Springboard Capstone Project

Help DonorsChoose.org connect donors with projects they care about
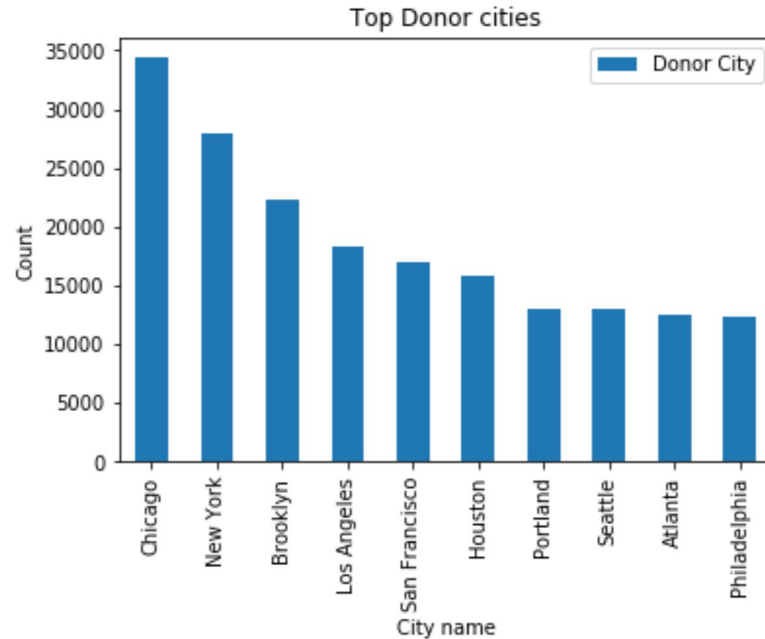
By Reena Turak

Challenge

❖ Founded in 2000 by a Bronx history teacher, DonorsChoose.org has raised $685 million for America's classrooms.

❖ Teachers at three-quarters of all the public schools in the U.S. have come to DonorsChoose.org to request what their students need, making DonorsChoose.org the leading platform for supporting public education.

❖ To date, 3 million people and partners have funded 1.1 million DonorsChoose.org projects.

❖ The team at DonorsChoose.org needs to be able to connect donors with the projects that most inspire them.
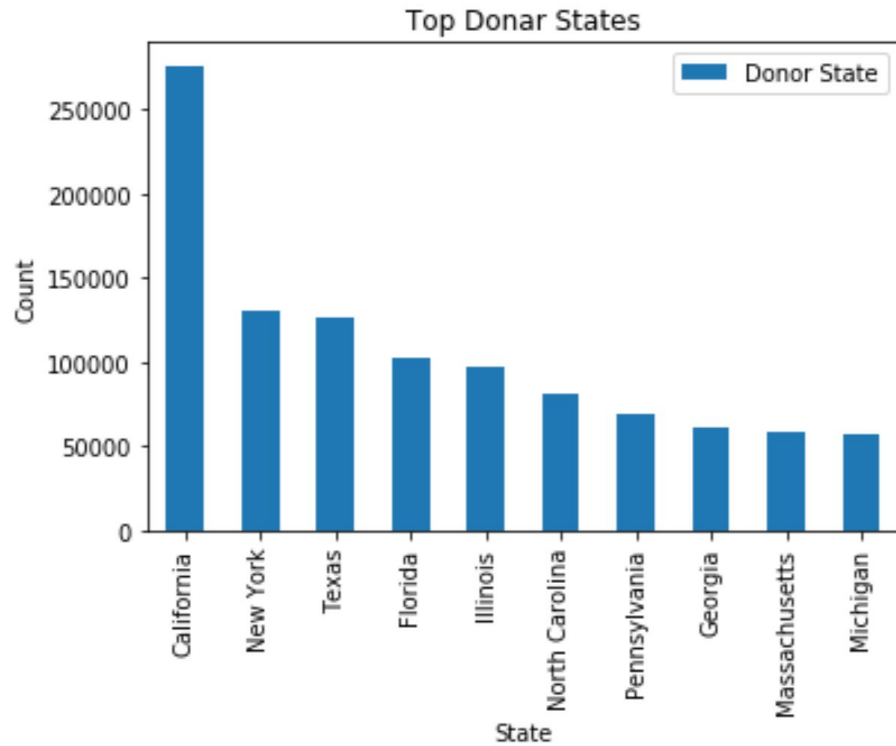
Introduction to Data

❖ Donorschoose.org, in partnership with Google.org, invited community to participate in Kaggle Data Science for Good competition.

❖ Data contains 4.7 million records of donation transactions

❖ In addition, more detail data for Donors, Projects, schools, resources and teachers.
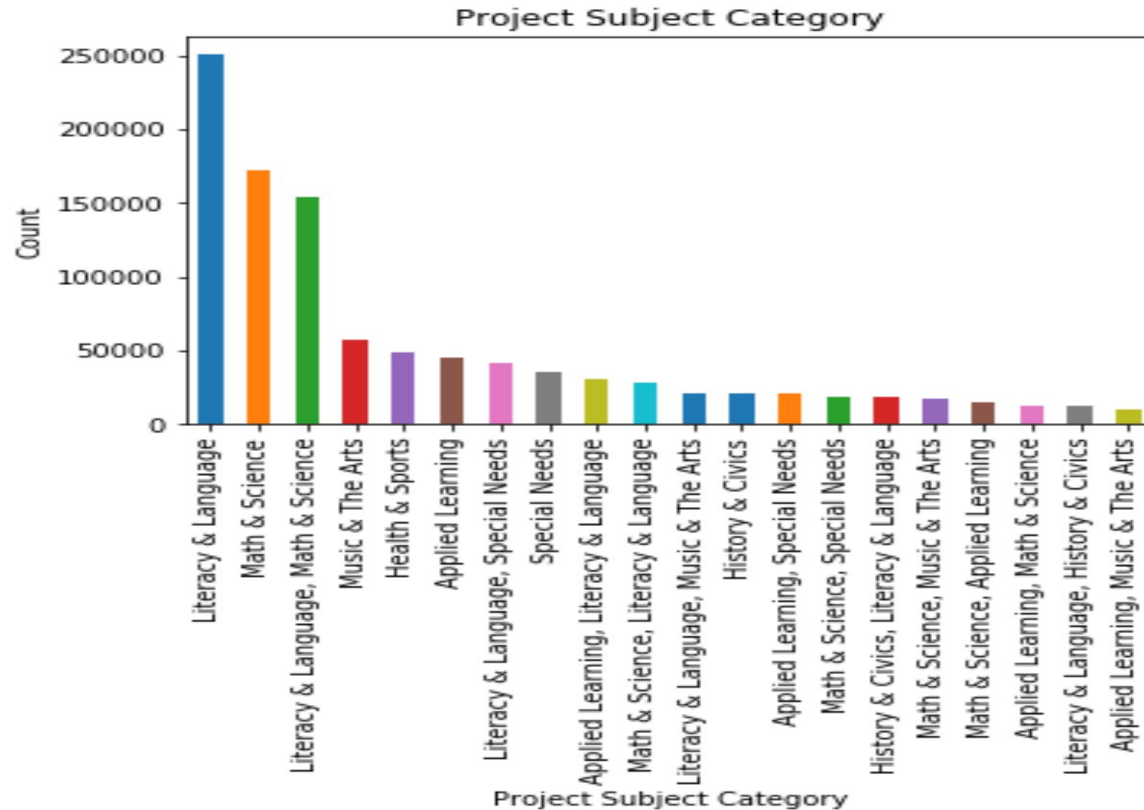
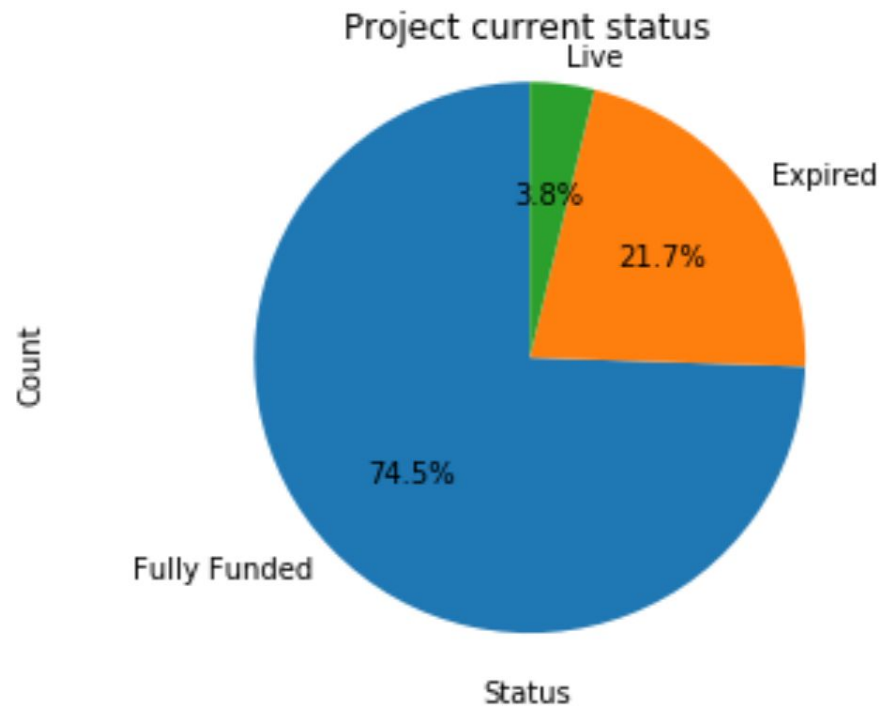# Exploratory Data Analysis

Chicago has highest Donors

Top Donor state is **California**

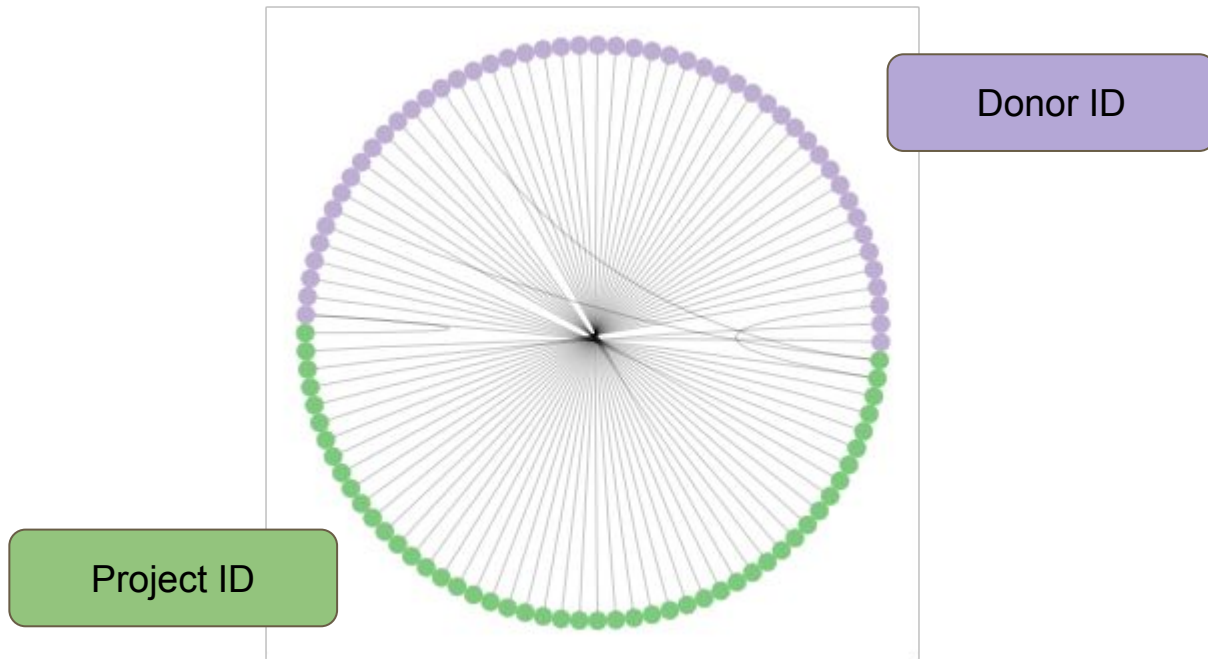**Literacy and Language** are most popular project categories

Project Subject Category

74.5% projects are fully funded

Project current status

Modelling

Bipartite Graph recommendation system

Donor ID

Project ID

Training set : Donor A

Project ID	Project Title
5c0a745e7221fd694c06e0f34294c04a	Teamwork, Dreamwork
4aee9b49ee844c634953875e7e6d0159	Writing Our Way to Reading
20c30ca10ed8b35fadbf26f61123bfa4	Math Centers - Making Our Time Count!
77742c8884292abd56aefa421b130ea7	Let's Get First Grade Readers On Track!
e3e02800b34312518abaea6d9ec7f1de	Class Carpet
b2543024beec38f6db1251102b253444	Can You Hear Me Now?
19fba96918fe3aac349ead3dcdcfa574	A Great Space to Collaborate
1f8e73b1191c650f7cfa77d126518b43	Kids Just Want to Move!
8c01cea288819f6e3403a245d518c746	Technology for Us

Test set vs Recommended

Actual Projects	Recommended Projects
Research Tech	Block Out Distractions!
iPad Fundraiser	Robotics Reality
Goodies Galore!	Goodies Galore!
Positioning Ourselves for Success!	Research Tech
None	iPad Fundraiser
None	Learning Tablets for Title 1 Students
None	Writing Is The Paining Of The Voice-Voltaire
None	Social Skills Library
None	Tech For Ms. M's Room

Evaluation Metrics

❖ Precision@k = (# of recommended items @k that are relevant) / (# of recommended items @k)

❖ Recall@k = (# of recommended items @k that are relevant) / (total # of relevant items)

❖ F1 score is interpret as a weighted average of the precision and recall.

Precision @10 is **0.33**

Recall@10 is **0.75**

F1 Score @10 is **0.46**

Conclusions

❖ Leveraging GPU and on demand cloud power would have led to better accuracy and speed to train and test the model with different variations.

❖ Adding weighted edges to the graph would increase accuracy

❖ There are many projects whose project title are not exactly matching though the meaning or purpose of the project is same. So using NLP to match and find similar projects like this would be helpful to increase accuracy of recommend projects.

Additional References

- ❖ [Github repo](#)

- ❖ Detailed Report

- ❖ Email : reenaturak@gmail.com