Presentation1:

Your presentations meet my imagination in the literacy rate. It’s good to learn from your slide about how to clean the unneeded data. However, I highly recommend that you can modify the label, x-axis, y-axis, and the main size to present a better resolution in your slide.(I have suggested them in presentation 2, maybe you can learn something from that). Also, on pages 9 & 10, they are two ways that you can point out which bar you want to focus on. For one thing, you can input labels and arrows to highlight one specified bar. Also, you can put the names on the x-axis in the specified country you want to figure out. There is one way I self-learned from my individual project when there have dozens of bars in a plot. You can create a column where other rows are empty, but just include something you want to print on the x-axis. As a result, when you use the code: names.arg, because other bars are empty in character, it will show the specified character you include.

Presentation2:

I like your presentation on exploratory data analysis in the predictive maintenance data set. Your presentation utilizes all the skills you learn from the course. However, I recommend you change something a little bit.

First of all, on page 4, you have better tell us the background information in slide three. For example, because it’s a very new topic and maybe viewers do not understand them very well. So, the best way is to tell us the baseline (maybe just draw a baseline to tell us if all the four models are better than the baseline or not). In the left-hand plot, it seemly that model3 is better than the others, but in the following slide, you will point out that model3 is not a good model to be used. That is why I get confused.

Second, On pages 7-8, you should make the average and failure at the same level, then you can do a comparison (try to use breaks = seq(15, 90, 5) to unify the y-axis.

Third, I recommend you slightly modify your output picture in labels, x-axis, main topic, and so on. For example, being a viewer, I can hardly see the label very well. If you use Rstudio, you can adjust the output width and height, combining with the code function, such as cex.axis= 1.5, cex.lab= 1.5, cex.main =1.5, to amplify your wording in the R curve).

Chart, histogram

Description automatically generated

Finally, you should make a short conclusion such as the below:

Table

Description automatically generated

Presentation 3:

I totally agree with your presentation. It combines all the needed parameters NBA games will use. Also, the visualization has a very good resolution that viewers like me can figure out without explanation. However, I suggest you should detail look at the regression line again. If you have included all the dots in the scatter plot on page 9, the linear regression line will not look like that. It highly deviates from the clusters and the centerline of the data. Also, you can analyze Kobe's more advanced data, such as OFFRTG, DEFRTG, and so on. Maybe in such areas, Kobe will take advantage of Michael.

Presentation4:

I like your presentation on baseball. I like to use data analytical skills in baseball because baseball is an area that provides a huge amount of raw data for analysts to use. Your data cleaning presentation is clear, but just like I say, it’s post-arranged data. Hence, it’s difficult to find rows that include missing data. You use SB as your main topic discussing if there has something different between above-average and below-average payroll. That looks cool. However, if you keep going the topic in advanced parameters, such as UBR+wSB, maybe you can realize that some teams are good at SB but not good at getting the score back. That increases my curiosity about understanding the SB trend by year in the two leagues. Because in my stereotypes, MLB runners avoid themselves doing that risky action now compared to the past, but the result is beyond my expectations.

Chart

Description automatically generated

Last, I like your conclusion, that’s why analysts can survive in this area. Finally, maybe DH rules are essential factors why NL teams have more wins than AL teams.

Presentation5:

I enjoy your opening in this presentation. You clearly describe how to get the data. However, you should include where is zone A~C. because most of us live in California, maybe some classmates will have an interest in knowing where the zone is. Also, you should take care that in a frequency plot with different datasets, be sure to make the same bar.

Presentation7:  
Interesting topic. Your presentation tells me a lot about Spotify. You clearly illustrate the box plot in energy scale with different genres. Your visualization is also good. However, I think you can plot a frequency plot or a bar plot to present how many counts in different genres. It looks like the country has a different behavior when compared to others. I guess that there are only a few counts in this genre. This strange behavior also can be found in the histogram barplot: Instrumentalness and Genre Relationship. Because if the data set is small in the ‘country’ genre, the high instrumentalness scale may come from the small data, too. As a result, the dataset will give us the wrong direction.

Presentation10:

I enjoy watching your presentation. You illustrate an interesting topic. First, you describe how you choose your data to draw a linear regression line on pages 5 and 6. Then, you compare two different equations and illustrate that there are similar to the slop when rollback the data. Second, I like your presentation on the Spine chart on page 8. It clearly tells us the different patterns in Animation and Fantasy/Sci-fi movies. It is logical, though. Finally, on page 10, your final slide tell me that action movies do not earn a lot even if they get more money. However, there still have something to improve. First, on pages 5 and 6, I highly recommend you change money types from $ to million $. As a result, the visualization will be better. I consider you can draw a profit line. That is to say, above this line, the movies make a profit whereas the movies lose money when they are below the line. Also, on page 9, if you want to make cost and profit side by side in a picture, you had better align the y-axis value. As a result, viewers will not detail check the cost and the profit (For example, the action movie costs 60 million on average and the profit is eighty. But in the barplot, people will get enough information in a quick look.

Presentation11:

Your presentation is so impressive. It is easy to understand what you are presenting. Also, your visualization graph is good even if your graph is simple. But they still have something to improve (no visualization, your presenting is what should learn and then follow). First of all, on page 7, your linear regression line looks like combining male and female data in a plot. However, in order to do a comparison, you need to include more detail, such as the female and male regression model, and P-value and R square to assist us in doubling-check the data. Second, your slide on pages 8 and 9 seems to contradict page 7. On page 7, your conclusion is happiness correlates to attractiveness. But on pages 8 and 9, Even though Asians and Blacks have an opposing effect, both of them cannot explain the curve on page 7. I recommend you detail check the logic then the presentation will be perfect.