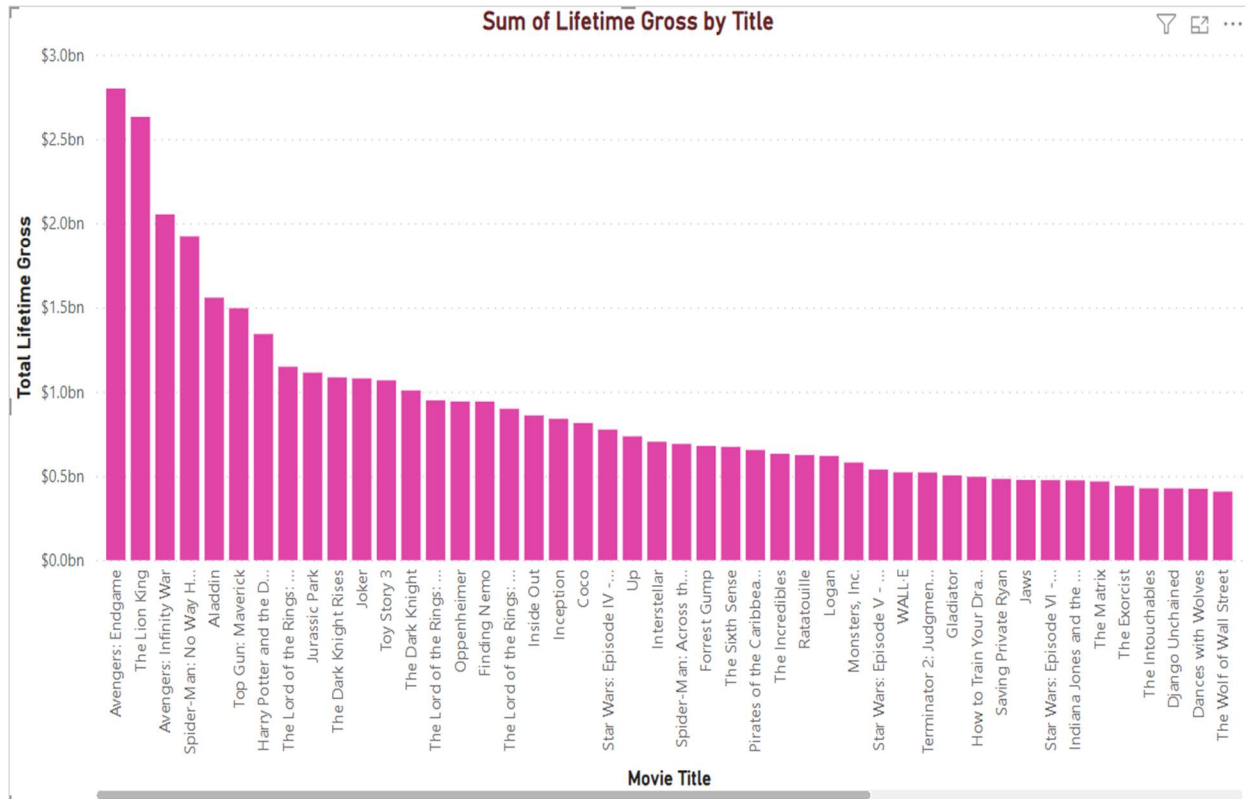


PROJECT 2

Question 1: What is the movie title with the highest lifetime gross, and what is the corresponding highest lifetime gross value?

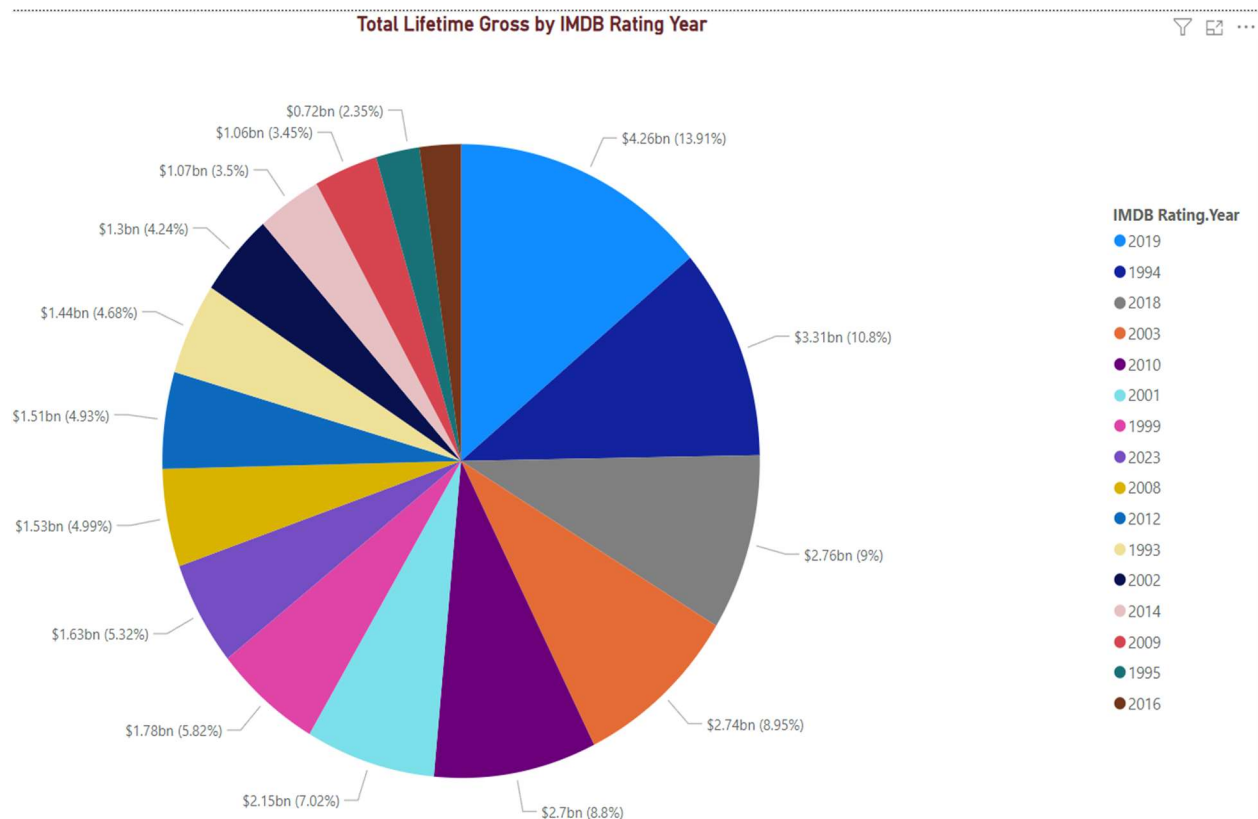


Answer:

"Avengers: Endgame", which had the highest box office receipts for the relevant years, is depicted in the above graph. For enhanced display, only a few movie titles have been selected in ascending order on the graph. Quite noticeable that action/sci-fi films usually make more than other genres. "Avengers: Endgame" is the highest-grossing movie, with an estimated box office receipt of \$2,294,805,697. To keep the visualization simple, only a few movie titles were selected based on their lifetime gross income. Among all the movies on the above list, the film "Shutter Island" made the least money. The \$294,805,697 total income for "Shutter Island" indicates a decline in demand for thriller films among younger viewers.

The entire gross lifetime revenue is displayed on the Y-axis, and the movie titles were selected based on their reducing lifetime gross. I decided to use a bar chart as the visualization technique since it effectively shows how income varies between different movies.

Question 2: What is the annual aggregate lifetime gross for all movies each year?



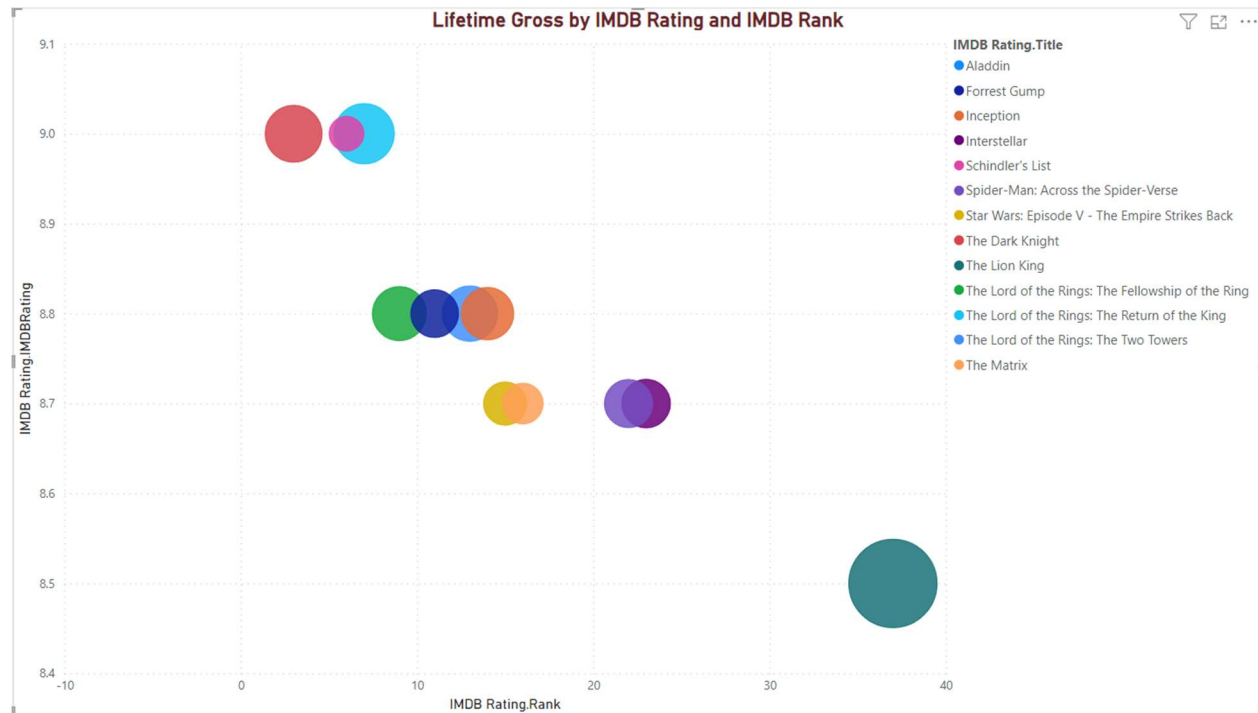
Answer:

This pie chart gives a snapshot of how much movies have earned over their lifetimes, year by year. Each slice of the pie is a different year, and the bigger the slice, the more profit movies made that year. So, the slice is bigger, it means that year had a lot of hit movies or just a bunch of films that did well.

It's a handy way to see which years movies did great in terms of earnings. From the above graph, we can conclude that for the year 2019, 3 movies made \$4.26 bn in total which is the highest of rest and so on. It could be because of big blockbuster releases, trends in the industry, or other major events.

For people in the movie business, this chart can be a guide. It will help them figure out when to release movies, how to market them, and what kind of movies to make. Basically, it's a bird's eye view of how the movie world has been doing money-wise over time.

Question 3: Is there a discernible impact on a movie's lifetime gross based on its IMDb rating and rank?



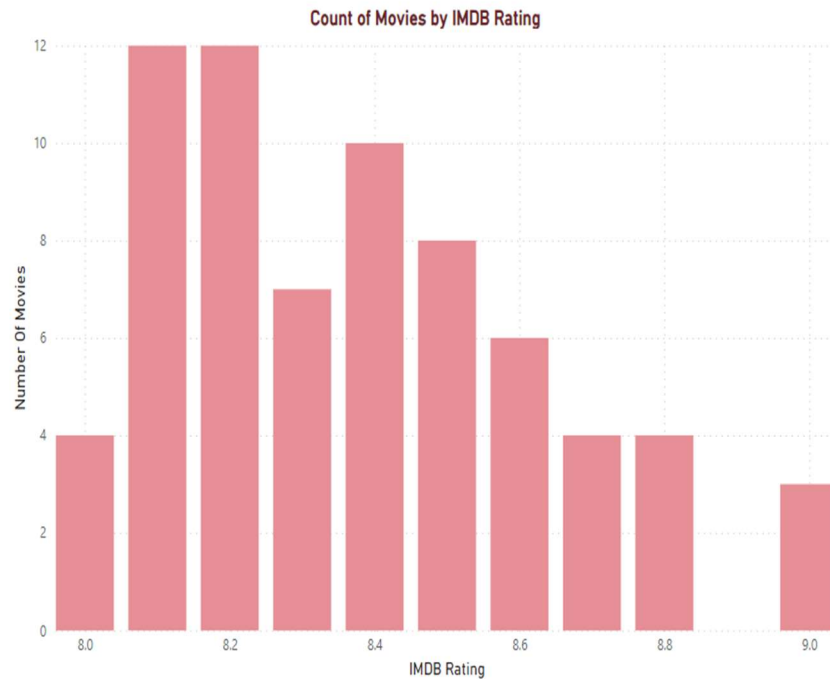
Answer:

The scatter plot visualizes the relationship between the lifetime gross, rank, and rating of movies. The horizontal axis represents the rank of the movies, while the vertical axis showcases their rating. Each point in the plot corresponds to a specific movie, with its position determined by these two variables. The size of the point can be adjusted to reflect the lifetime gross, with more vibrant or larger points indicating higher grossing movies.

We can identify clusters or trends that might suggest a correlation between rating and financial success. For instance, if most high-grossing movies are also highly rated and occupy the top-right quadrant, it indicates that audience preferences align with box office success. Conversely, movies that are outliers - highly rated but with low gross or vice versa - provide intriguing case studies for further analysis.

So, if we analyze the above graph, the movies with high ratings are having slightly high lifetime gross than the one with low rating. But the movies having the same ratings are earning almost the same, except some outliers like the 'Lion King' and 'Schlender's List'.

Question 4: How many films boasted IMDb ratings surpassing 8.5 out of 10?



Answer:

The column chart has different rating scores on the horizontal axis (from lowest to highest) and the number of movies on the vertical axis. Each column represents a specific rating, and the height of the column tells you how many movies received that rating.

In the above graph, we can easily get the count of movies which got rated 8.5 i.e., 8 movies. This chart helps us quickly understand the general quality of movies based on how frequently certain ratings are given out.

I used a column chart to visually show how movie ratings are spread out. Even though it's straightforward, it's perfect for quickly spotting the most common ratings and seeing any that stand out from the rest.

For this visualization project, I aimed towards plotting easy to analyze graphs without missing important factors. I wanted to form questions which can be asked generally and frequently when it comes to Movies. For my first visualization, I have used the clustered column chart providing a visual representation of the economic success of individual movies over their lifetime. I represented each movie with a distinct bar and its height indicating the total gross revenue. The X-axis enlists the movie titles, and Y-axis gives a range for total gross revenue. The graph facilitates a direct comparison between movies. I chose this format because it is particularly useful for stakeholders and audiences to gauge the financial performance of movies in relation to one another.

The second visualization, I plotted a pie chart which aims to represent the aggregate lifetime gross of movies made each year, providing a visual perspective on the annual contribution to the industry's revenue. Each color shade represents a year, and the tooltip gives information about number of movies contributing, total gross revenue and the year. And legend displays the color codes used for each year. The reason for presenting the data in this format is that it allows stakeholders, researchers, and the public to quickly grasp which years were particularly lucrative for the film industry and identify potential patterns or trends.

Third one, I chose a scatter plot because it offers an intuitive way to observe the relationship between two variables in this case, I picked rank, and rating against the lifetime gross of movies. By plotting each movie as an individual point determined by its rank and rating, stakeholders can discern patterns, correlations, or anomalies in the data. X-axis have rank and Y-axis notes rating given. And the size of each point speaks about the gross revenue made by that movie. Filter has been provided for TOP 10 ranked movies. This visualization is particularly helpful to understand if higher-rated movies tend to generate more revenue or if there are any outliers that defy the general trend. And the observation has been noted below the graph in this report. Finally, I utilized a column chart which allows for an intuitive visual representation of the distribution of movie ratings. Though it's simple, it's the ideal format to effortlessly discern which ratings are most prevalent and identify any outliers.

Power BI is a cool tool from Microsoft that lets us create detailed visuals from our data. It's great because we can drag and drop things, making it easy to build reports and dashboards. Plus, if we want to dive deeper into specific data points, the drill-through feature is super handy. It provides a wide range of visualizations, custom visuals, and Data model flexibility. Along with that the tool also gave an option of fetching data from web page and perform pre-processing tasks on the same. The Query Editor in Power BI is a useful tool for jobs involving data integration because it can extract and modify data from a wide range of sources. A positive point about UI aspect is without requiring any coding knowledge, users can easily design visualizations with the help of the straightforward drag-and-drop interface. And negative, have to deal with several layers of fields, filters, and images can make the interface feel crowded. Compared to the web service, Power BI Desktop offers a more sophisticated experience. Also, with a lot of data, it can sometimes slow down. It's perfect for piecing data stories together and sharing insights with teams. But all in all, to paint a picture with data, Power BI is a solid pick.