IRS990 Analysis

Rahul Bazaz | Snehal Pandit | Yash Nanavati

Objective

Evaluate the performance of nonprofits based on quantitative metrics such as program expenses ratio, liabilities to assets ratio, working capital ratio and surplus margin.

Purpose

• Each nonprofit can compare itself to other nonprofits and identify areas for improvement.

• Tax authorities and government officials can use this analysis to get a better understanding of how certain nonprofits are using the money saved from their tax-exempt status.

Dataset

• Our dataset is a collection of electronic 990 documents available to the public through Amazon Web Services. These are used by United States IRS to gather financial information about nonprofit organizations.

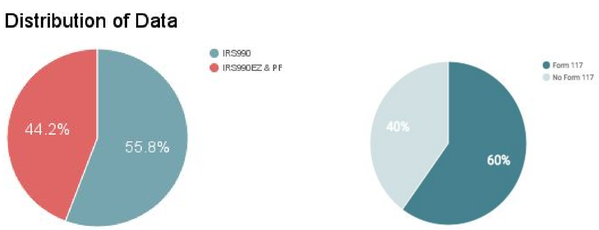
• It contains detailed information regarding the nonprofits themselves, how their information was filed, their revenues and their expenditures. Nonprofits fill out one of three kinds of forms: IRS990, IRS990PF or IRS990EZ. From that ones that fill out IRS990, a majority of them fill out a Form 117 that requires them to divide their assets into unrestricted net assets, temporarily restricted net assets and permanently restricted assets. The ones that do not fill out this form, only list their total assets. **In our analysis we are only looking at nonprofits that fill out an IRS990 and Form 117**

• **Program Expenses** which were calculated as Program Expenses/Total Functional Cost, this ratio reveals what percentage of the nonprofit's total costs are going towards its actual mission

•**Liabilities to Assets Ratio** is calculated by dividing the total liabilities by the total assets.

•**Working Captial Ratio** is calculated using the formula Unrestricted_Net_Assets + Temporarily_Rstr_Net_Assets / Total_Functional_Expenses.

•**Surplus Margin** (also known as a profit margin) is net income (i.e., the change in net assets) divided by total revenues.



Organizations that fill out an IRS990 have Gross Receipts ≥ USD 200,000 and Total Assets ≥ USD 500,000

Organizations not included in our analysis:

• Any religious organizations, like the Salvation Army, which are exempt under Internal Revenue Code from filing the Form 990

• Private foundations receive the majority of their money from only one individual, family, or corporation. On the other hand, public charities have a broad-base of support from the general public as well as variety of other funding sources. Private foundations are only required to file a Form 990-PF which differs from the document public charities file.

• Small organizations that fill out a form 990-EZ are also not included because Form EZ requires less financial reporting than the Form 990, and as such, we would lack important data needed in our analysis.

Preparing the Data

• Dividing the entire dataset into: small, medium, large and national based on their total expenses.

  • Small: Less than 1 million

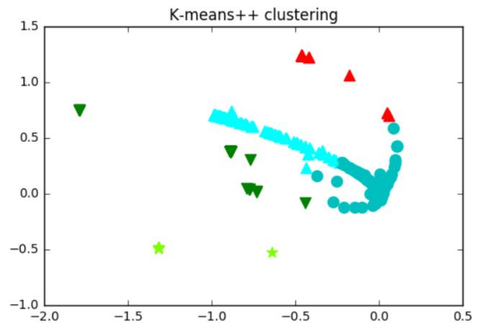
  • Medium: Between 1 and 10 million

  • Large: Between 10 and 50 million

  • National: Greater than 50 million

• Normalized the data by subtracting each column from its mean and dividing it by its standard deviation.

• Performed a TruncatedSVD on the data in order to reduce its dimensionality.



KMeans

• Initialized 5 cluster centers, assigning points to each cluster and repeating this process trying to minimize Euclidean distance from points to cluster centers with each iteration.

•Outliers were calculated as being the point that was the furthest Euclidean distance away from its cluster center

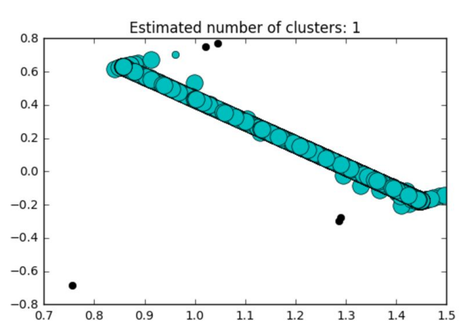
OneClassSVM

Classifies data as financially efficient or inefficient given training data of only efficient nonprofits. It defines a boundary based on the training data and classifies the data as positive(efficient) or negative(inefficient).

DBSCAN (Density-Based Spatial Clustering of Applications with Noise)

•Find core samples of high density and expand clusters from them

•The minimum number of samples in a neighborhood for a point to be considered as a core point was set at 10 and the maximum distance between two samples within the same neighborhood was set at .1.



Kmeans Results

Finds outliers on entire set of IRS990s, forms clusters on a combination of metrics. The results of this analysis are not necessarily financially inefficient, they are just performing differently as compared to their contemporaries.

Color	EIN	ProgExp%	Liability/Asset	WorkCapRatio	SurplusMargin	TotalExp
Turquoise	521221108	0.400001	16883.31579	-1.915329	-9.796983	167472
Light Green	232787307	1	2728.935484	-234.905556	0	360
Green	481252775	0	0	0	0.008741	0
Cyan	561949970	1	0	1	0	1
Red	352090479	0.85741	0.217362	298.930677	-12877.66667	4818

DBSCAN Results

Results of this method are similar to that of Kmeans. The results are based on scaled down version of all the metrics and therefore gives us a good overview of all the outliers.

EIN	ProgExp%	Liability/Asset	WorkCapRatio	SurplusMargin	TotalExp
680383921	0	0	172324.3	0	10
311040228	0	0.540167	149559.47	1.000333	100
166050703	1	33.472201	-148963.5636	0	220
943152652	1	254.451508	-7.144765	-250550	250551
232799695	0.985825	0.426822	2.387819	-54849.57143	202757

OneClassSVM Results

Result of Small Businesses: Low Program_Expenses Percentage and high Working Capital. This highlights how they are spending a majority of their money on administrative and purchase of assets rather than directly to their programs

EIN	Program_Exp	Liabilities_To_Asset	Working_Capital	Surplus_Margin	Total_Expenses
7.53E+08	0	0.00427391	48.9994	0.516164	19209
2.03E+08	0.00990099	0.00414066	1704.09	0.979933	808
7.53E+08	0	0.00427391	48.9994	0.516164	19209
2.37E+08	0	0.029641	37.2855	-20.4117	23039
4.22E+08	0	0	146.259	0.467074	7849

Result of Large Businesses: High Liability to Asset Ratio and negative Surplus Margin. These are growing businesses that are spending all of their surplus revenue on purchasing assets

EIN	Program_Exp	Liabilities_To_Asset	Working_Capital	Surplus_Margin	Total_Expenses
6.07E+07	0	0.644904	0.654848	-0.296143	1.96E+07
1.31E+08	0.786773	0.798405	-0.551577	-1.0185	1.28E+07
1.13E+08	0.731517	2.05646	-0.37316	-0.0310824	1.03E+07
6.07E+07	0	0.644904	0.654848	-0.296143	1.96E+07
1.31E+08	0.786773	0.798405	-0.551577	-1.0185	1.28E+07