Homeland Security

Office of Inspector General
Office of Audits

Leadership Perspective on Using Advanced Data Analytics

Presentation by:
Maryann Pereira
DHS OIG, Office of Audits

Misconception of Analytics

- Data Analytics is THE buzz word of the day
- Everyone thinks oh we have analytics now, everything will be faster, better, bigger
- Some actual quotes I've heard:
  - "Just attach AI to it"
  - "Machine learning can replace most processes"
  - "Once we have the software or contractor, we will see immediate benefit"
- Not reality – major benefits BUT research and knowledge are KEY
- Example using LWA

Lost Wages Assistance Audits (OIG-22-69 and OIG-22-73)

We conducted these audits to determine to what extent FEMA ensured that states and territories distributed; and DHS employees received supplemental Lost Wages Assistance (LWA) to eligible recipients.

- FEMA's LWA provided up to $44 billion in financial relief.
- Since the enactment of the Coronavirus Aid, Relief, and Economic Security (CARES) Act, states have experienced a surge in fraudulent unemployment claims involving cyber scams and identity theft.
- Unemployment benefits have the highest rate of fraud across the federal government.
- We reviewed the LWA payments dispersed by 21 participating states and territories which accounted for 81 percent of the total LWA benefits obligated.
- Using advanced data analytics, we identified more than $3.7 billion in improper payments and potentially fraudulent payments.

Advanced Data Analytics

Office of Inspector General
Homeland Security

Benefits of Using Advanced Data Analytics

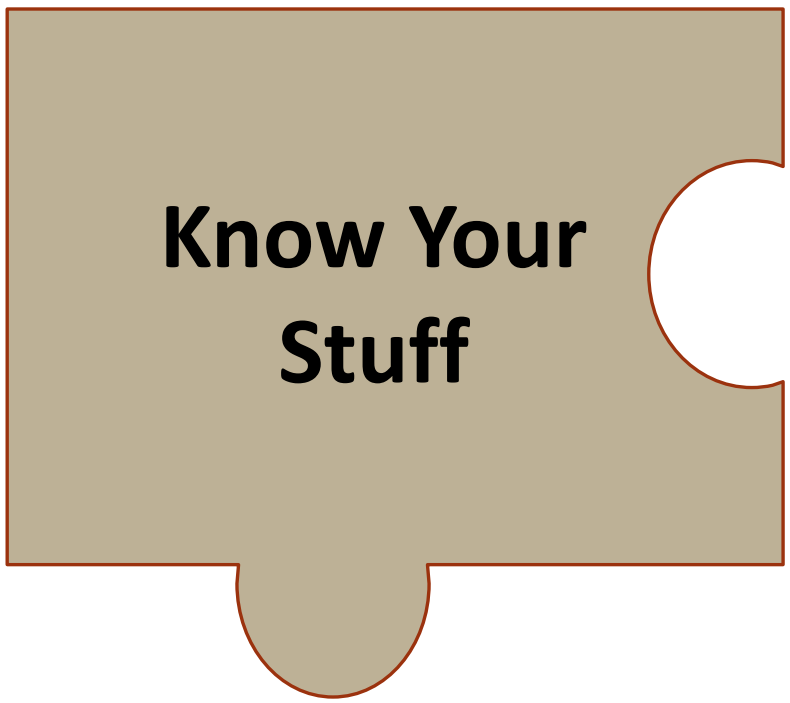
Able to analyze large quantities of data where manual audit analysis is prohibitive and subject to human error	Increase Audit/Inspection/Investigative scope	Efficiency, reduced timelines and work hours
Increase in savings without need for projection	Software advances, low/no-cost - Microsoft Access, Audit Command Language, Tableau, IDEA, Microsoft Power BI, etc.	Visualization – interactive heat maps, graphs, charts, etc.

While the benefits are many, **ADVANCED DATA ANLYTICS DOESN'T REPLACE HARD WORK AND AN INQUISTIVE NATURE**

Steps BEFORE Applying Analytics

Office of Inspector General
Homeland Security

Puzzle Piece One



- Critical thinking to answer the objective
- Must KNOW the criteria
- Mastery of the program under audit
- Identify the significant risks
  - Internal Controls
  - Fraud Risk Assessment – use what is out there, DODIG Fraud Portal, public facing website
- Identify Data Sources
- Example from our assessment of LWA, self-certification

Steps BEFORE Applying Analytics

Office of Inspector General
Homeland Security

Puzzle Piece Two

- On our LWA Audit we worked with:
  - Dept of Labor OIG
  - Pandemic Response and Accountability Committee
  - DHS OIG Office of Investigation
  - DHS OIG Office of Innovation

- Benefits of this collaboration
  - Sharing of known fraud indicators
  - Identifying what ‘data’ exists
  - Brainstorming ideas on how to ‘mine’ the data

Insights from Oversight Partners

Steps BEFORE Applying Analytics

Office of Inspector General
Homeland Security

Puzzle Piece Three

- Working with our Oversight Partners to get data:
  - Each of our partners had data we needed to test transactions
  - Developed Data Use Agreements
  - Transferred data securely
- A few examples:
  - Personnel rosters: home and work addresses
  - Listing, by individual, for assistance applications and payment transactions
  - Payroll data showing employment status
  - Geolocation data showing the IP address of the individual applying for assistance, this included city, state, zip, latitude, longitude

Steps BEFORE Applying Analytics

Puzzle Piece Four

- Understanding the data
- Wrong conclusion
- Independent testing

Know your Data

Office of Inspector General
Homeland Security

Personnel Data	employee rosters, employee residence, and work addresses
LWA applications and transactions	a listing, by individual who applied for and/or received unemployment and lost wages assistance
Payroll	employee payroll and time-keeping information to help identify employment or unemployment status
Bank Accounts	bank account and routing numbers for an employee's payroll, in other words, where did they deposit normal salary paychecks, and for lost wages and unemployment assistance payments was the same bank account used
IP addresses	Geolocation data (city, state, zip, latitude, longitude) associated with the IP address used to file for lost wages or unemployment assistance

	A	B	C	D
1	**Table**	**Column**	**Datatype**	**Description**
2	Categories	CategoryID	Int	Primary key
3	Categories	CategoryName	NVarChar (15)	Category name
4	Categories	Description	NText	Provided by marketing dept.
5	Categories	Picture	Image	Deprecated
6	CustomerCustomerD	CustomerID	NChar (5)	Link to Customers table
7	CustomerCustomerD	CustomerTypeID	NChar (10)	Link to CustomerDemographics
8	CustomerDemograp	CustomerTypeID	NChar (10)	Primary key
9	CustomerDemograp	CustomerDesc	NText	
10	Customers	CustomerID	NChar (5)	Primary key

Notional Example

Steps BEFORE Applying Analytics

Office of Inspector General
Homeland Security

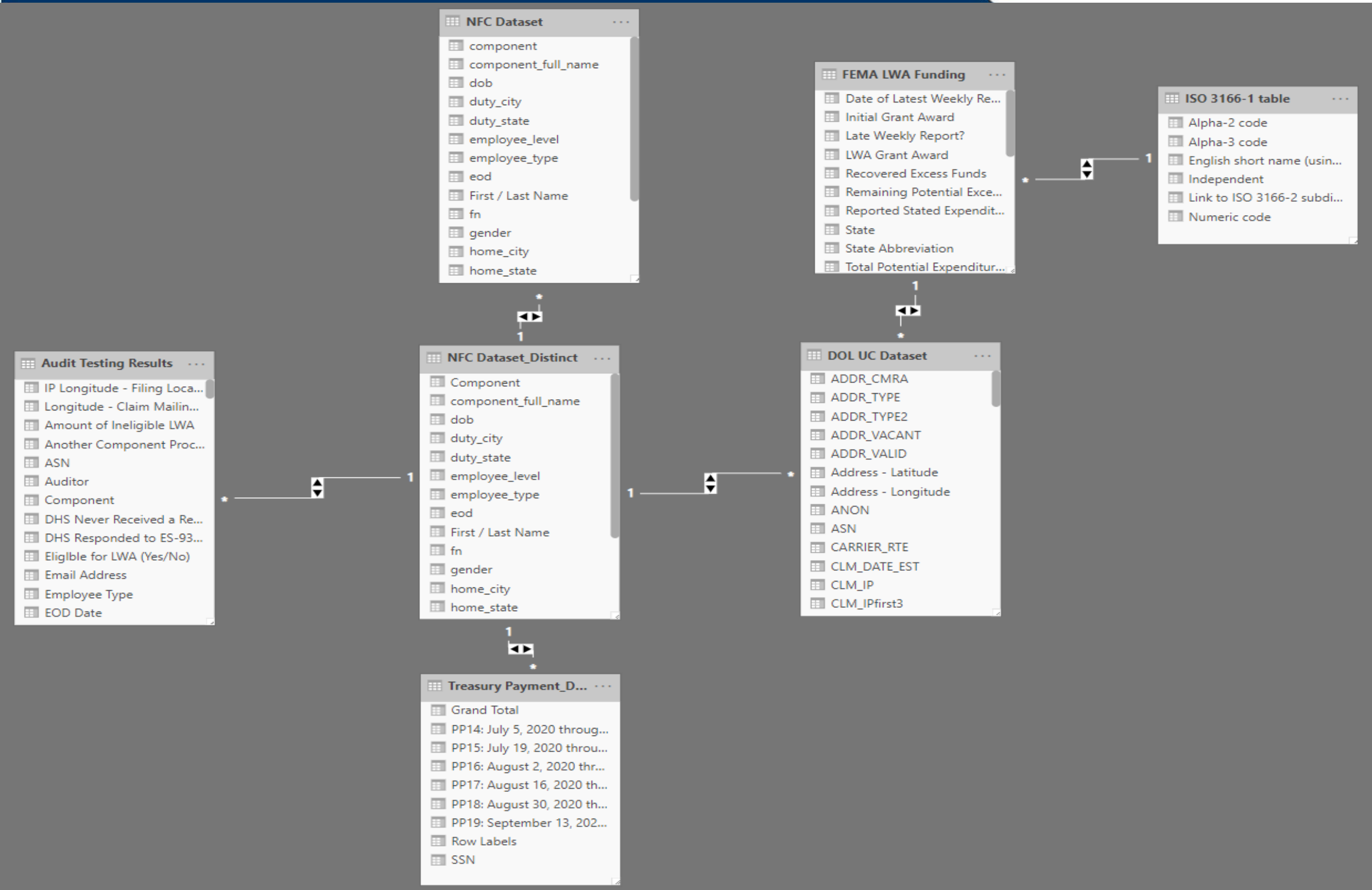
Puzzle Piece Five

Normalize and Combine Data

- Relating
- "Like" fields
- LWA combined 7 different datasets

Related Tables

Office of Inspector General
Homeland Security



Applying Analytics

Puzzle Piece Six

- 14 fraud risk indicators run against 7 datasets
- Risk scoring of each indicator
- Examples
  - Individuals not unemployed
  - Bank Account anomalies
  - Location anomalies
  - Multi-state claims
  - Multiple claims, unique IP address
  - Fictitious email addresses
- Indicators ARE NOT necessarily fraud

Fraud Risk Indicators

Office of Inspector General
Homeland Security

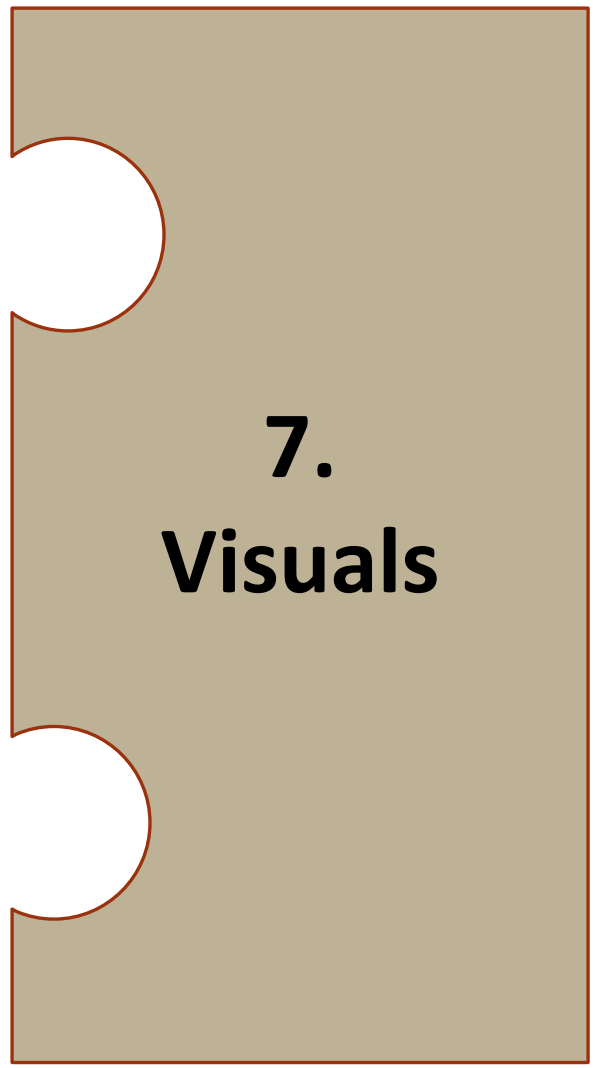
Fraud Risk Indicator		DHS Component										
No.	Description											Total
1	Claim Name Not Associated with Social Security Number	20	0	19	1	10	0	23	4	10	4	**91**
2	Out-of-State Claim	142	5	133	1	46	9	146	19	39	15	**555**
3	Claim Home State Discrepancy	143	5	132	1	49	9	141	17	40	14	**551**
4	Non-DHS Former Employer	103	1	140	1	31	4	163	9	31	10	**493**
5	Questionable Email Address	116	4	169	0	36	5	162	8	44	7	**551**
6	Another DHS Component Listed	5	0	22	0	3	1	27	0	2	1	**61**
7	Multistate Claim	57	1	141	0	29	2	95	9	24	6	**364**
8	Multiple Claims from IP Address	50	1	86	0	10	1	56	5	11	1	**221**
9	Multiple Claims Using Bank Account	1	0	10	0	1	0	6	0	0	0	**18**
10	Filed Claim Outside U.S.	4	0	1	0	0	0	1	0	1	0	**7**
11	Filed Claim Outside State	67	3	130	0	19	3	101	11	29	4	**367**
12	Filed Claim using DHS Network	1	0	26	0	1	0	6	0	1	0	**35**
13	DHS Employment and Bank Account	9	0	4	0	4	0	147	1	2	1	**168**
14	Multiple Claims with Same Mailing Address	28	0	74	0	11	0	45	4	11	4	**177**

Source: DHS OIG analysis of DOL UI and LWA program data

Creating Visuals

Puzzle Piece Seven

- Let's create visuals using our data



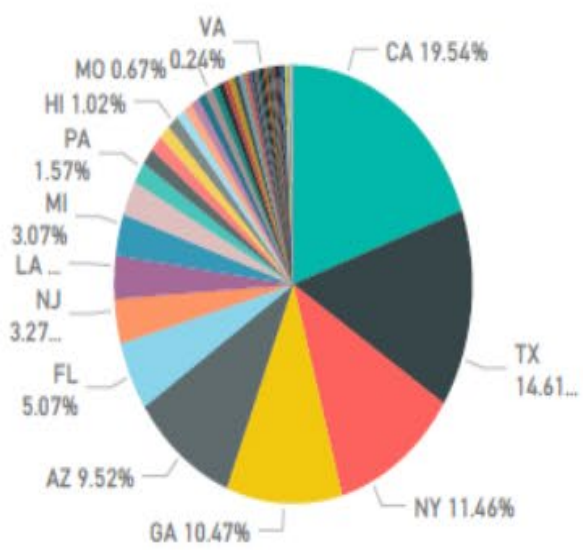
LWA – Data Visuals

Office of Inspector General
Homeland Security

LWAP Amount

$3,306,976

DHS LWA Benefits by State



DHS LWA Benefits by Component

Component	LWAP Amount
	$1,570,560
	$1,108,156
	$310,170
	$136,140
	$103,350
	$29,700
	$23,400
	$14,400
	$7,500
	$3,600
Total	$3,306,976

Major Fraud Risk Indicators

23.0%
% Claims with Questionable Email Addr...

23.2%
% Claims Filed in Another State

15.3%
% Claims Filed From Another State

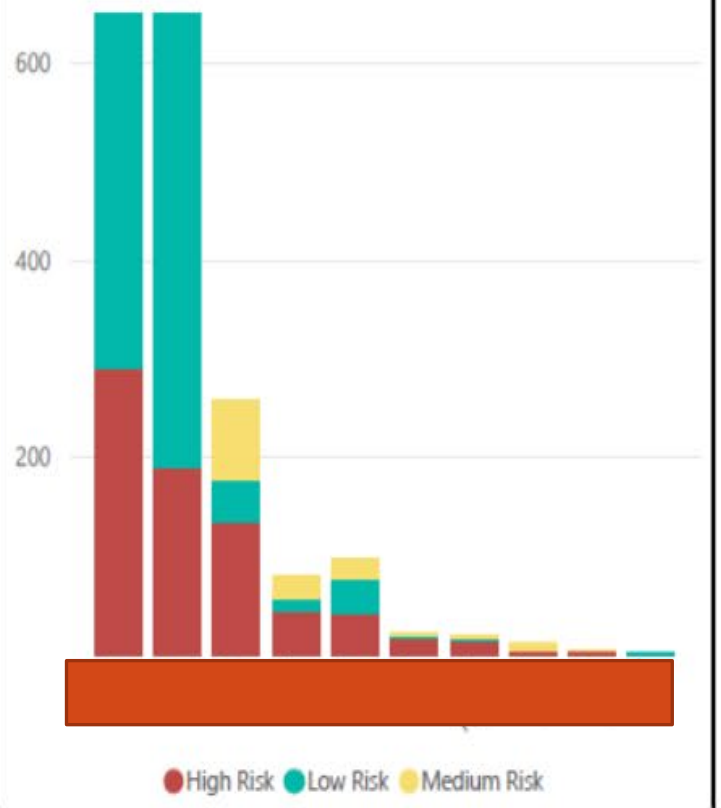
9.2%
% Claims Filed with Same IP Address

8.1%
% Claims Filed in Multiple States

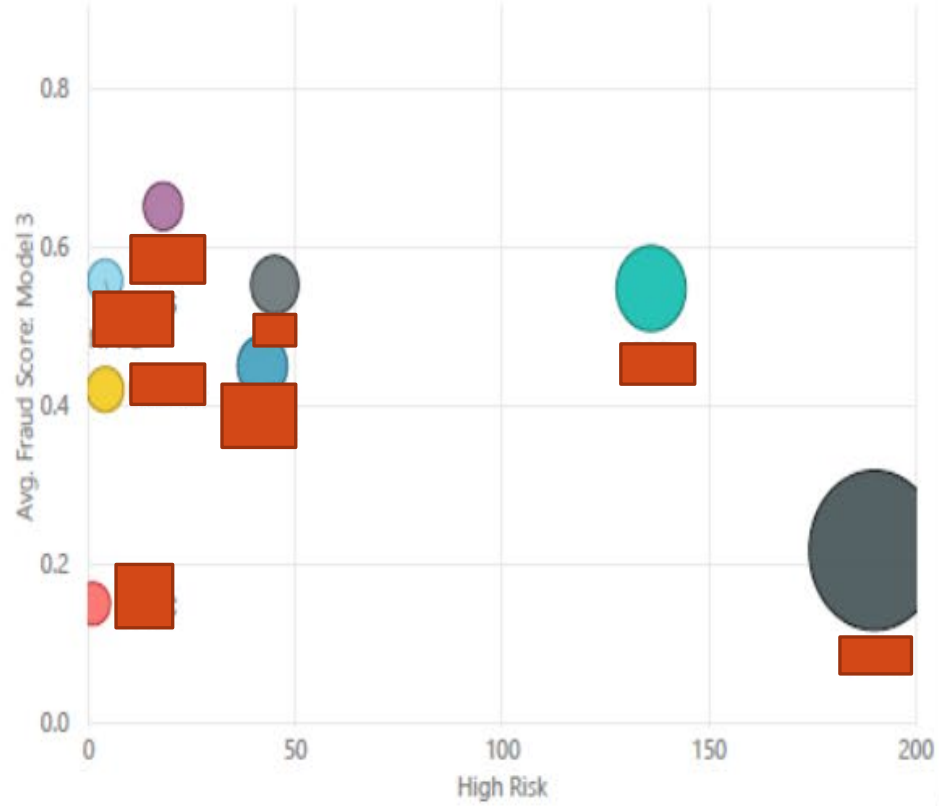
High Risk Claims by Component

Component	High Risk	Total Claims	% High Risk
	292	769	38%
	190	1123	17%
	136	259	53%
	45	81	56%
	42	97	43%
	18	23	78%
	15	20	75%
	4	12	33%
	4	6	67%
	1	3	33%
Total	**747**	**2393**	**31%**

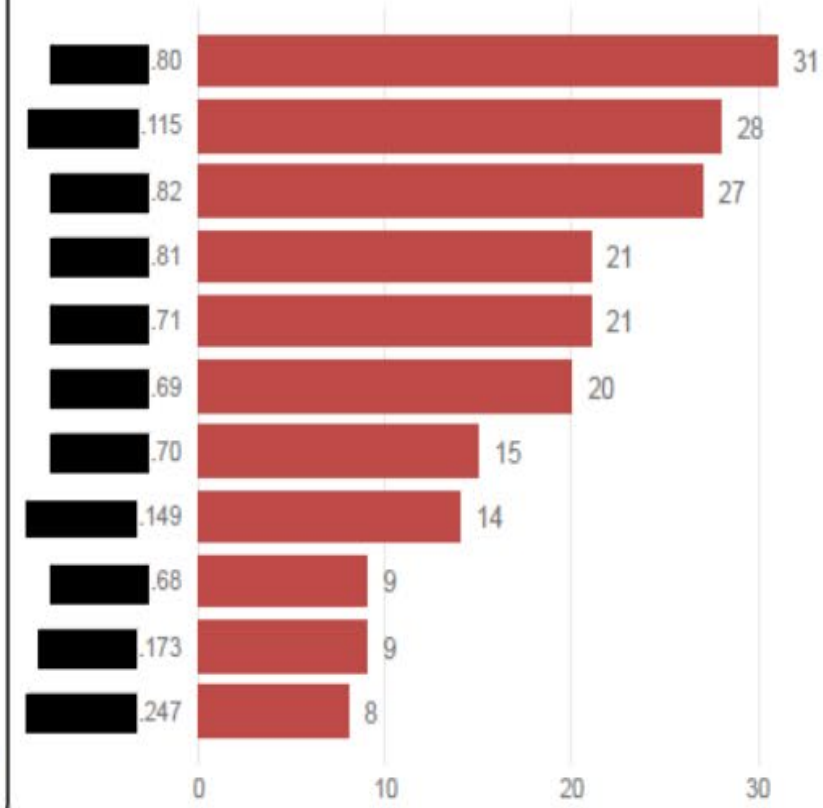
Breakdown of Claim Risk by Component



Avg. Fraud Score by No. of High Risk Claims



Top 15 IP Addresses Used to File UC Claims for DHS Employees



Summary

Seven Puzzle Pieces to Data Analytics

1. Know Your Stuff
2. Insights from Oversight Partners
3. Working with Our Partners to Get Data
4. Know Your Data
5. Normalize and Combine Data
6. Analytics
7. Visuals

THANK YOU

ANY QUESTIONS?