

# Large Scale Lookups From an End-User and a Programmer Perspective

Peter Stegehuis and Naxin Zheng  
Westat, USA

**IBUC - October 2023**



# Overview

- Lookups in Blaise
  - Trigrams
- Location-specific lookups
- Adding to the look and feel

# Blaise lookups with trigrams

- Creating trigram lookup file
  - Datamodel with trigram secondary key
  - Three characters snippets, lookup file indexed
- Using in a lookup
  - Interviewer types in search string
  - That gets chopped up the same way
  - Immediate results
  - More 'resistant' to typos than e.g. SQL 'LIKE' search

# Location-dependent lookup files

For many purposes a lookup will be the same for all interviews

- Example: Prescribed medicines

In our case, we needed to show only relevant medical providers

- For an interview in New York, a Family Doctor in California is not relevant
- Subdivision or filtering needed
- Seven million provider records!

# Implementation

Use zip code (postal code), and include all records within a 100 mile radius

- Pro: very localized
- Con: There are 42 K zip codes in the US

Too many lookup files, and too big (20 TB) to create all and install on interviewer's laptop

- Solution: create on laptop as needed

# Issues

Even such a local lookup file can get big

- Possibly hundreds of thousands of records
- Takes time to create

Manipula: superfast in processing text files

- Funny finding: two-step process faster than one single setup
- Size of a source file: too big for in-memory processing

# Presentation

SELECT PROVIDER

DETAILS: JOHNS HOPKINS UNIVERSITY JOHNS HOPKINS UNIVERSITY INTERNATIONAL STD LABORATORY; 855 N WOLFE ST RANGOS 520; BALTIMORE; 4106140932; Clinical Medical Lab

Type	AHA	Name 1	Name 2	Address	City	Phone Number	Taxonomy
F		JOHNS HOPKINS UNIVERSITY	JOHNS HOPKINS UNIVERSITY INT	855 N WOLFE ST RANGOS 520	BALTIMORE	4106140932	Clinical Medical Laboratory
F	AHA	JOHNS HOPKINS HOSPITAL		600 N WOLFE ST	BALTIMORE	8556623017	General Acute Care Hospital
P		MATTHEW CURTIS WOODS		600 N WOLFE ST THE JOHNS HOPKINS HOSPITAL	BALTIMORE	4109555000	Student in an Organized Health Care Educatio
P		ADAM IDDRISS		JOHNS HOPKINS HOSPITAL 600 N WOLFE ST BLALOCK 658	BALTIMORE	4125088573	Student in an Organized Health Care Educatio
P		PEYI SU		600 N WOLFE ST THE JOHNS HOPKINS HOSPITAL	BALTIMORE	4109555000	Student in an Organized Health Care Educatio
F		JOHNS HOPKINS UNIVERSITY		600 N WOLFE ST CMSC 309	BALTIMORE	4109557858	Psychiatric Hospital
P		MANOJ MADDALI		600 N WOLFE ST THE JOHNS HOPKINS HOSPITAL	BALTIMORE	4109555000	Student in an Organized Health Care Educatio
F		JOHNS HOPKINS UNIVERSITY	JHU NEUROLOGY BEHAVIORAL H	600 N WOLFE ST	BALTIMORE	4109555000	Psychiatry & Neurology - Neurology
F		JOHNS HOPKINS UNIVERSITY	JHU OTOLARYNGOLOGY HEAD A	600 N WOLFE ST	BALTIMORE	4109336401	Otolaryngology
P		MARIJA VASILJEVIC		600 N WOLFE ST THE JOHNS HOPKINS HOSPITAL	BALTIMORE	4109555000	Student in an Organized Health Care Educatio
P		ALEX JANG		600 N WOLFE ST THE JOHNS HOPKINS HOSPITAL	BALTIMORE	4109555000	Student in an Organized Health Care Educatio
P		CHENGCHENG GUI		600 N WOLFE ST THE JOHNS HOPKINS HOSPITAL	BALTIMORE	4109555000	Student in an Organized Health Care Educatio
F		JOHNS HOPKINS UNIVERSITY	JHU SURGERY	600 N WOLFE ST	BALTIMORE	4109508400	Surgery
P		MOON JEONG LEE		600 N WOLFE ST THE JOHNS HOPKINS HOSPITAL	BALTIMORE	4109555000	Student in an Organized Health Care Educatio
P		STEPHEN LESCHKE		600 N WOLFE ST THE JOHNS HOPKINS HOSPITAL	BALTIMORE	4109555000	Student in an Organized Health Care Educatio
P		LANA LEE		200 N WOLFE ST JOHNS HOPKINS SCHOOL OF MEDICINE	BALTIMORE	4109552910	Student in an Organized Health Care Educatio
F		JOHNS HOPKINS MEDICAL INSTITUTIONS		600 N WOLFE ST	BALTIMORE	4439979688	General Acute Care Hospital
F		JOHNS HOPKINS UNIVERSITY FM AND R DEF		600 N WOLFE ST STE 160	BALTIMORE	4105022447	Rehabilitation Hospital
P		JONATHAN ALBERT LESTER		600 N WOLFE ST THE JOHNS HOPKINS SCHOOL OF MEDICINE	BALTIMORE	4109555000	Student in an Organized Health Care Educatio
F		JOHNS HOPKINS UNIVERSITY	JHU ONCOLOGY	600 N WOLFE ST	BALTIMORE	4109337400	Internal Medicine - Medical Oncology
F		JOHNS HOPKINS UNIVERSITY	JHU ONCOLOGY HOSPICE AND P	600 N WOLFE ST BLALOCK 359	BALTIMORE	4109558306	Internal Medicine - Hospice and Palliative Med
P		MEGAN GILMARTIN		JOHNS HOPKINS HOSPITAL 600 N WOLFE ST	BALTIMORE	4106143840	Registered Nurse
P		EVA MARIA LUDEROWSKI		600 N WOLFE ST THE JOHNS HOPKINS HOSPITAL	BALTIMORE	4109555000	Student in an Organized Health Care Educatio
F		JOHNS HOPKINS HOSPITAL		600 N WOLFE ST	BALTIMORE	4106143234	Rehabilitation Hospital
P		AMINA WHITE I AVIGNE		600 N WOLFE ST THE JOHNS HOPKINS HOSPITAL	BALTIMORE	4109555000	Student in an Organized Health Care Educatio

1:776

Search for: johns hopkins wolfe 855

Show only facilities

SEARCH TIPS: \* Fewer words often work better than lots of words-but you must type at least 6 letters

\* Unusual or unique aspects of provider name or address finds better matches

\* Common identifiers (e.g., medical center, office, associates, group, health) can increase the number of 'extra' matches displayed

\* If searching by address - include street number and name, but not office, suite, or room number

\* If searching by phone number - do not include hyphens

\* Review the DETAILS line at the top to verify correct selection

SELECT

No Match

Cancel

# Presentation

- Checkbox for filtering
- Details of selected record at the top
- Interviewer instructions to guide successful search

# Conclusions

- The use of trigram lookups in Blaise is very powerful
- Lookups with thousands of records can be complicated for field interviewers
  - make lookups as user-friendly as possible
  - with options and pointers for efficient use
- It pays off to find the most efficient way to create lookups, at least if you have to create many
  - Sometimes better results with what may seem a counter-intuitive method