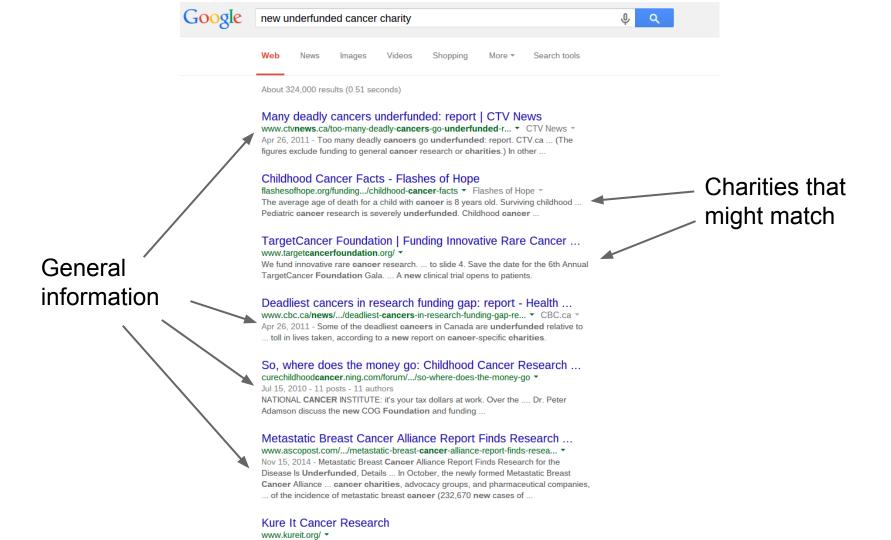
CharityChooser

Kristy Simmons
Insight Health Data Science
Summer 2015

...but maybe you want to give to a more established charity.

of contributions...

I want to donate to a cancer charity that's pretty new and doesn't get a lot



My Charities Top Ten Lists

Hot Topics Tips for Donors

Methodolo

CHARITY SEARCH enter charity name, keywords, or EIN Name Only SEARCH | Advanced | A-Z Direct **Charity Directory** Advanced Search A - Z Directory Keywords / Charity Name 4 Star Charities Name Only cancer New Ratings EIN (Employer Identification Number) **New Charities** Select Category/Cause: **Browse Categories** Health **Donor Advisories** All CN Watchlist Location: Interactive World Map All ▼ State and/or City: Advanced Search O Zip Code: in this vicinity: this ZIP only ▼ Additional Search Fields: (available to registered users) Overall Rating: up to \$3.5M Size (Total Expenses): CEO Salary All Yes Donor Privacy Policy All Regional National International Scope of Work Combined Federal Campaign (CFC) Charity Receives \$0 in government support Has Results Reporting Data

SUBMIT

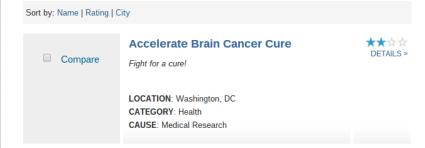
Rated Organizations (73)

Unrated Organizations (100+)

Results 1 - 20 of 73

«1234»

Click on a charity's name to view their complete rating report.



Accelerate Brain Cancer Cure

Fight for a cure!



DEMO

www.kristinadsimmons.com/charitychooser

Data



- BBB accreditation
- Year founded
- Board size
- Staff size
- Tax-exempt status



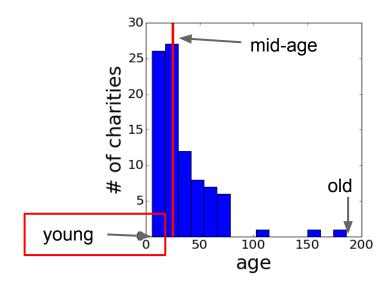
Number of Twitter followers



- Overall score
 - Financial
 - Accountability and transparency
- Total expenses
- Total contributions
- Breakdown of expenses
 - Program
 - Administration
 - Fund-raising

Database and features

- 880 organizations
- 36 diseases
- Ranked using 17 features
- "Ideal" charity constructed from user input and feature distributions



Charity rankings

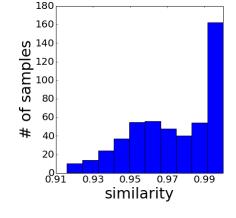
- Compare each organization to the ideal one
- Challenge: combination of binary and numerical data

- **Solution**: Gower similarity
- Compute similarity on each feature:
 - binary: agree or disagree
 - numerical: absolute difference
- Combine to get global similarity
- Choose up to 10 top charities

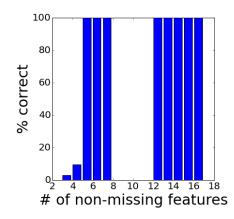
feature	ideal	actual	similarity
tax-exempt	yes	yes	1
age	6	24	
normalized age	0	0.1	0.9
		Total	0.95

Validation

- How close are my recommendations to the input charity?
 - Similarity of 0.9+



- If I put in a specific charity's features, does my algorithm recommend that charity?
 - If target has 5+ non-missing features: 100%



About me...



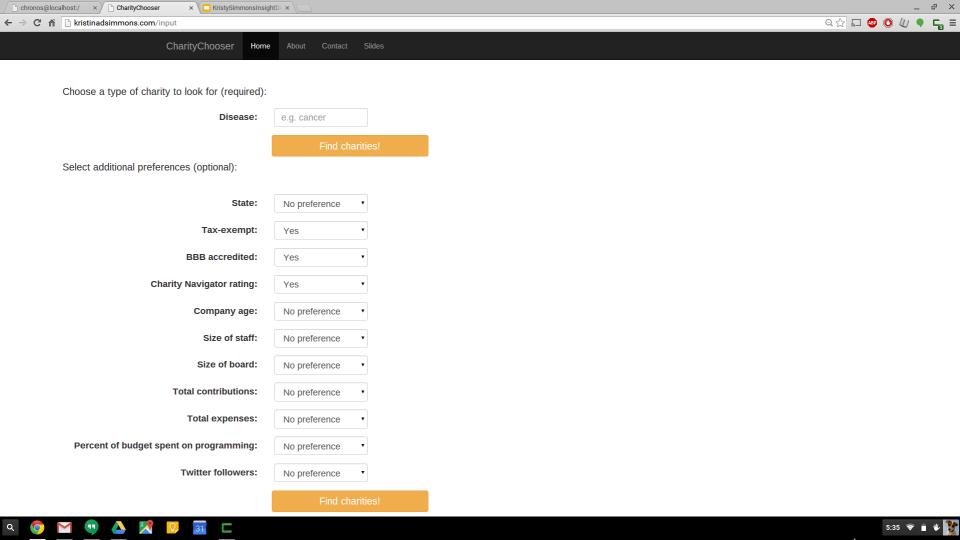


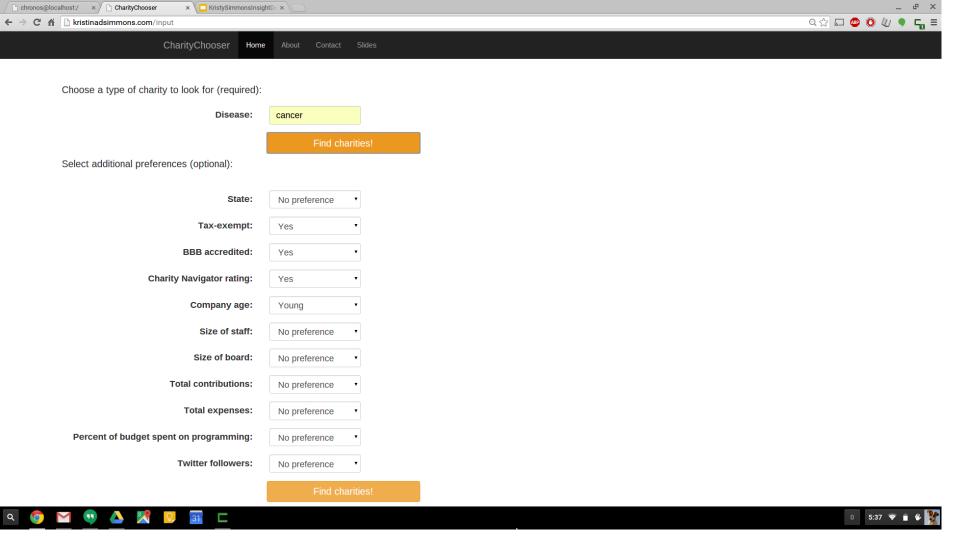
The Neuroscience Graduate Group





DEPARTMENT OF SURGERY







Address City Rank Name 1500 Manhattan Beach, 1 **Pancreatic** Rosecrans CA 90266 Cancer Action Avenue. Network Suite 200 **Breast** 60 East 56th 8th Floor 2 Donate Cancer Street. Research Foundation ZERO: 515 King Alexandria, VA 3 The End of Street Suite 22314 420 Prostate Cancer 14 New York, NY 4 Ovarian Donate Pennsylvania 10122 Cancer Plaza, Suite Research 1710 Fund 27 Union Suite 304 5 Uniting Against Square West Lung Cancer Pediatric 5550 West Suite 300 6 Donate Executive Dr. Cancer























			More information	Less	information	Ch	arity Naviga	ator Scores						Bre	akdown of Budg	jet	
Rank	Name	Address	City	Tax- exempt?	BBB Accredited?	Overall	Financial Health	Accountability and Transparency	Year Founded		Board Size	Total Contributions	Total Expenses	Program	Administrative	Fund- raising	Twitter Followers
1	Pancreatic Cancer Action Network	1500 Rosecrans Avenue, Suite 200	Manhattan Beach, CA 90266	Yes	Yes	96.84	95.54	100.00	1999	104	7	\$23,168,567	\$20,555,919	80.4%	9.6%	9.9%	
2	Breast Dor Cancer Research Foundation	60 East 56th Street,	8th Floor	Yes	Yes	95.89	94.19	100.00	1993	25	13	\$55,890,938	\$66,418,878	91.9%	2.4%	5.5%	9,616
3	ZERO: The End of Prostate Cancer	515 King Street Suite 420	Alexandria, VA 22314	Yes	Yes	94.58	92.34	100.00	1996	20	13	\$47,781,593	\$46,376,704	98.4%	0.7%	0.8%	5,943
4	Ovarian Cancer Research Fund	Pennsylvania Plaza, Suite 1710	New York, NY 10122	Yes	Yes	93.82	91.27	100.00	1994	12	9	\$7,642,696	\$8,633,550	86.2%	6.3%	7.3%	5,000
5	Uniting Against Lung Cancer	27 Union Square West	Suite 304	Yes	Yes	89.76	85.53	100.00	2002	8	11	\$3,027,879	\$2,799,727	78.7%	7.2%	13.9%	1,504
6	Pediatric Cancer Foundation	5550 West Executive Dr.	Suite 300	Yes	Yes	90.70	86.85	100.00	1991	7	24	\$1,767,485	\$1,600,434	78.6%	5.6%	15.7%	6,148
7	Prostate Cancer Foundation	1250 Fourth Street	Santa Monica, CA 90401	Yes	Yes	94.53	93.38	96.00	1993	34	27	\$50,028,788	\$45,650,436	83.8%	5.5%	10.5%	6,720
8	Cancer	One	55 Broadway Suite	Yos	Yos	95.76	9/1 01	100.00	1053	18	37	\$16 <i>4</i> 72 823	\$18 585 0/15	83.2%	7.6%	0.1%	11 38/

























































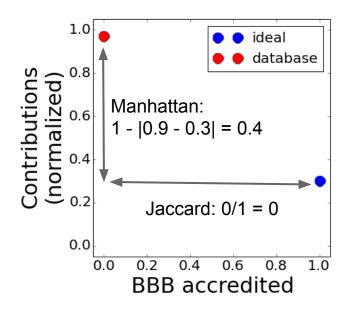




Charity rankings (old)

- Compare each organization to the ideal one
- Challenge: combination of binary and numerical data

- **Solution**: Gower similarity
- Compute similarity on each feature and combine
- Binary data:
 - Jaccard similarity coefficient
- Numerical data:
 - 1 Manhattan distance



Gower similarity = (0.4 + 0)/(1 + 1) = 0.2

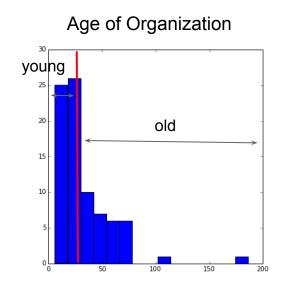
Charity rankings

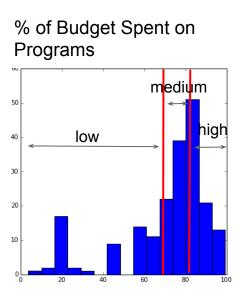
- Compare each organization to the ideal one
- Challenge: combination of binary and numerical data
- Solution: Gower similarity
 - Compute similarity on each feature and combine
 - 1 Manhattan distance for numerical data
 - \blacksquare s = 1-abs(x_data-x_ideal)
 - Jaccard similarity coefficient for binary data
 - s = 1 if x data = 1 and x ideal = 1
 - s = 0 if x_data = 1 xor x_ideal = 1
 - s = NaN if x_data = 0 and x_ideal = 0
- Return up to 10 most similar charities

Database and features

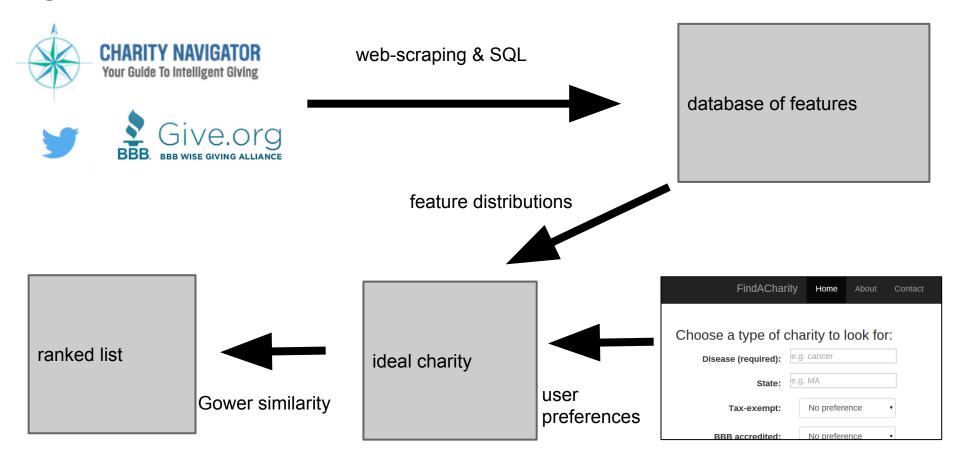
772 organizations

- 117 (15%) accredited by Better Business Bureau
- 203 (26%) reviewed by Charity Navigator





Algorithm



Service and Other

Zoos and Veterinary Services Health (All)

Addiction and Substance Abuse

Diseases and Disease Research

Health Care Facilities and Programs

NTEE Code:

(To select multiple choices, hold Ctrl and click with mouse.)

A Arts, Culture, and Humanities

A01 Alliance/Advocacy Organizations

A02 Management & Technical Assistance

A03 Professional Societies, Associations

A05 Research Institutes and/or Public Policy Anal

Income:

You may use a 'k' for thousands, an 'm' for millions, or a 'b' for billions (e.g., 120k, 3.5M, etc.)

and 2m Between 0



Donate to GuideStar

Welcome to GuideStar! Sign Out

Update Nonpi

Search GuideStar

Advanced Search Refine Your Search

Home > Search Results

Your search for cancer produced 4,378 results.

Viewing 1 - 25 of 1,000 results

Related Keywords: prostate cancer research, ovarian cancer research, ofthe lung cancer research, lung cancer research, cancer research ... more >



Bear Necessities Pediatric Cancer Foundation

AKA Bear Necessities PCFI Chicago, IL

bearnecessities.org

Forms 990 & People & External Summarv Financials Programs Docs Governance Perspectives

CuidoCtor Cummoru

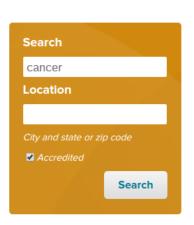




for **DONORS** for **CHARITIES** news & **UPDATES** about us **BBB WGA**

Purpose

Your search for **cancer** returned 49 results.



CATEGORY SEARCH

Charity - American Indian

Charity - Animal Protection

Charity - Arts & Culture

Charity - Blind & Visually Impaired

Charity - Cancer

Charity - Child Sponsorship

Charity - Children & Youth

Charity - Civil Rights

Charity Community Dovolonment 9

CHARITY DISTANCE Colorectal Cancer Coalition aka C3: Colorectal Cancer Coalition 1414 Prince Street, Suite 204 Alexandria, VA 22314 ▶ Click here to view the full report Category: Cancer CureSearch for Children's Cancer aka National Childhood Cancer Foundation 4600 East West Highway, Suite 600 Bethesda, MD 20814 Click here to view the full report Category: Cancer

Uniting Against Lung Cancer

Click here to view the full report

American Cancer Society

250 Williams Street NW

27 Union Square West

New York, NY 10003

Category: Cancer

Suite 304

& Staff The following information is based on the ACS' audited financial statements for the fisca

Governance

Tax

Status

Financial

Fund Raising

Source of Funds

August 31, 2010.

Conclusion

Contributions	302,652,000
Legacies and bequests	43,893,000
Contributed services, merchandise and other	21,328,000
in-kind contributions (at fair value)	
Exchange transactions, net	13,360,000
Investment income	6,569,000
Grants and contracts	5,697,000
Other revenue	2,839,000
Change in value of split-interest agreements	39,000
Loss on disposal of fixed assets	-259,000
Total Income	\$396,118,000

Programs

Governance Tax **Programs Fund Raising** Conclusion **Purpose** Financial & Staff Status

Year, State Incorporated: 1913, New York



Validation

- Does my algorithm rank organizations correctly?
 - Unlabeled data
 - Not clear how to define "correct"
- If I rank two charities similarly, do people tend to give to both?
 - No data that I could find to check this
- How close are my recommendations to the ideal charity?
 - Could measure this using random input
 - Uses the same algorithm to produce results and to validate
- If I put in a specific charity's features, is that charity the top recommendation?
 - Could measure this systematically
 - Assumes that features are chosen and coded reasonably