ML TEAM - WEEK 2

Hola peeps, I hope your data exploration is going fine. So last week we cleaned the data and got a few inferences from it. Now let's make it more presentable and readable in this task.

Task 2

This task is all about data visualization. Wonder what data visualization is?? It's the pictorial representation of text or numerical data. Python has many libraries such as pandas, NumPy seaborn, matplotlib etc to visualize data. In this task, we will use plotly and seaborn to perform various techniques to explore data using plots.
As mentioned before, every question is provided with hints and reference materials, you are also free to use other resources to solve the questions.

Keep learning, and Happy coding ;D

Questions:

1. **Find the average runtime of all movies**

   (Hint: Use the .mean function on the specific column)
   Websites for reference:
   How to get the mean of a specific column in a dataframe in Python?

2. **Plot a boxplot between the release_year and revenue_adj columns.**

   (Hint: Use boxplot function from the seaborn library (release_year on the x-axis and revenue_adj on the y-axis)
   Websites for reference:
   Boxplot using Seaborn in Python - GeeksforGeeks

3. **Construct a scatter plot for the column's revenue and vote count.**

   (Hint: Use scatter function from Plotly library - Revenue on the x-axis and vote_count on the y-axis)
   Websites for reference:
   Scatter plot using Plotly in Python - GeeksforGeeks

4. **Plot a line graph depicting the number of movies released every year.**

   (Hint: Use group by function and count function.)
   Use the line function from the Plotly library
   Websites for reference:
   pandas.DataFrame.groupby — pandas 1.5.1 documentation

   Line Chart using Plotly in Python - GeeksforGeeks

5. **Find a top director(the one who has made a lot of revenue)**

   (Hint: Use group by function first and then use sort_values function.)
   Websites for reference:

   Python | Pandas Dataframe.sort_values() | Set-1 - GeeksforGeeks

6. **Plot a bar plot to find the most-rated movie.**

   (Hint: Use the bar function from the Plotly library)
   Websites for reference:

TECH DYSSEY

Python | Pandas Dataframe.plot.bar - GeeksforGeeks

THANK YOU!