* Create a report in Microsoft Word, and answer the following questions:

The data provided will help identify which sectors the campaign succeeded or failed in, as well as the factors influencing these outcomes, such as the launch date, the country where the campaign took place, and the goals that were set.

Given the data provided, what are three conclusions that we can draw about crowdfunding campaigns?

The conclusions I derived from the data provided are mentioned below:

**Key Conclusions from the Data**

1. **People Are More Inclined to Contribute to Film and Theatre Projects**
   1. The data reveals that film and theatre projects consistently attracted a significant portion of the funds raised. These categories outperformed others, suggesting a strong public interest in entertainment, particularly in performances such as films and live theatre. This trend indicates that campaigns within these sectors have a broader appeal, as people enjoy both the cultural experience and the opportunity to support creative arts projects. Campaign creators in these sectors can capitalize on this strong interest to maximize funding and engagement.
2. **The Success of Campaigns is Influenced by the Month**
   1. The analysis shows that June and July had the highest success rates for crowdfunding campaigns, whereas December experienced the lowest. This trend appears to be influenced by seasonal factors. During December, people are more focused on personal expenses due to the holiday season, which may limit their willingness to contribute to crowdfunding campaigns. On the other hand, campaigns launched in June and July, outside of major holidays, seem to benefit from more discretionary spending. This finding emphasizes the importance of timing in campaign planning, and it suggests that campaign creators should consider launching campaigns in months with historically higher success rates.
3. **The Higher the Goal, the Lower the Success Rate**
   1. Campaigns with higher funding goals generally had lower success rates. This trend suggests that setting an overly ambitious goal can create doubt among potential backers about the feasibility of the campaign. Backers may feel that the project is unlikely to reach its target, leading to lower contributions. To ensure greater success, campaign creators should aim to set realistic, achievable goals that inspire confidence and encourage participation. Setting more modest goals that are clearly attainable is more likely to lead to a successful campaign.

**2. Limitations of the Dataset**

While the dataset provided is useful, it has some limitations:

a. **No guarantees for future success**  
Even if the campaign succeeds, the data does not offer insights into the factors that would contribute to the project's long-term success. There is no guarantee that the project will be successful in the future.

b. **Lack of backer background information**  
Crowdfunding success is often influenced by factors such as ethnicity, gender, age group, location, and income level. However, the dataset does not include this information, making it difficult to fully understand the preferences and motivations of individual backers.

c. **Absence of detailed financial breakdowns**  
The data does not provide specific details on how the funds are allocated, limiting the ability to analyze how the money is being used within the campaign.

* What are some other possible tables and/or graphs that we could create, and what additional value would they provide?

To dive deeper into the analysis, here are the tables and graphs you could create, along with their added value:

1. **Bar Graph: Income Level vs. Age vs. Money Contributed**
   * **Value**: This graph would show how contributions vary across different age groups and income levels. It would help identify which income and age brackets are most likely to back a campaign, providing clear insights into the target audience. Understanding these patterns would enable future campaign creators to tailor their marketing efforts and incentives to specific demographics more effectively.
2. **Bar Graph or Pie Chart: Total Money Contributed by Category and Subcategory**
   * **Value**: This would display how contributions are distributed across different campaign categories and their respective subcategories. It helps identify which areas of the campaign (e.g., product features, rewards, or specific causes) attract the most funding. This data would be valuable for future campaign creators in focusing on the most popular aspects and ensuring their efforts are directed toward high-interest areas.
3. **Table: Success Rate Based on Campaign Duration**
   * **Value**: This table would show the correlation between campaign duration and its success rate (e.g., the percentage of funding achieved relative to the target goal). This analysis would help us to understand whether longer campaigns tend to perform better or if a shorter, more focused campaign strategy might be more effective. It provides strategic insights into planning future campaigns and setting realistic timelines.
4. **Bar Graph: Success Rate by Country and Category**
   * **Value**: This graph provides insights into which countries have the highest success rates in specific categories (e.g., technology, art, health). It can help identify global trends and preferences, showing which regions are more likely to back certain types of projects. This information could guide campaign creators in tailoring their campaigns to regions with higher potential for success.
5. **Line Graph: Funding Over Time**
   * **Value**: This line graph would track the funding progress throughout the campaign, highlighting trends over time. It could reveal whether certain tactics (e.g., social media pushes, influencer collaborations) or events (e.g., limited time offers, milestones) impact funding significantly. Understanding these trends helps campaign creators optimize their timing and promotional strategies for maximum engagement and contribution.

* Use your data to determine whether the mean or the median better summarizes the data.

For this specific data I think the median better summarizes the data. The mean is much higher in both the successful and the failed outcome, indicating that there are some extremely high values that are affecting the mean. The median, however is not influenced by these extreme high values. So, the median being 201 in successful outcome and 115 in failed outcome should be taken into consideration for the backers count for both the cases.

* Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?

The variance and standard deviation for successful campaigns are 1606217 and 1267 respectively and for the failed campaigns it is 924113 and 961 respectively indicating that the successful campaigns have more variability in the number of the backers. The higher variance and the standard deviation means that the significant fluctuation in the number of backers across successful campaigns is likely due to a few high-performing campaigns. In some successful campaigns there are significantly large number of backers which influenced the overall variability. In the case of failed campaigns there is low variability which means most of them had fewer backers.

Yes, this makes sense. Successful campaigns tend to have more variability because some campaigns might have massive support, while others may only exceed their goals by a small margin. This high variability in the number of backers results in a larger standard deviation.

Failed campaigns, on the other hand, are generally less variable because many campaigns fail to raise much funding, so the number of backers tends to be relatively low across the board. While there are still some outliers, the overall variability is lower.