

Kickstarter Exploratory Data Analysis

Kickstarter is a website and online fundraising platform on which you can raise funds for a project or new business venture. To raise funds, you describe your project, set a goal for the amount you will raise, and ask “backers” to invest in your project. If you reach your goal, then you receive all the funds received. Otherwise, your project will not be funded, and all the money is returned back to the backers. Read the following link to better understand how the website works: <https://help.kickstarter.com/hc/en-us>. In this assignment, you will analyze data from past Kickstarter projects that attempted to raise funds.

Data Description

The dataset in this project is obtained from Kaggle (see <https://www.kaggle.com/kemical/kickstarter-projects?select=ks-projects-201801.csv>). Each observation provides information about one kickstarter project. Following is the description of the variables in the dataset.

IMPORTANT: Use the dataset that is available on elearning, and do not directly use the dataset from the above link. The dataset on elearning has been cleaned for your convenience.

- ID: Project ID Number
- Name: Project name
- Main Category: Project main category (decided by project creators)
- Category: Project subcategory within main category (decided by project creators)
- Launched: When project was launched on Kickstarter
- Deadline: Deadline by which fundraising goal should be met
- Goal: Project fundraising goal
- Pledged: Amount of money pledged by the deadline
- State: Whether project was successful (reached fundraising goal), failed (did not reach fundraising goal), or cancelled (cancelled early by the project creator)
- Backers: Number of people who have contributed to the project by paying an amount of money
- Country: Where the project is located
- USD_pledged_real: conversion in US dollars of the pledged column
- USD_Goal_real: conversion in US dollars of the goal column

Assignment Description

In this assignment, you will conduct exploratory analysis to develop insights that could be managerially relevant. The overall objective is to understand how fundraising projects have been performing on the platform, what factors that may be influencing funding success and money raised, and identifying interesting trends, patterns or other insights that could be of interest to managers in this setting. Do spend time understanding how this platform works before you analyze the data. You may also read some online articles about it. This assignment is mainly about exploratory and descriptive analysis and simple hypothesis tests. You are not required to build more advanced models.

When exploring the data and identifying insights, ask yourself “so what” and “who will be interested in knowing the answer”. Insights could be relevant for the project creator trying to raise funds, for kickstarter as the platform, or even for the backer to know which projects to back. This will help you understand the practical relevance of the insight. Discuss these aspects in your report along with your findings and insights.

Here are some examples of insights that could be managerially relevant:

- From Kickstarter's perspective, which categories are more successful in reaching the goal? Which categories raise more money? Which categories are performing poorly? What can you say about the subcategories?
- Which categories have more competition (more projects attempting to raise funds at the same time)? What role does competition play in success of projects being funded?
- Do you notice significant differences in the projects and performance across countries? What are the similarities and differences across countries? Do you notice significant trends over time? (note: do you think this data can be used to understand the variation across countries, and time trends – why / why not?)
- What guidance can you provide a project creator about how to choose a deadline, or what goal amount to set? If the project creator is trying to decide which category or subcategory to develop a new project in, what would you recommend?
- How do projects that are successful differ from projects that fail? How does this vary with the project category?
- How does the success of funding depend on how many similar projects have been funded recently?

IMPORTANT: The above are only examples. These are not mandatory questions to answer. Since this is exploratory analysis, by its very nature the project is open-ended. The above questions are only to provide you some ideas. You need to explore data and present your most interesting findings. The above questions may or may not lead to interesting insights. The grading will be based mainly on the quality of the analysis and results, including your creativity in identifying questions and conducting analysis, and how you present and discuss the results. So do not blindly keep adding analysis and results – it is not about the quantity or number of pages / graphs / figures.

Project Deliverables

A well-written report with graphs and tables as suitable and necessary, describing the key insights or results from your analysis and their practical relevance. Present three to five key or most interesting insights from your analysis. Each insight might be backed by multiple different analysis, tables, figures etc. as may be relevant. The scope of this project is exploratory and descriptive analysis and simple hypothesis tests. You are not required to build more advanced models to back your insights. However, you can mention any limitations of your exploratory analysis in fully verifying the insight.

Follow the following structure for your report:

1. Executive summary (2 – 3 paragraphs): Summarize the main insights from your analysis
2. Introduction: Describe the platform and how it works, what you found interesting about the platform and some key questions you were interested in exploring with this data.
3. Insights – Describe three to five key insights from your analysis. Show how each insight is supported by the data using suitable tables, graphs, simple analysis etc. If relevant, mention any limitations of the current analysis and what further analysis could be conducted. (Do not include code or other technical details here. This is a portion that must be readable by a manager.)
4. Appendix – Provide details about how you carried out each step of the analysis presented in the main report. For example, some steps may have been performed in SAS, others in R etc. This needs to be clearly documented and backed by relevant program code or other details. You can include the program code in this appendix.