

## **1 Dataset Overview & Basic Statistics**

1. How many unique movies are in the dataset?
  2. What is the overall distribution of movie runtimes?
  3. How many movies are marked as "Adult" (isAdult=1) vs. non-adult (isAdult=0)?
  4. What are the top 10 most common movie genres?
  5. What is the distribution of averageRating? Are ratings skewed toward high or low values?
  6. How many movies have received over 100,000 votes (numVotes > 100,000)?
- 

## **2 Budget & Gross Revenue Analysis**

7. What is the distribution of movie budgets?
  8. What is the distribution of gross revenue?
  9. What is the average budget and gross revenue of movies?
  10. Is there a significant correlation between budget and gross revenue?
  11. Do high-budget movies tend to have higher ratings?
  12. Which are the top 10 highest-grossing movies?
  13. Which are the top 10 most profitable movies (profit = gross - budget)?
  14. Are movies with a higher number of votes (numVotes) more profitable?
- 

## **3 Genre-Based Analysis**

15. Which genre has the highest number of movies?
  16. What are the average budget and gross revenue per genre?
  17. Which genre has the highest average rating?
  18. Are there genres that tend to have longer runtimes?
  19. Which genre has the most audience engagement (numVotes)?
- 

## **4 Rating & Popularity Analysis**

20. What is the correlation between numVotes and averageRating?
  21. Do high-rated movies necessarily have a high number of votes?
  22. What are the top 10 highest-rated movies?
  23. What are the bottom 10 lowest-rated movies?
  24. Do longer movies tend to have higher ratings?
- 

## **5 Director-Based Analysis**

25. Which directors have directed the most movies?
26. Which directors have the highest-grossing movies?
27. Which directors have the highest-rated movies on average?
28. Are there directors known for making high-budget movies?
  
29. Are there any extreme outliers in budget, gross revenue, or rating?
30. Do outliers in budget affect movie performance?