

Data Divers Part 2: I have this data, now what?

Jamie Black, MPH

Josh Dowell, MPH

Kiera Schultz, MPH

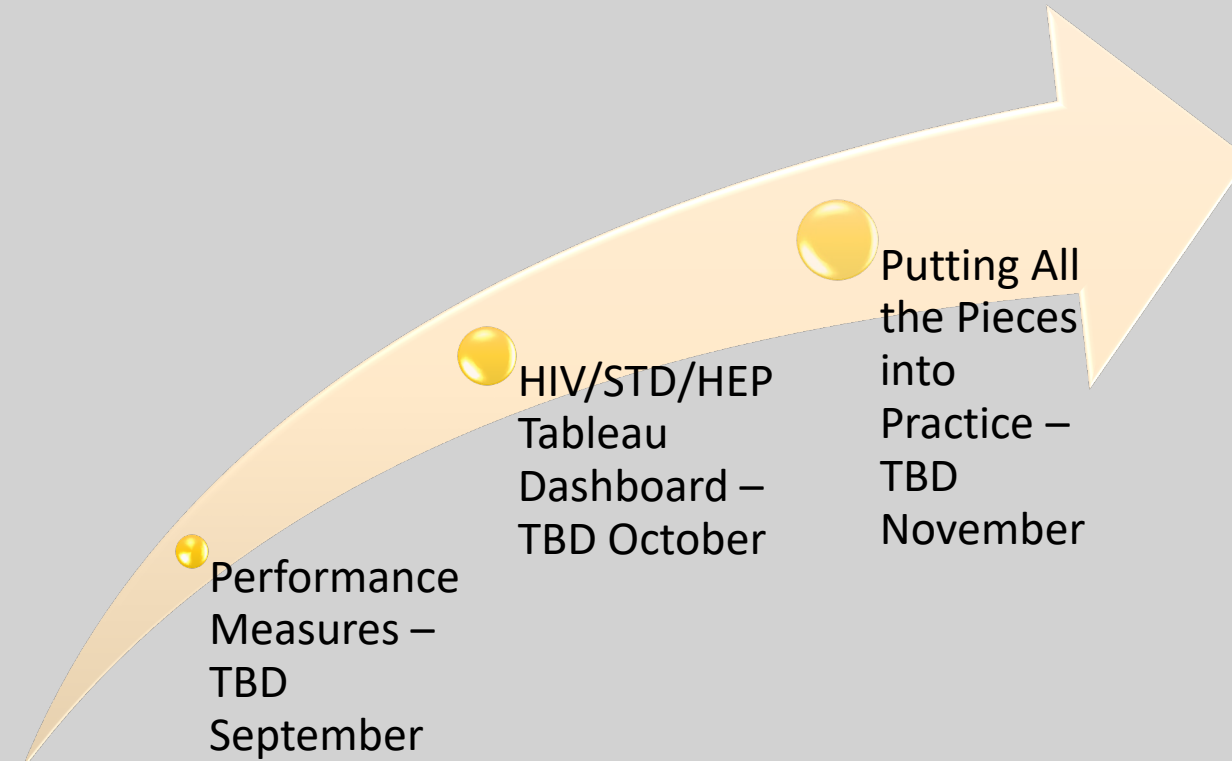
Colin Stretch, MPA

Conner Tiffany, MPH

August 24th, 2021

A large yellow triangle is positioned in the bottom right corner of the slide, pointing towards the top right.

Future Data Diver Topics





Objectives

- What can the data tell us
 - Aggregate vs. Disaggregate data
- What can we build with the data
 - Outbreak tracker
- How do we know the data is “good”
 - Principles of data cleaning
- How can I circumvent anxiety related to data analysis
 - Tips and tools

Recap from Presentation 1



Review of common epidemiology / data definitions

- Calculations for prevalence, incidence
- Priority population overview
- Surveillance vs Services data

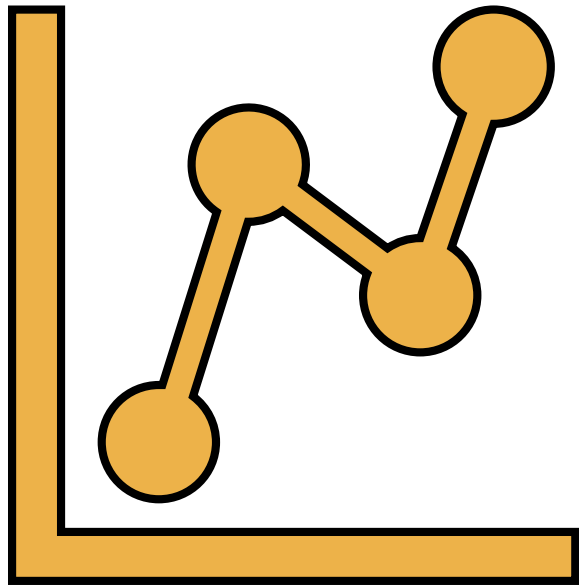


Review of data systems

- Surveillance Systems
- Services Systems
- Importance of defining data system

Anti-stigma Language

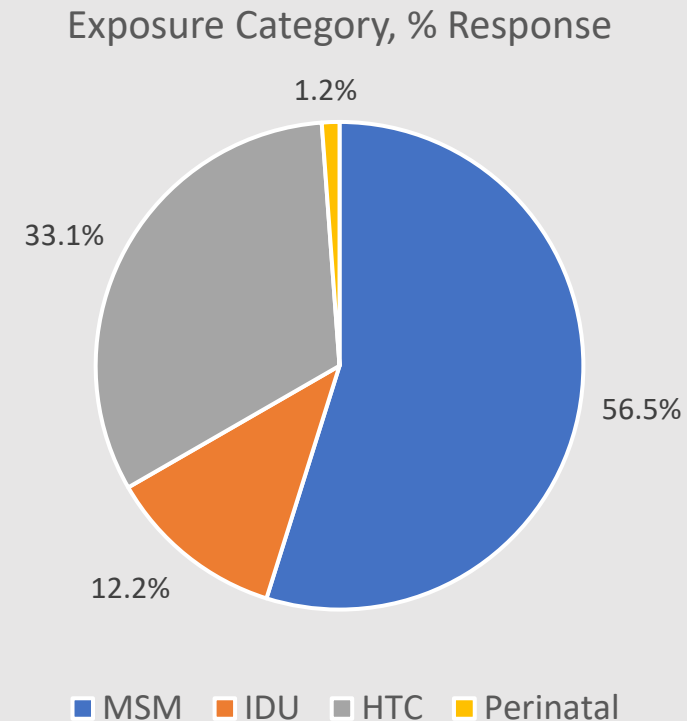
- Person-centered language
- Importance of stigma reduction
- Avoid terms such as:
 - “Infected”
 - “Addict”
 - “Clean”
 - “Risky” or “promiscuous”
- For more resources, reach out to us



Let's take a look
at some data

Aggregate Data- HIV Risk Factor, Indiana

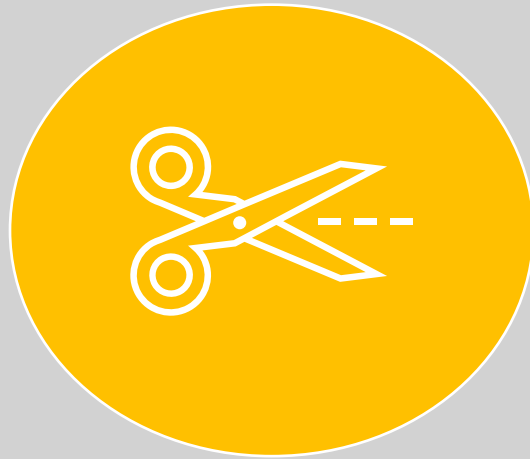
Exposure Category	PLWH
MSM	5747
IDU	415
HETEROSEXUAL (HTC)	2700
MSM & IDU	469
IDU & HTC	491
MSM & HTC	888
MSM & IDU & HTC	200
PERINATAL	157
OTHER	31
NO IDENTIFIED RISK	1468
NO REPORTED RISK	352



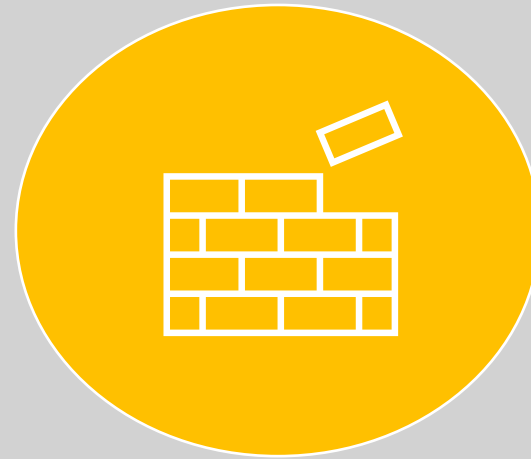


What does
this data tell
us?

Aggregate vs. disaggregate data



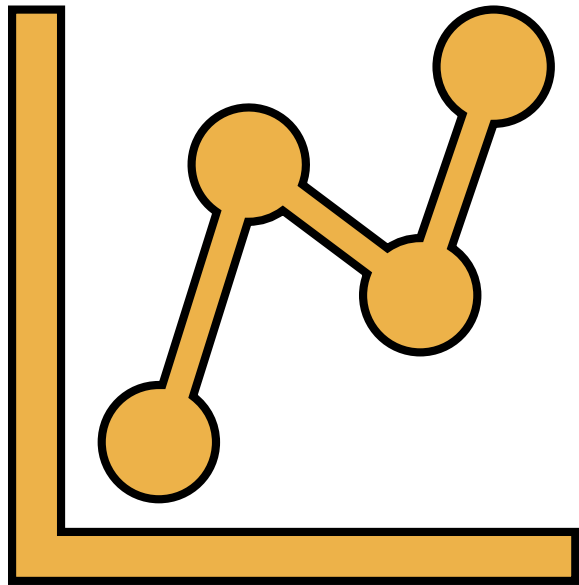
Disaggregate:
Separate into smaller units



Aggregate:
Compile, summary data

Disaggregate units

- Demographics
 - Race/ethnicity
 - Age
- Geographic location
 - ZIP region
 - DIS district
- Risk factor
 - MSM
 - IDU
- Comorbidities
 - HIV and HCV
 - Chlamydia and HIV
- Service Category
 - VLS and Food Bank
 - PLWH and screening
- PLWH and Hep B Vaccine

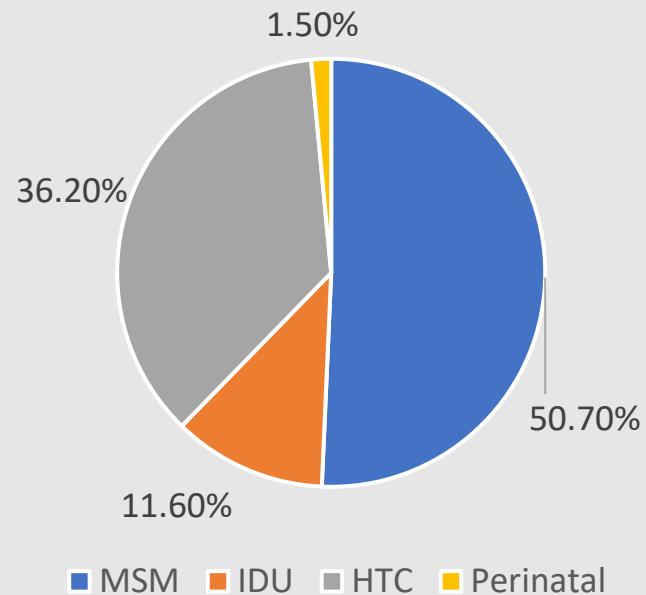


Let's take a look
at *disaggregate*
data

HIV Disaggregate Data – Risk Factor / ZIP Coalition region

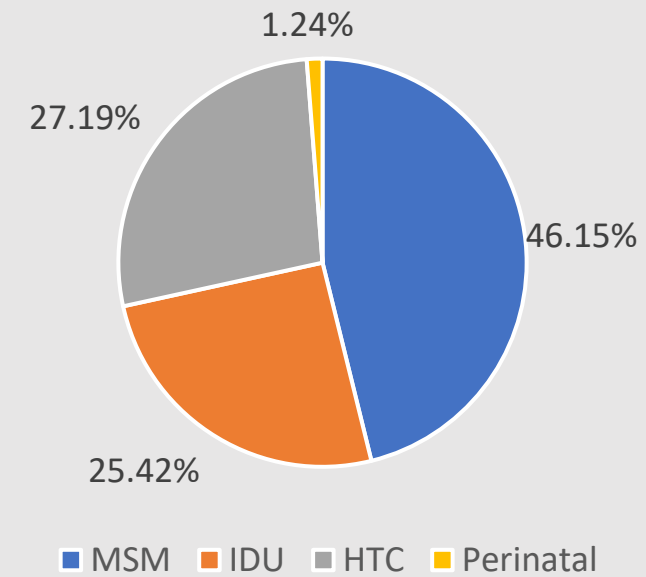
ZIP Coalition 1 -

Exposure Category	PLWH
MSM	599
IDU	57
Heterosexual (HTC)	341
MSM & IDU	34
IDU & HTC	67
MSM & HTC	130
MSM & IDU & HTC	22
PERINATAL	21
NO IDENTIFIED RISK	296
NO RISK REPORTED	66



ZIP Coalition 10

Exposure Category	PLWH
MSM	345
IDU	118
HETEROSEXUAL (HTC)	120
MSM & IDU	37
IDU & HTC	82
MSM & HTC	48
MSM & IDU & HTC	8
PERINATAL	8
NO IDENTIFIED RISK	70
NO REPORTED RISK	22



STD Disaggregate Data – Gender, Race, and Age

Rate, Gonorrhea Cases	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
MALES	90.4	84.2	96.2	99.8	104.1	115.0	135.9	171.6	171.8	161.3
White	17.7	18.9	24.8	33.0	39.2	42.6	52.6	64.8	69.9	65.9
0-9	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0
10-12	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
13-24	49.0	61.9	65.4	90.5	94.6	109.3	121.2	132.3	135.6	122.9
25-34	46.5	44.2	71.4	80.4	111.1	115.7	145.2	185.7	213.6	192.2
35-44	16.1	13.0	21.2	31.0	41.7	42.3	63.1	93.5	99.4	100.4
45-54	7.6	4.8	9.1	17.3	18.1	21.3	29.7	39.1	39.8	40.5
55-64	0.9	1.4	3.9	5.2	7.0	7.2	11.0	12.8	14.9	20.2
65+	0.0	0.9	0.0	1.1	1.3	1.8	2.5	2.6	2.6	2.7
Black or African American	672.4	599.2	682.1	685.0	691.8	726.7	808.3	1021.9	941.7	929.7
0-9	7.8	1.9	0.0	0.0	0.0	1.9	0.0	0.0	0.0	0.0
10-12	6.3	6.2	0.0	0.0	6.5	0.0	0.0	0.0	0.0	0.0
13-24	1895.2	1668.8	1781.0	1727.2	1738.7	1772.0	1913.7	2329.0	2009.5	2043.5
25-34	1367.6	1199.0	1432.1	1511.4	1476.4	1624.8	1846.0	2420.8	2340.1	2091.6
35-44	454.4	384.6	503.1	525.6	541.0	594.5	769.3	860.6	814.9	945.9
45-54	116.8	170.2	201.3	202.6	221.2	229.1	231.8	356.0	415.4	379.3
55-64	31.5	45.1	93.7	55.7	101.1	97.6	73.3	178.6	129.7	155.0
65+	15.3	5.0	14.5	9.2	4.4	0.0	16.2	46.8	11.2	25.1
Other Race	367.4	330.4	291.0	183.4	128.5	218.9	299.9	378.2	417.1	278.3
0-9	3.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.8
10-12	0.0	25.4	0.0	0.0	0.0	0.0	10.9	0.0	0.0	0.0
13-24	1027.0	834.2	692.1	397.9	248.6	428.3	558.3	698.0	737.3	428.7
25-34	608.0	483.2	543.4	367.6	285.6	487.2	678.9	848.2	887.2	708.2
35-44	184.6	308.3	208.8	138.0	126.9	198.7	258.9	344.0	490.8	324.1
45-54	137.2	167.4	119.9	74.1	55.7	107.4	188.9	237.0	273.4	146.5
55-64	29.3	54.8	65.2	74.7	47.3	56.7	86.6	92.6	96.5	54.9
65+	0.0	19.3	0.0	32.4	0.0	13.7	25.4	70.9	55.1	41.0

Per 100,000

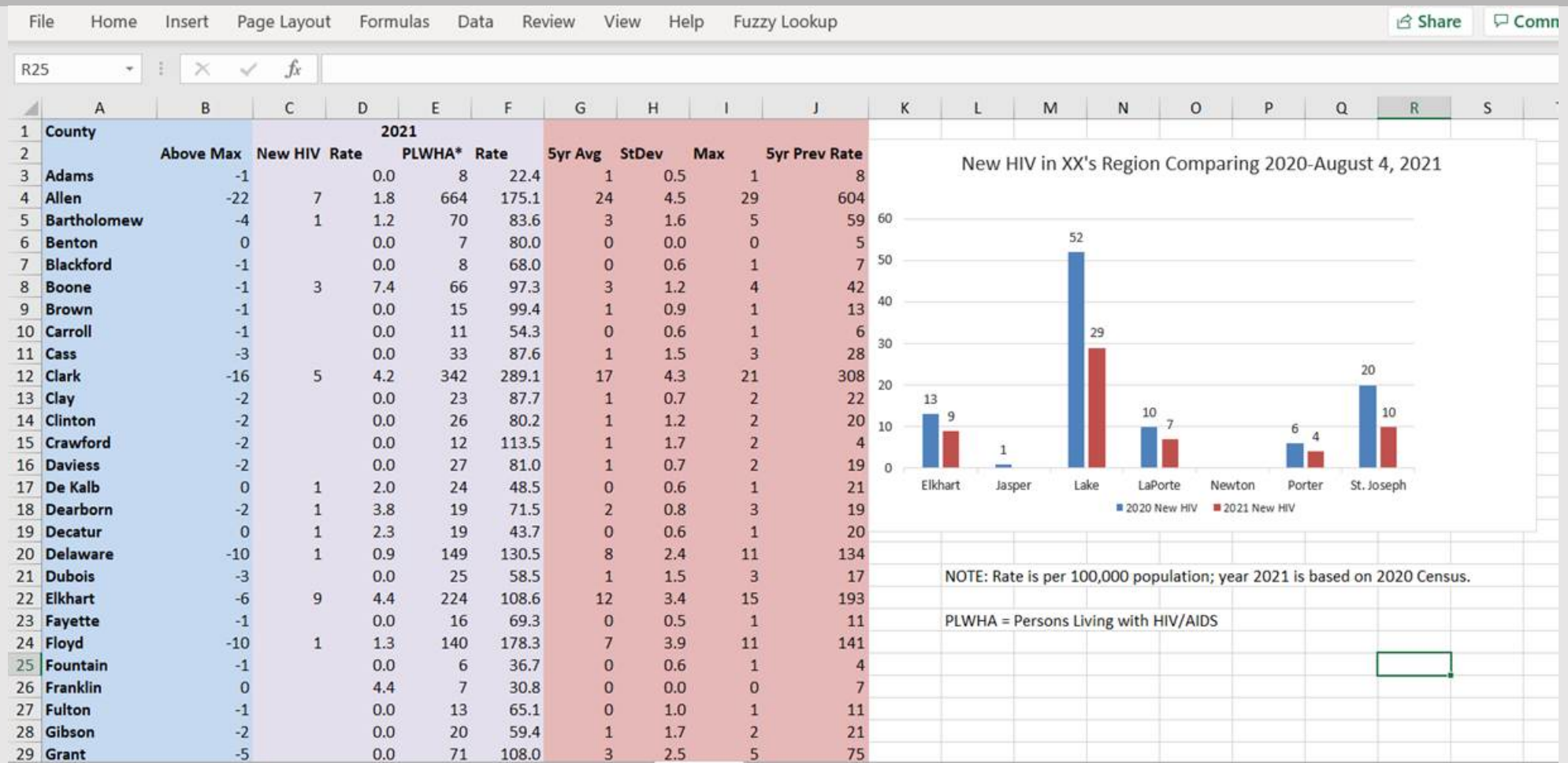
HCV Disaggregate Data - Age

Acute/Chronic Probable/Confirmed Newly Reported HCV Cases per 100,000 population

Year	0-17	18-29	30-39	40-49	50-59	60-69	>69
2010	2.2	85.7	119.0	153.5	204.0	72.3	19.9
2011	2.1	104.7	118.2	125.3	177.4	69.5	19.7
2012	2.6	122.5	139.9	118.6	156.3	74.6	19.1
2013	1.9	142.9	140.2	95.2	138.0	71.9	16.2
2014	u	175.1	176.1	104.3	132.0	80.8	14.0
2015	2.4	202.7	199.9	118.0	130.1	92.9	15.6
2016	5.0	221.7	254.9	132.3	148.8	115.9	19.2
2017	4.1	218.0	275.6	142.9	150.4	137.9	26.6
2018	1.5	189.4	277.7	144.2	141.0	136.1	26.5
2019	u	131.5	226.7	124.9	102.8	116.5	28.0

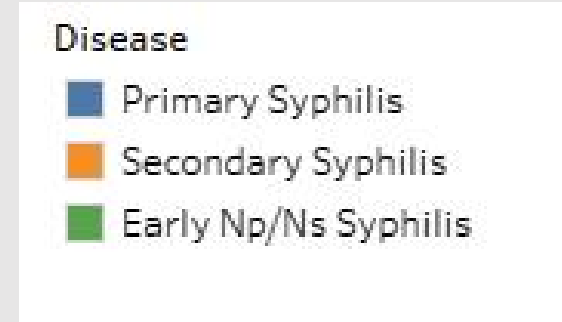
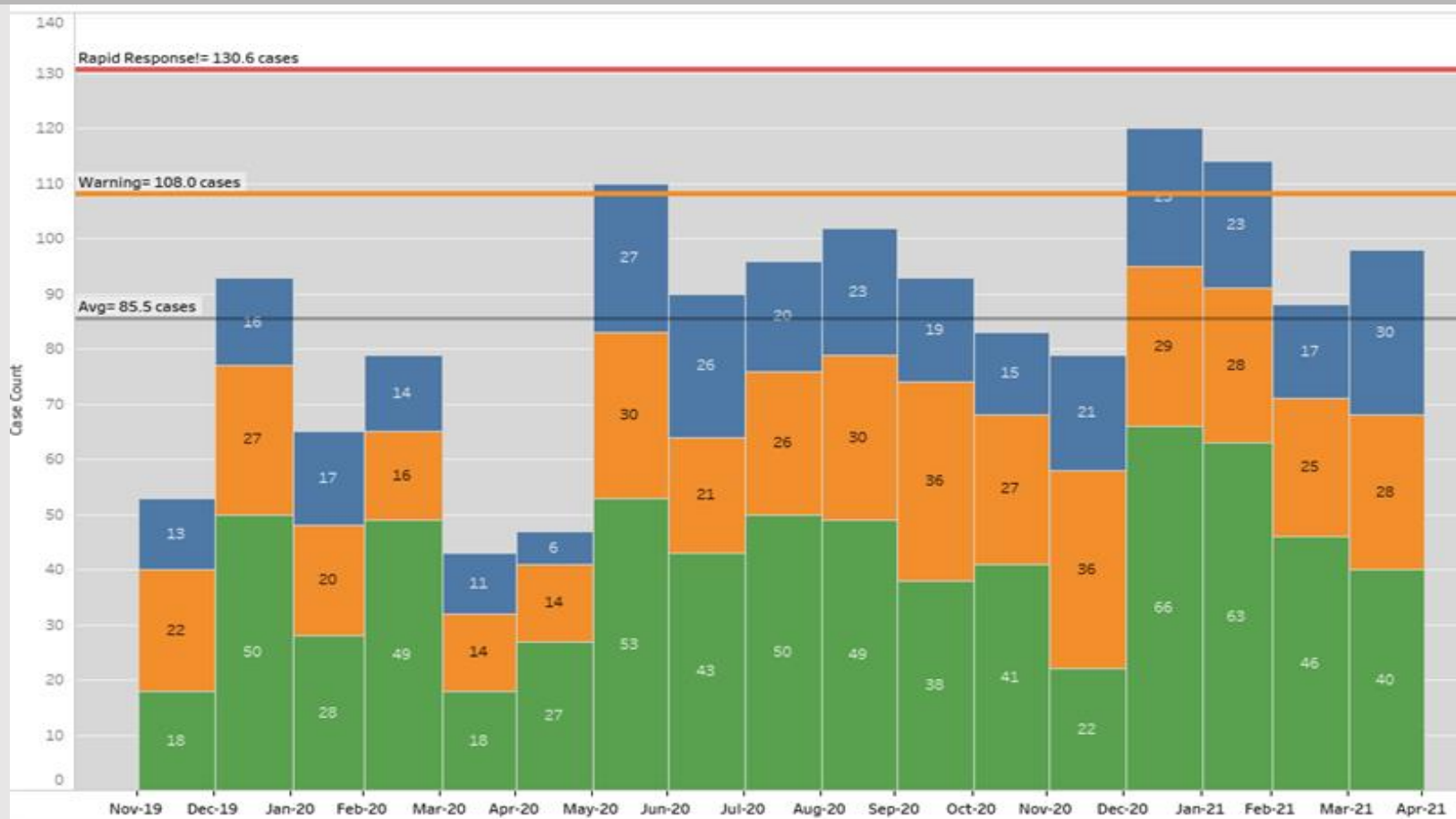
u = unstable rate

Outbreak Tracker* - HIV



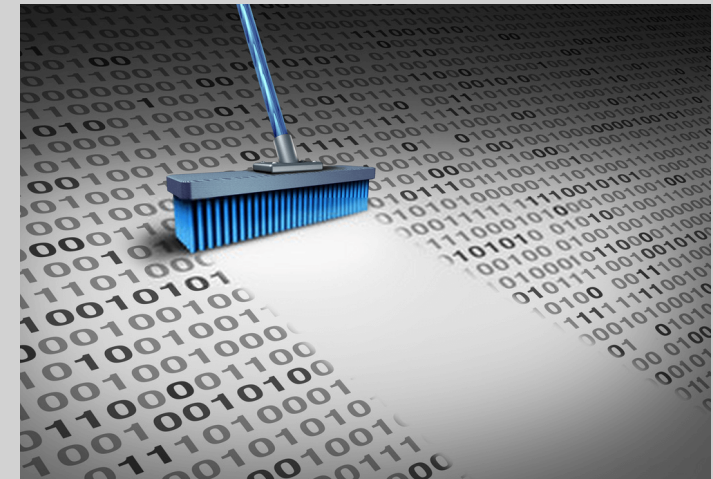
* This data is not representative of true rates/counts

Outbreak Tracker - STD



Data Cleaning

- Data cleaning starts **BEFORE** receiving data
 - What kind of lab tests are valid?
 - How do we get all valid results?
- Multi-step process that includes:
 - Deduplicating
 - Transforming
 - Integrating
 - Verifying
- Necessary to maintain high quality data



01

Play around
with the data

02

Remember
your goal(s) and
objective(s)

03

Transform data
into visual
representations

04

Ask for help –
we are here

Mitigating Data Anxiety

Mitigating Data Anxiety

© MARK ANDERSON

WWW.ANDERSTOONS.COM



"After analyzing all your data, I think we can safely say that none of it is useful."



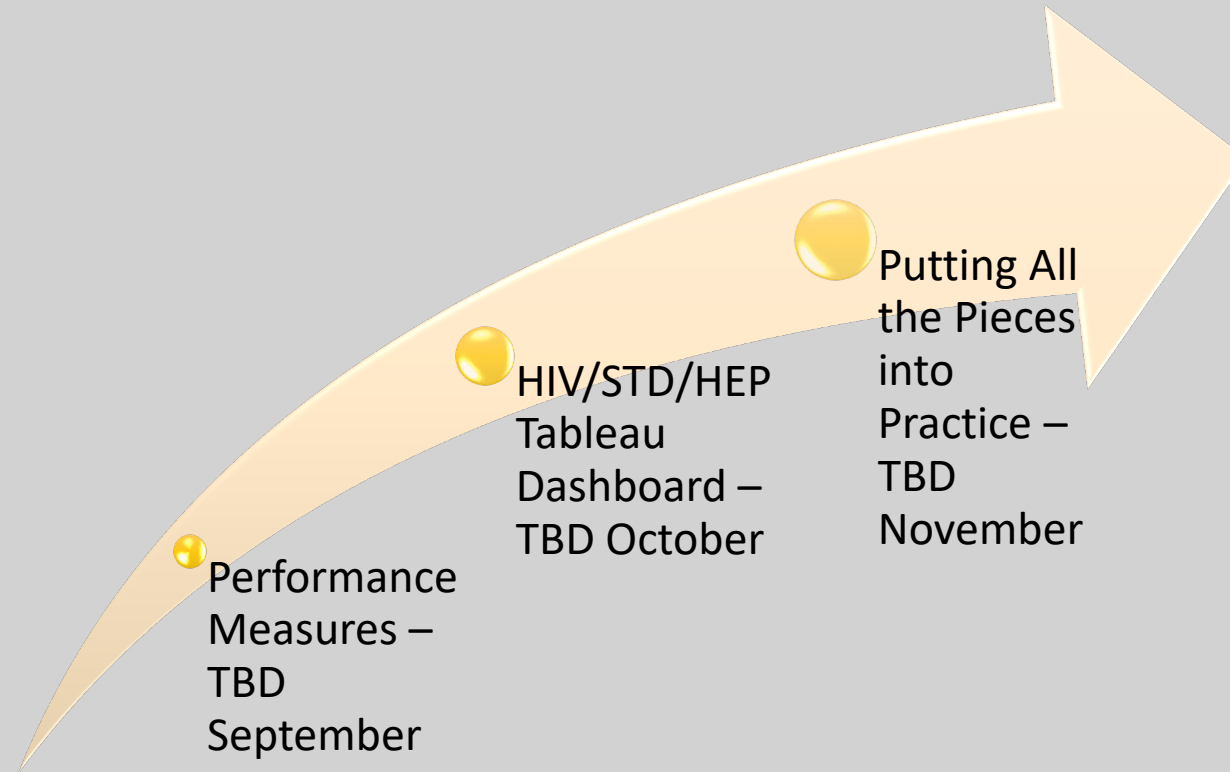
**I want you to keep
calm and disaggregate
your data!**

(and ask for our help if needed!)

Summary

- What can the data tell us
 - Aggregate vs. Disaggregate data
- What can we build with the data
 - Outbreak tracker
- How do we know the data is “good”
 - Principles of data cleaning
- How can I circumvent anxiety related to data analysis
 - Tips and tools

Reminder – Upcoming Data Diver Topics



Follow-Up Survey

<https://redcap.isdh.in.gov/surveys/?s=NDKMAEKWLN>

Contact Information

- Jamie Black, MPH: jamblack@isdh.IN.gov
- Josh Dowell, MPH: jdowell@isdh.IN.gov
- Kiera Schultz, MPH: kschultz2@isdh.IN.gov
- Colin Stretch, MPA: cstretch@isdh.IN.gov
- Conner Tiffany, MPH: ctiffany@isdh.IN.gov