge question. And now we have two

SANCHEZ-MARTINEZ:	variables. So now you have recovery. You start at A. You go this way. You reach B. You do recovery there. And then you run back.
	So you have recovery here and here. And we're asking about times here. So we're saying, what's the probability with the same situation of 7-minute headway and a fleet size of 10 that this trip can depart on time?
	So for that-- there are different approaches for this. How many of you tried the question? And how many of you think that they solved the question? Most of you, OK.
AUDIENCE:	Now with that answer.
GABRIEL SANCHEZ-MARTINEZ:	So it's actually not that difficult. Essentially what you have is a situation where you have different cases, right? So if you run here, and you reached terminal B early before your half cycle time is completed, then you will recover at B until you reach your half cycle time. And in that case, the only thing that matters is the running time coming back.
	But there is a case where you run here and it's late. And you depart immediately from B. So for those cases, you have to compute the probability that the combination of running this way and that way is less than 70.
	So you have those two cases. One way of doing that is to take all the running times that apply. I think this said at AM peak. So direction one is on column A. There are 200 of these. Direction two is on row one.
	So I've done the transpose and put it on a row. And then, for each of the cells, we can compute the combined running time. So notice that, for the first one, we're saying, max of 35, which is the half cycle time-- I decided here arbitrarily to split the whole cycle time of 70 minutes in half. So I'm saying the half cycle for the first run is 35 minutes.
	And so the time of departure from B is going to be the maximum of 35 for cases where it was faster than 35, or the running time if it was longer than 35. And then we add the running time in the return direction. So we combine these two distributions-- that's called the convolution-- by adding all of these up. And all of these are possibilities. All the combinations of these observations are possibilities.
	And then we repeat the same thing. So we say, let's assume this is 50%. And let's calculate the percentile on all the cells in that matrix and ask the spreadsheet to provide or solve for the

percentile that gives you 70. And the answer is 0.418, so a little over.

Then there was an 11th question, which is very important. It's not a number question here. But it said, there is this person who is watching these vehicles recover at the terminal. And this person is annoyed that service is not frequent enough. And these drivers are wasting time. And they're sitting there not working.

And that person proposes that you should run-- this person is an activist. And they've gotten on the bus. And they've measured the average time. And they say, you can run this much faster. You don't have to wait. They're budgeting for so much time, so you should run this at average with average running times.

And it's your turn to sort of argue back. So hopefully you said that, no, recovery time is important. If you don't have recovery time, then many trips will have-- will start late. And if you're trying to take your bus based on a schedule, then it'll never be there on time. And that would be-- that would mean that you have to wait a lot longer. And it will be very annoying for everyone.

So that's it for Assignment 1. I hope you understood all of this. If you have questions, let me know. I will grade it and have it back to you as soon as possible.