**[Crowdfunding Dataset]**  
**Date:** November 21, 2024  
**Author:** Peanut Gallery

**Introduction**

This report presents an analysis of the Crowdfunding Dataset. The focus is on key conclusions drawn from the data, limitations within the dataset, and suggestions for improving data collection and analysis to extract greater insights.

**1. Conclusions**

1. Likelihood of Campaign Outcomes:
   * Campaign success rates are consistent across most categories, apart from Games. This is notable because it leads me to believe that while some categories may be more popular than others, overall other factors than Category or Genre contribute more towards a campaign’s success or failure.
2. Theatre Campaign Success:
   * The Plays sub-category shows a significant number of crowdfunding campaigns, with strong support from theatre enthusiasts who actively fund these projects.
3. Games Category Analysis:
   * Both sub-categories within Games show higher-than-average failure rates, an anomaly compared to other categories where success rates are typically higher.

**2. Limitations in Dataset**

1. Data Completeness:
   * The dataset has not been updated since 2020, making it difficult to account for recent trends or shifts in crowdfunding dynamics over the past four years.
2. Potential Sample Bias:
   * Theatre campaigns comprise almost one-third of the dataset. This raises concerns about potential bias in data collection, possibly influenced by the data source's audience demographics or focus.
3. Measurement Error:
   * The United States accounts for 76.3% of the data. This overrepresentation skews the dataset leaving smaller sample sizes from other countries highly sensitive to outliers and less statistically reliable.

**3. Possible Ways to Improve the Data**

1. Outcomes based on Goals:
   * Map dollar amounts of the stated goal to the statistical likelihood of getting funded. A line graph showing rising dollar amounts of goals on the X-axis and percentage funded on the Y-axis could help identify patterns relating to outcomes based on price of goal.
2. Effect of Highlighted Campaigns:
   * Analyzing how being a staff pick or spotlighted affects the likelihood of getting funded. Compare campaigns with staff pick and spotlight statuses to those without any highlights. A three-way clustered column pivot showing outcomes categorized by highlight type could reveal any correlation, providing valuable insights for potential crowdfunders.
3. Backer Volume Trends:
   * Examine the number of backers and average contributions per backer. This information could offer insight into engagement levels and the typical funding amount per supporter. We can create a scatter plot with number of backers on the X axis and average dollar amounts per backer on the Y axis. Still colored for outcomes this reveals if there is a correlation between the number of backers and average dollar amounts and how that relates to successful outcomes.