

The Zombie Eligibility Problem:

**One solution to improve
health data integrity**





I am **Carrie Minerich**

- El Paso, TX is homebase but I'm currently in Western WA.
- I have 4 years of professional experience in data and software
- I hold a MSc from the University of Idaho





1

Identify the problem

What is a zombie eligibility?

Privacy Expectations

- Protect Springbuk and PHI
- Fictional database entries
- Sanitized Python code

Health data is
sensitive and
complex

3 common data types

- **Medical Claims**
- **Prescription Claims**
- **Enrollment** or Eligibility
 - Employee
 - Members
 - Members: dependents & spouses included on a health plan
 - Member count includes employee count

Review of a health data **pipeline**

Brief overview of data pipeline

- Cloud storage
- File contents to raw database
- Raw tables through Python to warehouse
- QA checks and sync to app

Eligibility data

Known Eligibility Data Issues

- Duplicate data
 - Existing logic to detect duplicate files / individuals regardless of data type
 - Blacklist data source if receiving data from two carriers and delete blacklisted dup data
- Ghost Eligibility
 - We have claims but no eligibility file for an individual
- Termed by Absence
 - If individual absent from eligibility file, we disenroll that individual

The path to building an internal **pipeline**

External Data Company

EDC is a company that has a well established health data pipeline.

Direct Data Consumption

Springbuk is developing health data pipelines for direct data consumption

Springbuk licensed EDCs pipeline to process our clients data while we built our own pipeline

*all Springbuk intellectual property is protected during this presentation

What exactly is a **Zombie** Eligibility?

Once a carrier's pipeline is complete, we stop employers on that carrier from the EDC feed and start them on DDC feed

Individuals are still enrolled on the EDC feed, causing inflated eligibility counts. Can't delete, must disenroll before blacklisting

This dupe eligibility issue is unique to the migration of clients from the external data feed to our in-house data pipeline

Why does it **matter?**



Inflated enrollment counts results in overcharging clients

Accurate health data builds trust in other forecast and predictive health modeling for a population



2

Identify a solution

How to accurately disenroll a zombie individual

Define the solution

- What are the specific parameters that define an individual as a zombie
- What parameters define accurate disenrollment date

What is a **Zombie** Individual?

- EDC data source has a null disenrollment date
- Individual must also be present in DDC
 - Disenrollment date n/a

Determine EDC disenrollment date

- Ensure EDC disenrollment date is chronologically after the EDC enrollment date
- Set disenrollment date:
 - Disenroll EDC individual the month before lowest enrollment date on DDC
 - OR
 - If enrollment dates are identical for both EDC and DDC, set disenrollment date as one day after enrollment date

Example 1: Individual defined as having a 'zombie' eligibility entry

employer_ id	individual_id	first_name	last_name	enrollment_ date	disenrollment_ date	data_source
101	987654	Peppa	Pig	01/01/2018	06/01/2018	EDC
101	987654	Peppa	Pig	07/01/2018		Anthem

Peppa converted to DDC pipeline on July 1, 2018
However Peppa was never disenrolled from the EDC pipeline after the migration

Example 2: Individual defined as having a 'zombie' eligibility entry

employer_id	individual_id	first_name	last_name	enrollment_date	disenrollment_date	data_source
102	123456	Sonic	Hedgehog	01/01/2019	12/01/2019	EDC
102	123456	Sonic	Hedgehog	01/01/2020	12/31/2020	Anthem

Sonic is disenrolled from the health plan on our DDC feed
Sonic is still showing as enrolled in EDC

Example 3: Individual defined as having a 'zombie' eligibility entry

employer_id	individual_id	first_name	last_name	enrollment_date	disenrollment_date	data_source
103	456789	Spongebob	Squarepants	01/01/2020	01/02/2020	EDC
103	456789	Spongebob	Squarepants	01/01/2020	06/30/2020	Anthem
103	456789	Spongebob	Squarepants	01/01/2020		Anthem

- Spongebob is still enrolled in the EDC pipeline
- Spongebob is currently enrolled in DDC
- Same enrollment date for all rows require us to set disenrollment date for EDC as 01/02/2020

Note same enrollment date regardless of data source

Other factors to consider during implementation

- Accuracy and precision > speed
- Documentation
- Error handling
- Sandbox testing
- Peer Review

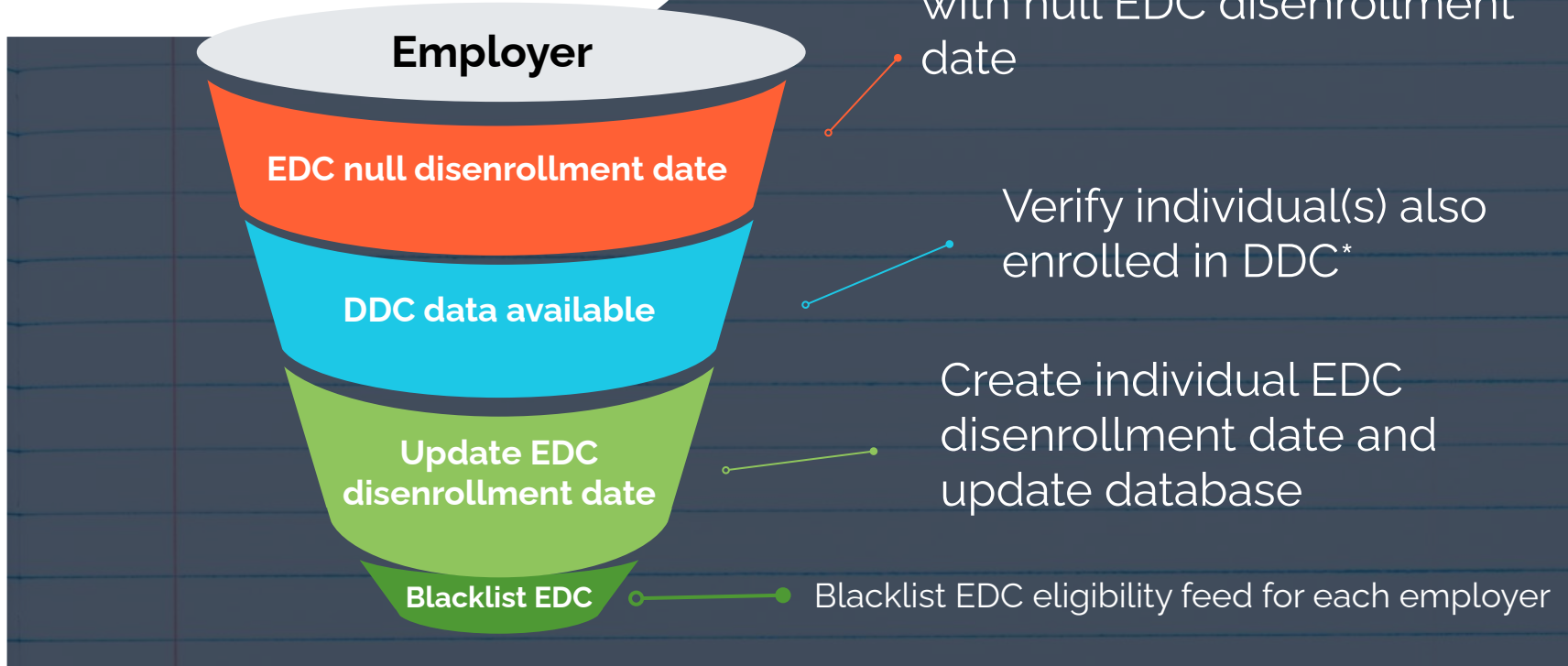


3

Implement Solution

Python and SQL

Solution Funnel



Python script

Brief presentation of
production script





4

Results and future uses

How to prepare for the next zombie eligibility

Overall results

Wide reaching impacts

9.2%

197 of 2000 employers had eligibility counts reduced

65,397

Total count of members disenrolled from EDC

2021 results

A Focused Analysis

19.4%

38 of 197 employers had eligibility counts reduced > 15% in 2021

7,047

Count of members disenrolled from EDC

Review of how **2021 zombie disenrollment results** are presented

	Employer Name	Member Count Pre	Member Count Post	Disenrolled	% enrollment improvement
No Change:	Piglets LLC	100	100	0	0%

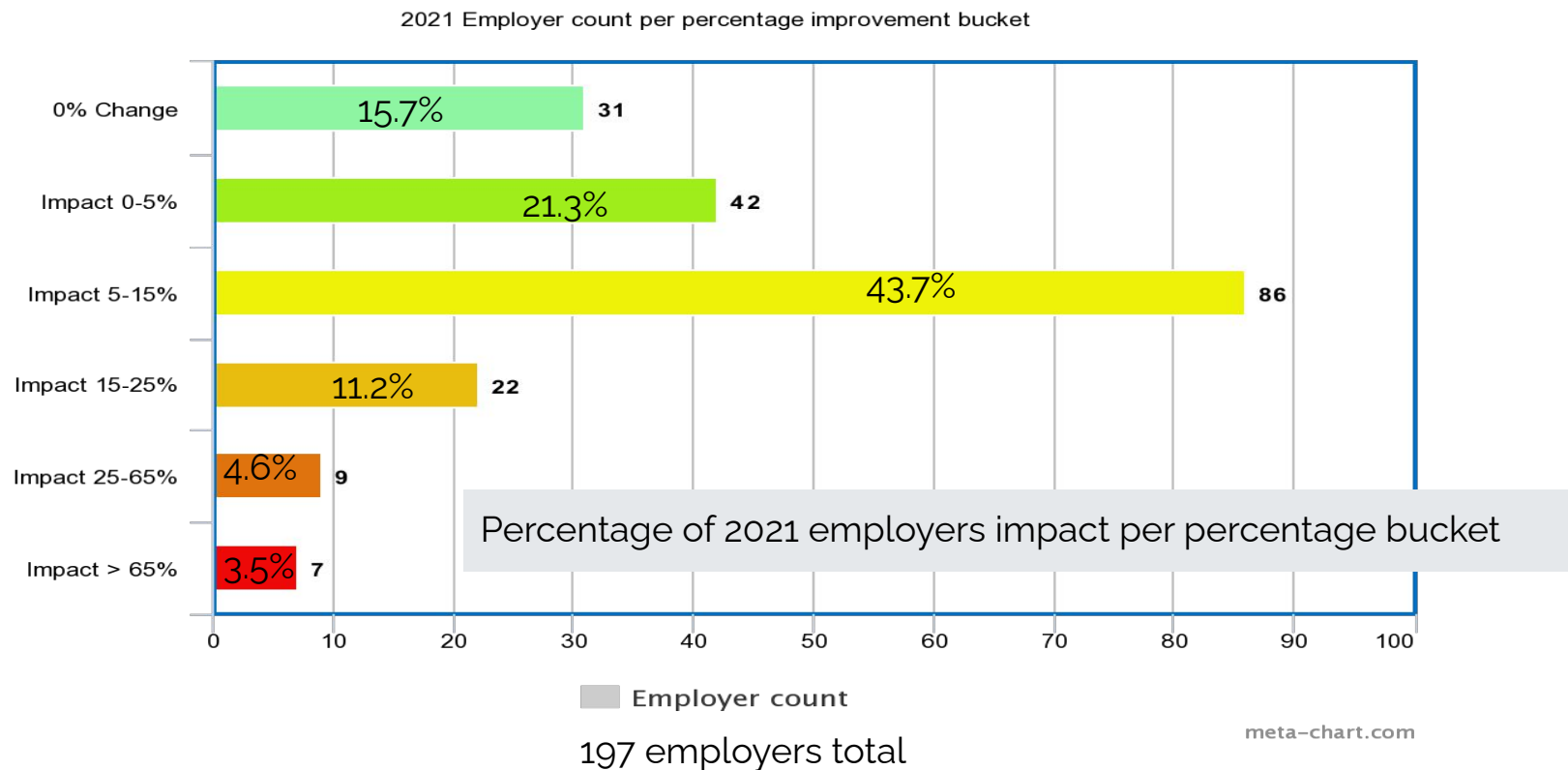
Review of how **2021 zombie disenrollment results** are presented

	Employer Name	Member Count Pre	Member Count Post	Disenrolled	% enrollment improvement
No Change	Piglets LLC	100	100	0	0%
Zombie Impact 5-15%	Hedgehog Holdings	101	89	12	13.5%

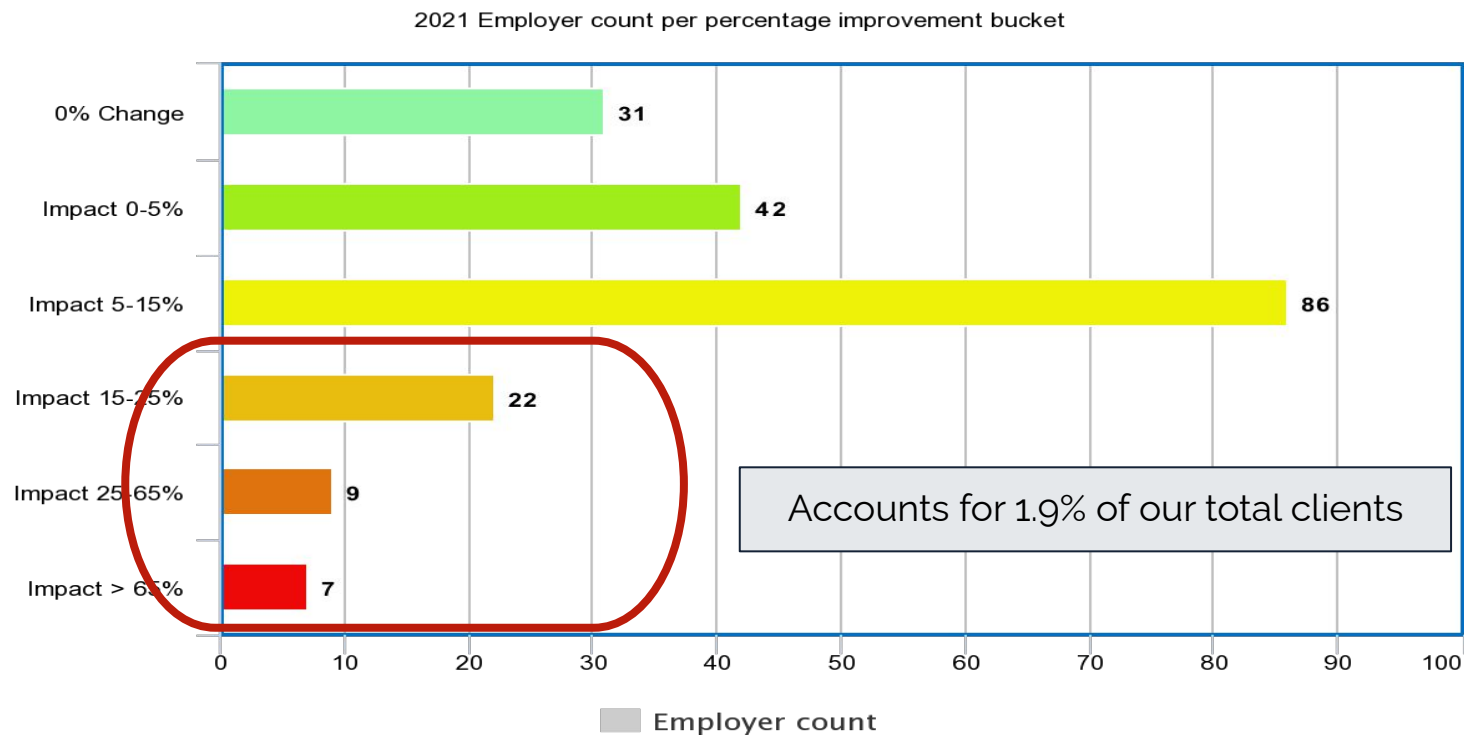
Review of how **2021 zombie disenrollment results** are presented

	Employer Name	Member Count Pre	Member Count Post	Disenrolled	% enrollment improvement
No Change	Piglets LLC	100	100	0	0%
Zombie Impact 5-15%	Hedgehog Holdings	101	89	12	13.5%
Zombie Impact > 65%	Spongebob Ocean Imports	390	182	208	114.3%

Count of employers per percentile **2021 eligibility decrease**



Count of employers per percentile 2021 eligibility decrease



197 employers total

meta-chart.com

Significance

Health data expects no more than 15% discrepancy between reported enrollment counts and control reports

Employers reported with > 15% eligibility count reduction in 2021 is reported to internal account managers for follow-up with client

Future Applications

Health Data:

- Similar inflated eligibility count issues have been observed that will benefit from this project

Other Industries:

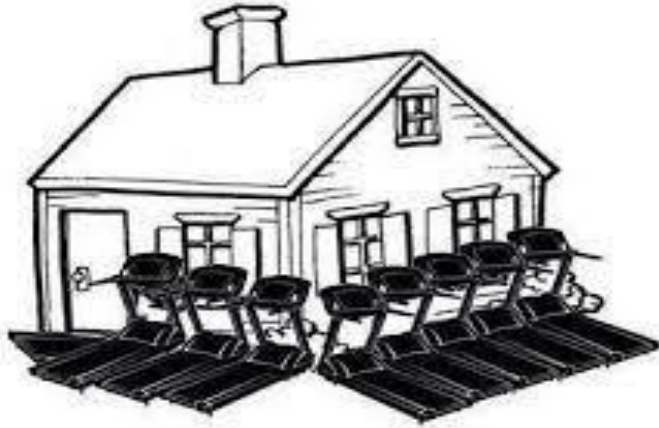
- Solution framework to remove duplicate database entries when you cannot delete before blacklisting



**"Oh, I do it to blend in. You know.
Zombies don't mess with other
zombies."**

–Bill Murray, *Zombieland*

The zombie treadmill defense



Thanks!

Questions?

Let's connect on LinkedIn or cminerich@fastmail.com

END

Count of employers per percentile bucket with 2021 eligibility decrease

2021 eligibility decrease: pre to post disenrollment	Number of employers disenrolled	% of disenrolled employers impacted 2021
No Change	31	15.7%
0-5%	42	21.3%
5-15%	86	43.7%
15-25%	22	11.2
25-65%	9	4.6%
> 65%	7	3.5%
	197	100%