[Moderator] 10:00:21

This is the next slide. Our distribution of movie ratings. You guys can go and

[P13] 10:00:52

A question. Like the rest four smaller graphs you have action comedy drama science fiction And the fifth one is just of what?

[P13] 10:01:07

Is it a specific genre or just like of everything combined.

[P14] 10:01:21

What did you say?

[P13] 10:01:24

No, I'm saying like the small four smaller crops, right? There is a particular genre for you know movie ratings.

[P13] 10:01:30

Like what kind of distribution of ratings also but what what's the genre of it Yeah, but the bigger graph there is no genre so is it for… does it include everything?

[P13] 10:01:41

Like all films or

[P14] 10:01:41

Yeah, I think so. I think.

[P13] 10:01:47

Then don't you think the frequency of movie should also have been different like look at the y-axis for like the small four smaller so the frequency is different because There are four different journals.

[P13] 10:02:01

Each has like you know different frequency of films But if you combine all of them together the frequency has to be different right

[P14] 10:02:10

Hmm.

[P14] 10:02:21

Yeah, no, I'm still, I also have the same question. It looks like… the biggest graph it's not the biggest graph it doesn't include everything because if it does, the sample size will be too small.

[P13] 10:02:37

Yes.

[P14] 10:02:42

And especially we can see so many zero rating movies for the four different yarns.

[P14] 10:02:50

In the biggest graph, there's almost zero

[P13] 10:02:54

Yeah, I think… either there is something wrong with the data set or someone like whoever made the graph forgot to put, you know, what kind of genre it is.

[P14] 10:03:09

Yeah, but there's a common trend in the shape of the distributions The four genres, they all have four outlier group, the zero rating people are extremely angry or extremely angry and also the shape they are all like kind of left skewed

[P13] 10:03:35

Yeah. I think, yeah, maybe because like, you know. Average because it's the votes are out of 10, right? So they're rating it out of 10, I'm guessing.

[P13] 10:03:45

So I think depending on it would make sense because most people would rate any film like most of the films even when they say average rating it always lies somewhere between what six to eight is what they consider an average rating.

[P13] 10:04:02

Probably like seven so almost every donor has the highest bar somewhere like between six and eight.

[P14] 10:04:20

Yep. And if you want to compare the different yarnas would be really hard Because they're put into four different panels.

[P14] 10:04:33

And… if you can… If somehow we don't put the histogram, if we only overlay the density and group them by different colors will be easier to compare.

[P13] 10:04:34

Yeah.

[P13] 10:04:47

That's true. Like instead of just, you know.

[P13] 10:04:53

So yeah, like. Remove all the bars histograms.

[P13] 10:04:57

All you need is just like…

[P13] 10:05:05

Yeah, the density and and maybe you can you know you can like Because if you look at you can maybe like I think you can combine all the five graphs into one This is all the histograms and you know just get the densities

[P13] 10:05:24

That should because the density is pretty much tell you that like more or less 90% of the entire story of a graph in this case and you can just minimize the space, put them all in one graph and make life easier.

[P13] 10:05:39

Instead of making five graphs.

[P14] 10:05:39

Yeah. Yeah, you use box plots or use the overlaid line uh

[P13] 10:05:45

No, like even this is fine like a line plot is like this just over take the five densities just simple line plots do them just five different color styles or like line styles with different markers and just put legend

[P13] 10:06:01

Which line is for which density which line is for which line film genre you got like pretty much of it sorted

[P14] 10:06:16

Yeah, they probably want to show um the count the con count This is very different between the movies.

[P14] 10:06:25

So…

[P13] 10:06:25

Yeah, that will be there but yeah but yeah essentially what is also like not Exactly.

[P13] 10:06:32

The difference is not too much. Because somewhere like on the left the biggest drop is 1200 to zero the ranges.

[P13] 10:06:43

The other graph for action is pretty much the same. Look at the drama one and the drama and comedy or you can say slightly different, like slightly more and science fiction is slightly less so but yeah like you can essentially fit into like

[P13] 10:07:01

All five graphs into one graph pretty easily i would say and you don't need and make it actually less confusing if you combine them.

[P13] 10:07:13

And it also makes them easier to compare. Instead of looking at five figures separately and comparing them Let's look at one.

[P13] 10:07:23

And a single figure makes it easier to compare i would say.

[P13] 10:07:32

Because I don't think we care about, you know, the actual histogram like okay what's the frequency of movies we just care about the density.

[P14] 10:07:47

Yeah.

[P13] 10:07:49

We can just like now next next slide.

[Moderator] 10:07:53

Okay, I'll go to the next slide. Yeah.