1. **Given the provided data, what are three conclusions that we can draw about crowdfunding campaigns?**
   1. ***Theatre plays*** (plays being their sole sub-category) make up the **majority** of crowdfunding campaigns, with ***journalism*** campaigns being the **rarest**. Over 75% (763) of crowdfunding campaigns originated from the ***US***, whereas campaigns do **not** occur often in ***CH*** (China?)
   2. Overall, more crowdfunding campaigns **succeed** (achieve their initial goal) than those that **failed** (approximately 60:40 ratio). The only exception is campaigns that focused on ***games*** (23 failed, 21 successful, although 3 campaigns are currently live). To date, no *journalism* campaigns have failed (at least in the current dataset).
   3. Although crowdfunding campaigns seem to **succeed** more often if they launched during *June* and less often during *August*, the outcome of campaigns remain relatively stable regardless of the month they launched.
2. **What are some limitations of this dataset?**

* No descriptions of what some of the variables mean – For example, ***staff\_pick*** and ***spotlight*** seem to have true/false cells but it is unclear what they are referring to. This makes it a redundant variable in the dataset.
* **Blurb description** seems to be **very unhelpful** in describing the ‘product’ of the crowdfunding – the blurbs also does not seem to relate to their corresponding category/sub-category description. Thus text analysis to determine more successful project topics within sub-categories would be unusable here.
* Some formatting annoyances
  + The use of **UNIX time** rather than defaulting to standard date formatting
  + Crowdfund projects **not listed in order of date released** (oldest first as Id 1 etc.)
    - Having ID number start at 0 rather than 1 may also cause issues in data cleaning – formulas may determine it as a null value.
  + **Lack** of **currency standardisation** – Perhaps all foreign projects could have their currencies be converted into USD for better ‘average donation’ comparisons.

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

* If all currencies were converted to USD and standardised, creating a graph of **average donations** and **categories** may provide information in the type of field/domain (category – like food/music/technology) that people are more willing to provide monetary support more.
  + Sub-analyses in the **outcome** of campaigns (successful/failed etc.) between donations/categories can be used as another source of information to partly determine **levels of** **interest** for a particular domain.
* Creating a table describing the **median** of ***goal*** between each ***parent category*** may help in determining whether particular categories require more funds to get started/be successful (median to account for potential outliers)
  + Sub-analysis of ***outcome*** for this can then provide a more accurate estimation of funds that crowdfunds of each category needs to be successful.