



INSIGHT
PHILANTHROPY
RESULTS

EXPLORE

PD25

New Depths

August 19-22, 2025

Hilton Baltimore Inner Harbor Hotel, Baltimore, Maryland

A Journey in Name Matching Techniques

FROM EXCEL TO PYTHON

EXPLORE
PD25
New Depths

While you're waiting, complete your session evaluations in the mobile app!

SESSION OBJECTIVES:

- Attendees will learn to use available tools and new techniques to find name matches between lists and their database with increased efficiency.
- Attendees will develop an increased capacity for solving problems independently and within organizational constraints

This is not a step-by-step instruction guide:
look for email follow ups

WHO AM I?

Jack Pitfield

Lead Analyst,

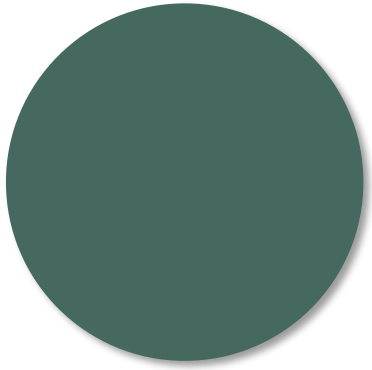
Fundraising Performance Analytics

University of Chicago

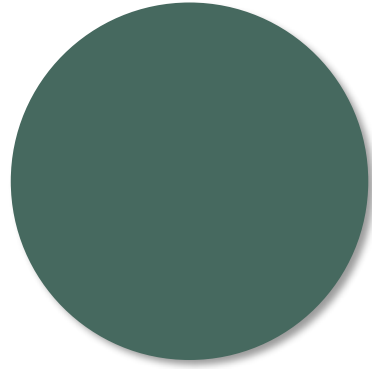
jpitfield@uchicago.edu



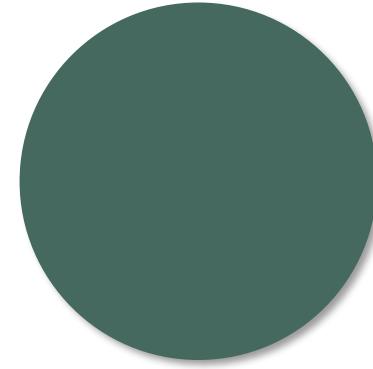
TEAM & PARTNERS



Jake Leslie
Senior Analyst



Teo Icliyurek
Associate Director



Prospect Development and
Decision Support
University of Chicago



LET'S START WITH A REQUEST

EXPLORE
PD25
New Depths

MANUALLY SEARCHING 2,500 NAMES WASN'T GOING TO WORK



OTHER SOLUTIONS

- Name Matching Utility
- Tool within CRM

AM I MISSING ANYTHING?



Why don't your lists come with my database ID?

DEFINE THE PROBLEM

EXPLORE
PD25
New Depths

HOW DO YOU MATCH AN UNSTRUCTURED LIST OF NAMES WITH THE NAMES IN YOUR DATABASE?



ELEMENTS OF THIS PROBLEM

Examples:

- Event attendance
- Performance attendance
- Publicly available lists
- Donor Lists
- Dinner Lists

Problems:

- Lists don't come with my database IDs
- Data Inconsistencies:
 - Names don't match
 - Duplicate names
- Scale:
 - Time consuming
 - Feels like computing

EXAMPLE DATA

Event List:

Name
Christina Patterson
Susan Valencia
Sarah Douglas
Cory Cowan
Sydney Jackson DDS

Find a match
(If there is one)

Database List:

CRM ID	Name	Other data
0011U00K2U6TV9C4JR	Terry Rodriguez	...
0011U0FDIBGMZY95LD	Jasmine Powell	...
0011U0CPWOLGPQLOFF	Brian Martinez	...
0011U0V3WLP47HUMCT	Kimberly Hunter	...
0011U034KSY80KWBYP	Jennifer Lawrence	...



Starting in my comfort zone

EXCEL

EXPLORE
PD25
New Depths

VLOOKUP / XLOOKUP

=VLOOKUP(What you want to look up, where you want to look for it, the column number in the range containing the value to return, return an Approximate or Exact match – indicated as 1/TRUE or 0/FALSE)

Event List		Database	
Names	Matched ID	Names	ID
Eric Russo		Pamela Anderson	1
Teresa Maldonado		Tammy Allen	2
Rachel Wise		Timothy Clay	3
Marilyn Hughes		Tam Allen	4
Tammy Allen		James Walters	5
		Russell Thompson	6
		Tammy Allen	7

Link: <https://support.microsoft.com/en-us/office/vlookup-function-0bbc8083-26fe-4963-8ab8-93a18ad188a1>

WHY THIS DOESN'T WORK

- Limited to first result or highest match
- Scale becomes difficult to manage
- Repeatability in Excel is challenging

Event List		Database	
Names	Matched ID	Names	ID
Eric Russo		Pamela Anderson	1
Teresa Maldonado		Tammy Allen	2
Rachel Wise		Timothy Clay	3
Marilyn Hughes		Tam Allen	4
Tammy Allen		James Walters	5
		Russell Thompson	6
		Tammy Allen	7

Vlookup Limitations with Exact Match

Partial Match Link: <https://www.excel-university.com/vlookup-hack-9-partial-match/>

WHY THIS DOESN'T WORK

- Limited to first result or highest match
- Scale becomes difficult to manage
- Repeatability in Excel is challenging

Event List		Database	
Names	Matched ID	Names	ID
Eric Russo		Pamela Anderson	1
Teresa Maldonado		Tammy Allen	2
Rachel Wise		Timothy Clay	3
Marilyn Hughes		Tam Allen	4
Tammy Allen		James Walters	5
		Russell Thompson	6
		Tammy Allen	7

Vlookup Limitations with Approximate Match

Partial Match Link: <https://www.excel-university.com/vlookup-hack-9-partial-match/>

CLEANING OUR DATA

EXPLORE
PD25
New Depths

IDENTIFY UNIQUE IDENTIFIERS

- Email
- Phone numbers
- Address
- Credit Card Details

Separate these for a more direct matching approach (using vlookup for example)

SOME HELPFUL STEPS

1. Remove punctuation
2. Remove spaces
3. Remove capitalization

Email Follow Up #2:
Cleaning Data with Python

Jack Pitfield' 1. 2. 3.
 Jack Pitfield JackPitfield jackpitfield

```
graph LR; A["Jack Pitfield'"] -- "1. Remove punctuation" --> B["Jack Pitfield"]; B -- "2. Remove spaces" --> C["JackPitfield"]; C -- "3. Remove capitalization" --> D["jackpitfield"];
```

(PRE)FILTER YOUR DATABASE NAMES

- Who are you trying to match?
- Can you filter by region?
- What about by wealth rating?



Something a little more powerful

PYTHON (OR R)

EXPLORE
PD25
New Depths

DO I REALLY NEED TO USE CODE?

- Open source (your only cost is time)
- Incredibly flexible
- Scalable
- Great documentation and opportunities for learning

Email Follow Up #1:
Getting Started with Python

Quick note:
See resources
at the end of
the deck



HOW DOES THIS WORK?

- Generate two lists
- Use an algorithm to compare similarity
- Set a matching threshold
- (optional) set a max number of results

Email Follow Up #3:
Name Matching Scripts

Excalidraw diagram would be great
here



NAME MATCHING ALGORITHMS

1. Fuzzy Matching
2. Vector Matching
3. Levenshtein Distance

Similarity Algorithm	Similarity	Match*
Cosine	82.50%	true
Jaccard	70.00%	false
Jaro	85.50%	true
Jaro Winkler	85.50%	true
Q-gram	82.40%	true
Sorensen Dice	82.40%	true

Distance Algorithm	Distance	Match*
Damerau Levenshtein	2	true
LCS Edit	3	true
Levenshtein	2	true
OSA Damerau Levenshtein	2	true

Phonetic Algorithm	Phonetic Encoding 1	Phonetic Encoding 2	Match*
Cologne	05684	45684	false
Soundex	E452	C452	false
Metaphone	LNMSK	KLNMSK	false

A note on testing: you need to test on full sample sets otherwise it's easy to get overconfident.



WHAT DOES THE CODE LOOK LIKE:

Import Libraries

Import Tables

Clean Names

Matching Code

Match!

Clean Results

Export Results

[illegible]

OUTPUT

List Name	Match Score	Match Name	Match Details
Christina Patterson	0011U0FG5LQX1CU6WY	Christian Pearson	...
Christina Patterson	0011U00GEKCCX8DSL	Christian Pearson	...
Christina Patterson	0011U0HY3LEQ5JQKD8	Christopher Watkins	...
Christina Patterson	0011U0BIS8VNM3VHKC	Christopher Watkins	...
Christina Patterson	0011U0R09J9R7ORVYX	Christopher Watkins	...
Christina Patterson	0011U0R6LIK9FSDT4	Christina Tate	...
Christina Patterson	0011U0H5FBYK5ZAKBI	Christian Yates	...
Christina Patterson	0011U0ZYRNKH4HDL8P	Stephanie Armstrong	...
Christina Patterson	0011U00R1W6PNIFG4Q	Stephanie Armstrong	...
Christina Patterson	0011U0GBKI243VMQAP	Stephanie Armstrong	...

Notice my sample dataset generated records with duplicate names



MAKING THIS MORE COMPLEX

- Match across other matching fields as available
 - Name to name
 - Job title to job title
 - Address to address

Or keep it a little simpler:

Similarity Score (.8-1)

+1 Same Country

+.2 Employment Information

+1 Last Name is the same

The value of manual review

RESULTS

EXPLORE
PD25
New Depths

OUR FIRST TEST:

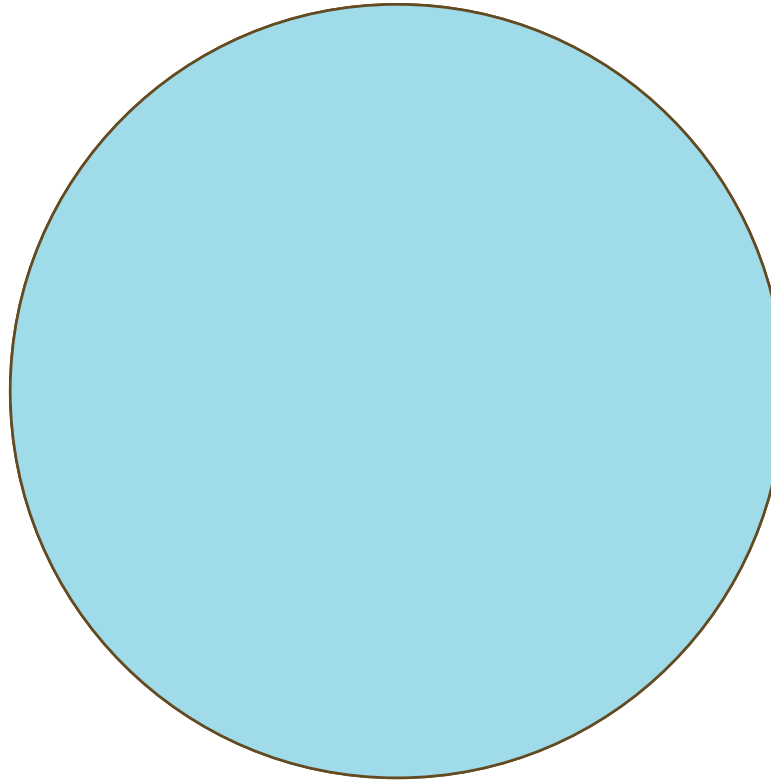
● 2.5K Event Attendees to match

* Not to scale



OUR FIRST TEST:

● 2.5K Event Attendees to match

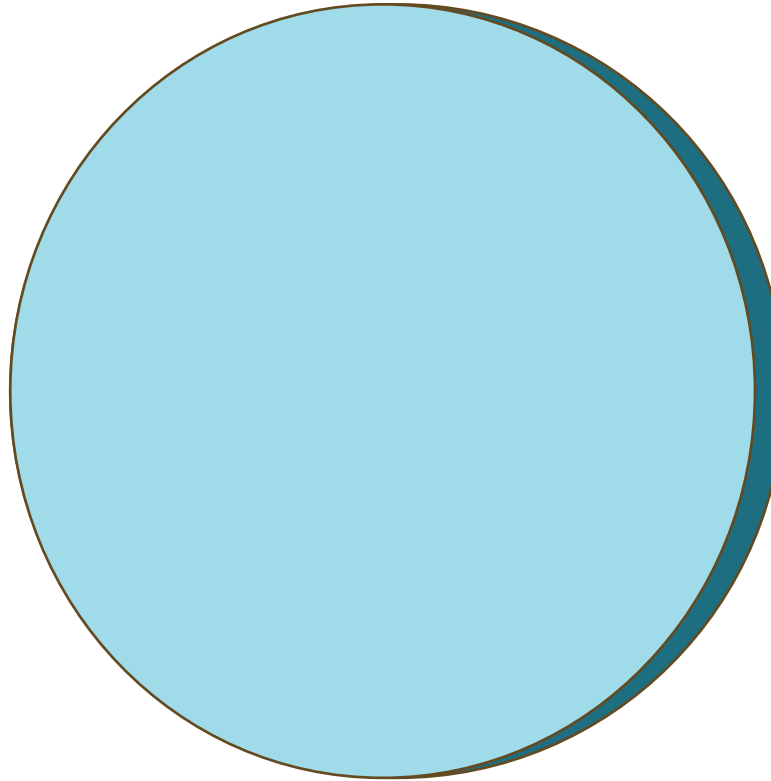


Generating 6.5M
matches

* Not to scale

OUR FIRST TEST:

1.25K Attendees had already been reviewed

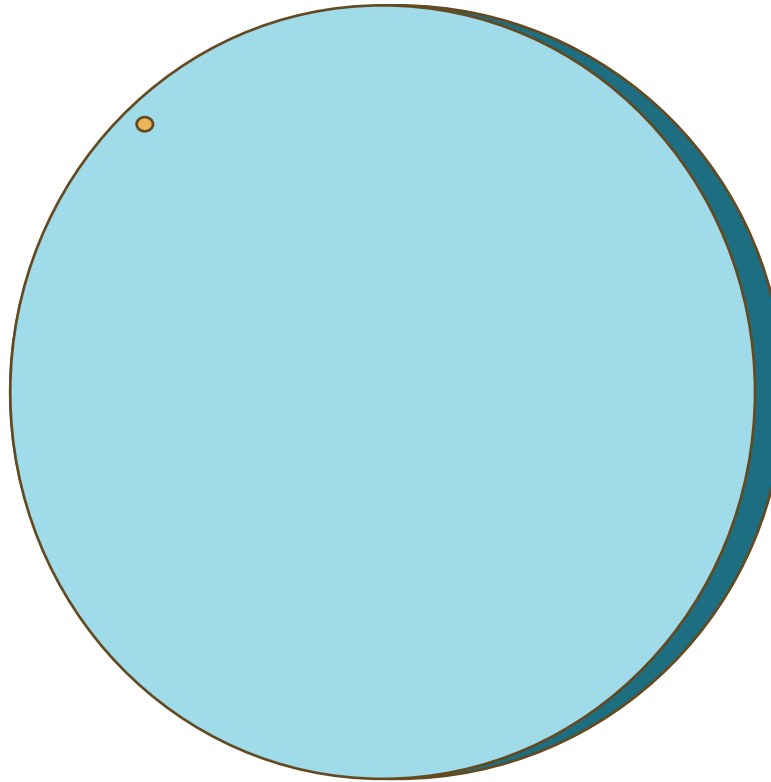


Leaving 6.4M matches

* Not to scale

OUR FIRST TEST:

1.25K Attendees had already been reviewed



After manual review, 60 new matches were identified

* Not to scale

USING OUR SAMPLE DATA

Similarity Score	Count
.95-1	60
.9-.95	29
.85-.9	19
.8-.85	9



MORE COMPLEX

Similarity Score+:

Similarity Score (.8-1)

+1 Same Country

+.2 Employment Information

+1 Last Name is the same

Similarity Score+	Count
3-3.5	22
2-2.5	90
1-1.5	5



FINAL THOUGHTS

EXPLORE
PD25
New Depths

FOLLOW UP

Look for a 5-email series to help you get started with python and name matching.

- 1: Getting Started with Python
- 2: Cleaning data
- 3: Name matching algorithms
- 4: Loading a local LLM
- 5: Setting up GUI

There are a lot of available resources (look for slide in the back), so these are far from your only option, but it's a good place to start.

FEEDBACK REQUEST

- Was this a helpful topic? Do you expect to use it in the future?
- Did you enjoy a more technical presentation? Would you like to see more of this at APRA?



THANK YOU!

Please complete your session
evaluations in the mobile app.

QUESTIONS?



RESOURCES

Getting started with Python

- Automate the Boring Stuff with Python
- Python for Everybody
- Python Basics

Data Analysis with Python

- Python for Data Analysis
- Exploratory Data Analysis with Python Cookbook
- Data Science for Fundraising (its in R)

Development Environments:

- Anaconda
- PyCharm
- Marimo

Getting Started with Local LLMs:

- Easier than you think (Medium)

TIPS FOR AN ACCESSIBLE PRESENTATION

1. Use Slide Layouts

- Use the built-in slide layouts in PowerPoint. These templates are structured to work with assistive technologies like screen readers.

2. Add Descriptive Alt Text to Images

- Right-click on images, charts, and other visuals, select **Edit Alt Text**, and provide a concise description.
- If the image is decorative, mark it as such.

3. Ensure Sufficient Color Contrast

- Use high-contrast color combinations between text and background.
- Tools like the **Accessibility Checker** (under **Review > Check Accessibility**) in PowerPoint or online contrast checkers can help verify contrast ratios.

4. Use Meaningful Link Text

- Avoid vague text like "Click here" for hyperlinks. Instead, use descriptive text.

5. Structure Content Clearly

- Use headings and bullet points to organize content logically.
- Avoid long blocks of text.

6. Avoid Animations and Transitions

- Minimize the use of flashing animations, which can be distracting or trigger seizures.

7. Test with Assistive Technology

- If possible, test your presentation using screen readers or other assistive tools to ensure compatibility.

Recommended Font Sizes

- **Titles/Headings:** At least **24–32 pt.**
- **Body Text:** At least **18 pt.**
- **Footnotes or Supplemental Text:**
No smaller than **14 pt**, but avoid small text unless absolutely necessary.