



## **MGT 201: Winter 2023 (Homework #4)**

Due on Monday, April 10<sup>th</sup>, 2023 (11:59pm)

Submit the Jupyter Notebook (ipynb) file and the HTML or PDF file that shows your code and output. When you write your program, please add applicable comments with the hashtag (#) and/or with markdown cells so the grader can follow your work.

---

As an analyst for the movie industry, you are presented with a data file that contains information on over 200 movies that were released in 2006 and 2007.

The data file consists of:

- Index, Movie (i.e., title of movie), Distributor, Genre, 7 Day Gross Sales, 14 Day Gross Sales, Total US Gross Sales, International Gross Sales, US DVD Sales, and Budget

You are required to provide the following:

1. Create a dataframe that counts the number of movies by genre and distributor. Calculate the total number of movies.
2. Which distributor distributed the highest percentage of action movies? Which distributor distributed the highest percentage of adventure movies?
3. For each distributor (i.e., there are a total of 28 distributors) show the total number of movies distributed by genre. What is the total number of movies distributed by each distributor? What is the total number of movies by genre?
4. For Buena Vista, which genre represented the highest percentage of genres distributed?
5. What is the average US Gross sales for all movies for each distributor? What is the average US Gross sales for all movies for each genre? What is the average US Gross sales of all movies?
6. What is the average US DVD sales for all movies for each distributor? What is the average US DVD sales for all movies for each genre? What is the average US DVD sales of all movies?