Mike的data visualization 3D图建议

Wed, Mar 19, 2025 4:37PM • 30:16

**SUMMARY KEYWORDS**

data visualization, 3D plot, target distance, reach distance, error rate, speed, duration, max speed, average speed, standard deviation, research question, MATLAB, heat map, polynomial regression, graduation planning

**SPEAKERS**

Speaker 3, Speaker 2, Speaker 1

**Speaker 1** 00:32

Okay, so let's see the graph. So as you suggest that I look more into the distance. So I plot the distance, and I need to say sorry, because last time I forget the correct distance. So this time, I plotted more clear. And so after transfer from the pixels, the first this plot is target distance in millimeters. So we can see that actually it's, it's, it's 100 203 100, which is 10 centimeters, 20 centimeter and 30 centimeters. So it's three clusters. What's the X axis? The what the x axis is is, it's just trials. It is Charles numbers and and the y axis is a, is a target distance in millimeters. Y

**Speaker 2** 01:48

axis is not target distance. The y axis is reach distance. Oh, no, no, you randomized

01:55

target distance. Yeah.

**Speaker 1** 01:59

Oh yeah. It's a little, it's a it's a little bit random, nice rice, yeah, so, like,

02:06

no data in there. This is your experiment.

02:12

There's no human data in that graph. That's just your experiment.

**Speaker 1** 02:17

This is human data. You told me the y axis is target, yeah. This is Yeah. You said, right, yeah, yeah. So this is totally 60 plus 1010, 240 trials here, and with six blocks, four blocks and each block say 60, twice, access again. It's just child numbers. It doesn't y axis, y x, y x is a target distance.

02:48

So there's no human data. There no

**Speaker 1** 02:51

human data, yeah, in this one no, yeah, that's just telling, yeah, yeah, yeah. And here, here, the the y axis is the same as the previous one. It, it is. It is also the target distance in millimeter. Yes, and duration

03:14

was the duration constant.

**Speaker 1** 03:18

Duration constant. Do?

03:25

What do you mean? How long

03:27

did the reach last?

**Speaker 1** 03:31

Oh, we didn't. We didn't limit that but, but you miss the lifespan of each child

**Speaker 2** 03:39

the longest time from start to finish. We're interested in reach noise as a function of reach distance and speed or duration.

03:50

Yeah, so you're

03:52

giving me reach distance, but you're not telling me duration.

03:58

I didn't put duration here, but

04:02

actually

**Speaker 1** 04:04

you can, you can refer to this one. So this one I plot the the x axis is an arrow, arrow, again, arrow, and Y axis is max speed, and I use different color to plot those three cluster of of this distance. Yeah, so, of course, there's a tendency not

04:35

so max speed

**Speaker 2** 04:39

depends on reach distance and when error does not

**Speaker 1** 04:49

for the arrow one, here's a one for Arrow. So

04:53

you had error on the previous one.

**Speaker 1** 04:57

Yes, yes. So. This one is gain error, which is a rich direction error, and this one is orthogonal error. And this one, I combine the orthogonal and the gain error. If you combine them, yeah, I combine them. So this is

05:18

a scatter plot of one against the other.

**Speaker 1** 05:22

This this is two error. Two errors. One, one is independent, yeah, and there's a 10. Indeed. There's a tendency like you see the black one black is black is less distance and more more light. So more dark, dark is a is a less distance, small distance and light, light color is a longer distance. So you can, we can see that Yeah, seems

05:56

to be a little bit of an increase in error with distance,

**Speaker 2** 06:01

yeah, which could be an effect of distance, or could be an effect of speed.

**Speaker 1** 06:08

You miss, uh, sorry, you said effect by spit, or not. Effect by spit. Well,

**Speaker 2** 06:15

you just showed me that they they tend to go faster for the longer distance reaches. Which means that the size of the error goes either with reach or with top speed, since top speed and reach go together and reach distance go together.

06:36

You mean this plot?

**Speaker 2** 06:39

Yes, that plot says further further distances give you somewhat greater errors. Previous plot says further distance gives you a higher speed, from which I conclude that higher speed, higher error and higher speed go together. So I don't know from this whether it's speed that drives error or whether it's distance that drives error. That's our question. Lisa, I think that was our question, but you don't seem to be thinking about that question. What are you doing?

**Speaker 1** 07:18

Speed drives error? Uh, you

**Speaker 3** 07:22

means

**Speaker 1** 07:26

I should combine these two plots, but, but actually I I combined these three elements here, the speed, the error and and the distance

07:43

Yes, I know that. Yeah.

**Speaker 2** 07:53

So the lighter points have more horizontal scatter,

07:59

yeah, and the lighter points are greater

**Speaker 1** 08:02

This is not our Oh, yeah.

**Speaker 2** 08:07

And the lighter points have a higher speed, yes, which means more error goes with greater distance, and more error goes with higher speed, yes. Now we'd like to know what's the rule error as a function of speed and distance, or error as a function of duration and distance? That's the rule we're trying to figure out. And unless you vary speed at a given distance, you won't know do.

**Speaker 1** 08:46

Oh, I kind of understand, so you miss, actually, I should, so I should control, like, like, the speed, like, like, use one speed. I'm

**Speaker 2** 09:01

asking, you know, I assumed you were going to repeat in your setup the experiment that xenon and Marissa ran, I don't know, a year, year and a half ago, which was to vary distance and to vary speed separately, so that for a fixed distance you would have slow reaches fast, reaches medium reaches speed. And for a long distance you would have slow reaches medium speed, speed reaches, and high speed reaches. And so then you'd have error as a function of two variables. And I'd like to see what that looks like.

**Speaker 1** 09:45

But how to control speed like, as you said, How to way they did

**Speaker 2** 09:50

it was with a demand. The way you're doing it is by having this slowly or quickly disappearing target, right?

**Speaker 1** 09:56

Yeah, ah, you miss control. In controlling the shrinking speed of the time. Yeah, this is what

**Speaker 2** 10:04

we talked about many months ago. It's like you've forgotten what the point of this experiment is. Okay, how about you? Tell me what's your research question?

**Speaker 1** 10:21

So previously, is going to see whether the speed, whether speed, will affect the error in the under the under the same distance, and means you need

10:39

to vary speed.

10:44

Yeah, correct, yeah, yeah,

**Speaker 1** 10:47

okay, but, but now we will also want to look at whether the variance of distance will affect error.

**Speaker 2** 10:58

Right? Variance of distance will affect her, Don't you mean just whether the distance will affect her. What's your question? What do you know I'm asking you, sorry you don't ask me. What's your research question? Yeah.

**Speaker 1** 11:22

So that's what I said. So the variance of distance will affect the error. What do you mean

11:28

by the variance of distance?

**Speaker 1** 11:32

Different different distance, like, like three cluster of different distance.

**Speaker 2** 11:39

Okay, so first of all, English, or maybe logic, you're interested in whether the variance of distance affects error, or whether distance of the Reach, reach length affects error. Which one

**Speaker 1** 11:58

so? So actually means like, the target distance, it's different from it's different from the reaching distance. So actually, there are two elements.

**Speaker 2** 12:14

Yeah. I mean, most reaches are pretty close to the target. So if you double the target distance on average, you're gonna double the reach distance. So I repeat, could you give me in one sentence, what's your research question?

**Speaker 1** 12:33

I've already said that I know, I know that's that's a good point. I i Yeah, I think I didn't. So here I just plot the target distance. I didn't plot the reach distance. Or

**Speaker 2** 12:48

my question, very simple question, what is your research question? What question are you trying to answer with this experiment?

**Speaker 1** 12:58

I want to see what will affect the error rate

**Speaker 2** 13:03

amongst what possible things that could affect it? What are the things that you'd like to animal analyze? So,

**Speaker 1** 13:11

yeah. So firstly, it's a speed. Speed includes the average speed and max speed, and also the distance. Distance includes the target distance and reaching distance.

**Speaker 2** 13:27

Okay, how are you going to separate target distance from reaching distance when they are extremely correlated?

**Speaker 1** 13:42

Separate. Um, I didn't think about this one. Yeah. So,

**Speaker 2** 13:51

okay, so now let's, let's now talk about this. Okay, you're interested. Let's just first. Let's skip reaching distance for the moment, and I'll come back to it in a second, but target distance and could be speed, duration, could be max speed, could be average speed, okay, so for any one of those, The result is a three dimensional graph. The X axis is target distance. The y axis is something about speed or duration. Could be duration could be average speed. Those are obviously related. It could be max speed. Okay, so those are your two independent variables, your dependent variable, the Z axis is average error,

14:51

average, absolute, average

14:52

error,

14:56

okay, average, average error

**Speaker 2** 14:58

over trial. Yes, you're gonna average, okay, so think about what that graph will look like. Look at me. You got an axis which is speed a distance. You got an axis which is something about time duration or max speed or mean speed, and then you got a surface which is error, okay, that 3d plot will have a shape. It will probably it might be a plane, it might be curved. It might have a locus where error is constant. I don't know, but that's what the output of your experiment should be. It doesn't necessarily have to be plotted that way, but that's the idea. I mean, you could plot it on a 2d plot, where you have three target distances and you have error as a function of

16:02

reach duration for three target speeds,

**Speaker 2** 16:07

or error as a function of target distance for three Max speeds, but you'd have to impose Those Max speeds with your shrinking target

**Speaker 1** 16:21

shrinking. Yeah, yeah. So, uh huh. That's

**Speaker 2** 16:26

what you're trying to achieve. And you need an experiment that will give you that 3d surface and allow you to fit functions to that 3d surface. You keep giving me plots that hopelessly mix these things together so that it's impossible to tell. I

**Speaker 1** 16:43

don't understand why the the third, which is the Z axis is, should be the average error. So actually, arrow should, should also change random. Look at

**Speaker 2** 16:54

the plot that's on the screen right now. Horizontal position is error. Yeah, yeah, any given target distance, and for any given what's the y max speed, that's a horizontal slice, and you have a bunch of dots. All that matters is, on average, how far are those dots from zero, otherwise known as standard deviation. So all I really want you to plot is not the raw single trials, but the standard deviation as a function of the independent variables you'd like to study. I'm having to read into these plots and compute standard deviations by i in order to determine it right? The green points look like they're wider than the black points. All that matters is width. You need to compute the width possibly bend. But you need to compute the width as a function of what's the color of the point and how on the eye, on the axis, is the point.

17:59

So can it be understand that.

**Speaker 1** 18:03

Now I should actually explore that under the under the under the, for example, under one distance and one speed. What's the standard deviation of the error? And the plot all

**Speaker 2** 18:18

the like to compute that. And then you'd like to plot that as a function of distance and as function of, for example, max speed. Yes, that's what you're trying to do. That's the 3d plot.

**Speaker 1** 18:28

And then, and then put all the all the speed and all the distance together

**Speaker 2** 18:36

in, oh, you don't put them together. You keep them separate. Each data point is for a fixed combination of distance and speed, and then you plot them as a function of distance and speed. Don't pull them. I don't know what you meant by put them together, but you definitely don't pool them. You plot them. Okay,

**Speaker 1** 19:02

what's the difference? What's the difference? You sell like for them, average

**Speaker 2** 19:07

together, yeah, yeah. Like binning. If you average the trials for all the distances together, then you can no longer learn about the function of distance. You got to keep them

**Speaker 1** 19:22

separate. So

**Speaker 2** 19:26

would you please try to produce a plot that actually clearly shows us something about your research question you just told me your research question is, what is error as a function of, for example, speed and target distance? So I want to see a plot that shows me for different speeds and for different target distances. What's the error? You're plotting the raw data here. I don't want the raw data. I want the error. Four which means that for a particular speed and distance, for example, the black points that are near 100 that are near have trouble reading your plot that are near 800 is that pixel, pixels per second or something.

**Speaker 1** 20:21

It's not pixel. It is transferred to the to the millimeters.

**Speaker 2** 20:28

Wait, the y axis is millimeters per second, yes, so small, I can't read it,

**Speaker 1** 20:33

yeah? Sorry. Millimeters per second. Yeah, it's millimeter.

**Speaker 2** 20:38

Yeah. Okay, so 800 millimeters per second. Black points. There are a bunch of them near the 800 line, and they have a bunch of errors. Take all those points in the slice and compute a standard deviation. Do that for the black points near 750 the black points near 700 the black points near 650, etc. Now, black means a particular reach distance. It's a huge range. That's black,

21:11

yeah.

**Speaker 2** 21:15

Anyway, the point is, and I don't love the fact that it's a huge range, because that means, you know, you're really blurring the crap out of your target distance axis. But regardless, I want to know standard deviation of error as a function of, for example, max speed and target distance. So for each max speed and target distance, I'd like to know what's the standard deviation of the endpoint.

21:54

That is your research question. So please plot it

**Speaker 2** 21:58

the way you've plotted here. You're giving me raw data, it's impossible to see what the effects are,

22:04

okay, okay, so

**Speaker 2** 22:06

go back to the drawing board and try to make some plots that make sense for your research question.

**Speaker 1** 22:15

Okay, but I think, I think I'm more understanding now that but, but I'm thinking how to, how to plot, how to plot the standard deviation of the error, and with, with the other two dimensions. Do you think I should put the SD of arrow to the z axis? Like, like, use different color you've been listening? Yes, I'm listening. So it's used like just described exactly

22:47

so by color or speed.

**Speaker 2** 22:49

Z axis is for each bin

**Speaker 1** 22:52

duration or speed, bin

**Speaker 2** 22:56

of distance and a bin of a range of speeds. So that's a square on the on the xy plane. Though you plot a data point in Z, that's the standard deviation of all those errors.

**Speaker 1** 23:12

So z you want to like, see, like the colored color. I definitely this

**Speaker 2** 23:18

is a 2d plot. I mean a plot, it's a surface above

**Speaker 1** 23:23

an X. You want to see the sweet, like, 3d plots like, like, okay, oh my god. I just want to say because I, I just used color.

**Speaker 2** 23:35

Literally, said this 10 minutes ago. I know if you didn't hear a word

**Speaker 1** 23:39

as I understand, I think color is also the Z axis. So what?

**Speaker 2** 23:47

Because color is the you use color when you have a discrete set of points that are plotted on the same plot. I'm saying, Yeah, take every single trial. Every single trial has a max speed and has a target distance and has an error. Okay, that's your raw data. Big, long list. Your raw data is a big, long list of

24:15

distance.

**Speaker 2** 24:18

I mean target distance or reach distance, whichever one max speed and signed error. Okay, that's the raw data I'm suggesting. You now collect all the data points based on a plane of speed and distance, and in that plane, you make a grid so it's got little grid boxes. In each grid box, there's a bunch of trials that landed in that grid box where the reach was in that range of distances and the speed was in that range of speeds, and they all have an error. Take all those errors and computer standard deviation. Deviation, and then above that square, plot a point at that standard deviation. So that'll give you a 3d thing. It will not lie on a nice surface. It'll be some noisy thing. And in MATLAB, you can spin it around and say, What does this surface, even though it's noisy, look like, does it look like a plane? Does it depend on the x value? Does it depend on the y value?

25:28

So it's like a heat map curve,

**Speaker 2** 25:32

such that there's a, you know, a set of things that give you the same value? My guess is it may look like a plane. If it looks like a plane, you can do linear regression in a plane through it, but first look at the data and see what they look like.

**Speaker 1** 25:48

Okay, so when you plot a 3d graph

**Speaker 2** 25:51

like that, you can actually rotate the graph in two dimensions with your mouse. MATLAB lets you do that so you can get structure from motion to allow you to see the 3d and allow you to judge perceptually whether the data look like they will be well fit by a plane.

**Speaker 1** 26:15

So the grid can also be 3d right? This three degree

26:19

is 2d you have two independent

**Speaker 1** 26:21

variables, yeah, but, but, how about those? The third axis, the

**Speaker 2** 26:25

third axis is the data standard deviation, yeah.

**Speaker 1** 26:31

So, so it's like a heat map, right?

**Speaker 2** 26:34

Well, you could plot it as a heat map. I strongly recommend you plot it as a 3d graph, because I want to see what the shape is, okay. And then you can do linear regression, which will be like fitting a plane. You can also get fancy and do polynomial regression, so you add not just y, a, z equals a, x plus b, y plus c, but you can get fancier and say Y, sorry, z equals a plus b, x plus c, y plus d x y plus e x squared, plus f x y squared. That's polynomial regression to degree two that would fit a curve. You can ask, did that curve thing fit any better than the plane that's easily answered with, for example, cross validation. Okay, so that's the analysis that makes sense here. Please do it, okay, okay, or when you have further questions about it, okay, okay,

**Speaker 1** 27:54

thank you. Actually, I did not try this plotting in MATLAB yet. But well, yeah, I will. I will do it. You're

**Speaker 2** 28:03

surrounded by people who know how to do that. But the first thing you need to do is take your raw data, Bin it by two of the independent variables, for example, max speed and distance, but another example is duration and distance, and another example is mean speed and distance, and for any one of those pairings, you get a grid that you impose on the data. First, you can plot all the data in that, in that 2d plot, ignoring error that'll tell you where the data lie. Then you put a grid over those, maybe a 10 by 10 grid, where most of the grid squares actually have a bunch of data points in them. And that means, for those bunches of data points in a grid square, you can compute an SD of error. Then you plot the 3d thing and look at it. Then you fit functions to the 3d thing, starting with a plane.

29:02

Okay. Okay, so try

**Speaker 2** 29:04

to do some of that and get back to me with some of those plots when you're ready.

**Speaker 1** 29:08

Okay, no problem. Okay, talk to you soon. Yeah, and, and also, last time you you asked about like, like for, for the thesis, right? So I want to tell you that I confirmed with the global OGs. Yeah, OGS, so it's an Office of Global student. So, so actually, I think my graduation time will be December, so it's end of the year. So, so it Okay, yeah, I my, my planning, uh, graduation time. But I think I will finish this project earlier. So maybe so. So then I will have time, yeah, so I will have time to for other speed

**Speaker 2** 29:58

has not been the quickest. Yes and so I will believe that when I see it

**Speaker 1** 30:06

Sure, yeah, okay, okay, okay, I already learned it Okay. Thank you. See you. Bye, bye.