**Title:** Project 2, Team 3

**Team:** Kimberly Childers, Joshua Hale, and Damarje Brown

**Project Description:** Project 2 required team members to work together and explore how to extract data utilizing Python and Pandas, transform and clean data, parse string data, place data in dictionaries, use lists to display readable code, and manipulate strings using regular expressions. Team 2 accomplished this by creating Category, Subcategory, Campaign, Contacts, and Crowdfunding data frames.

**Analysis:**

1) Category data frame was created. Reference category data frame file "category.csv" located in the “Resources (CSV Files)” folder on GitHub. The team created a “category\_id” and “category” column from the “crowdfunding.xlsx” data provided.

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

2) Subcategory data frame was created. Reference subcategory data frame file "subcategory.csv." The team created a “subcategory\_id” and “subcategory” column.

A table with a list of words

Description automatically generated

3) Campaign data frame was created. Reference campaign data frame file "campaign.csv." Columns were renamed from “launched\_at” and “deadline” to “launch\_date” and “end\_date.” Then both UTC times were converted to the datetime format.

A screenshot of a computer

Description automatically generated

4) Contact data frame was created. Reference contact data frame file "contacts.csv." The team chose to use option one which used Python dictionaries as opposed to regular expressions. Each row of the provided data was converted into a dictionary and pandas was used to make a data frame listing the contact\_id, name, and e-mail. The image below shows the names column split into first and last name.

A screenshot of a computer

Description automatically generated

5) Crowdfunding database was created. Reference crowdfunding database file "crowdfunding\_db." QuickDBD was used to an ERD of tables.

A screenshot of a computer

Description automatically generated

The team finally made a Postgres file and verified that all tables were created. All those files can be found on the GitHub folder titled “Screen shots of verifying data in pgAdmin.” The file contains screenshot and downloads of tables from pgAdmin4.

**Future Implications:**

Team 2 could utilize this data for future analysis to discover any trends such as creating another database to group successful campaigns by their category to observe categories with higher percentages. Automating scripts to pull campaign and contact data continuously would depict any trends over time on successful companies and assist with forecasting results. Finally, team 2 could improve the coding and make it more robust by accounting for decimal numbers. Data inputs currently characterized as integers should be changed to a float because monetary pledges will not always be whole number.