Dataviz and Storytelling

@bengarvey IREG 2018

https://docs.google.com/presentation/d/142BhRbLHMX_nAE8sbLN0rdYM0leO82egBOB52R-pxow/edit#slide=id.p

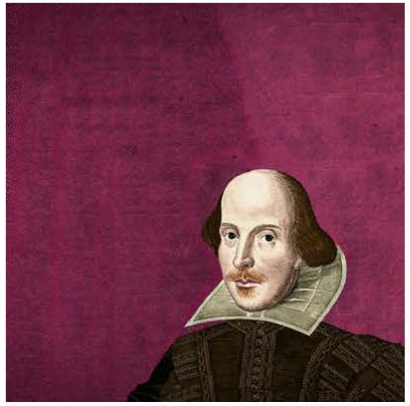
About me

Ben Garvey

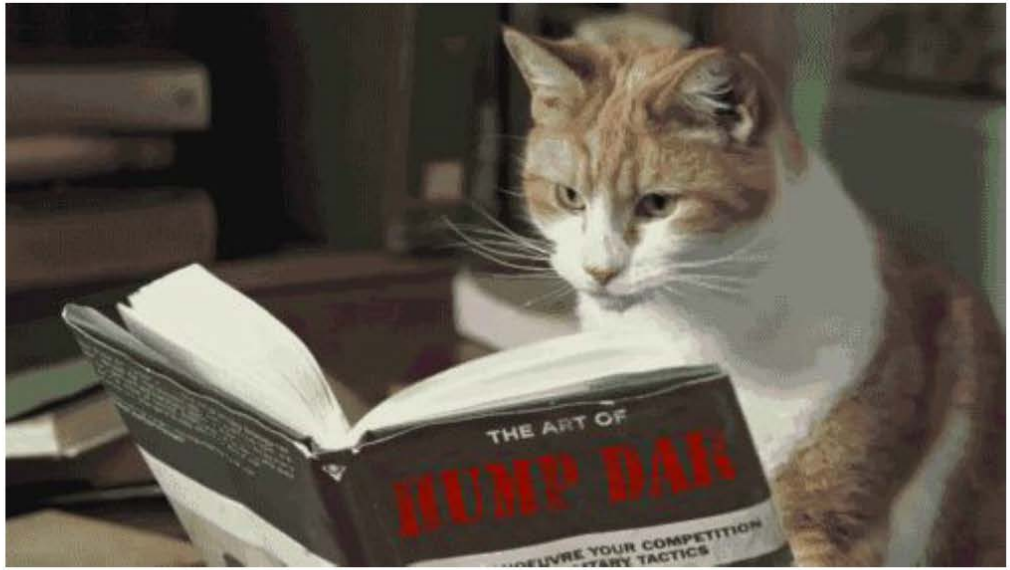
@bengarvey

Engineering Manager

Magento Business Intelligence (formerly RJMetrics)



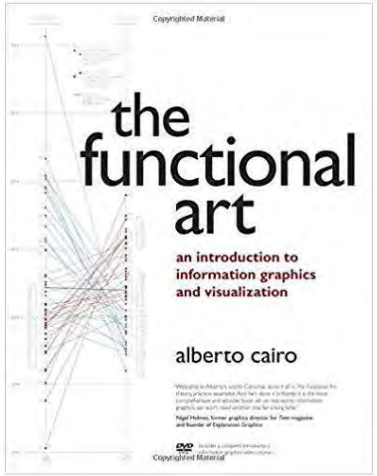

ALL RIGHT,
BUT YOUR NEXT STORY...

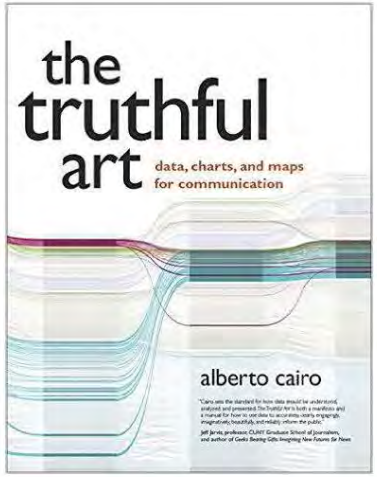

THE ART OF
HUMP DAY

Alberto Cairo ✓
@albertocairo

Following

"Storytelling" is a term that ought to be abandoned in journalism, #dataviz, data, etc. It has no meaning and leads to the wrong mindset





Alberto Cairo @albertocairo ·Sep 11

Replying to @albertocairo

...I'll rush to add that I've used it a lot, but when I saw its effects, I stopped

2 8

Alberto Cairo @albertocairo ·Sep 11

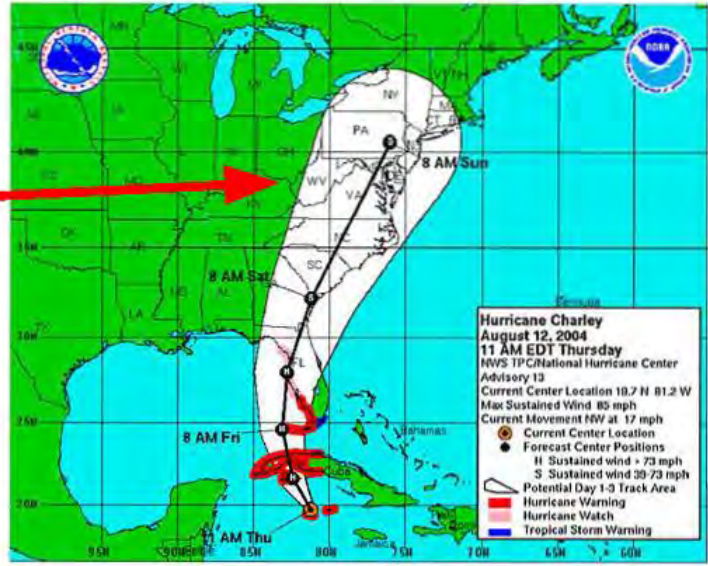
3) "Storytelling" is like the cone of uncertainty: you mean something when using it but people hear/see something entirely different

4 4 14

Alberto Cairo @albertocairo ·Sep 11

4) And as Donald Norman said, if most people misinterpret you, the problem isn't them. It's you, your design, and your words

8 5 25




LOUISIANA
MISSISSIPPI
ALABAMA
GEORGIA
TEXAS
Houston
Galveston
Corpus Christi
New Orleans
Biloxi
Mobile
Pensacola
Apalachicola River
FLORIDA
St. Petersburg
Key West
CUBA
MEXICO
Rio Grande
Gulf of Mexico
8PM Tuesday
8PM Monday
8PM Sunday
8PM Saturday
Hurricane
CAIRO
(category 5)


LOUISIANA
MISSISSIPPI
ALABAMA
GEORGIA
TEXAS
Houston
Galveston
New Orleans
Biloxi
Mobile
Pensacola
Corpus Christi
FLORIDA
St. Petersburg
Key West
CUBA
MEXICO
Rio Grande
Gulf of Mexico
Hurricane
CAIRO
(category 5)

Neil Halloran @neilhalloran ·Sep 11

Replying to @albertocairo

When the goal is making data-driven arguments appealing to wide audiences - not always the goal, but an important craft for our democracy 1/

1

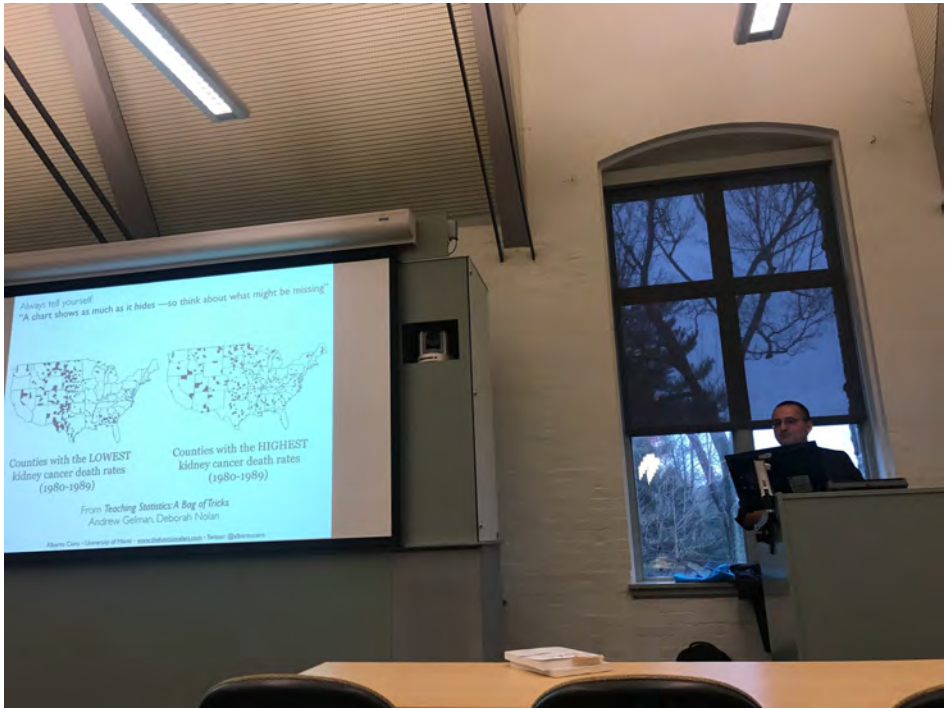
Neil Halloran @neilhalloran ·Sep 11

There is an imp. distinction between an argument and a story. Scientists, lawyers, academics make arguments. Storytellers have diff job /2

2 4 5


NOV | DEC | JAN | FEB | MAR | APR | MAY | JUN | JUL | AUG | SEP | OCT | NOV | DEC | JAN | FEB | MAR | APR
1944
fallen.io

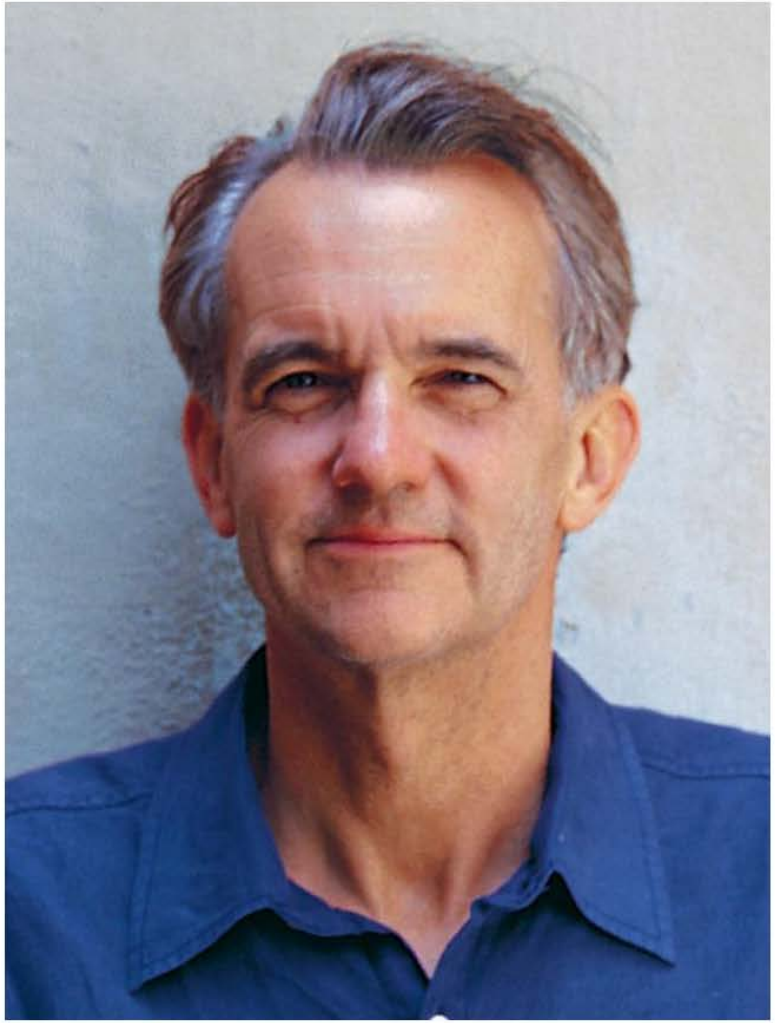
Who is right?



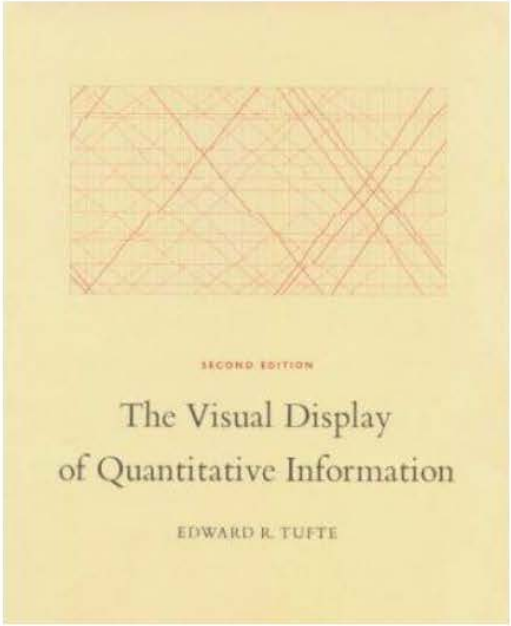
Me: “Is storytelling impossible or are we just very bad at it?

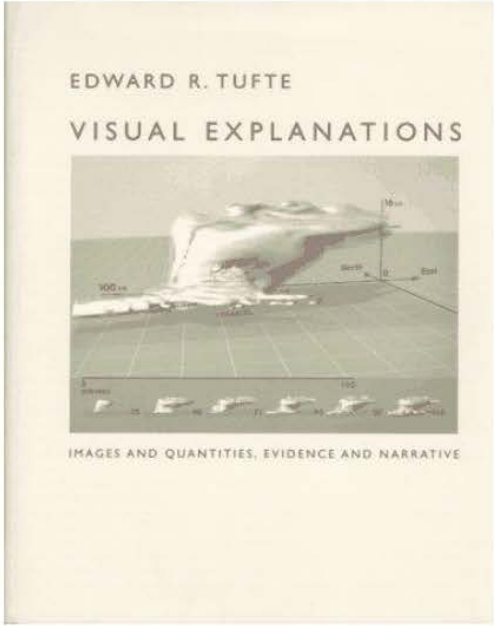
Alberto Cairo: “The latter.”

“Charts are not stories. They are visual arguments.”
- Alberto Cairo



“Strive to show causality”
- Ed Tufte





What is a story?

- An account of incidents or events
- A statement regarding the facts pertinent to a situation in question
- Anecdote; especially an amusing one
- A fictional narrative shorter than a novel
- The intrigue or plot of a narrative or dramatic work

What is a story?

- An **account of incidents** or events
- A statement regarding the facts pertinent to a situation in question
- Anecdote; especially **an amusing one**
- A fictional narrative shorter than a novel
- **The intrigue or plot of a narrative or dramatic work**

We have an **obligation** to
to use dataviz and
storytelling responsibly

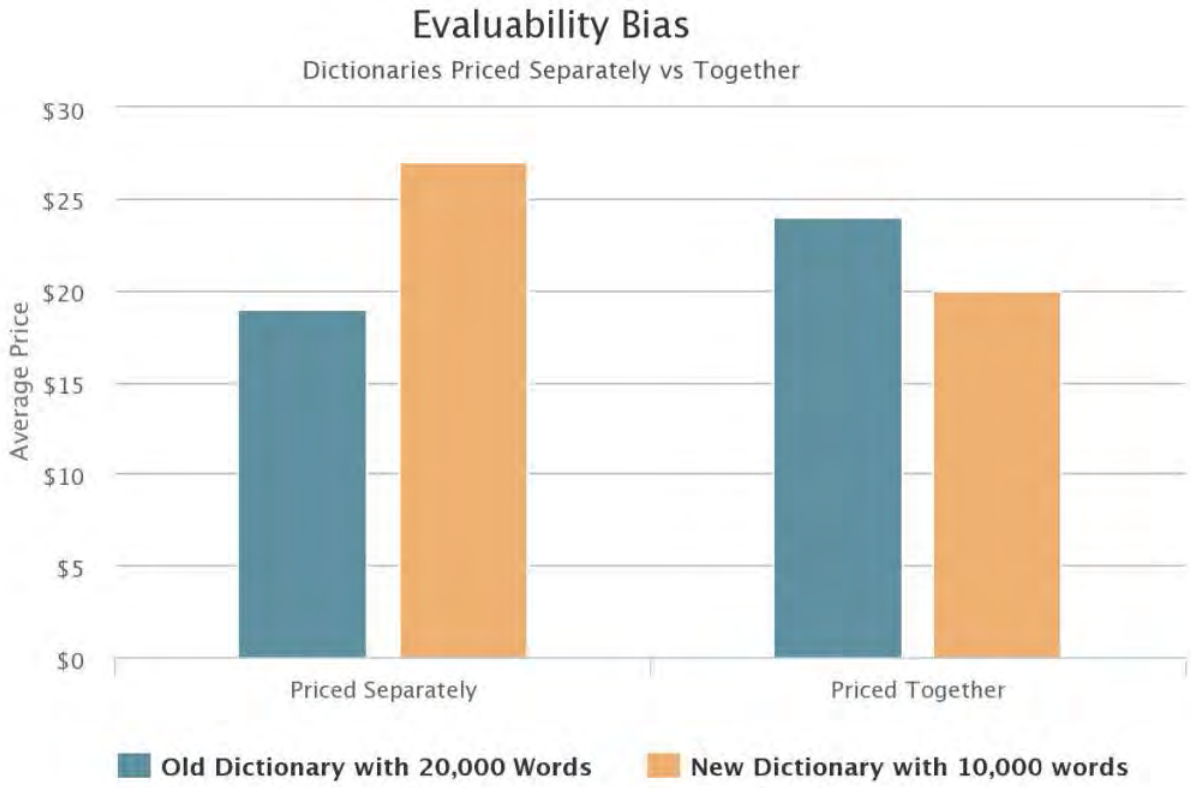
Why?
Because dataviz can easily exploit cognitive biases

Narrative Fallacy

The narrative fallacy addresses our limited ability to look at sequences of facts without weaving an explanation into them, or, equivalently, forcing a logical link, an arrow of relationship upon them. Explanations bind facts together. They make them all the more easily remembered; they help them make more sense. Where this propensity can go wrong is when it increases our impression of understanding.

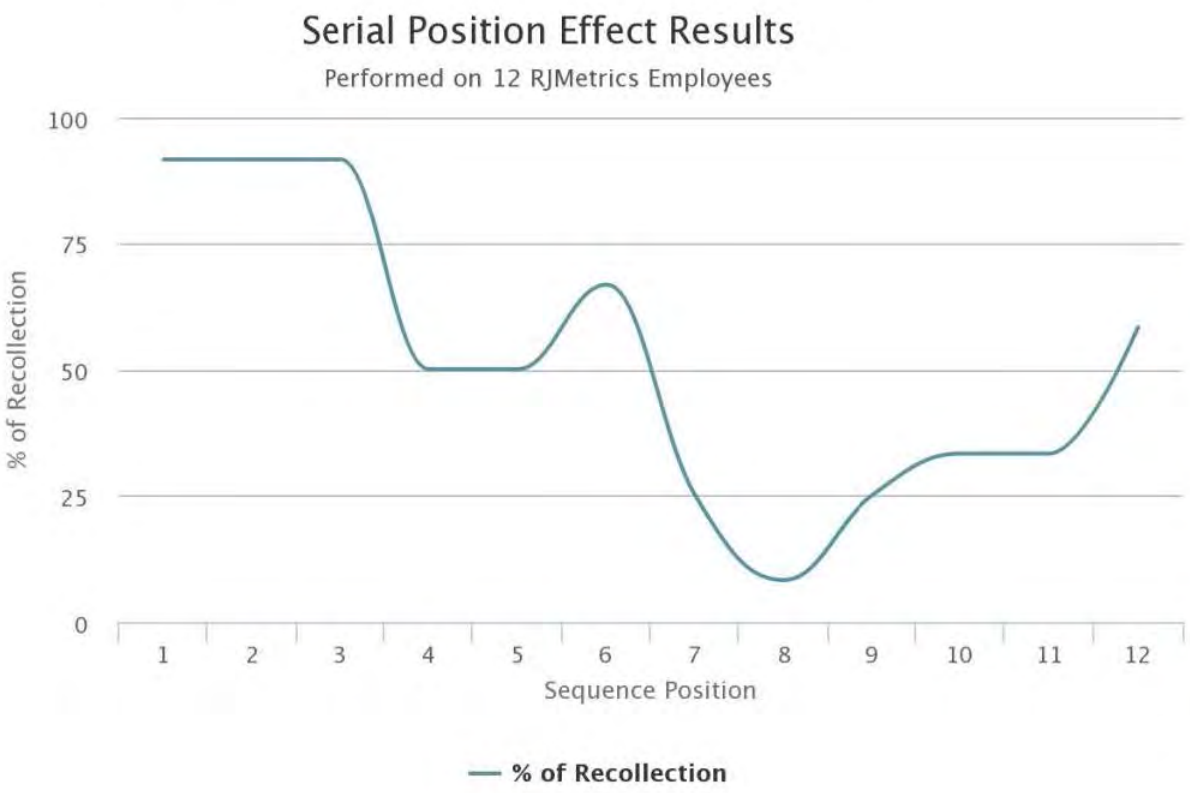
—Nassim Nicholas Taleb, The Black Swan

Evaluability Bias

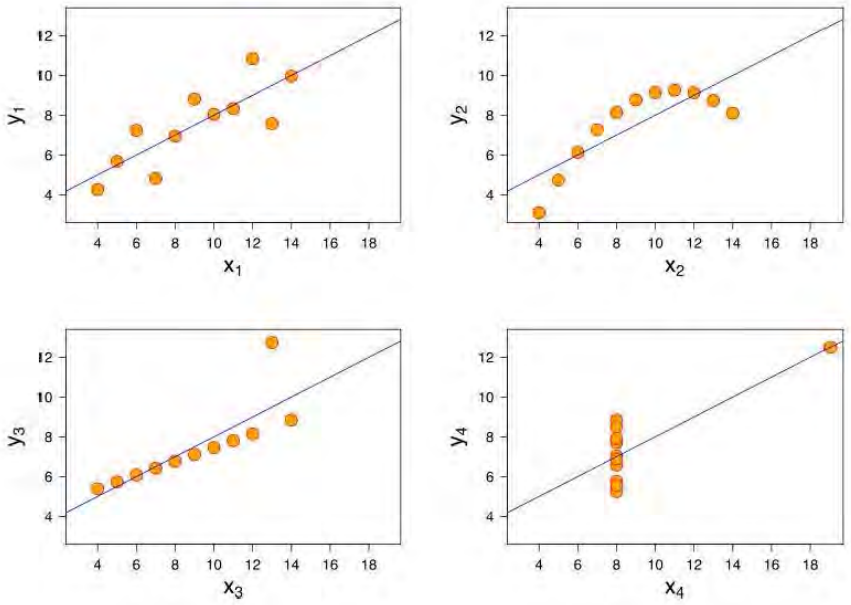


Christopher Hsee

Serial Position Effect



Hermann “in tha” Ebbinghaus



So how do we use dataviz and storytelling responsibly?

Follow Tufte's Design Principles

1. Enforce Visual Comparisons (ie. put stuff next to each other)
2. Show Causality ("correlation does not equal causation, but it helps" - Tufte)
3. Show Multivariate Data (ie. show the same data many different ways)
4. Integrate all visual elements (words, numbers, images)
5. Content-Driven Design (ie. high data density)

THIS IS A LOT OF WORK

Follow Cairo's Graphicacy Principles

1. Is the designer using appropriate data and disclosing its origin?
2. Are you reading too much into the graphic?
   "A chart shows as much as it hides —so think about what might be missing"
3. Is the data represented accurately? (axes, scales)
4. Is the graphic showing an appropriate amount of data? (distrust summaries)
5. Is uncertainty relevant? If so, is it revealed?

THIS IS ALSO A LOT OF WORK

Follow Ben Garvey's Principles

1. Be intensely curious
2. Don't let the cult-of-actionable discourage you
3. Ship early and get feedback

THIS IS EASY

Some of my
favorite examples

ELEIÇÕES 2010

Os sinais da bússola eleitoral

A disputa de 2010 foi parecida com a de 2006

Alberto Cairo, Alexandre Mansur, Carlos Eduardo Cruz Garcia, Elseu Barreira Junior, Marco Vergotti e Ricardo Mendonça

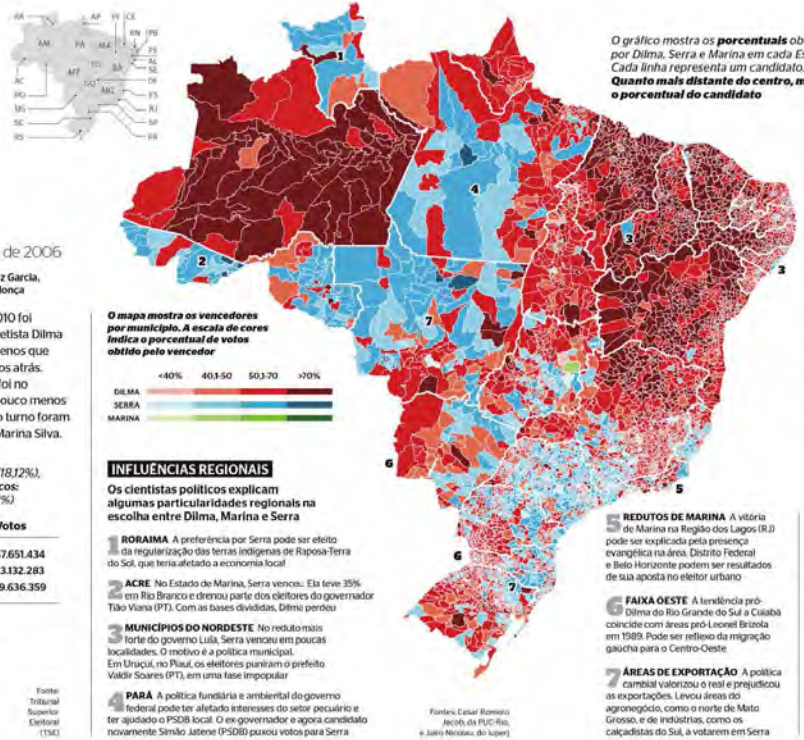
O PRIMEIRO turno da eleição presidencial de 2010 foi muito parecido com o da disputa de 2006. A petista Dilma Rousseff teve apenas 1,7 ponto porcentual a menos que o índice obtido pelo presidente Lula quatro anos atrás. A concentração maior de seus votos também foi no Nordeste. Desta vez, porém, a disputa foi um pouco menos polarizada. Os votos que provocaram segundo turno foram divididos entre o tucano José Serra e a verde Marina Silva.

***Eleitores:** 135.804.433, **abstenção:** 24.610.296 (18,12%), **votos válidos:** 101.590.153 (91,36%), **votos brancos:** 3.479.340 (3,13%) e **votos nulos:** 6.124.254 (5,51%)*

Candidatos	50%	Votos
Dilma Rousseff (PT)	46,9%	47.651.434
José Serra (PSDB)	32,6%	33.132.283
Marina Silva (PV)	19,3%	19.636.359

Outros candidatos	%	Votos
Plínio (PSOL)	0,87%	886.816
José Maria Eymael (PSDC)	0,09%	89.350
Zé Maria (PSTU)	0,08%	84.609
Levy Fidelix (PRTB)	0,06%	57.960
Ivan Pinheiro (PCB)	0,04%	39.136
Rui Costa Pimenta (PCO)	0,01%	12.206

Fonte: Tribunal Superior Eleitoral (TSE)



O mapa mostra os vencedores por município. A escala de cores indica o porcentual de votos obtido pelo vencedor

Fontes: Cesar Romero Jacob, da PUC-Rio, e Joiro Nicolau, do Iuperj

INFLUÊNCIAS REGIONAIS

Os cientistas políticos explicam algumas particularidades regionais na escolha entre Dilma, Marina e Serra

1 RORAIMA A preferência por Serra pode ser efeito da regularização das terras indígenas de Raposa-Terra do Sol, que teria afetado a economia local

2 ACRE No Estado de Marina, Serra vence.: Ela teve 35% em Rio Branco e drenou parte dos eleitores do governador Tião Viana (PT). Com as bases divididas, Dilma perdeu

3 MUNICÍPIOS DO NORDESTE No reduto mais forte do governo Lula, Serra venceu em poucas localidades. O motivo é a política municipal. Em Uruçuí, no Piauí, os eleitores puniram o prefeito Valdir Soares (PT), em uma fase impopular

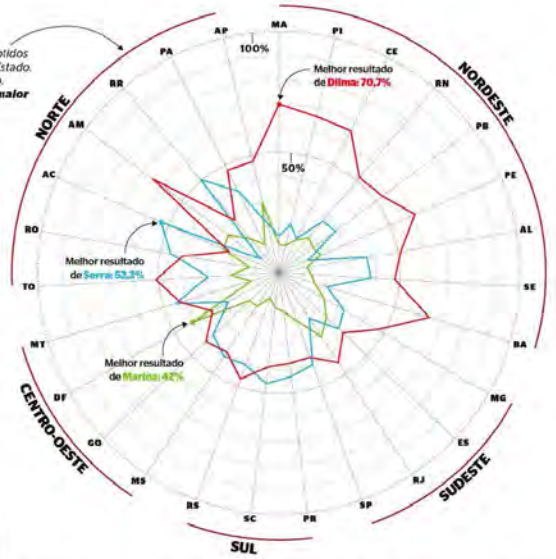
4 PARÁ A política fundiária e ambiental do governo federal pode ter afetado interesses do setor pecuário e ter ajudado o PSDB local. O ex-governador e agora candidato novamente Simão Jatene (PSDB) puxou votos para Serra

5 REDUTOS DE MARINA A vitória de Marina na Região dos Lagos (RJ) pode ser explicada pela presença evangélica na área. Distrito Federal e Belo Horizonte podem ser resultados de sua aposta no eleitor urbano

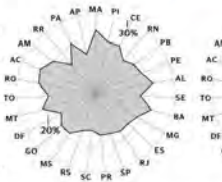
6 FAIXA OESTE A tendência pró-Dilma do Rio Grande do Sul a Cuiabá coincide com áreas pró-Leonel Brizola em 1989. Pode ser reflexo da migração gaúcha para o Centro-Oeste

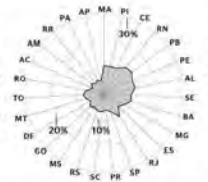
7 ÁREAS DE EXPORTAÇÃO A política cambial valorizou o real e prejudicou as exportações. Levou áreas do agronegócio, como o norte de Mato Grosso, e de indústrias, como os calçadistas do Sul, a votarem em Serra

*O gráfico mostra os **porcentuais** obtidos por Dilma, Serra e Marina em cada Estado. Cada linha representa um candidato.* ***Quanto mais distante do centro, maior o porcentual do candidato***



ABSTENÇÃO A taxa nacional foi de 18%, o mesmo padrão dos anos anteriores. Nos Estados, a abstenção variou de 14%, em Santa Catarina e Roraima, a 24%, no Maranhão





BRANCOS E NULOS O gráfico mostra que os índices de voto branco e nulo são maiores no Nordeste. O Estado com o maior índice é a Paraíba, com 13,2%. Com o menor índice é Roraima, 4,7%

DANÇA ESTADUAL

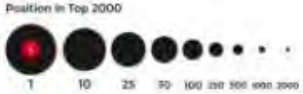
Na comparação com a eleição presidencial de 2006, PT e PSDB tiveram votação menor na maioria dos Estados. O PT enc olheu em 17 Estados. O PSDB, em 25. O motivo é a boa votação de Marina em várias regiões

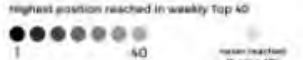
TOP 2000 ♥ 70's & 80's

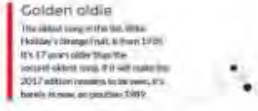
Since 1999 the 2000 most popular songs of all time, as voted by the show's audience, are played on Dutch national Radio 2 in a yearly marathon. The 2000 songs are on the air between noon on December 25th until New Year's Eve and over half of the Dutch population listens to the Top 2000 each year.

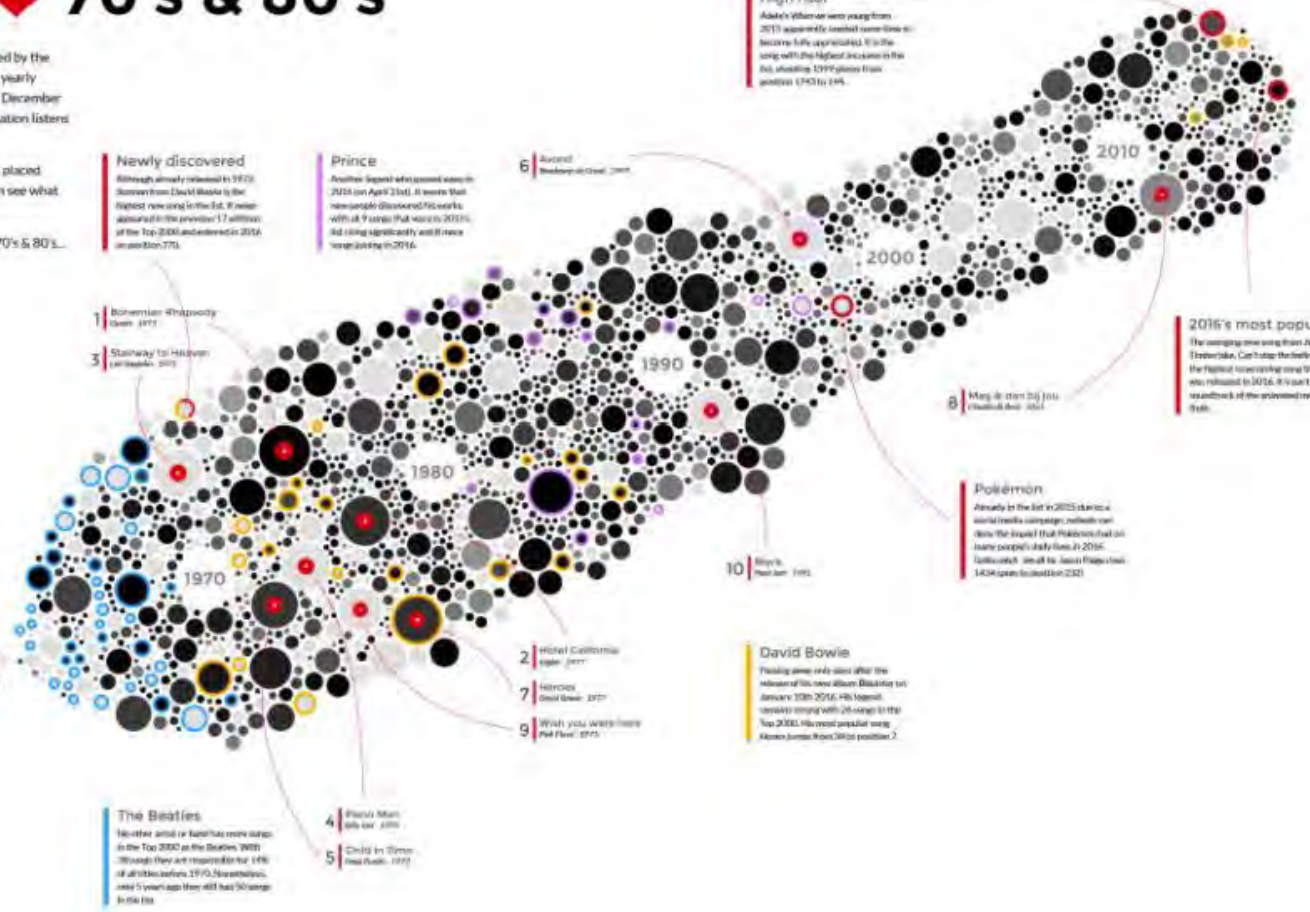
Each ● to the right represents a song in the Top 2000. It is placed according to its year of release. In the legend below you can see what the size and color of a song means.

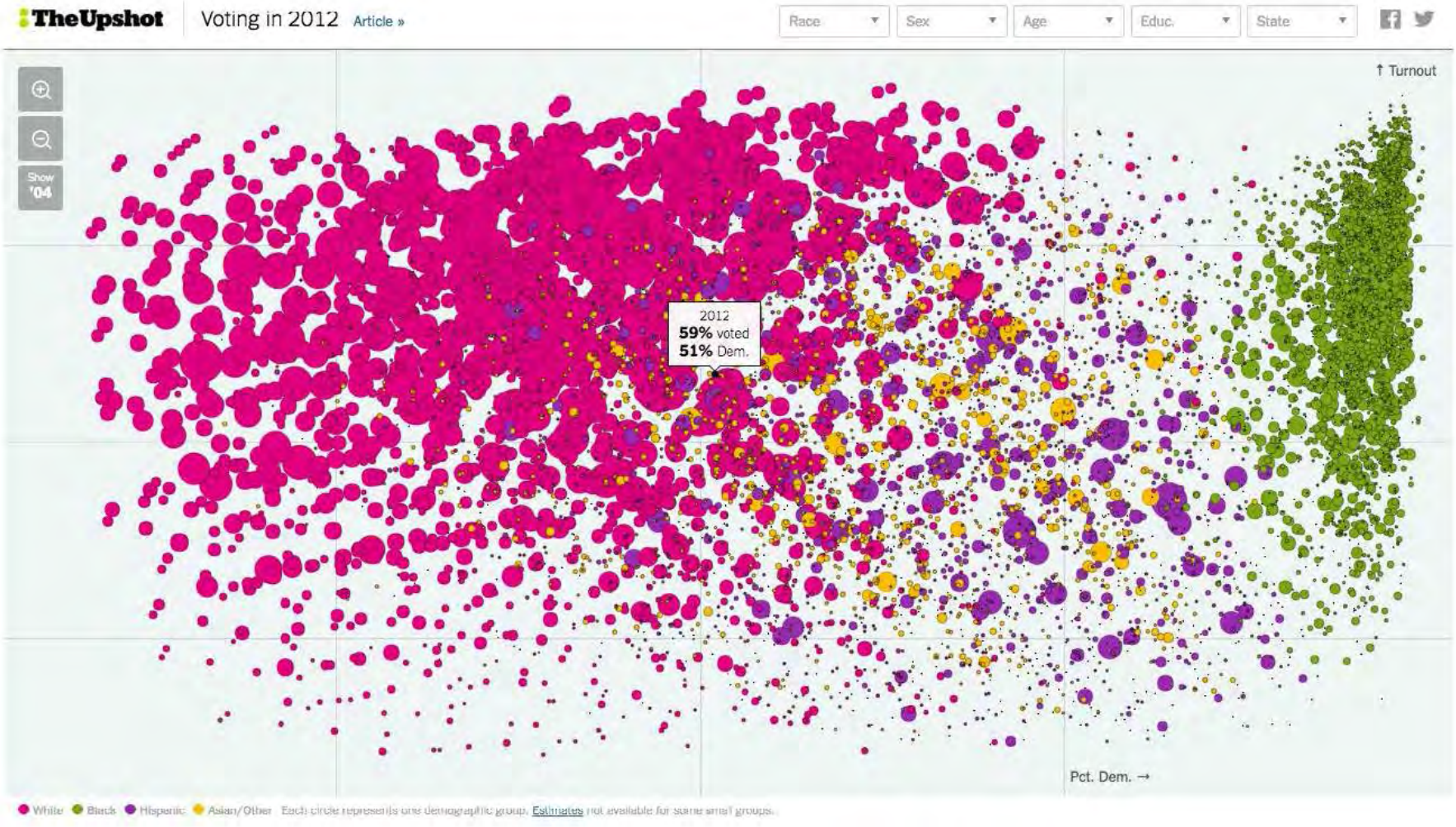
The bulk of the songs and most of the top 10 are from the 70's & 80's...










TheUpshot
Voting in 2012
Article »
Race
Sex
Age
Educ.
State
↑ Turnout
Show '04
2012
59% voted
51% Dem.
Pct. Dem. →
White
Black
Hispanic
Asian/Other
Each circle represents one demographic group. Estimates not available for some small groups.

Now, let's look at some mediocre projects that I've worked on

Last Updated on 2012-05-24 12:28:59

description	WL	id	30 day trend	hours	queue	HTW
Design Engineering	1	09125		1,640.00		139
Project Engineering	0	09110		1,541.00		78
Detail	0	09135		465.00		34
Electrical Engineering	0	09170		136.00		0
80 Ton X 8ft Version Brake (4-7)	1	00040.02		39.73		46
Tig Machine	2	00120		38.33		48
Laser	1	00295		35.20		37
Deburrit Deburring Dept. (6-10)	3	00060		17.50		55
Tape Prep/computer Bom	0	09180		12.78		18
Feacute #pg-6 Power Press 2-5	0	00020.01		7.60		2
Sheet Metal Assembly	1	00090		6.60		11
Jarvis Tappind Head Craftsman Drill Pre	0	00050.02		5.39		8
150 Ton X 10ft Cinn Brake (4-2)	1	00040		5.25		22
10 Ft Shear	0	00010		4.34		2
V & 0 (.o.b.d) Power Press 2-2	0	00020		3.40		6
Grind And Polish (6-5)	1	00065		3.27		21
Wiedematic	1	00290		2.71		0
Power Roller	1	00045		2.26		5
Drill Press Dept.	0	00050		2.13		1
Clausing Drill Press	0	00050.01		2.12		6
15 Ton X 4ft. Chicago Brake (4-1)	0	00040.01		1.88		1
Tumble	0	00055		1.02		0
Stamp Part Numbers	1	00035		0.90		19
Niagra #a 1 3/4 Power Press 2-19	0	00020.10		0.62		0
Timesave	0	00060.06		0.20		22
Thermwood	20	01186		91.36		57
Horizontal Milling	0	01141		35.26		0
Hydraulic Arbor Press	0	01125		24.54		5
Cnc Milling Center	0	01145		21.58		0
Hand Work	0	01175		15.97		20
Plastic Deburr	0	01176		14.50		14
Thermwood Programming	1	09186		9.74		87
Turret Lathe	0	01160		5.40		23
Table Saw Work	0	01190		5.35		1
Vertical Milling	1	01140		4.34		10
Sharp Lathe	2	01130		3.25		11
Bench Drill Press (tool Room)	0	01120		0.95		0
Spotweld	0	00125		0.61		0
Assembly	7	ZZZZZ		1,838.14		219

Resources	Description	30 day trend	Days	Queue
09180, 00295, 00060, 00040, 00040.02	Lasered + Bent		13.8	
0295, 00060, 00040, 00040.02, 00120, 00065	Lasered + Bent + Welded		19.0	
09186, 01186, 01175, 01176	Thermwood		16.4	
01150, 01140, 00060	Machine Shop		2.7	

BRETT MANDEL

DEMOCRAT FOR CONTROLLER

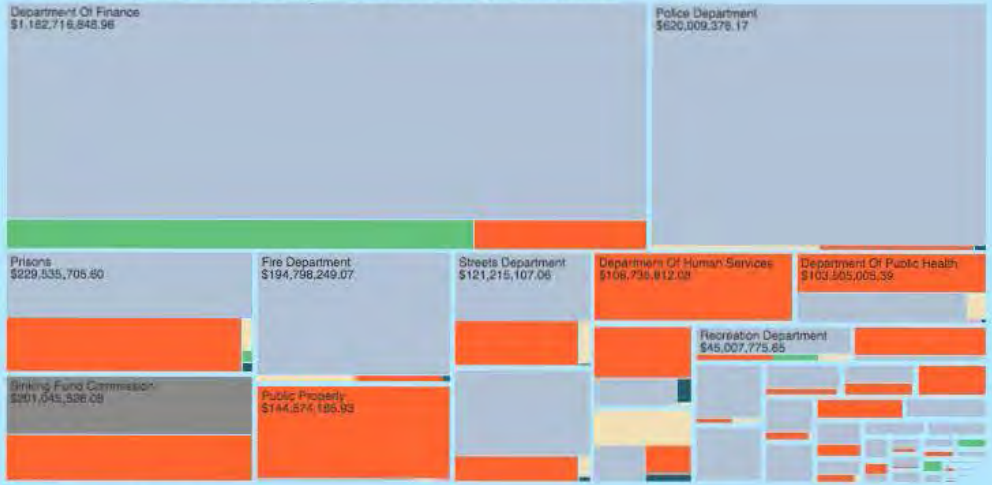
Fork me on GitHub

Mandel for Controller Bulldog Budget

About Help Make a Contribution Tell a Friend Contact

By Department By Category (New)

Home : Philadelphia General Fund Budget Fiscal Year 2012 : $3,493,484,651.90



Search for individual budget items

Search text Submit


2014

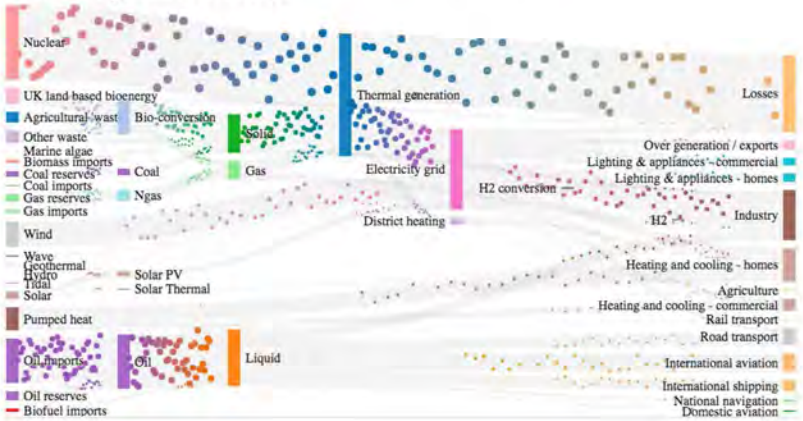
Semiotic is a javascript chart building library that combines React and D3

Semiotic Team @ NETFLIX

Elijah Meeks
@Elijah_Meeks,
author of
D3.js in Action
And many D3
blocks

Susie Lu
@DataToViz, author of
d3-annotations

Sankey Particles



d3.annotation()

AAPL stock example

Why I Like Semiotic

- Charts are minimalistic by default
- Makes some of the crazier D3 charts (networks, chords, sankey, etc.) easier
- Awesome team and company (Netflix) behind it
- Philosophically, it aspires to a sweet spot I agree with
- Data accessors make it easy to make changes quickly
- Annotations are a primary feature

Frustrations

- It's brand new and it shows (bugs, weird docs, new versions every day)
- Literally nothing on Stack Overflow about it.
- I suck at React (my fault)
- Like many javascript libraries/frameworks, it's very hard to tell what is wrong

The world of commodity data visualization seems convinced that it can enable data visualization by releasing ever more widgets.

This approach addresses a problem that no longer exists. Data visualization is not the question: *"How do I deploy as many charts as possible as quickly as possible."*
Rather, it's: *"How do I, in collaboration with fellow developers and stakeholders, create an analytical view into a dataset that best enables everyone to understand and navigate the domain area."* That's not done by enabling more and more charts, it's done by **facilitating information design**.

- Elijah Meeks

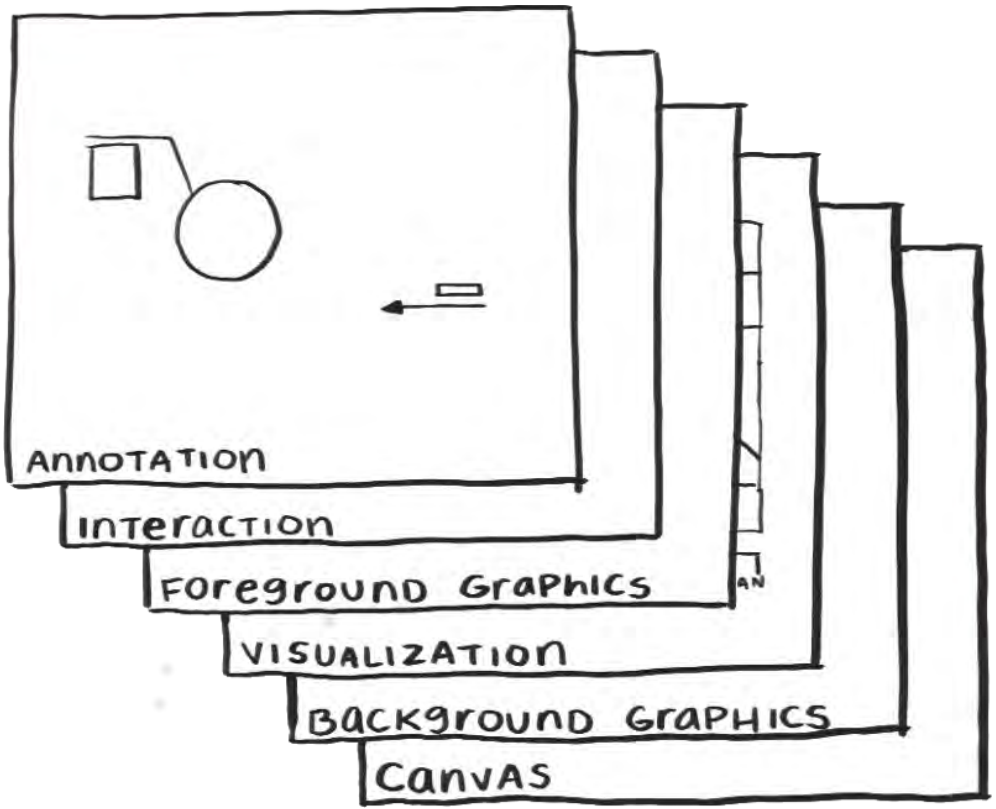
https://medium.com/@Elijah_Meeks/introducing-semiotic-for-data-visualization-88dc3c6b6926

Charting Technologies

Tradeoff between speed and flexibility

https://github.com/bengarvey/dataviz

Frames


ANNOTATION
INTERACTION
FOREGROUND GRAPHICS
VISUALIZATION
BACKGROUND GRAPHICS
CANVAS

SO
LET'S
MAKE
SOME
CHARTS

Simple XYFrame

```
<XYFrame
  size={[1000,400]}
  xAccessor="x"
  yAccessor="y"
  lineDataAccessor="data"
  hoverAnnotation={true}
  lines={deathDisplay}
  defined={d => d.y !== null}
  lineDataAccessor="data"
  lineType={{type:"line", interpolator: curveBasis}}
  lineRenderMode={d => d.renderMode}
  lineStyle={d => ({stroke: '#393e41', strokeWidth: "2px" })}
  margin={{ left: 60, bottom: 30, right: 100, top: 40 }}
/>
```

Simple XYFrame

Dual Axis XYFrame

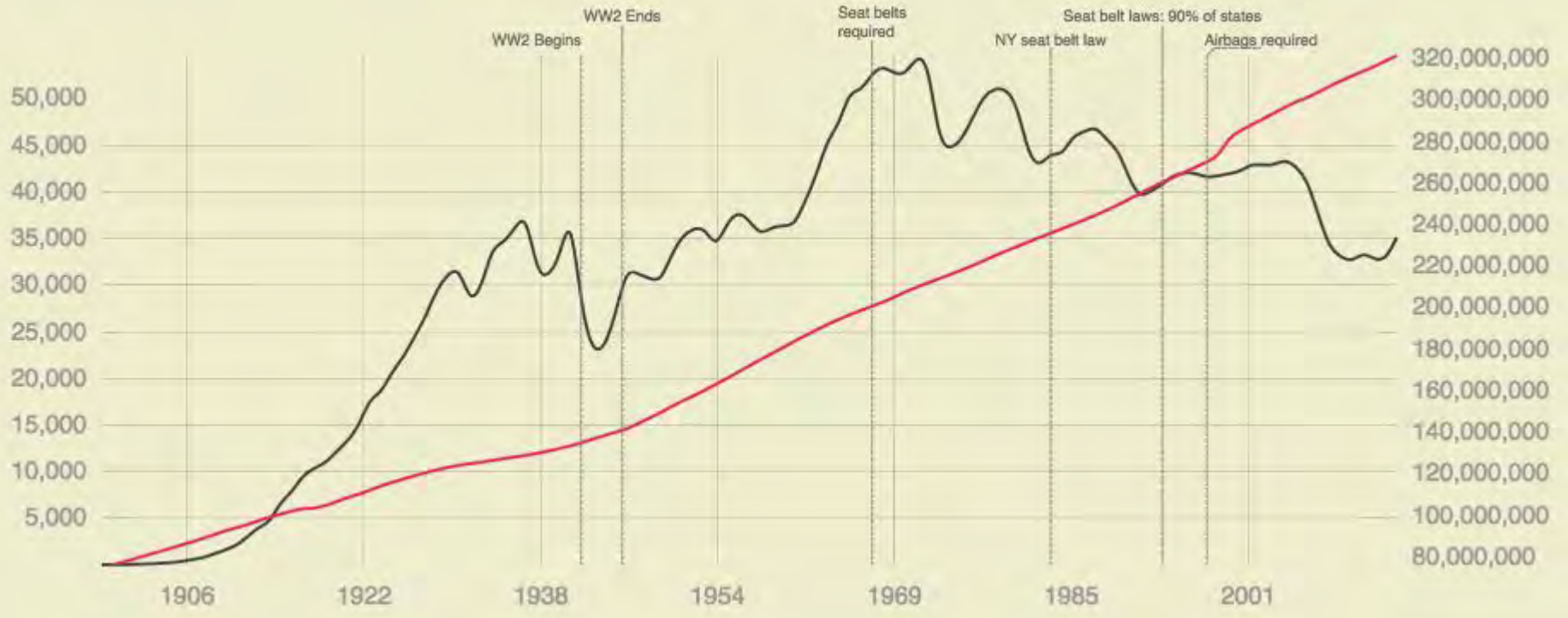
```
var sharedProps = {
  size: [1000,400],
  xAccessor: "x",
  yAccessor: "y",
  lineDataAccessor: "data",
  hoverAnnotation: true,
  lineType: {type:"line", interpolator: curveBasis},
  defined: d => d.y !== null,
  lineStyle: d => ({stroke: d.color, strokeWidth: "2px" }),
  margin: { left: 60, bottom: 30, right: 100, top: 40 }
};
```

Dual Axis XYFrame

```
<XYFrame
  { ...sharedProps }
  lines={deathDisplay}
  lineStyle={d => ({stroke: d.color, strokeWidth: "2px" })}
  axes={[
    { orient: 'bottom', ticks: 8, tickFormat: d => new Date(d).getFullYear() },
    { orient: 'left', ticks: 10, tickFormat: d => d.toLocaleString()}
  ]}
/>

<XYFrame
  { ...sharedProps }
  lines={popDisplay}
  axes={[
    { orient: 'bottom', ticks: 8, tickFormat: d => '' },
    { orient: 'right', ticks: 10, tickFormat: d => d.toLocaleString()}
  ]}
  annotations={popAnnotations}
/>
```

Dual Axis XYFrame

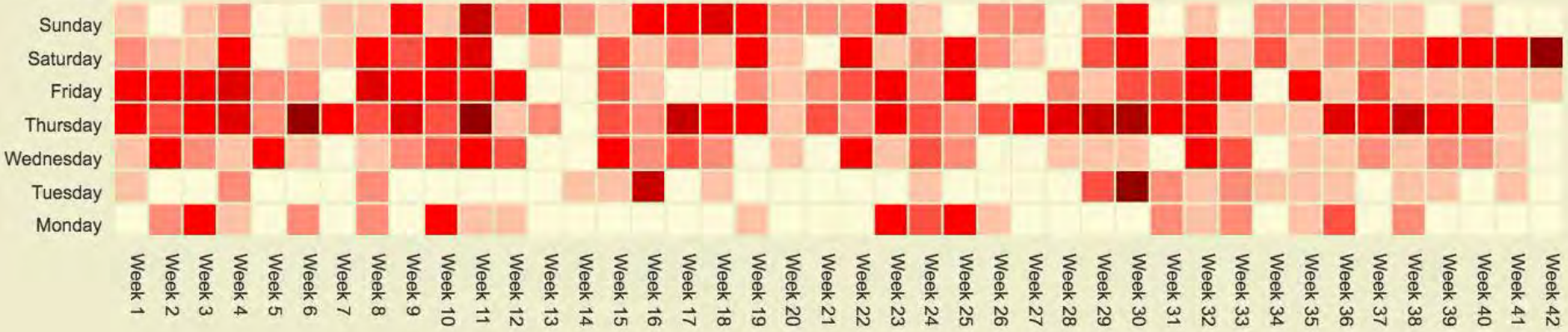


ORFrame

```
<ORFrame
  size={[ 1400, 350 ]}
  data={modified}
  rAccessor={() => 1}
  oAccessor={d => d.step}
  style={d => ({ fill: heatScale(d.value), stroke: "darkgray", strokeWidth: 0 })}
  type={"bar"}
  axis={daysAxis}
  hoverAnnotation={true}
  pieceHoverAnnotation={true}
  oLabel={d => <text transform="rotate(90)">Week {d+1}</text>}
  margin={{ left: 100, top: 10, bottom: 80, right: 50 }}
  oPadding={0}
/>
```

Pager Duty Incidents

Ahhhhhh!



NetworkFrame

```
  <NetworkFrame
      size={[ 1300, 500 ]}
      edges={network.links}
      nodes={network.nodes}
      margin={60}
      edgeStyle={(d) => ({ stroke: colors[d.relation], fill: '#a91a1a', fillOpacity: 0.25,
strokeWidth: '1px' })}
      nodeStyle={d => ({ fill: colors[d.side], r:"15px" })}
      networkType={{ type: 'force', iterations: 500, edgeStrength: 0.1 }}
      edgeType={'none'}
      nodeSizeAccessor={d => 7}
      zoomToFit={true}
      nodeLabels={d => d.name}
      nodeIDAccessor={"id"}
      margin={{left: 20, top: 20, bottom: 20, right: 50}}
   />
```

Star Wars Geneaology

How the force got with us

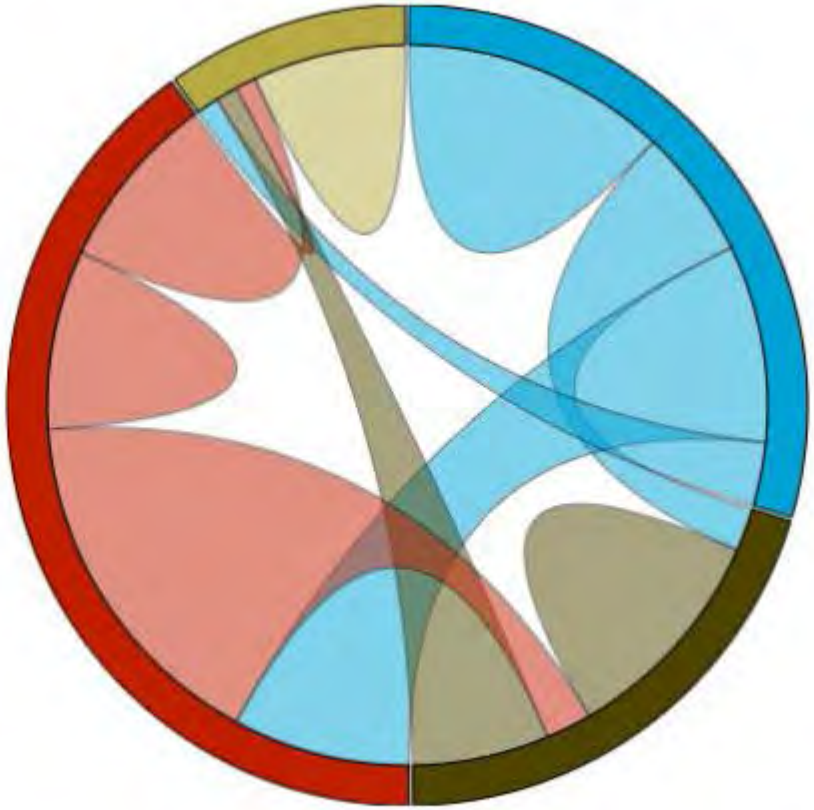
ResponsiveNetworkFrame

```
<ResponsiveNetworkFrame
	size={[360, 700 ]}
	responsiveWidth={true}
	edges={this.network.links}
	nodes={this.network.nodes}
	nodeStyle={d => ({
		fill: d.fill,
		stroke: d.stroke
	})}
	edgeStyle={(d) => ({ stroke: d.stroke, fill: d.fill, opacity: 0.5, strokeWidth: '1px' })}
	networkType={{ type: 'sankey', orient: 'justify', iterations: 500, nodeWidth: 100, nodePadding: 22}}
	nodeIDAccessor={"id"}
	zoomToFit={true}
	nodeLabels={d => d.name}
	sourceAccessor={"source"}
	targetAccessor={"target"}
	nodeSizeAccessor={d => d.size}
	margin={{left: 25, top: 20, bottom: 20, right: 25}}
/>
```

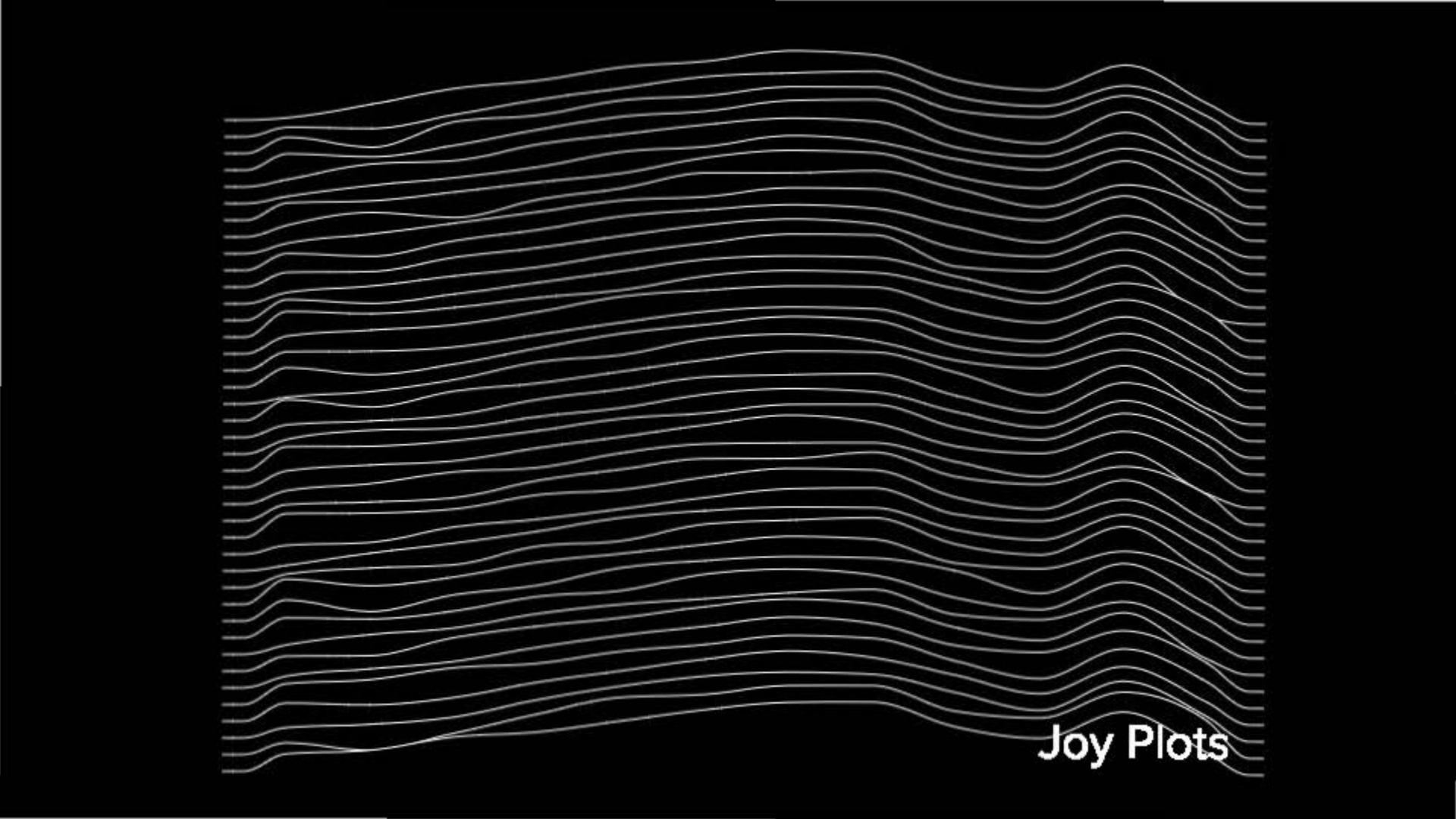
RJMetrics

Where are they now?

Many more chart types!



Chord Diagrams


Joy Plots

Using dataviz to understand problems

(and tell their stories)

We're going to talk about tornados and car accidents

(sorry if this gets morbid)

The Story:

Are tornadoes getting more deadly?

2011 Joplin tornado

From Wikipedia, the free encyclopedia

Coordinates: 37.060554°N 94.530938°W

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikipedia store

Interaction

Help
About Wikipedia
Community portal
Recent changes
Contact page

Tools

What links here
Related changes
Upload file
Special pages
Permanent link
Page information
Wikidata item

The **2011 Joplin tornado** was a catastrophic EF5-rated multiple-vortex tornado that struck Joplin, Missouri, late in the afternoon of Sunday, May 22, 2011. It was part of a larger late-May tornado outbreak and reached a maximum width of nearly 1 mile (1.6 km) during its path through the southern part of the city.[2] It rapidly intensified and tracked eastward across the city, and then continued eastward across Interstate 44 into rural portions of Jasper County and Newton County.[3] It was the third tornado to strike Joplin since May 1971.[4]

Overall, the tornado killed 158 people (with an additional three indirect deaths), injured some 1,150 others, and caused damages amounting to a total of $2.8 billion. It was the deadliest tornado to strike the United States since the 1947 Glazier–Higgins–Woodward tornadoes, and the seventh-deadliest overall. It also ranks as the costliest single tornado in U.S. history.

In a preliminary estimate, the insurance payout was expected to be $2.2 billion; the highest insurance payout in Missouri history, higher than the previous record of $2 billion in the April 10, 2001 hail storm, which is considered the costliest hail storm in history as it swept along the I-70 corridor from Kansas to Illinois.[5] Estimates earlier stated Joplin damage could be $3 billion. By July 15, 2011, there had been 16,656 insurance claims.[6]

Contents [hide]

2011 Joplin tornado

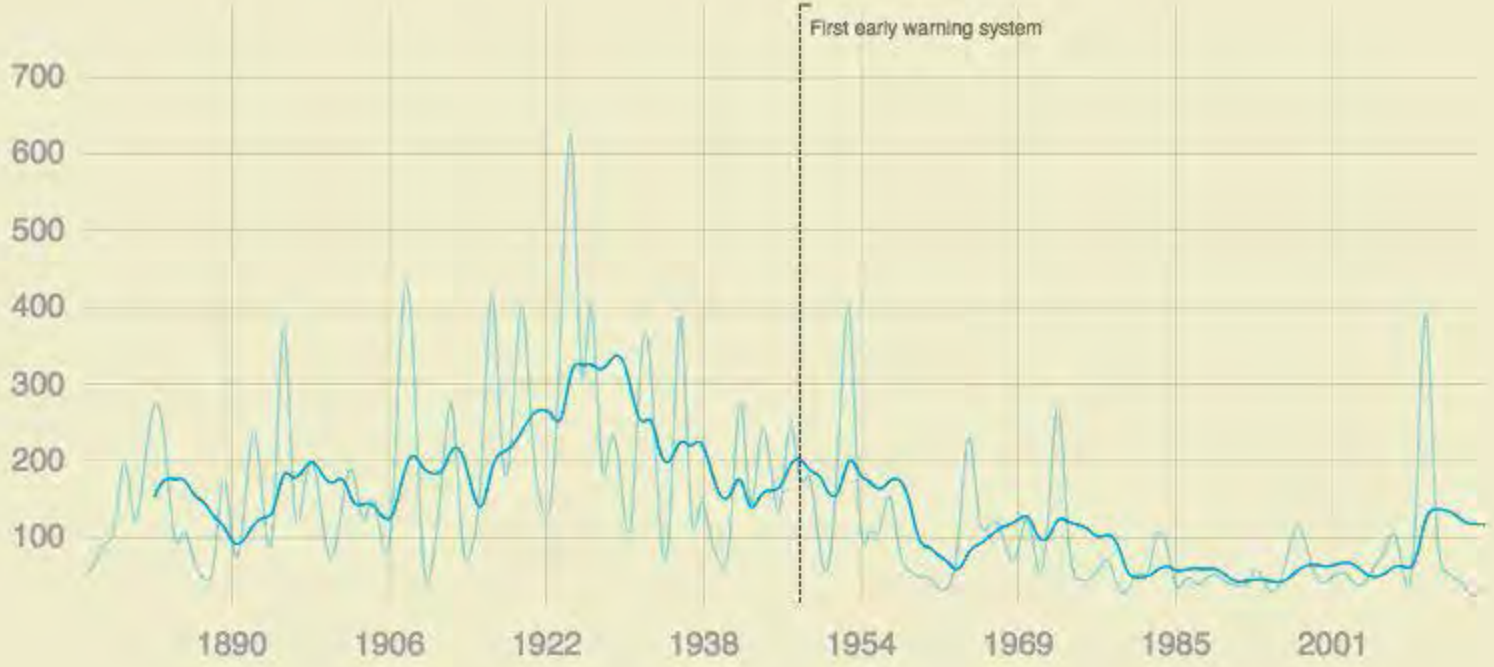
EF5 tornado

Devastation in Joplin shortly after the tornado

Formed	May 22, 2011, 5:34 p.m. CDT (UTC−05:00)
Duration	38 minutes
Dissipated	May 22, 2011, 6:12 p.m. CDT (UTC−05:00)
Max rating[1]	EF5 tornado

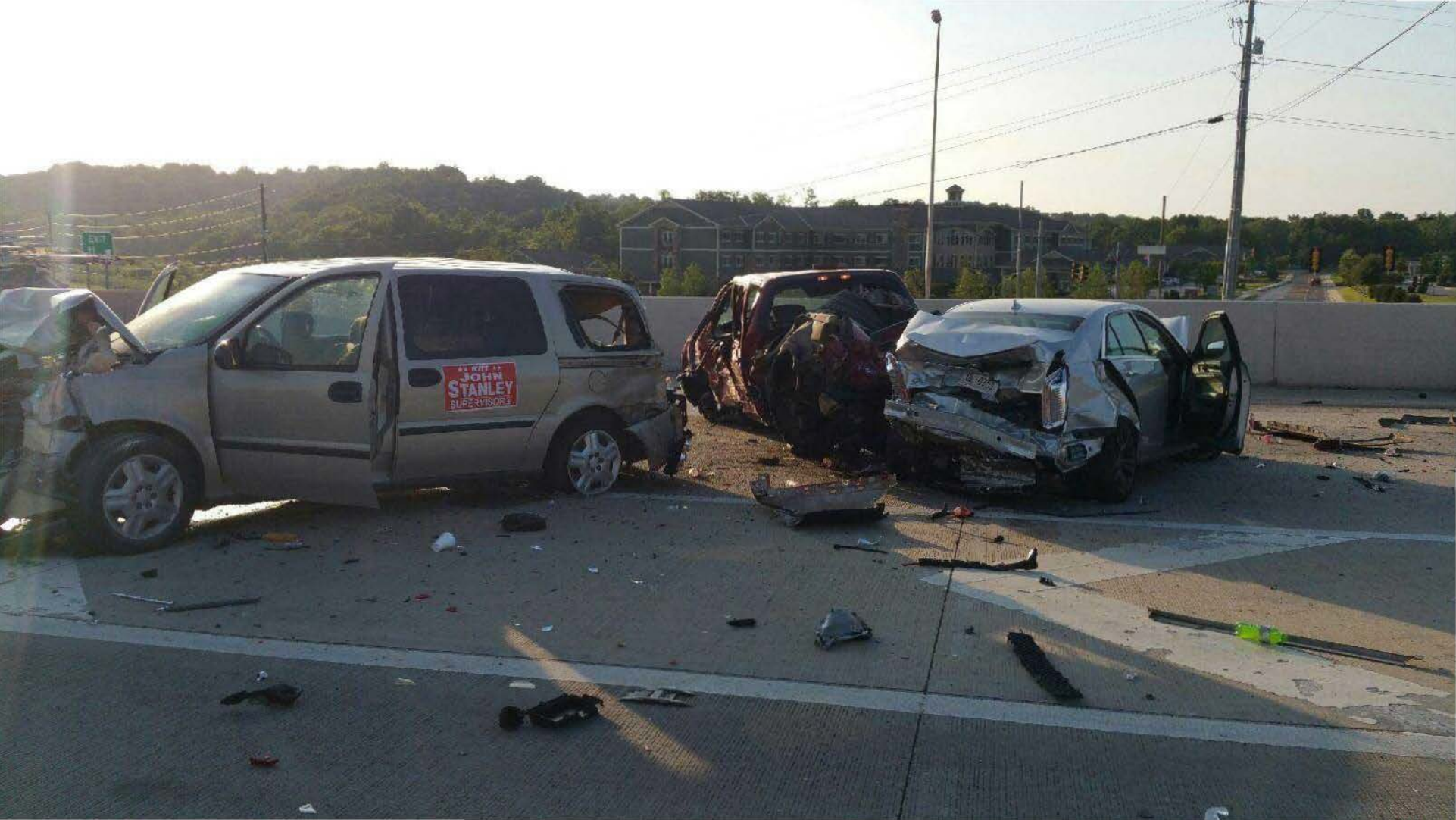
US Tornado Deaths 1876 - 2017

What happened in 2011?



The Story:

Why are automobile deaths decreasing?


JOHN
STANLEY

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikipedia store

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact page

Tools
What links here
Related changes
Upload file
Special pages
Permanent link
Page information
Wikidata item
Cite this page

Print/export
Create a book
Download as PDF
Printable version

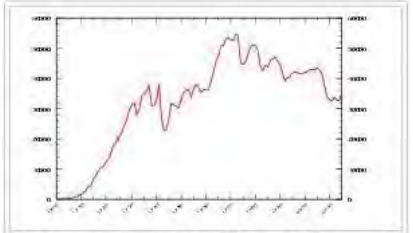
List of motor vehicle deaths in U.S. by year

From Wikipedia, the free encyclopedia

The table below is a **list of motor vehicle deaths in the United States by year**. According to data compiled by the National Highway Traffic Safety Administration (NHTSA), in 2016, 37,461 people were killed in 34,436 crashes, an average of 102[1] per day.

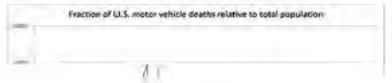
In 2010, there were an estimated 5,419,000 crashes (30,296 fatal crashes), killing 32,999 and injuring 2,239,000,[2] and around 2,000 children under 16 years old die every year due to traffic collisions.[3] Records indicate that there has been a total of 3,613,732 motor vehicle fatalities in the United States from 1899 to 2013.

Although the number of deaths, and deaths relative to the total US population, declined over most of the previous two decades, this trend reversed in 2015 and continued to move upward in 2016. From 1979 to 2005, the number of deaths per year decreased 14.97% while the number of deaths *per capita* decreased by 35.46%. The 32,479 traffic fatalities in 2011 were the lowest in 62 years (1949). Note: US motor death statistics reported by government only include those on public roads, they do not include parking lots, driveways and private roads.[4]



Motor vehicle deaths in the US since 1900

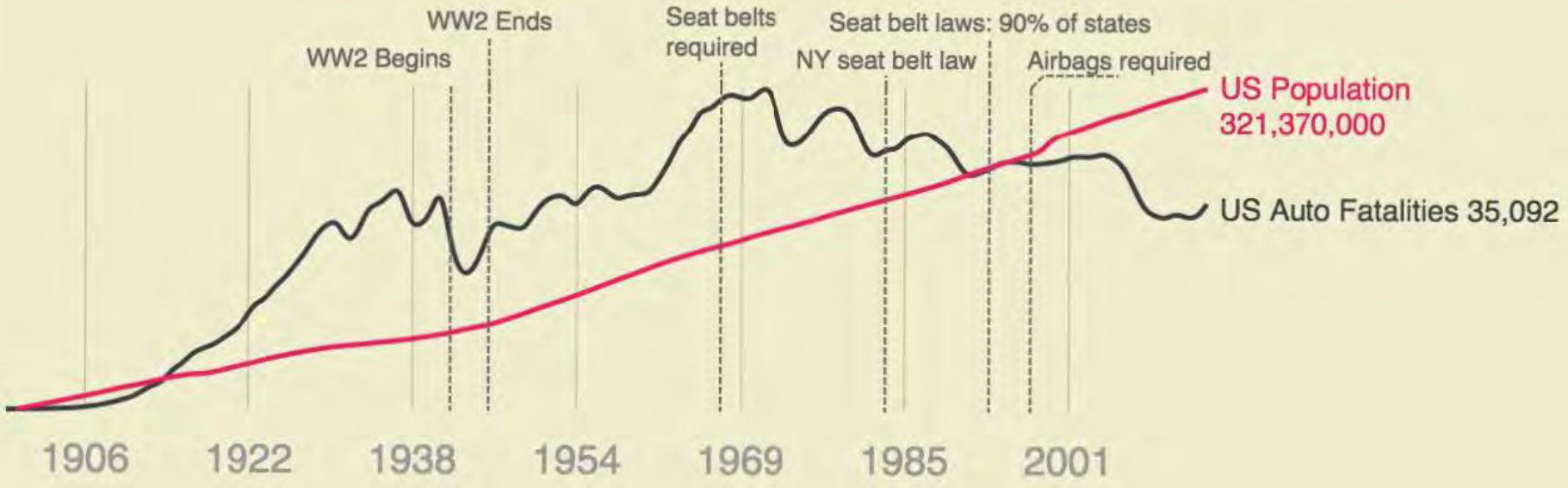
Annual US traffic fatalities per billion vehicle miles traveled (red) and miles traveled (blue) from 1921 to 2015.

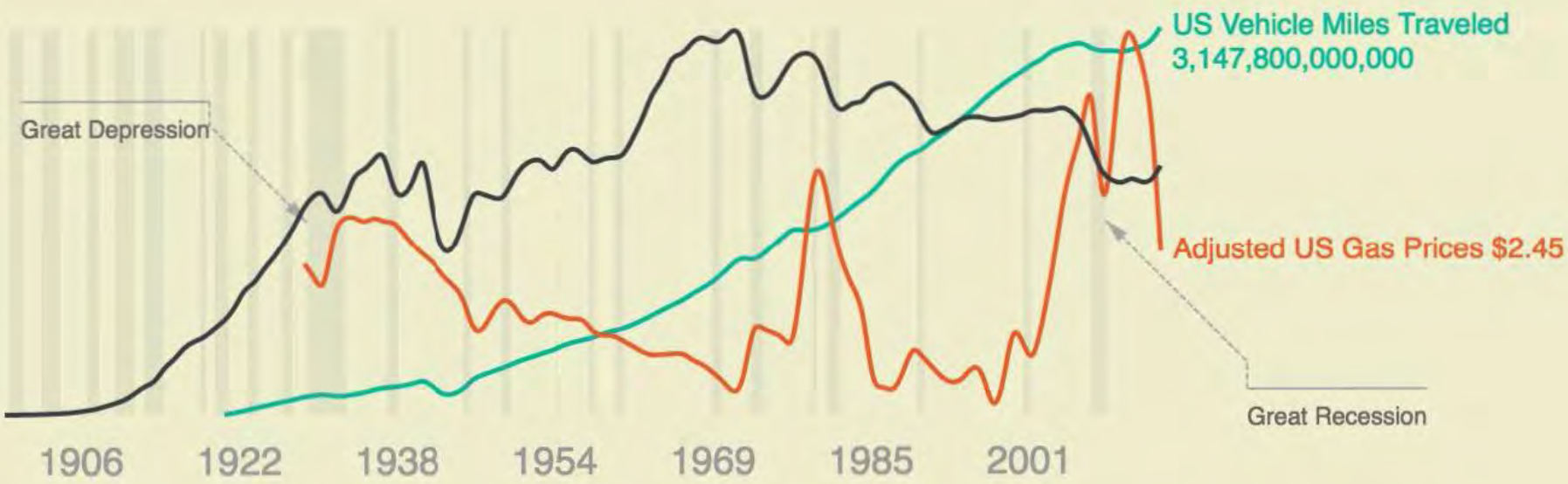


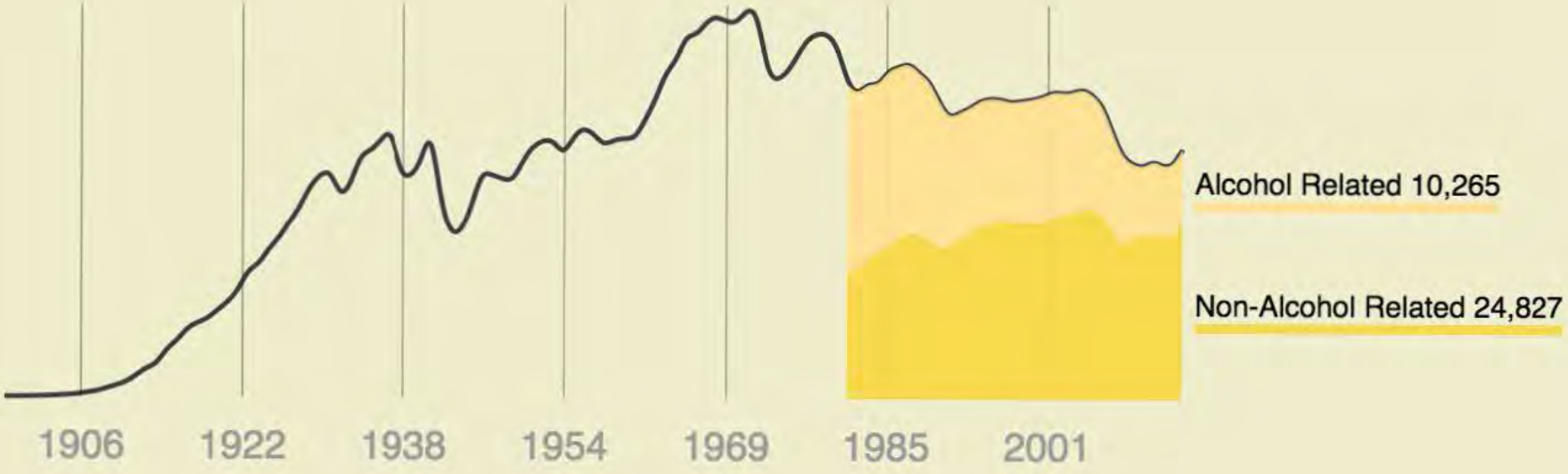
Contents [hide]

Motor vehicle deaths in U.S. by year [edit]

Year	Deaths	Vehicle miles travelled (billions)	Fatalities per 100 million VMT	Population	Fatalities per 100,000 population	Change (in percent)
1899	26[5]					

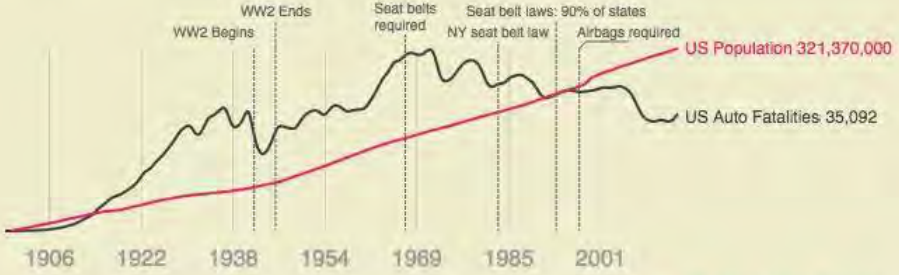

WW2 Begins
WW2 Ends
Seat belts required
Seat belt laws: 90% of states
NY seat belt law
Airbags required
US Population 321,370,000
US Auto Fatalities 35,092
1906
1922
1938
1954
1969
1985
2001


US Vehicle Miles Traveled
3,147,800,000,000
Great Depression
Adjusted US Gas Prices $2.45
Great Recession
1906
1922
1938
1954
1969
1985
2001


1906
1922
1938
1954
1969
1985
2001
Alcohol Related 10,265
Non-Alcohol Related 24,827

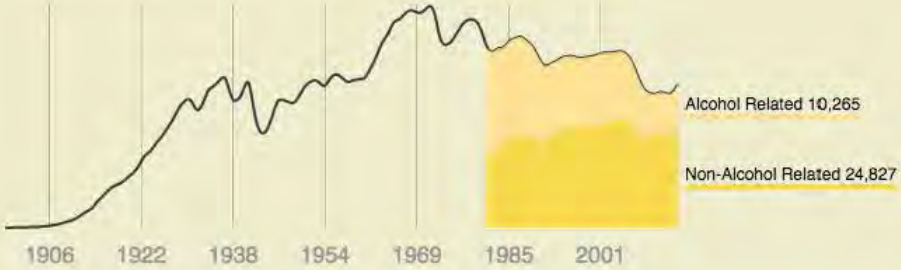
US Auto Fatalities 1899 - 2105

Why have deaths decreased since 1979?

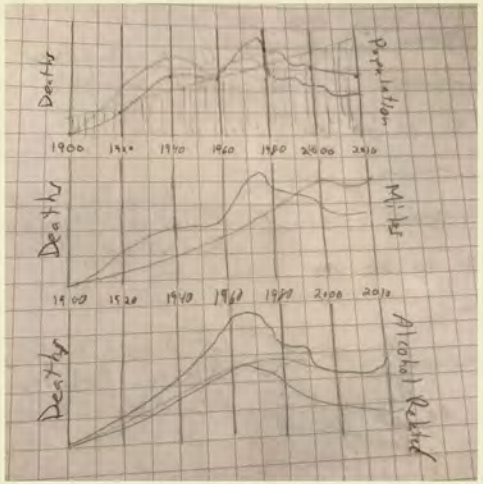


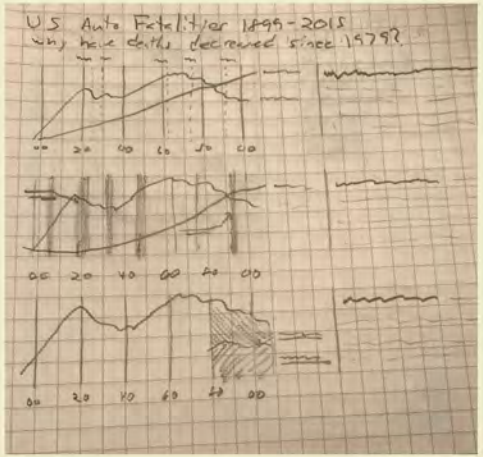
Since 1899 more than 3.6 million people in America have been killed in automobile accidents. After a peak of 51,093 deaths in 1979, the yearly deaths began to decline despite a growing US population. Common explanations are safety innovations such as seat belts and air bags, but they don't tell the entire story.

Spikes in gas prices tend to reduce auto deaths. Americans consistently drove more miles for decades... until 2008's Great Recession. Recessionary periods (shown in gray) related to gas prices seem to reduce auto deaths, even in cases where more and more miles were driven. The Great Recession of 2008 corresponds with an extraordinary drop. It could be due to more people working remotely, using public transit, carpooling, etc. The Obama administration also introduced it's popular Cash-For-Clunkers program, which put 690,114 new and safer cars on the road.



There are far fewer alcohol related deaths. Since 1982 the number of alcohol related deaths has dropped from 43,945 to 10,265 in 2015 (-76.6%). Harsher penalties and changing cultural norms are both possible factors in this reduction. Further gains, however, will have to come from the non-alcohol related deaths.

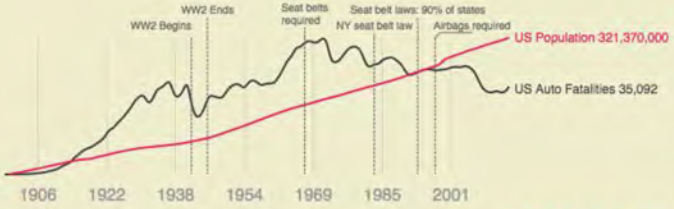




Get a sketchbook

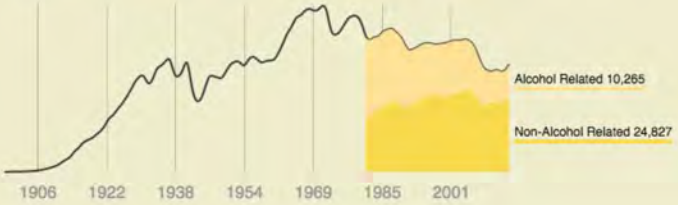
US Auto Fatalities 1899 - 2105

Why have deaths decreased since 1979?



Since 1899 more than 3.6 million people in America have been killed in automobile accidents. After a peak of 51,093 deaths in 1979, the yearly deaths began to decline despite a growing US population. Common explanations are safety innovations such as seat belts and air bags, but they don't tell the entire story.

Spikes in gas prices tend to reduce auto deaths. Americans consistently drove more miles for decades... until 2008's Great Recession. Recessionary periods (shown in gray) related to gas prices seem to reduce auto deaths, even in cases where more and more miles were driven. The Great Recession of 2008 corresponds with an extraordinary drop. It could be due to more people working remotely, using public transit, carpooling, etc. The Obama administration also introduced it's popular Cash-For-Clunkers program, which put 690,114 new and safer cars on the road.



There are far fewer alcohol related deaths. Since 1982 the number of alcohol related deaths has dropped from 43,945 to 10,265 in 2015 (-76.6%). Harsher penalties and changing cultural norms are both possible factors in this reduction. Further gains, however, will have to come from the non-alcohol related deaths.

The Story:

Who are the most popular characters on...


TWIN PEAKS
WELCOME
TWIN PEAKS
Population 51,201

Instagram

Search

kyle_maclachlan • Following

kyle_maclachlan Dress up as Coop (or your favorite #TwinPeaks character) to be entered to win a prize package of official Twin Peaks merch! To be considered, you must:

1. Post a photo of your Coop/Twin Peaks costume
2. Hashtag #CoopLovesCostumes

Contest open to US only. I'll be announcing the winner Nov. 2nd. But you can start posting your costumes now. Have fun!

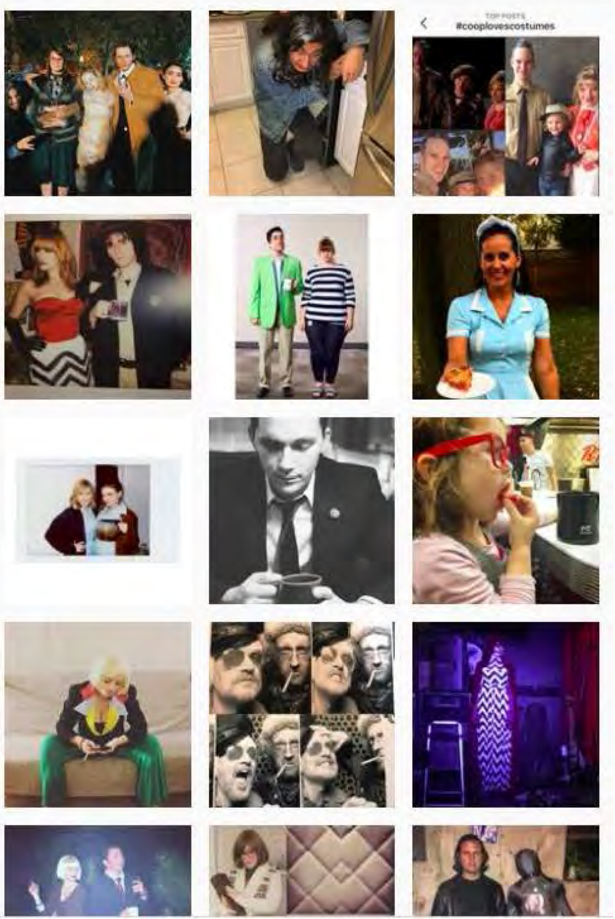
Load more comments

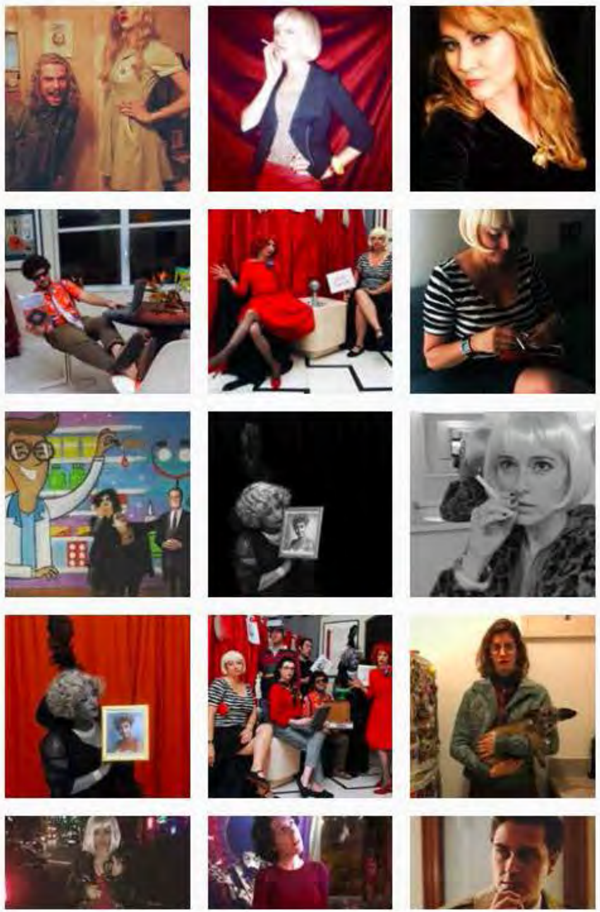
tuzantuzan Can you still enter if your Canadian and have a US address? Very cool contest by the way...#twinpeaksthereturn #wellworththewait

9,782 likes

OCTOBER 23

Add a comment...


TOP POSTS
#cooplovescostumes



Magento
Ben Garvey

More Unreads

musik
no-context
open-table-4
pets
philly
philly-games
phl-analysts
phl-case-swap
phl-cheese-wednesday
phl-clientfacers
phl-internal-analytic
phl-magento-bi-sales
phl-prossentials
pongadelphia
rjql-svc
slack-admins
social
spark
stitchgento
stitchgento-dev
striketeam
striketeam-firehose
t2
twinpeaks
unarmed
ux-analytics

Direct Messages
slackbot
Akash Agrawal
Akash Agrawal, Owen ...
anitaandrews
Ann Hudspeth

#twinpeaks

4 | 0 | Add a topic

Tuesday, October 31st

Ben Garvey 4:34 PM
@hroslin photo? I'm heading out

Harker Roslin 4:34 PM
YES

Harker Roslin 4:40 PM
uploaded this image: Screenshot 2017-10-31 16.40.31.png

2

Thursday, November 2nd

Ben Garvey 1:40 PM
If I had enough hours in the day I'd scrape 1000 photos from #cooplovescostumes and make a chart of the most popular costumes

Harker Roslin 1:44 PM
10/10 would retweet those charts

Ben Garvey 1:54 PM
There are some great longtail Pete Martell and Freddy costumes.

Harker Roslin 2:49 PM
Pete Martell is such a peach

Ben Garvey 2:51 PM
RIP

ugh, I want to do this analysis badly

And I want the Dale and Laura costumes broken down by category

Ben Garvey 1:40 PM ☆

If I had enough hours in the day I'd scrape 1000 photos from #cooplovescostumes and make a chart of the most popular costumes

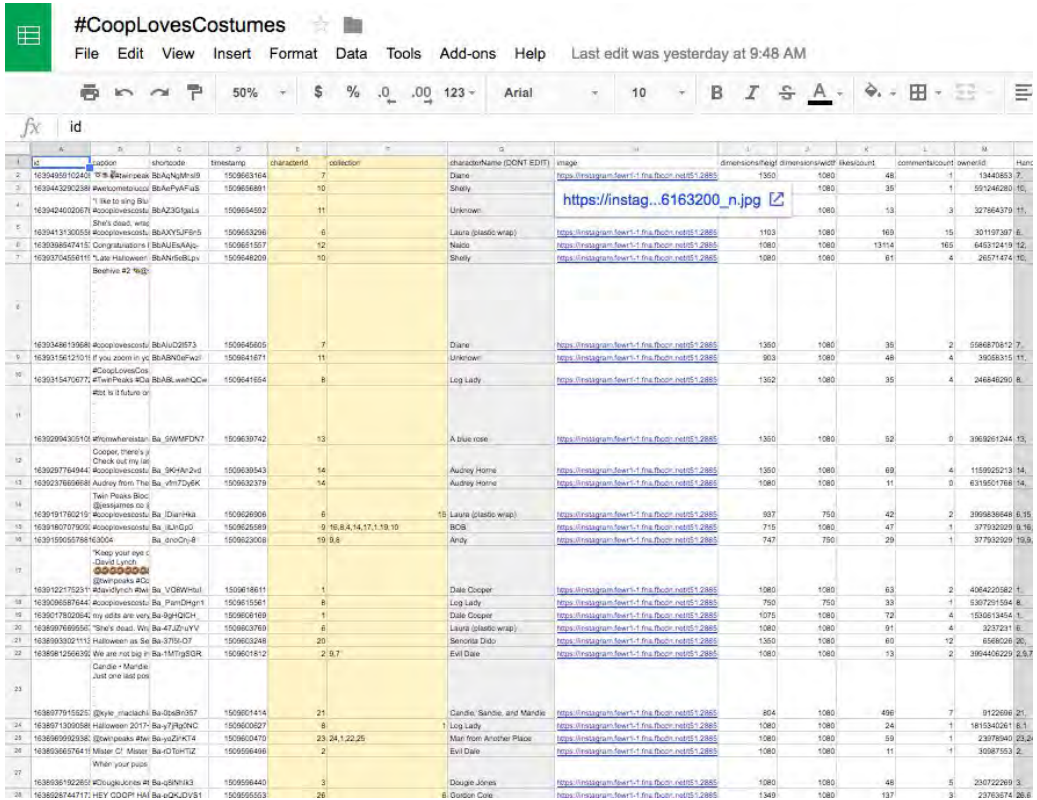
Ben Garvey 1:40 PM

If I had enough hours in the day I'd scrape 1000

B

If

I scraped metadata and urls for 2300 Instagram posts using an [open source tool] and then manually tagged 700+ of them with the costumes that were included in the photo.



Thursday 11:57

You're a lunatic with this spreadsheet.

- Shaun McAvinney, creator of the game, *Moxie: An Aspirational Horse*

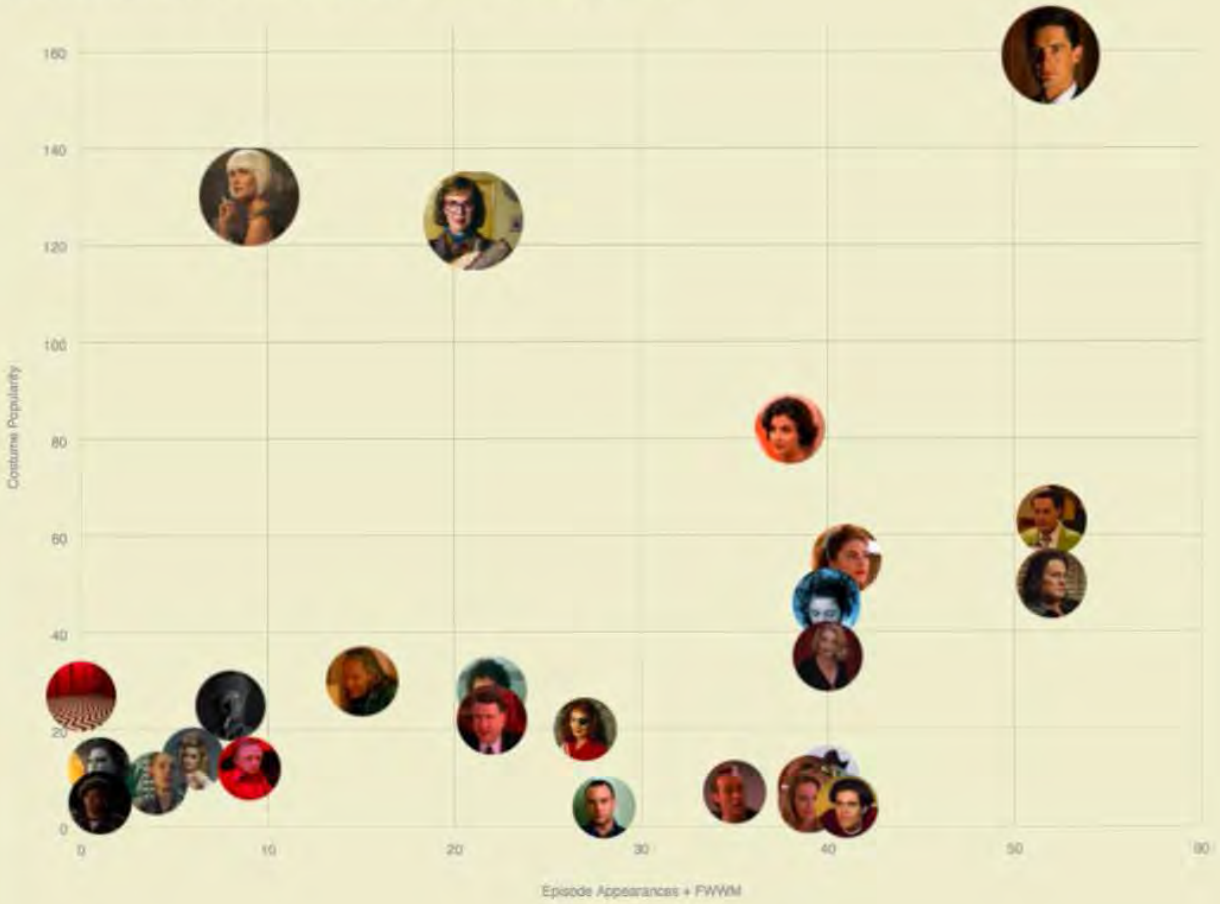
#CoopLovesCostumes

Which Twin Peaks costumes were the most popular?

On October 23rd, 2017 Kyle MacLachlan announced a Twin Peaks Halloween costume contest on Instagram. To be eligible, you had to tag your costume with #CoopLovesCostumes and omg the feed of costumes is incredible. But which characters were the most popular? To answer this question I scraped 2300+ photos and meta data from Instagram and so far have tagged around 200 photos. Here are the results so far

Popularity vs Episode Freq

Does the frequency of a character influence their costume popularity?



Fan Favorites

Few appearances but high popularity

Senorita Dido
Only appeared in one episode, but what an episode!

Diane
Clearly the most popular new character from season 3.

Candie, Sandie, and Mandie
Only appeared in 6 episodes.

No Shows

Frequently on the show, but didn't catch on for Halloween

Bobby Briggs
Appears in 41 episodes and I could only find 1 costume.

Donna Hayward
34 episodes and I only found 1 costume (and even that was debatable)

Sherrif Truman
Harry and Frank were in 40 episodes, but only a handful of costumes

Plot Device

A collection of dataviz projects by Ben Garvey

I. US Automobile Fatalities 1899 - 2015
II. Weight Over Time
III. Chart Creation Technologies
IV. Pager Duty Incidents
V. US Tornado Deaths 1875 - 2017
VI. #CoopLovesCostumes
VII. Star Wars Network (No Last Jedi Spoilers)
VIII. My Top 100 Favorite Movies
IX. My Steps
X. RJMetrics: Where are they now?

plotdevice.bengarvey.com

The Story:

Which Indego bike was ridden the most and why?

The quest for the most frequently used bike

Ben Garvey · Mar 5, 2018
@bengarvey

Fork of Mapping Indego Bike #2679 w/ Mapbox GL

Mapping Indego Bike #2679 w/ Google Maps

We recently created a new front end engineering challenge project at Magento BI and it uses the Philly's Indego bike service as the dataset.

I'm obsessed, so I wrote a script to parse the data and find the most popular bike.

Below is a trip history of the most frequently rode bike (#2679) from 2015-2017. It's been everywhere!

Philadelphia, PA

2679

Bike #2679 Stats

Trips
1975
Longest Trip
4.08 miles
Average Trip Length
1.01 miles
Total Distance Travelled
1985.79 miles
Longest Trip Duration
13.25 hours
Average Trip Duration
23.69 minutes
Total Time in Use
32.49 days
Most Common Stops
15th & Spruce - 147
Rittenhouse Square - 142
23rd & South - 138

Total Bikes
Mural Arts: 10
Normal: 2244

Avg Trips Per Day Per Bike
Mural Arts: 2.49 trips (+9.55%)
Normal: 2.27 trips

Avg Time In Service Per Bike
Mural Arts: 483.4 days (+28.15%)
Normal: 377.2 days

Are Mural Arts Bikes More Frequently Ridden?

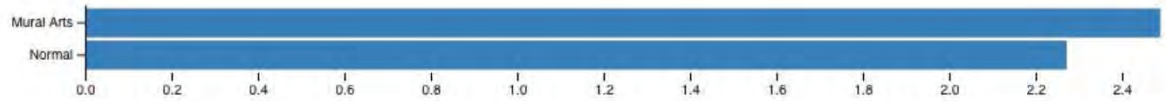
One surprising thing we found was that bike #2679 was painted a with a special design as part of the Mural Arts program. Was this the mostly frequenly used bike because it stood out from the rest? To find out, I calculated statistics on the 10 mural arts bikes and compared them to the normal bikes.

Total Bikes

Normal: 2244
Mural Arts: 10

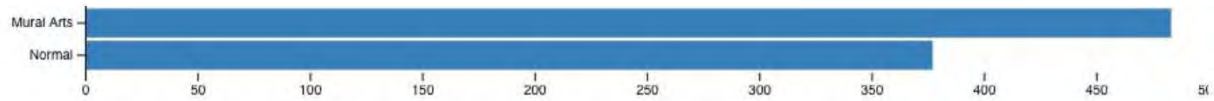
Avg Trips Per Day Per Bike

Mural Arts: 2.49 trips (+9.55%)
Normal: 2.27 trips



Avg Time In Service Per Bike

Mural Arts: 483.40 days (+28.15%)
Normal: 377.21 days



BILLYPENN

On top of Philly
and in your inbox every mo

NEW New nonstop flights to Mexico are a big deal for Philly immigrants

This sexy Mural Arts bike is the most popular ride in all Philly bike share

Indego's most-used bike has been on nearly 2,000 trips.

Bike No. 2679, Indego's most popular ride DANYA HENNINGER / BILLY PENN

https://beta.observablehq.com/@bengarvey/mapping-indego-bike-2679-w-google-maps

The End

Dataviz and
Storytelling

@bengarvey IREG 2018

Gerrymandering in Pennsylvania

Lee Hachadoorian, Temple University

March 14, 2018

Concerned Citizens for Democracy

CCFD is a Pennsylvania Non-profit Association founded by a group of lawyers, mathematicians and other concerned citizens fighting to put an end to partisan gerrymandering in Pennsylvania.

https://concernedcitizensfordemocracy.org/

Nonpartisan Issue

- By some measures Republicans may be “winning” the gerrymandering game, but states like Maryland stand out as Democratic gerrymanders
- Republican voters in Democratic districts
- New York State

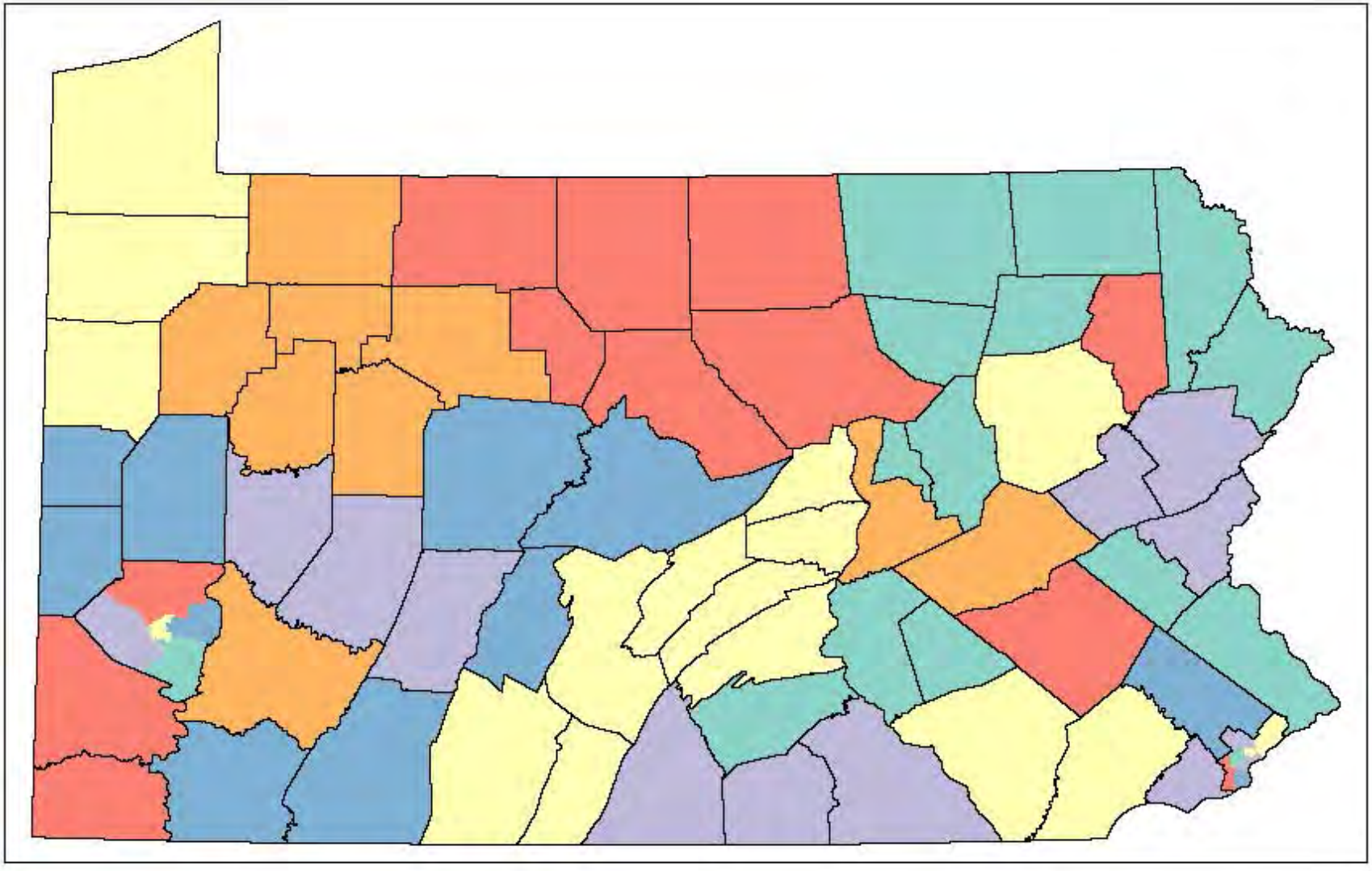
Important Dates

- *Gray v. Sanders* (1963), *Reynolds v. Sims* (1964) - establish "one person, one vote", strike down state unequal representation in state legislatures
- *Wesberry v. Sanders* (1964) - applies "one person, one vote" to Congressional districts

Since then, Supreme Court has moved closer and closer to requiring exact population equality for redistricting:

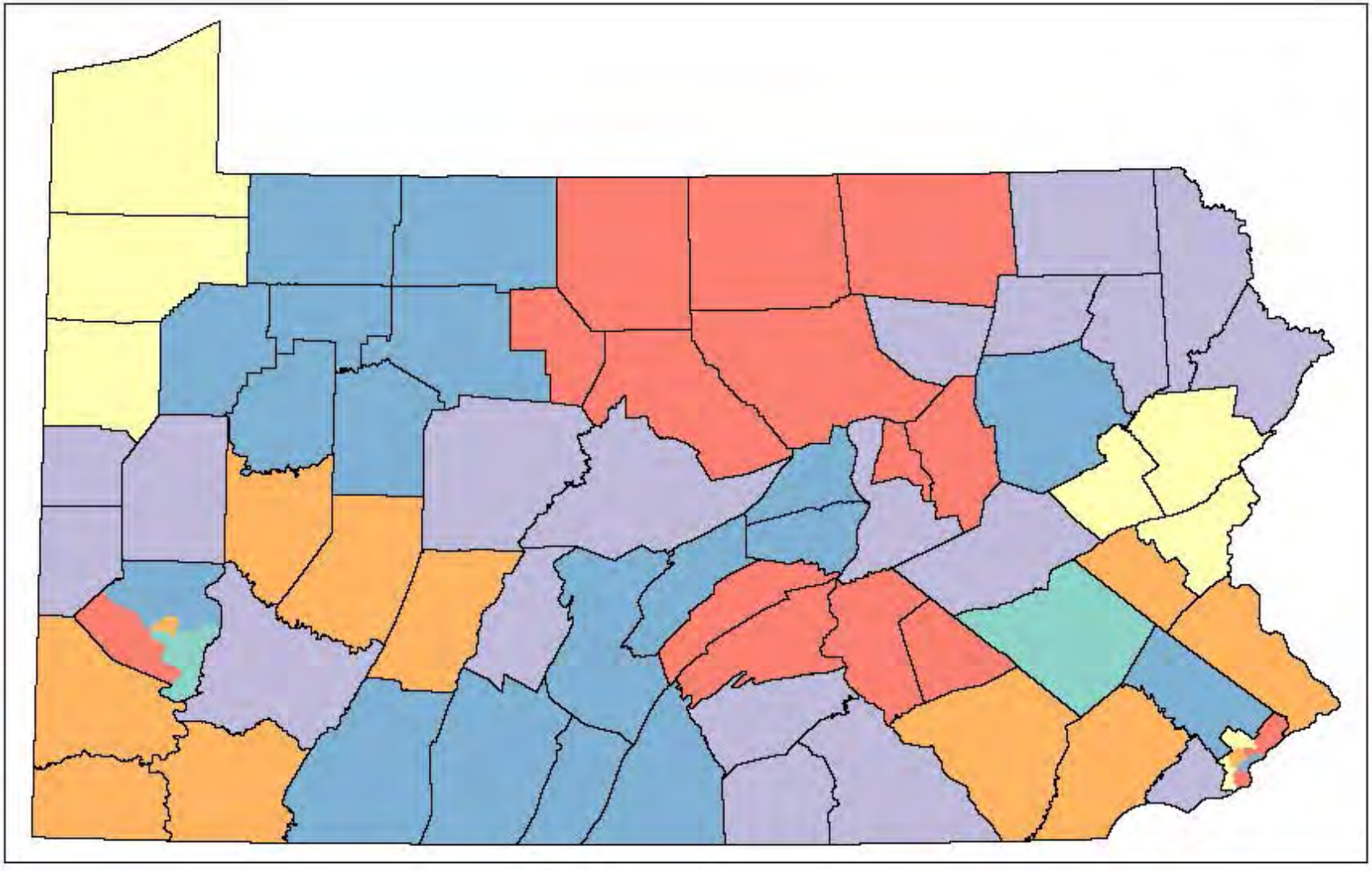
- Congressional districts may be suspect if $> 0.1\%$ population deviation
- State legislative districts may have up to 10% population deviation

1941



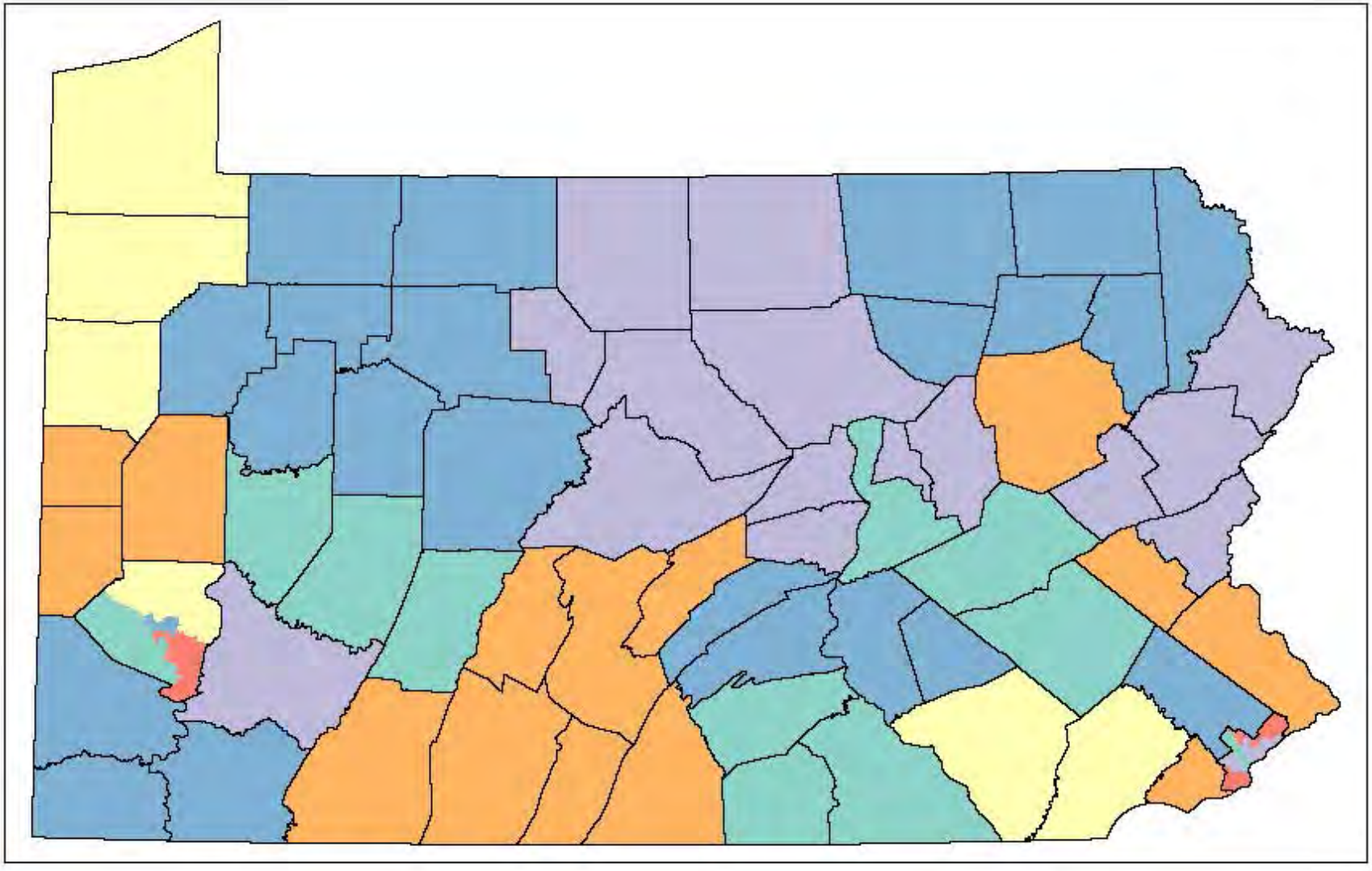
Compactness: min = 0.238, avg = 0.417

1951



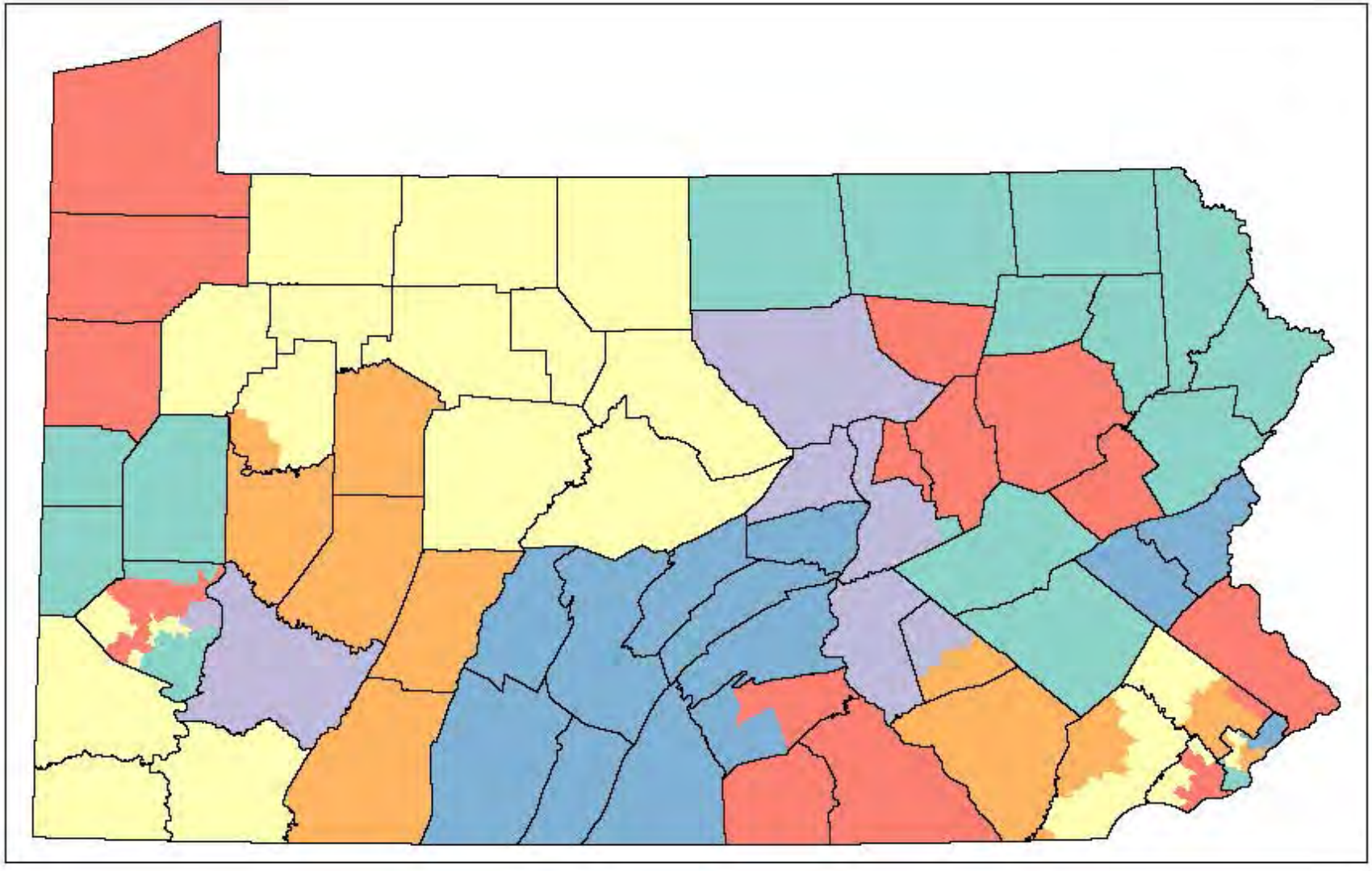
Compactness: min = 0.204, avg = 0.411

1961



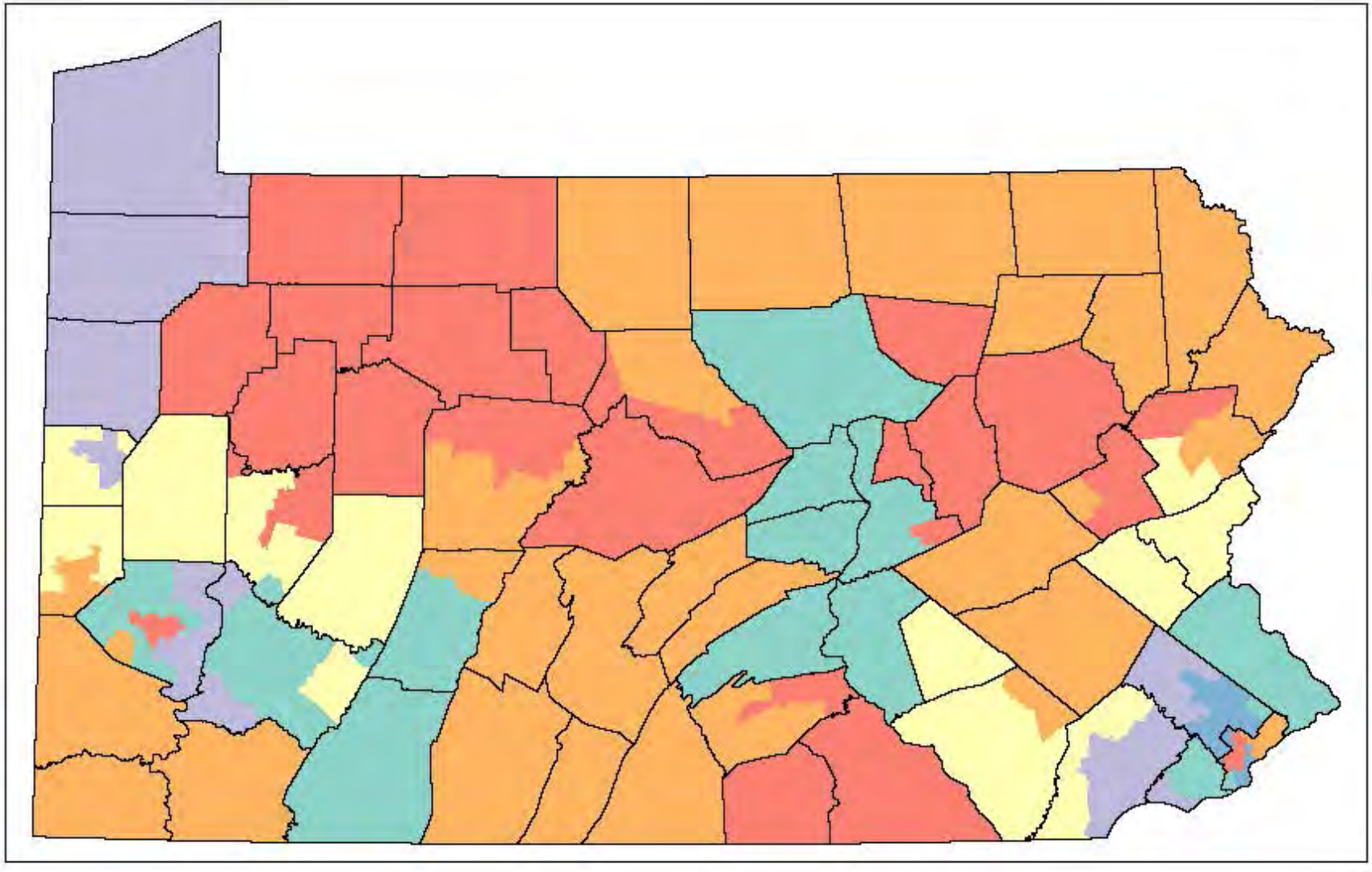
Compactness: min = 0.192, avg = 0.388

1971



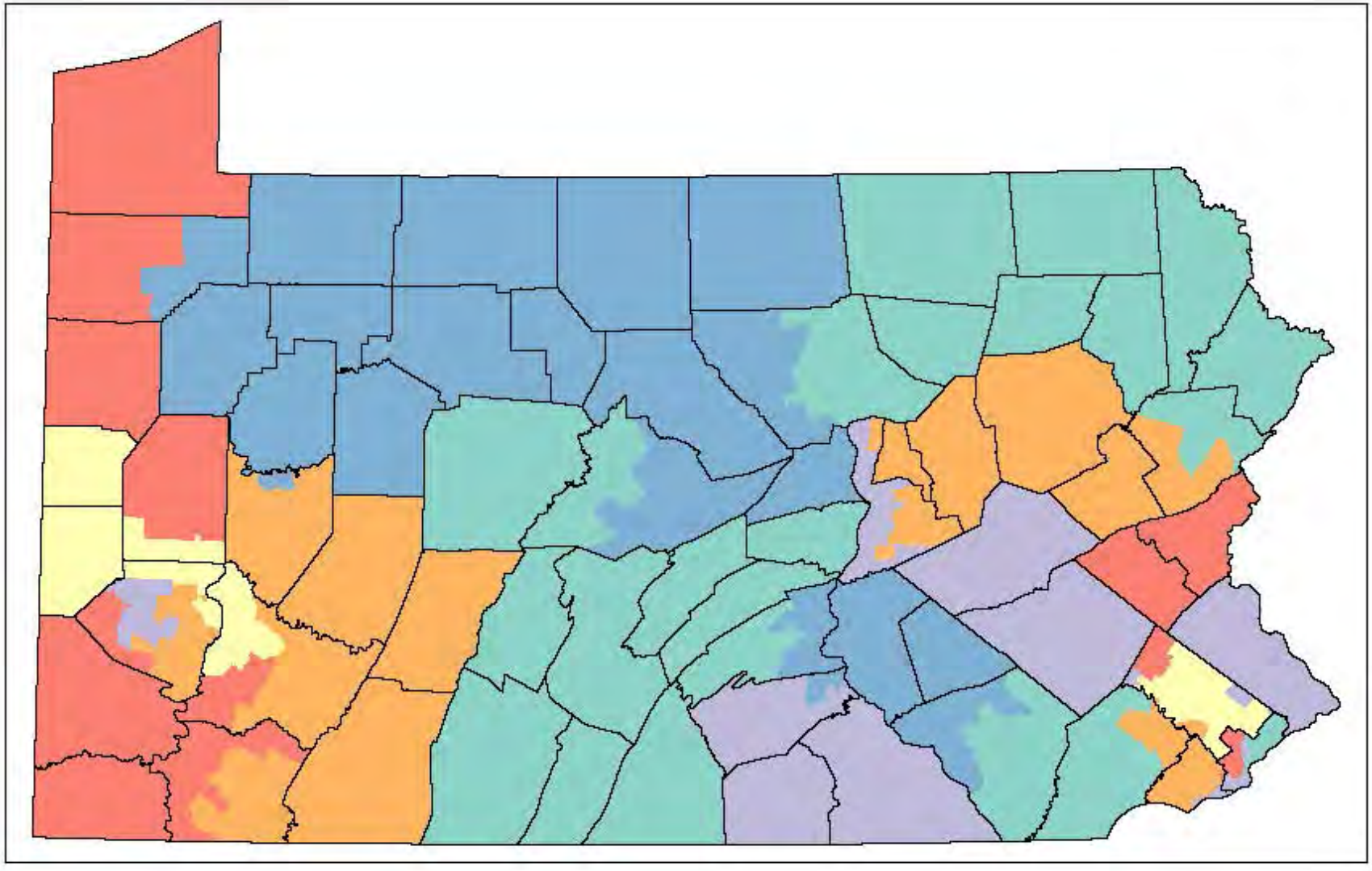
Compactness: min = 0.123, avg = 0.338

1981



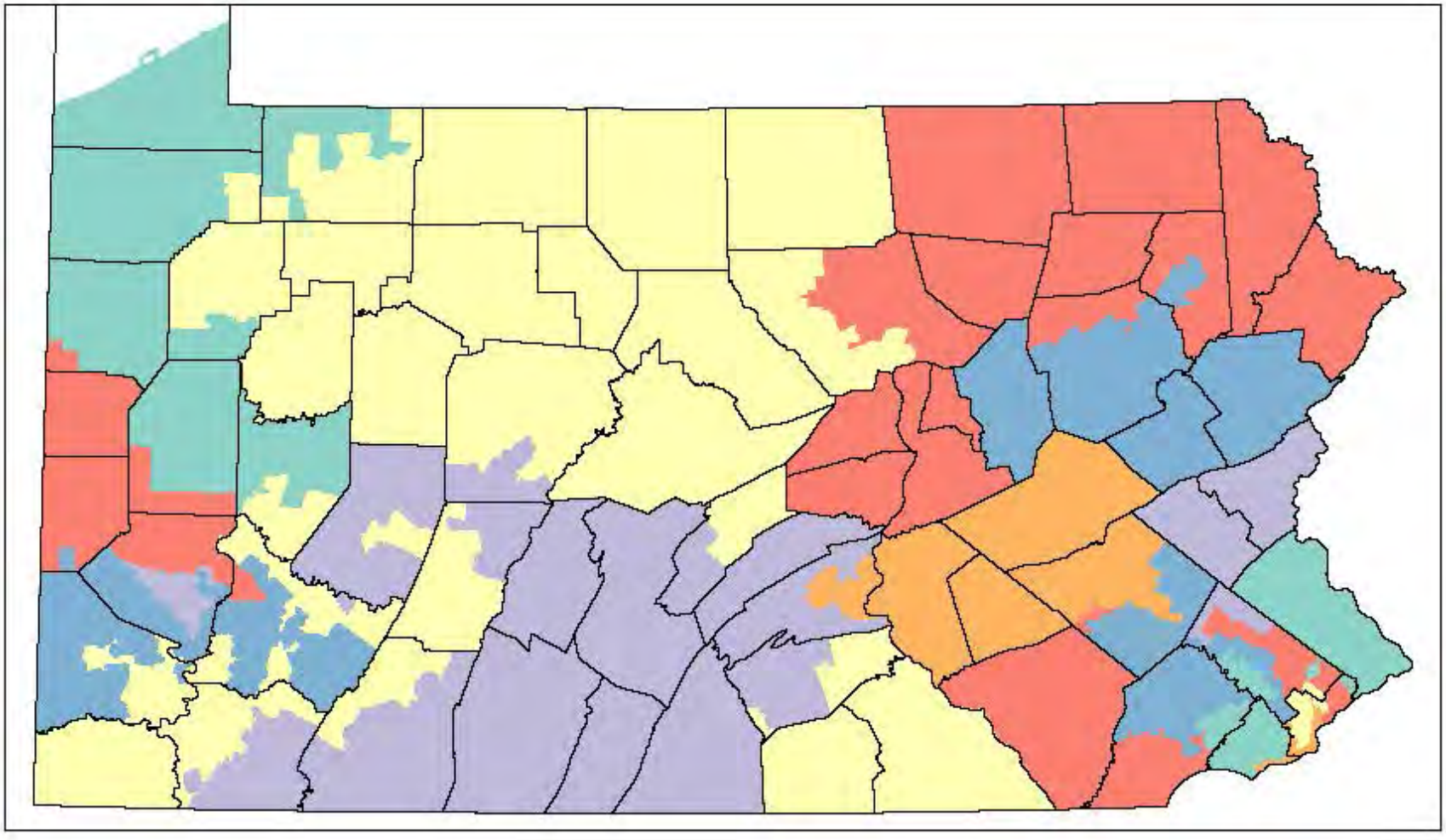
Compactness: min = 0.0917, avg = 0.252

1991



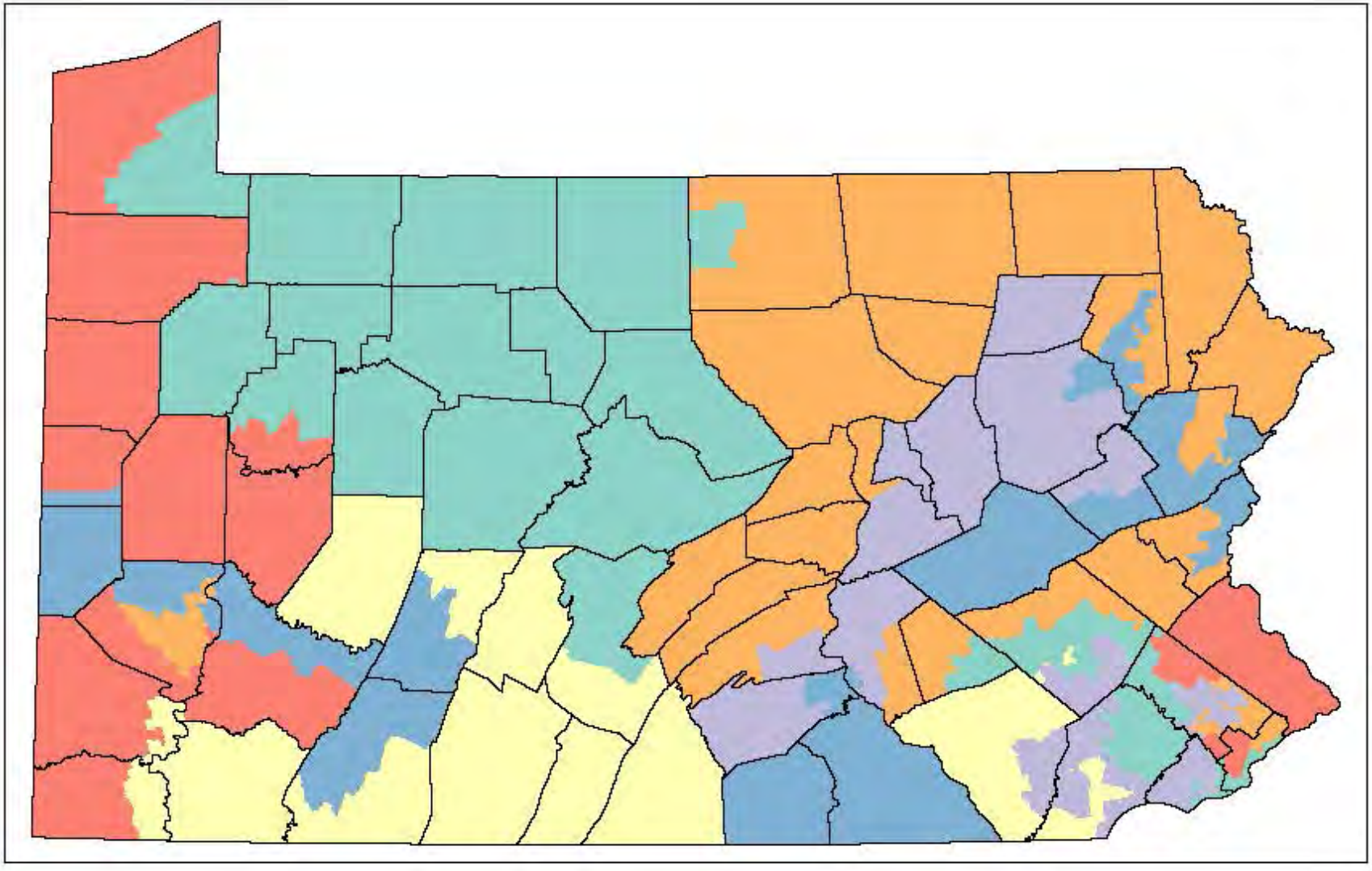
Compactness: min = 0.0983, avg = 0.259

2001



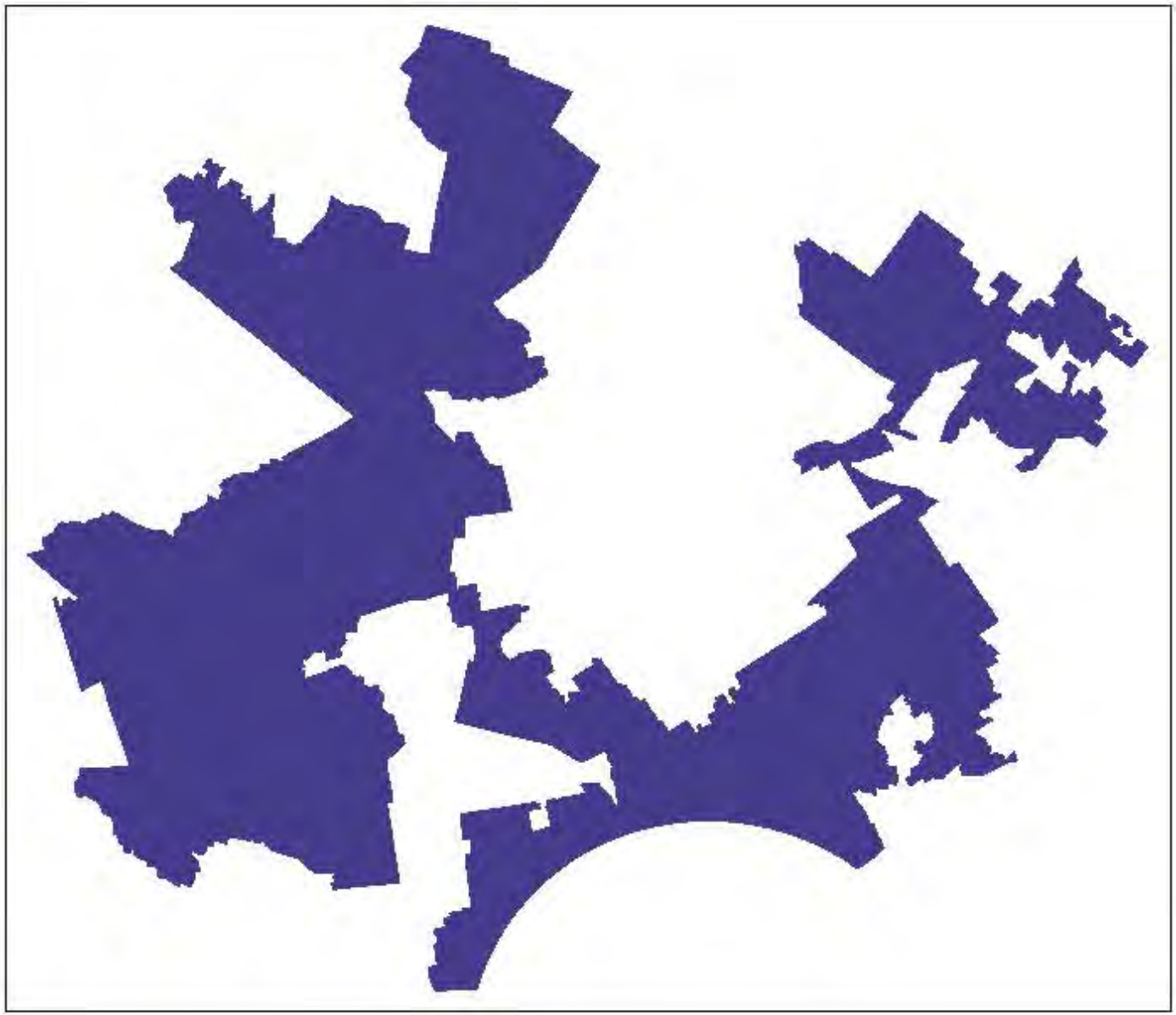
Compactness: min = 0.0483, avg = 0.185

2011



Compactness: min = 0.041, avg = 0.171

“Goofy Kicking Donald Duck”



Respecting Political Subdivisions

- 42 states require state legislative districts to account for internal political boundaries
- 19 states require congressional districts to account for internal political boundaries

Source: Justin Levitt, All About Redistricting, http://redistricting.lls.edu/where-state.php#bounds

Recent Court Cases

- *Agre v. Wolf* - Challenge under US Constitution Elections Clause "time, place and manner"
- *League of Women Voter v. Pennsylvania* - Challenge under Pennsylvania state constitution guarantee of "free and equal" elections

Respecting Political Subdivisions

> The Commonwealth shall be divided into fifty senatorial and two hundred three representative districts, which shall be composed of compact and contiguous territory as nearly equal in population as practicable. Each senatorial district shall elect one Senator, and each representative district one Representative. **Unless absolutely necessary no county, city, incorporated town, borough, township or ward shall be divided** in forming either a senatorial or representative district.
>
> —Pennsylvania State Constitution Article 2, Section 16

Rules Proposed in CCFD Amicus Brief

1. Keep small counties whole; split large counties *minimum* number of times necessary. (5% - 10% population deviation)
2. Add whole townships, boroughs, towns, or cities along the border. (2% population deviation)
3. Choose one and only one political subdivision at the border to split down to the block level. (± 1 person)

What Population Data Do We Use?

US Constitution authorizes a census for the purpose of apportionment, but is silent about its use for redistricting.

Problems with Census data:

- Undercount and overcount
  - Philadelphia County has a net undercount of 0.66% ($\sim$ 9,600 persons)
  - Franklin County has an extra 1.97% ($\sim$ 3,000 persons)
  - Undercount higher among African-Americans, renters, young males
  - Overcount higher among White non-Hispanics, homeowners, middle-aged females
- 0.7% of population is counted in the wrong Census block cluster

Ongoing Work

1. Appeal *Agre* to the US Supreme Court
2. Investigate (and challenge) state legislative districts for gerrymandering and conformance to rules.
3. Challenge ± 1 person.
4. Investigate implications for representation of racial bias in the undercount/overcount.
5. Contribute to multistate election data organizing efforts
6. Model election outcomes of hypothetical legislative/Congressional districts

CARTO and Mapbox Team Up

bit.ly/ireg-carto-mapbox

Andrew Thompson

Solutions Engineer

CARTO

CARTO IS THE PLATFORM FOR TURNING LOCATION DATA INTO BUSINESS OUTCOMES

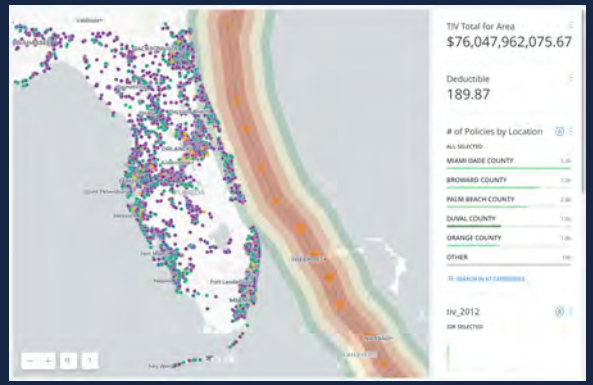
=

LOCATION INTELLIGENCE

MAKING LOCATION DATA UNDERSTANDABLE & ACTIONABLE FOR DIFFERENT USER TYPES

ANALYSTS & BUSINESS USERS

Out of the box location intelligence for analysts to create and use intuitive maps and map-based dashboards.

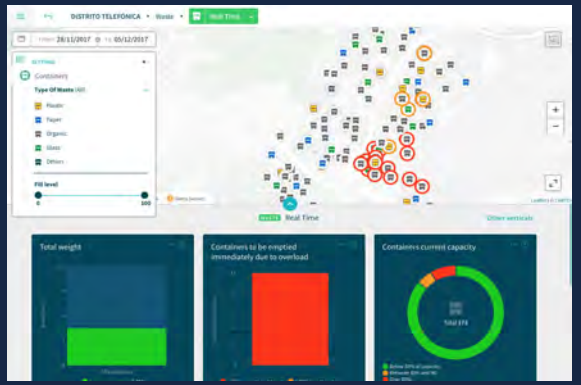


DATA SCIENTISTS

Powerful data science and analysis tools for understanding, predicting, and optimizing.

DEVELOPERS

Industrial grade APIs, SDKs and tools for developers to build world class geospatial apps.

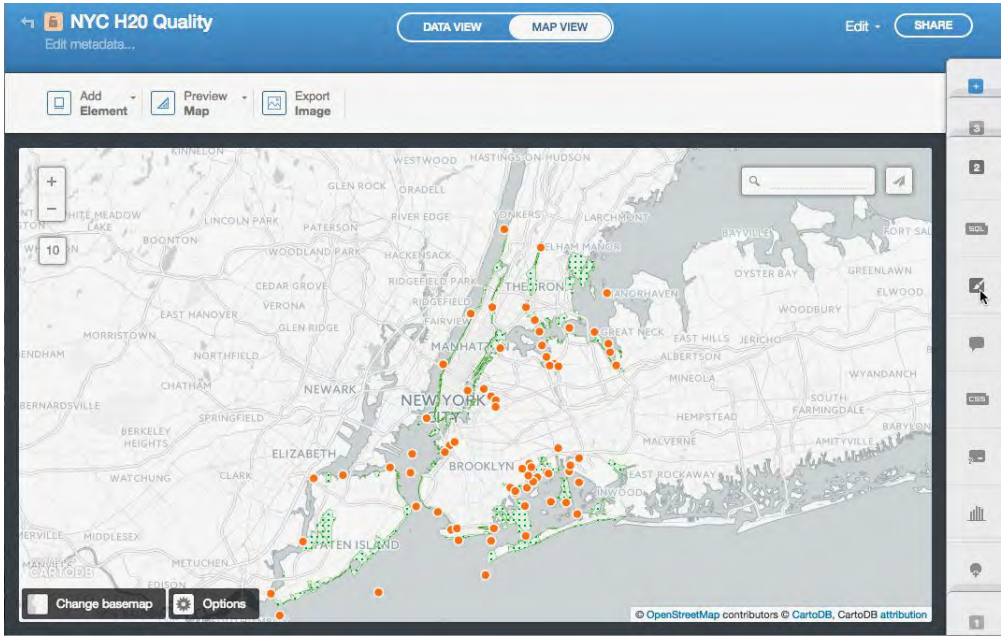


PUBLIC SECTOR CARTO CLIENTS


Andrew, you presented at IREG a year ago, right?

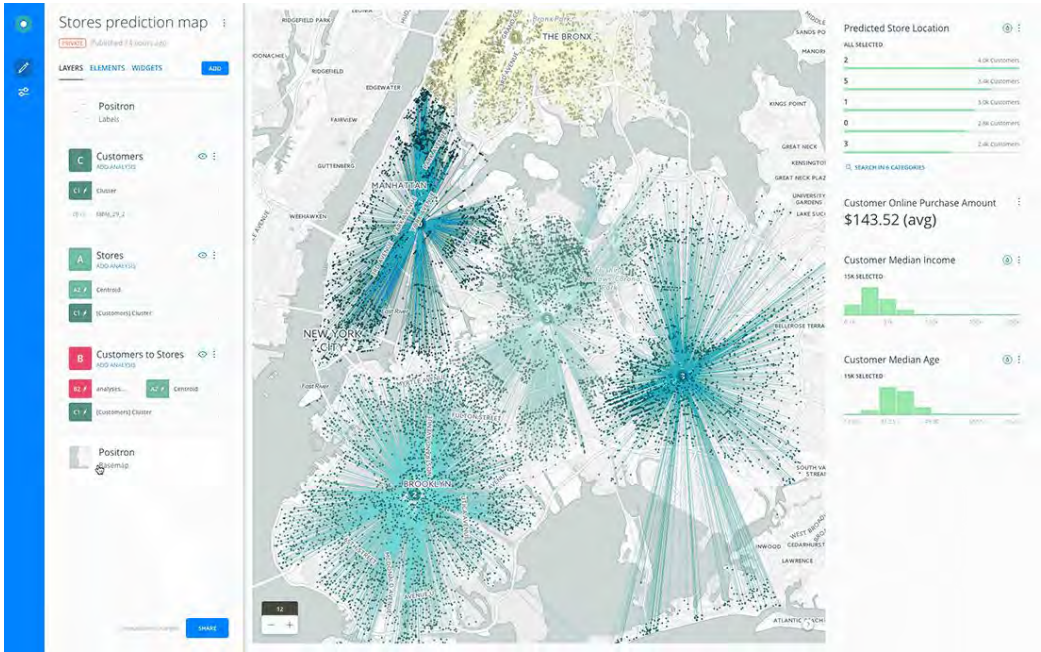
CARTODB

Old: “CartoDB Editor”



CARTO

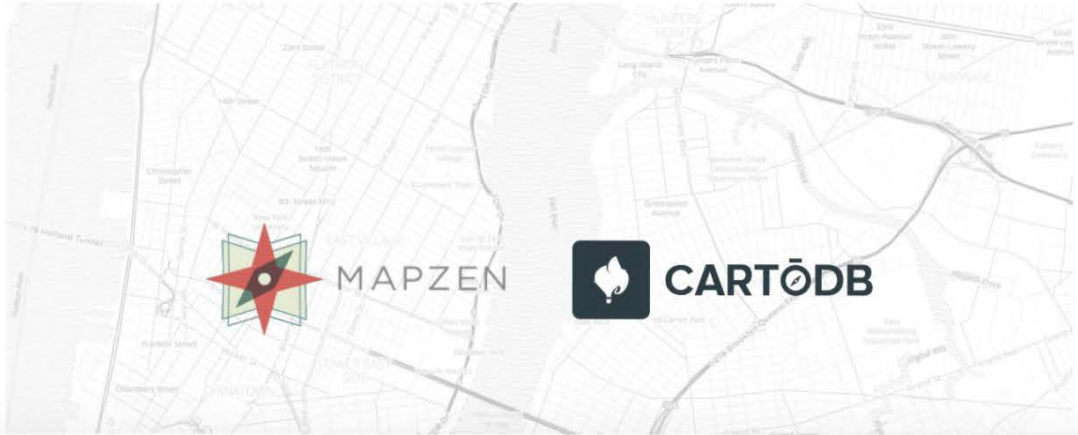
New: “CARTO Builder”




What else
happened way
back in 2016?

PARTNERING WITH MAPZEN TO PROVIDE NEXT-GENERATION LOCATION DATA SERVICES

Written by Javier de la Torre on Apr 21, 2016



Related Articles

FEB 20, 2018
Data Through Design Opening Reception: Kicking-Off NYC Open Data Week 2018 in Style

FEB 2, 2018
Discover Location Intelligence with CARTO at MWC 2018

Let’s talk Location Data Services...

- Geocoding and Routing are the foundation of spatial analysis
  - “Necessary ingredients”
- Location Data is Exploding
  - Self-driving cars, Internet of Things, Smart Cities, Mobile devices
- Driving Mapping Cars is *EXPENSIVE!*
  - Which makes proprietary geocoding data expensive...
- If the necessary ingredients stay expensive, “big data” analysis will be too
- The long arc of the Internet bends towards Free and Open
  - Costly, proprietary innovations become cheaper, open commodities

MAPZEN


Everything’s
great until...

SAMSUNG

Mapzen Shutdown :(

Unfortunately, we have some sad news. Mapzen will cease operations at the end of January 2018. Our hosted APIs and all related support and services will turn off on February 1, 2018. You will not be charged for API usage in December/January. We know this is an inconvenience and have provided a **migration guide to similar services** for our developer community. Our goal is to help as much as possible to ensure continuity in the services that you have built with us.

Fortunately, the core products of Mapzen are built entirely on open software and data. As a result, there are options to run Mapzen services yourself or to switch to other service providers.

Mapzen is shutting down its services.

READ MORE


Barely 2
months
earlier...

Mapbox Follow
We are a location data platform, changing the way people explore the world.
Oct 10, 2017 · 2 min read

SoftBank leads $164 million Series C Funding

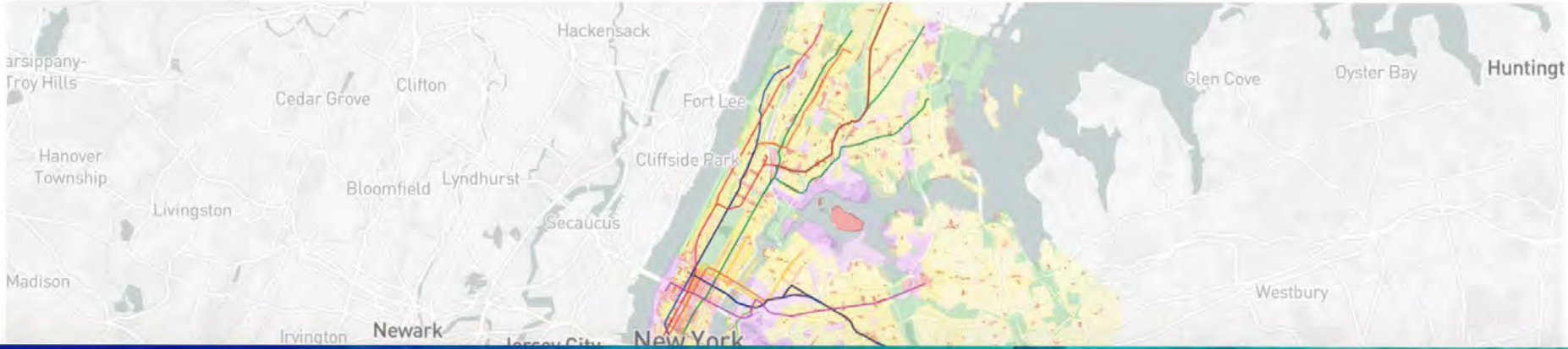
By: Eric Gundersen

We are mapping and measuring everything, live. About five minutes into the meeting with Masayoshi Son and his team, I knew SoftBank should be our partners. We didn't focus on "mapping" directly—rather we discussed how real-time location data will flow from decentralized networks of low powered mobile sensors that are now inside everything.


Out of
uncertainty,
comes
opportunity...

Partnering with Mapbox: A New Stack for Location Intelligence

Written by Javier de la Torre on Jan 16, 2018



Unnecessary confusion since forever ago

- Always focused on different but complementary things
  - CARTO: web-based spatial analysis and location intelligence
  - Mapbox: web-based basemaps and location data services
- Always was easy to use Mapbox and CARTO together
  - Many clients of both did/do so
- Yet, people new to mapping tech always had to ask “What’s the difference?”

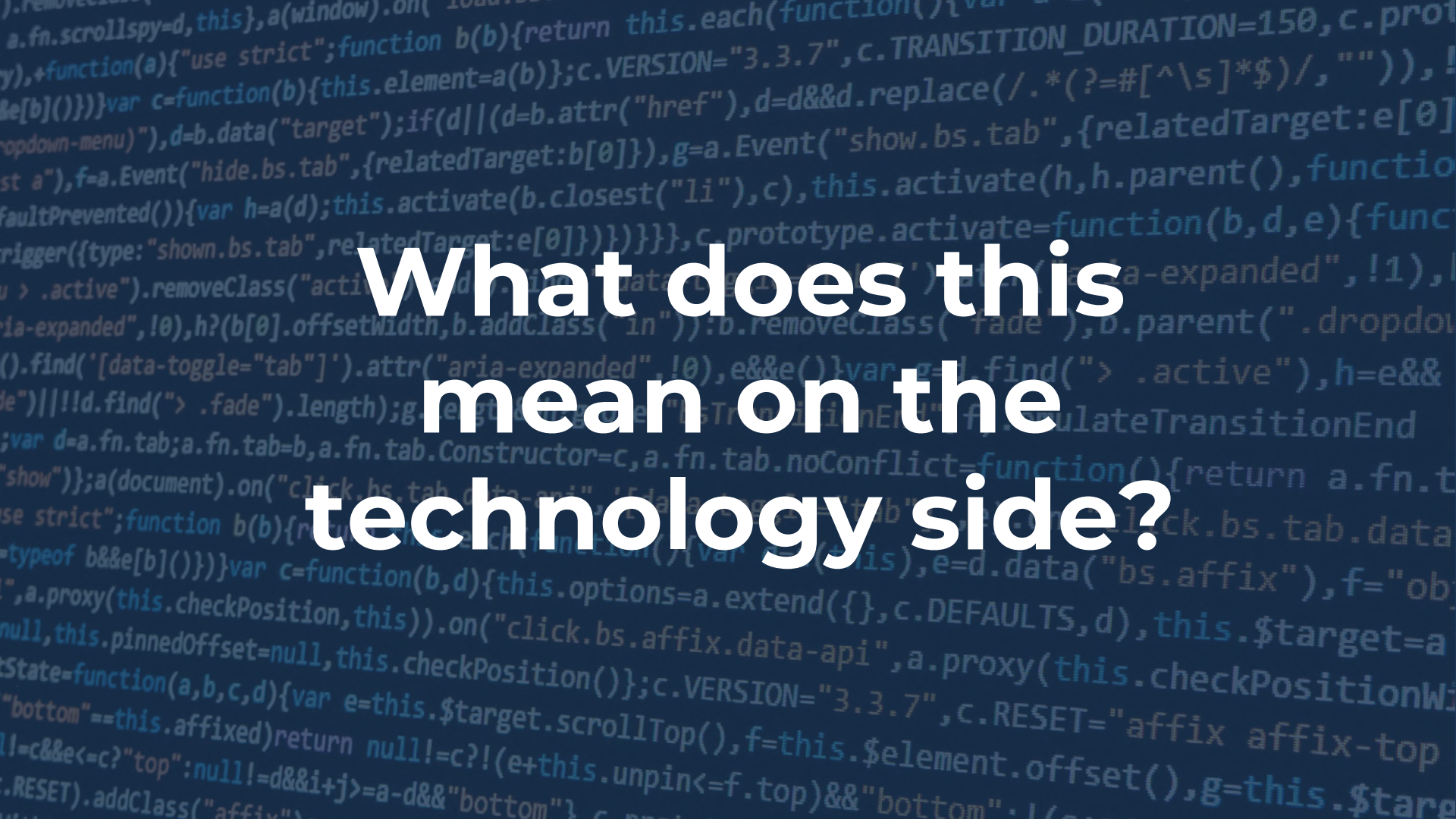
https://www.mapbox.com/help/carto/

The Legacy GIS Stack

- The last 20+ years...
- Closed proprietary code and data
- Desktop-first experience

The Modern Location Intelligence Stack

- The next 20 years!
- Open source code and data
- Web-first experience

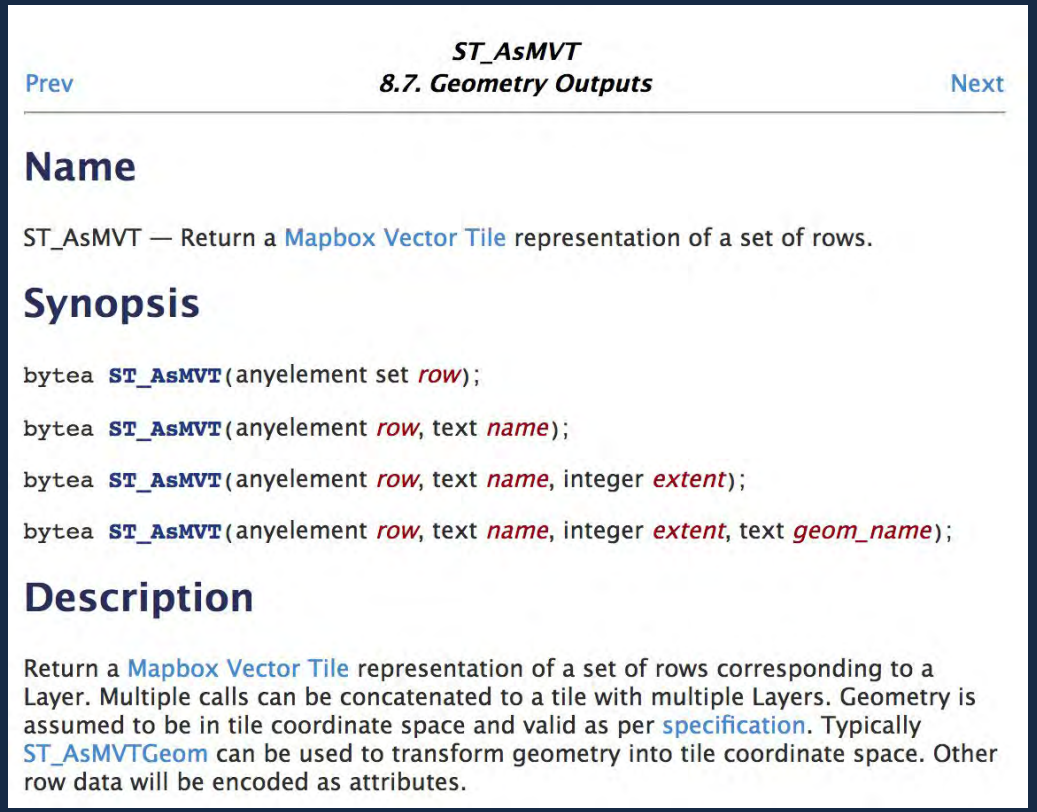

What does this mean on the technology side?

CARTO Supports Mapbox Vector Tiles!

https://carto.com/blog/using-mvt-in-carto/

Or...PostGIS Supports Mapbox Vector Tiles!

- CARTO sponsored creation of ST_AsMVT() in the brand-new PostGIS v2.4
  - This is how you open source!
  - Straight from DB to tile = fast!
  - Dynamic data? Just refresh your query



ST_AsMVT
Prev 8.7. Geometry Outputs Next

Name

ST_AsMVT — Return a Mapbox Vector Tile representation of a set of rows.

Synopsis

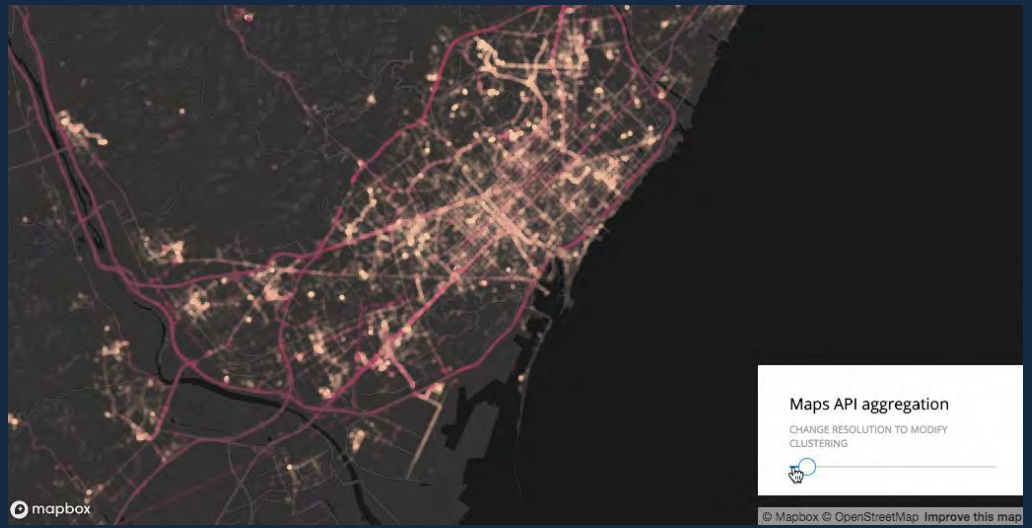
```
bytea ST_AsMVT(anyelement set row);
bytea ST_AsMVT(anyelement row, text name);
bytea ST_AsMVT(anyelement row, text name, integer extent);
bytea ST_AsMVT(anyelement row, text name, integer extent, text geom_name);
```

Description

Return a Mapbox Vector Tile representation of a set of rows corresponding to a Layer. Multiple calls can be concatenated to a tile with multiple Layers. Geometry is assumed to be in tile coordinate space and valid as per specification. Typically ST_AsMVTGeom can be used to transform geometry into tile coordinate space. Other row data will be encoded as attributes.

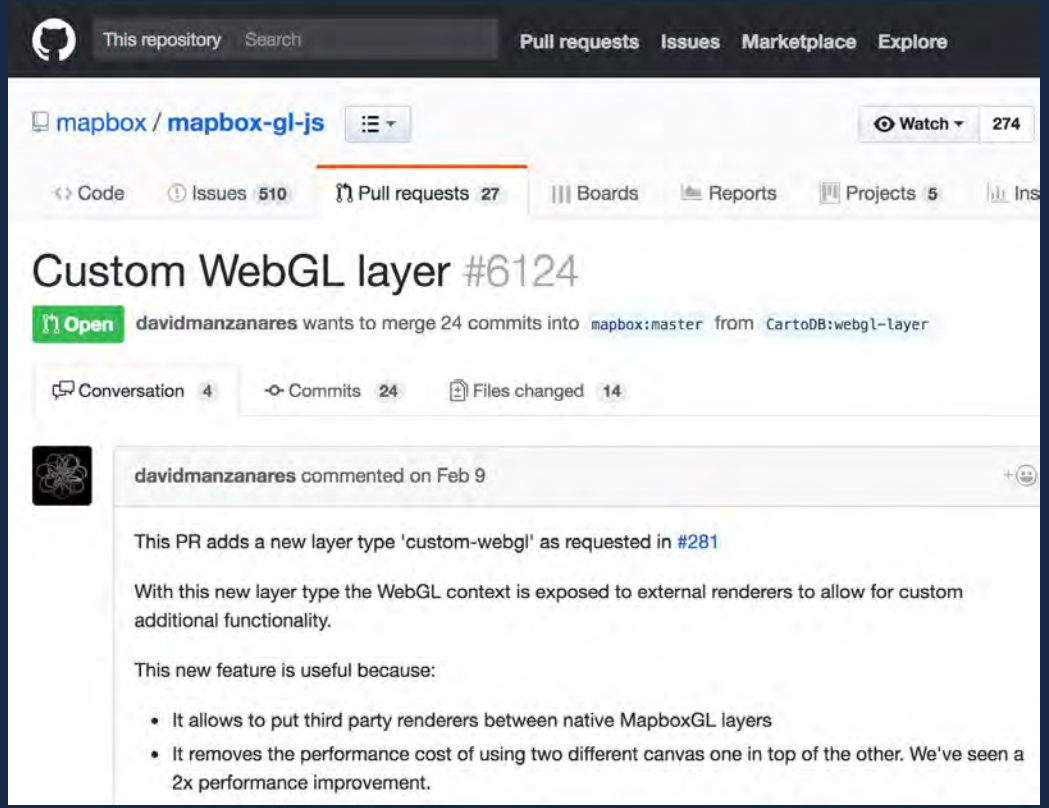
CARTO Makes MVT's Even Better

- Vector Tiles with 1 product versus Esri's 3
- Bringing our "Smart Aggregation" from raster tiles to vector tiles too
    - Principle: Don't render points on top of each other, that's a waste of time!
    - Getting better, but still cases raster > vector
        - FOSS4G Talk: "Raster is a Disaster, Vector is a Spectre" (slide, video)



Open Source Coordination

- Mapbox + CARTO engineering thinking about the future together
- Upcoming, To-be-named CARTO WebGL Javascript rendering library
    - Right now we're using MapboxGL too
- Pull Requests!
    - This is how two open source companies work in the open



Andrew Thompson
@andrewbt

22 Pull Requests. That's how many @CARTO teammates have made to non-CARTO-owned open source projects in roughly the last month, including @mapnikproject , @postgis and @OSGeo projects. Glad to be at a company that practices open source!

6:08 PM - 7 Mar 2018 from Philadelphia, PA

8 Retweets **36** Likes

2 8 36

Other Recent CARTO Nuggets?

bit.ly/ireg-carto-mapbox

CARTOFrames + Python SDK

Data Scientists work in Jupyter Notebooks (link)

New Basemap Designs

https://carto.com/basemaps
https://carto.com/blog/new-voyager-basemap/
https://carto.com/blog/inside/positron-dark-matter-new-look/

CARTO.JS V4 Beta!

Make your own filter widgets and JS Charts!
https://carto.com/documentation/cartojs/

Traffico and other Vertical Solutions

Ready-to-use application for traffic management
https://carto.com/solutions/traffico/

Thank You!

athompson@CARTO.COM

CARTO

LEVERAGING CENSUS DATA
FOR
MPO EQUITY ANALYSES

DELAWARE VALLEY dvrpc REGIONAL PLANNING COMMISSION

SHOSHANA AKINS
Public Participation Planner

KIM KOREJKO
Manager of Data Coordination

BEN GRUSWITZ
Senior Planner

LEVERAGING CENSUS DATA FOR MPO EQUITY ANALYSES

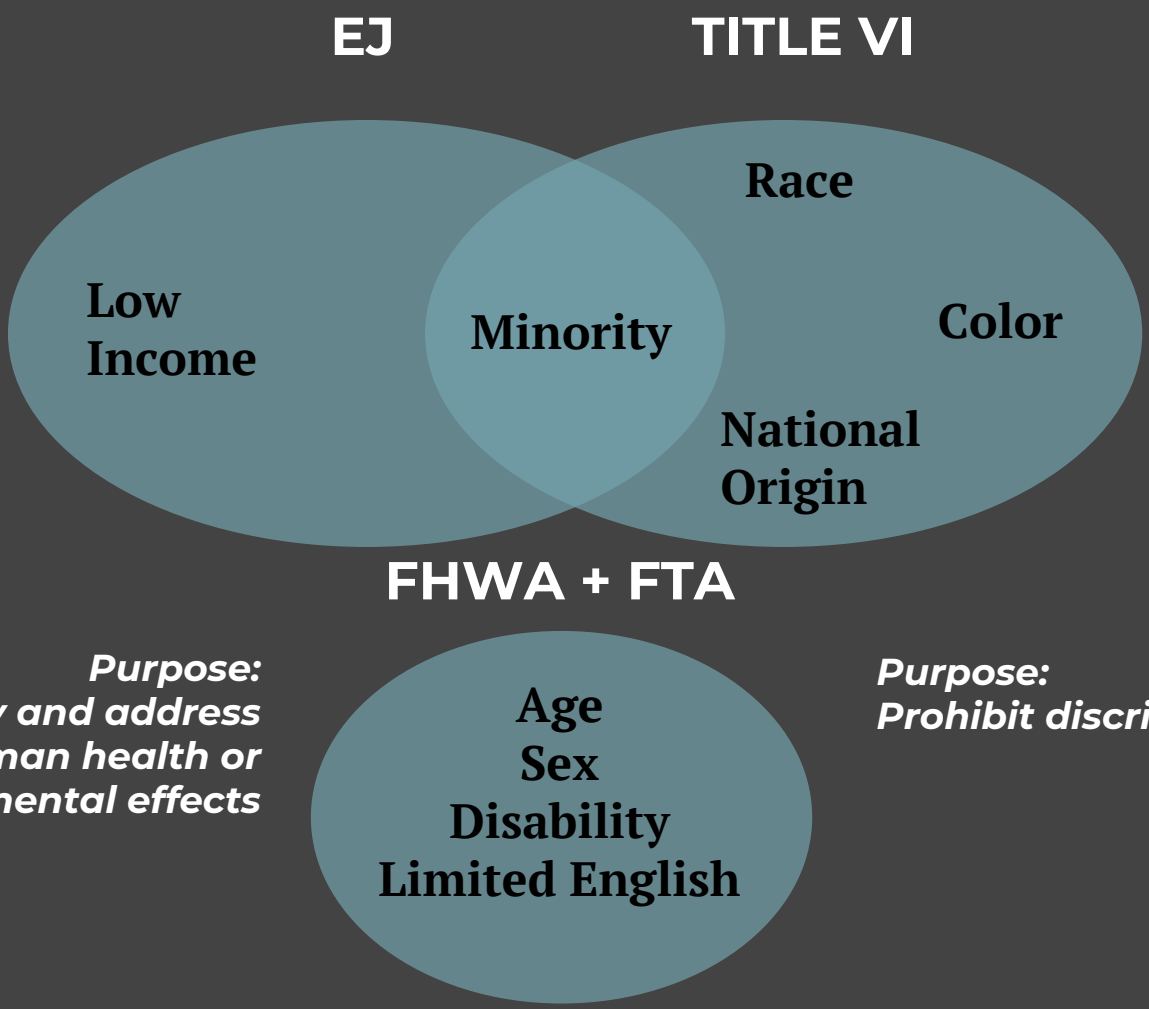
Overview

- DVRPC's Approach to Equity Analysis: *Indicators of Potential Disadvantage (IPD)*
- Understanding Equity Regulations for IPD 2.0
- Updating DVRPC's Methodology for IPD 2.0
- Lessons Learned: Beyond the Legislation
- IPD 2.1 Experiments:

IPD 1.0:
REGULATIONS + OVERVIEW

dvrpc

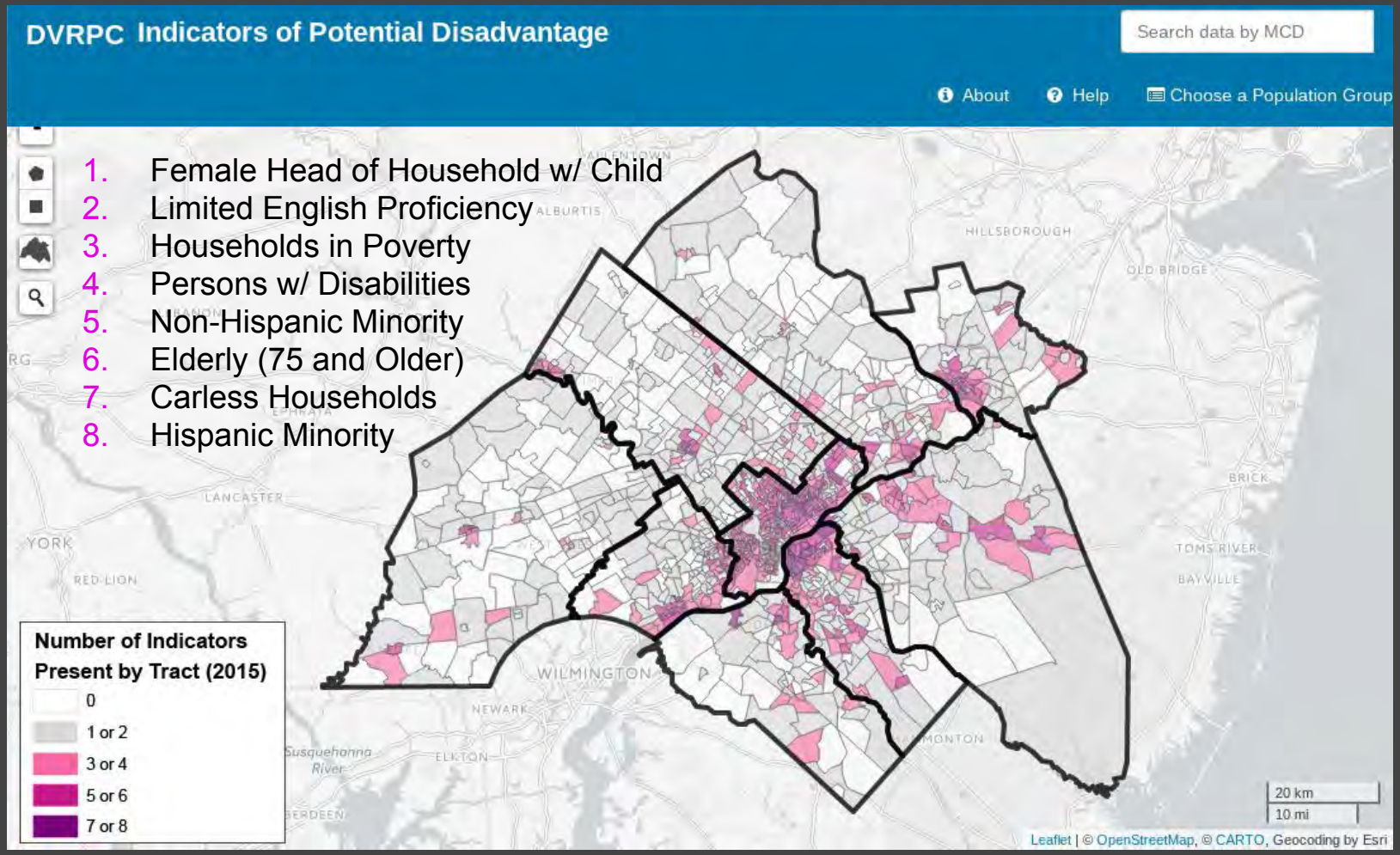
UNDERSTANDING EQUITY REGULATIONS



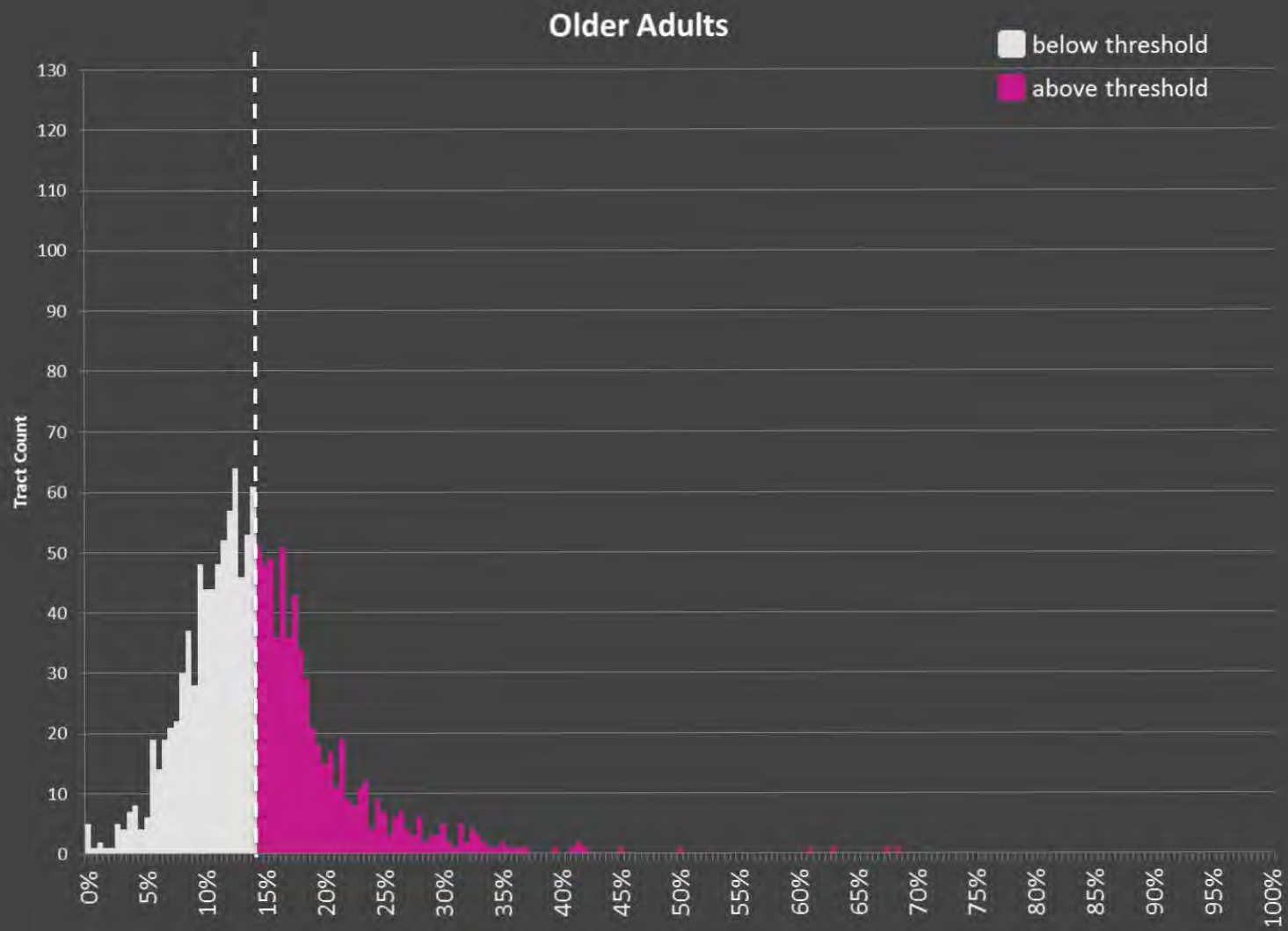
Purpose:
Identify and address adverse human health or environmental effects

Purpose:
Prohibit discrimination

INDICATORS OF POTENTIAL DISADVANTAGE 1.0



INDICATORS OF POTENTIAL DISADVANTAGE 1.0

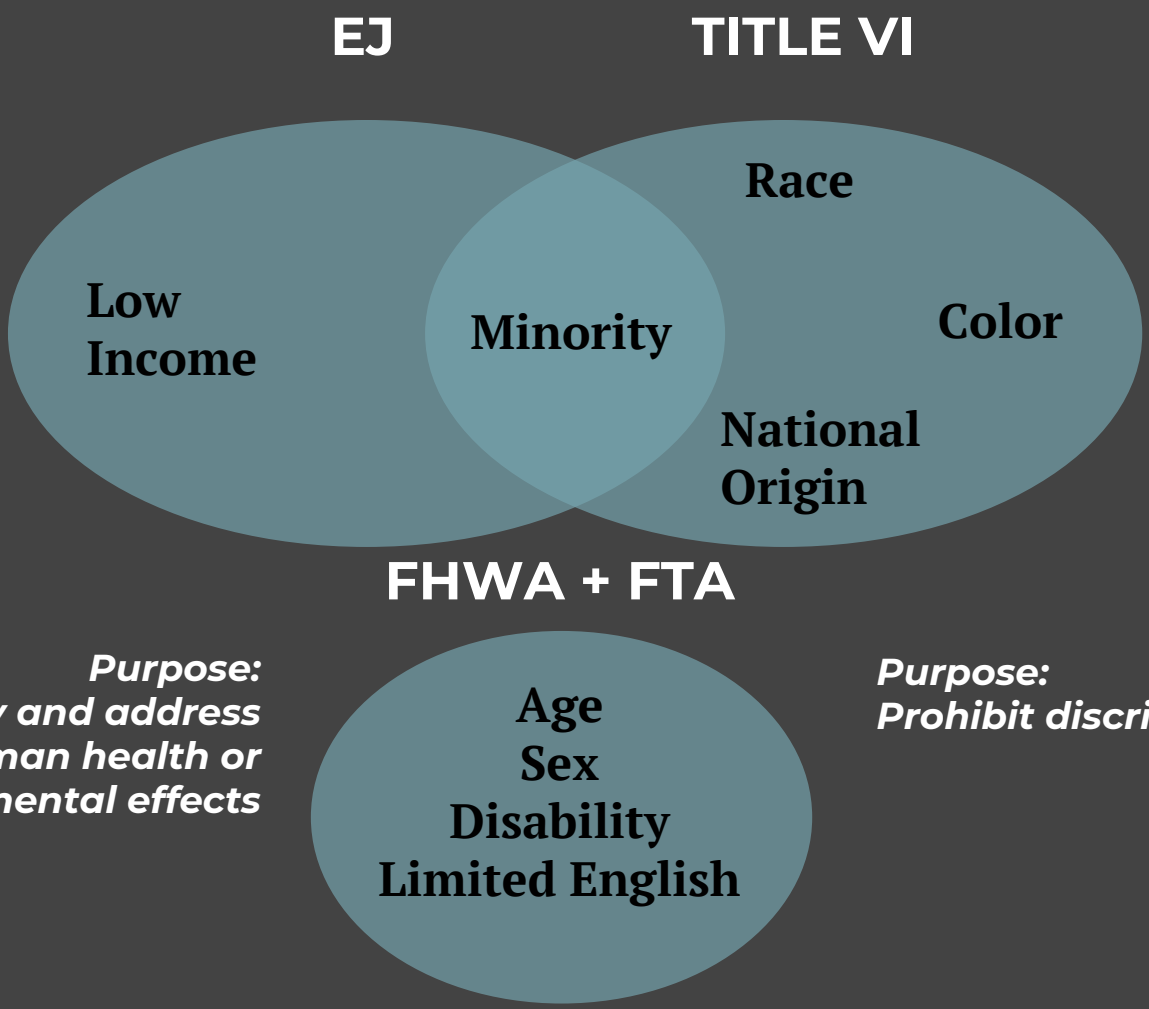


dvrpc

IPD 2.0: UNDERSTANDING EQUITY REGULATIONS

dvrpc

UNDERSTANDING EQUITY REGULATIONS



Purpose:
Identify and address adverse human health or environmental effects

Purpose:
Prohibit discrimination

ALIGNING INDICATORS WITH LEGISLATION

Current indicators	Title VI and EJ populations
• Elderly (75 and Older) • Female Head of Household with Child • Non-Hispanic Minority • Hispanic Minority • Limited English Proficiency • Persons with Disabilities • Households in Poverty • Carless Households	• Age • Sex • Minority • Race • Ethnicity • National origin • Limited English Proficiency • Disability • Low-Income

ALIGNING INDICATORS WITH LEGISLATION

Current indicators	Title VI and EJ populations
• Elderly (75 and Older) • Female Head of Household with Child • Non-Hispanic Minority • Hispanic Minority • Limited English Proficiency • Persons with Disabilities • Households in Poverty • Carless Households	• Age • Sex • Minority • Race • Ethnicity • National origin • Limited English Proficiency • Disability • Low-Income

ALIGNING INDICATORS WITH LEGISLATION

Current indicators	Title VI and EJ populations
• Elderly ~~(75+)~~ **(65+)**	• **Age**
• ~~**Female Head of Household with Child**~~	• **Sex**
• Non-Hispanic Minority	• Minority
• Hispanic Minority	• Race
• Limited English Proficiency	• Ethnicity
• Persons with Disabilities	• **National origin**
• Households ~~**in Poverty**~~	• Limited English Proficiency
• ~~**Carless Households**~~	• Disability
	• **Low-Income**

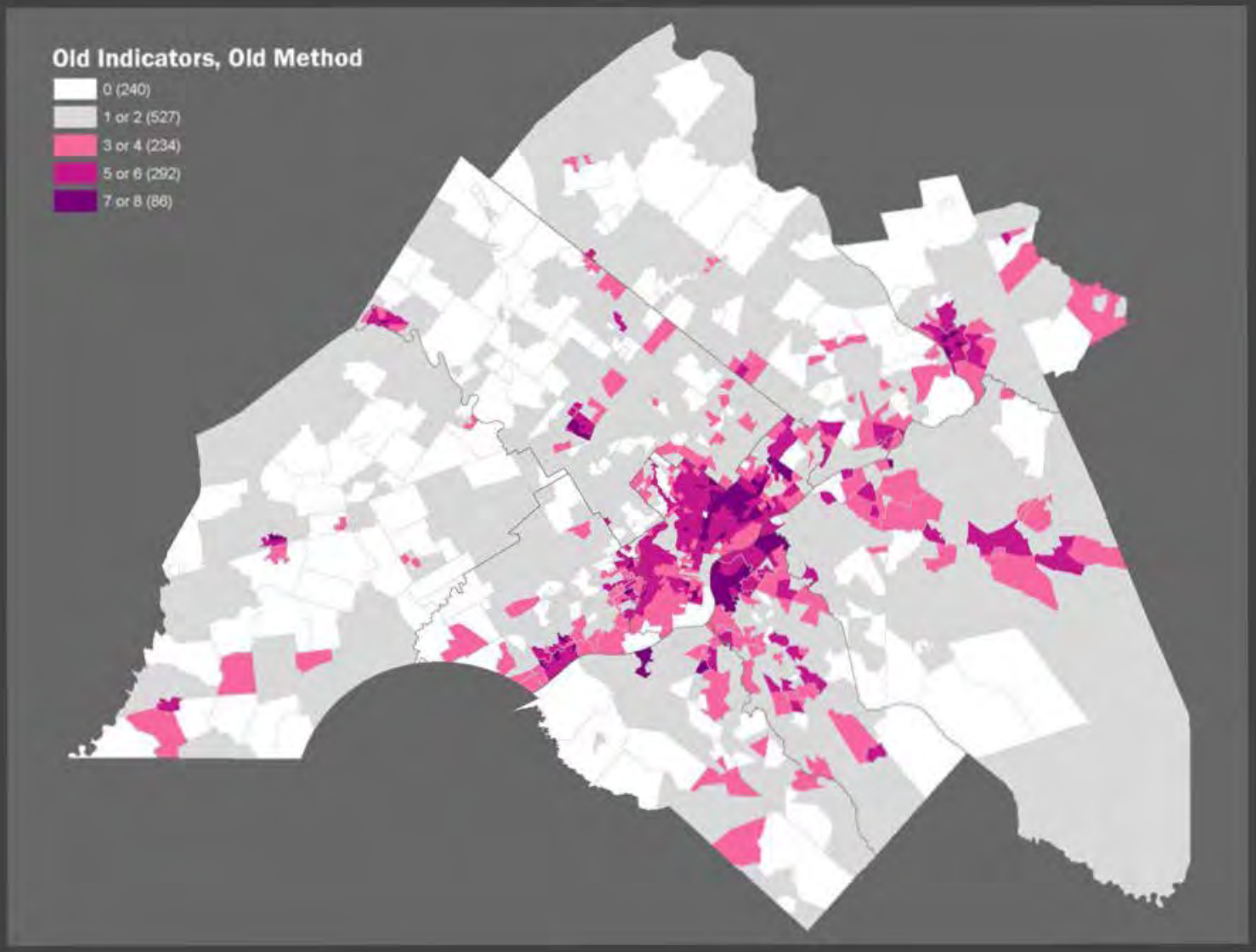
INDICATORS AND CENSUS TABLES

Indicator in IPD analysis update	ACS data table for indicator in IPD analysis	Protected class indicator represents
Youth	S0101: Age and Sex	Age
Older Adults	S0101: Age and Sex	Age
Female	S0101: Age and Sex	Sex
Racial Minority	B02001: Race	Race and Minority
Ethnic Minority	B03002: Hispanic or Latino Origin by Race	Minority and National Origin
Foreign Born	B05012: Nativity in the United States	National Origin
Limited English Proficiency	S1601: Language Spoken at Home	Limited English Proficiency, and National Origin
Disabled	S1810: Disability Characteristics	Disability
Low-Income	S1701: Poverty Status in the Past 12 Months	Low-Income

IPD 2.0:
UPDATING METHODOLOGY

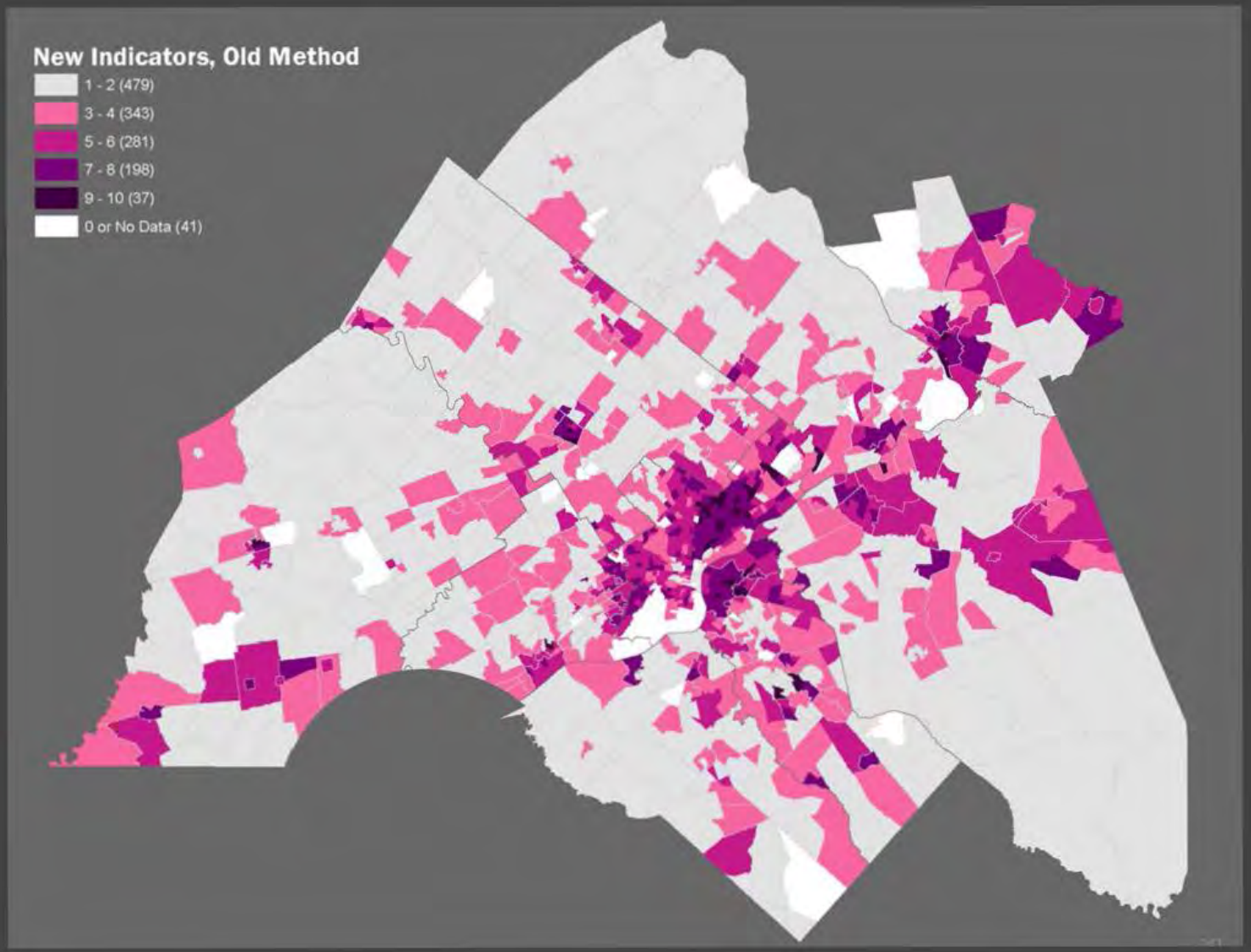
dvrpc

IPD 1.0 METHODOLOGY

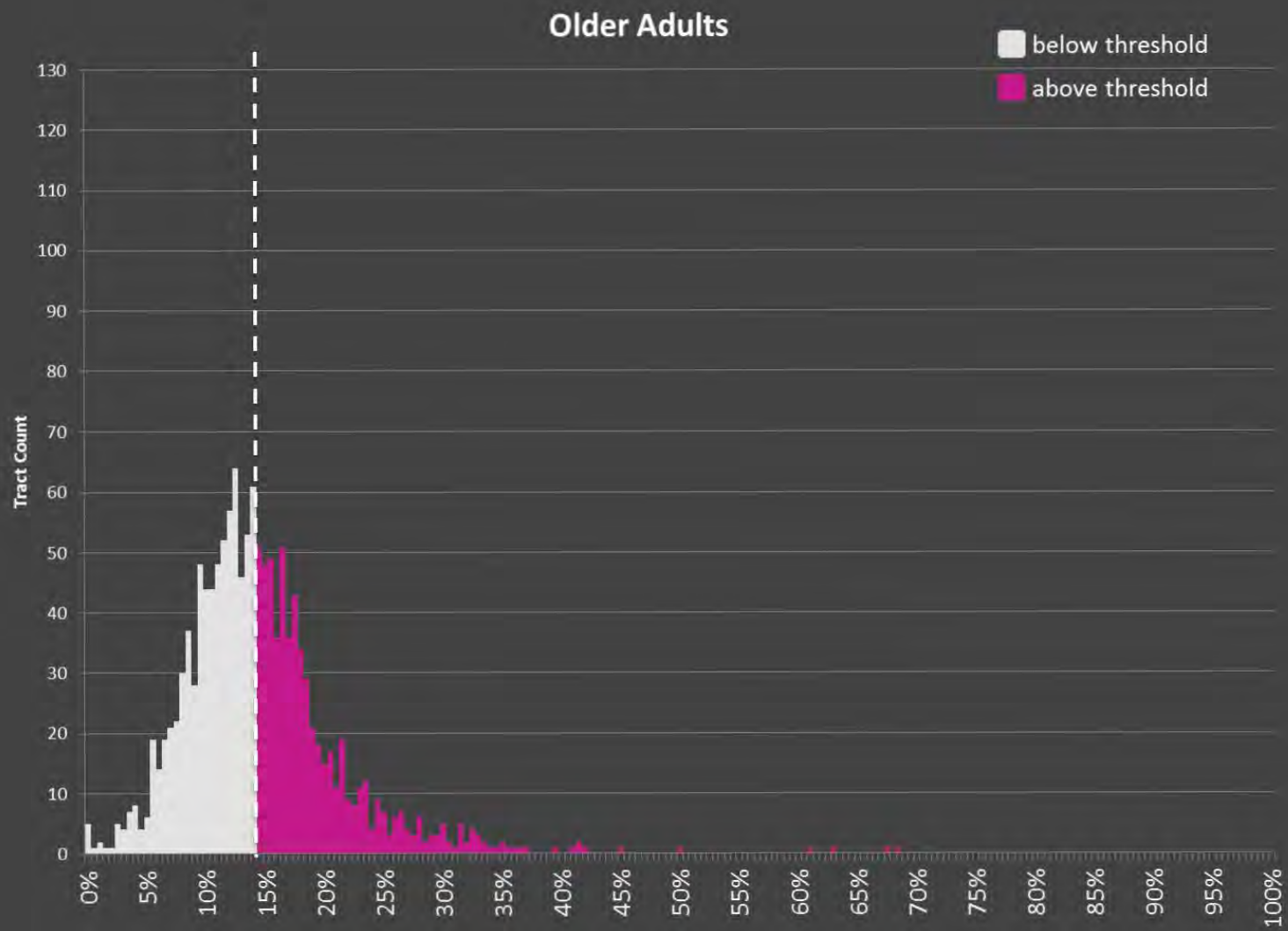


dvrpc

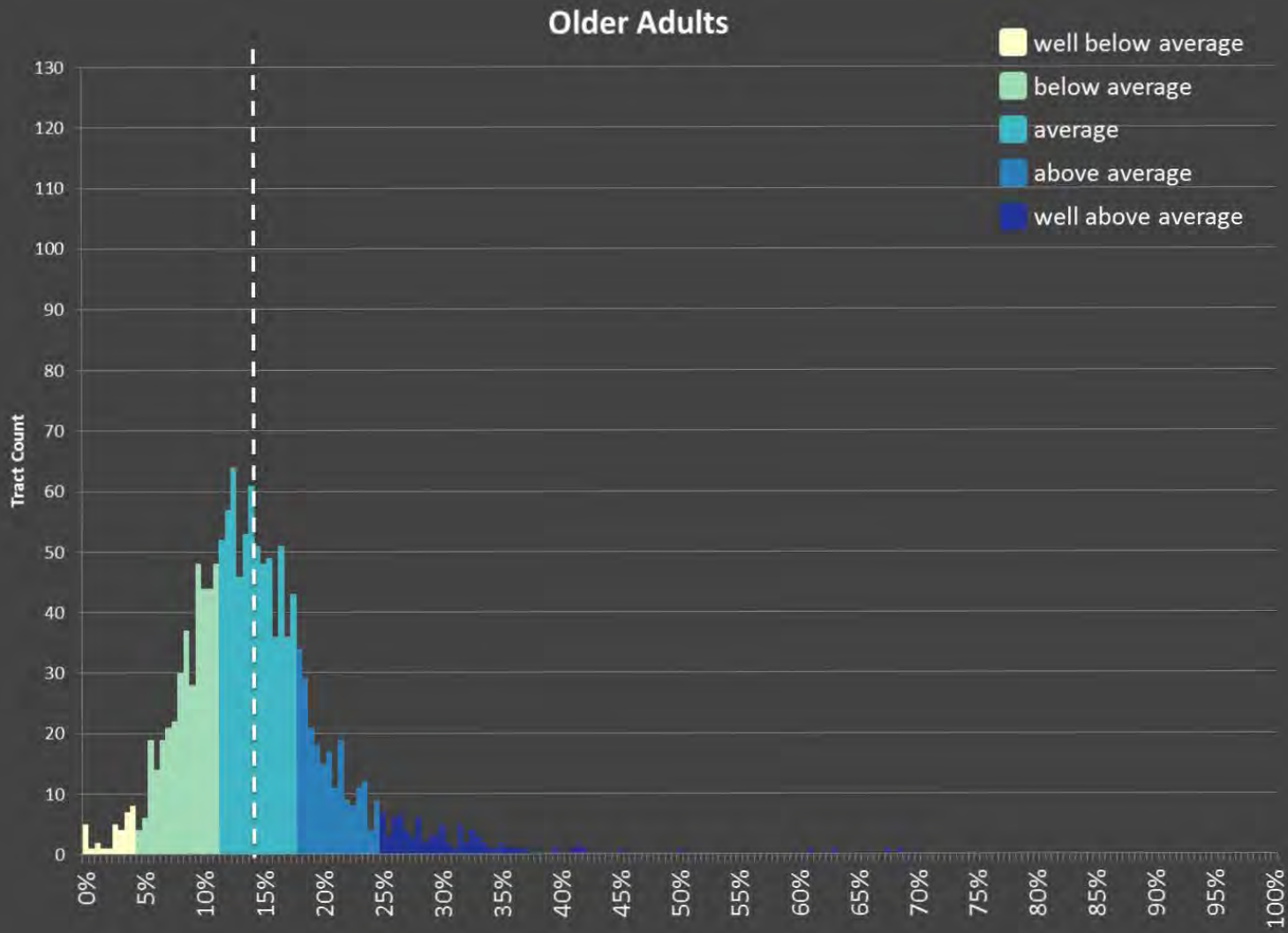
2.0 INDICATORS WITH 1.0 METHODOLOGY



1.0 METHODOLOGY



2.0 METHODOLOGY



IPD 2.0 METHODOLOGY: BINNING THE DATA

Youth

Older Adults

Female

Racial Minority

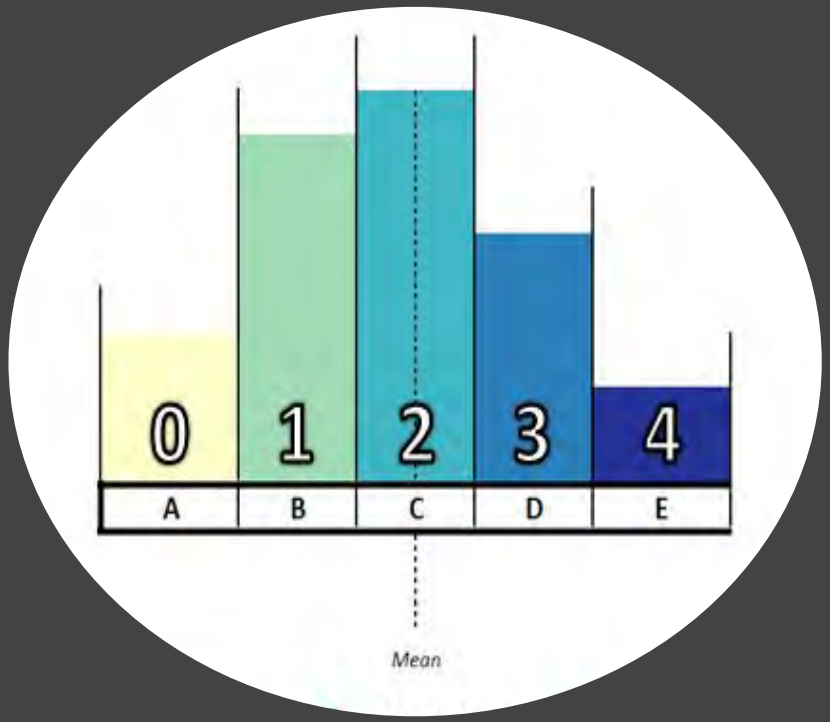
Ethnic Minority

Foreign Born

Limited English Proficiency

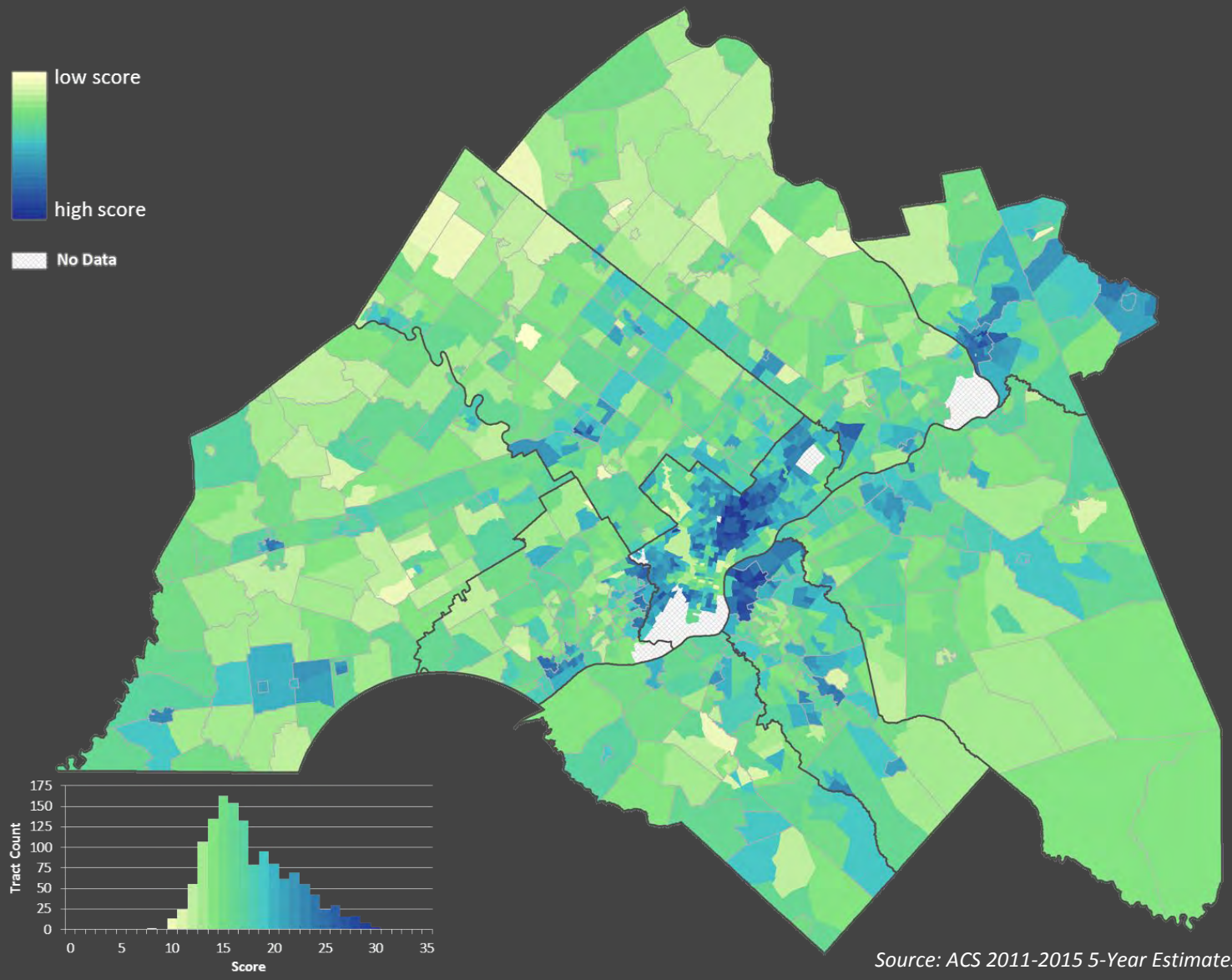
Disabled

Low-income



A - well below average
B - below average
C - average
D - above average
E - well above average

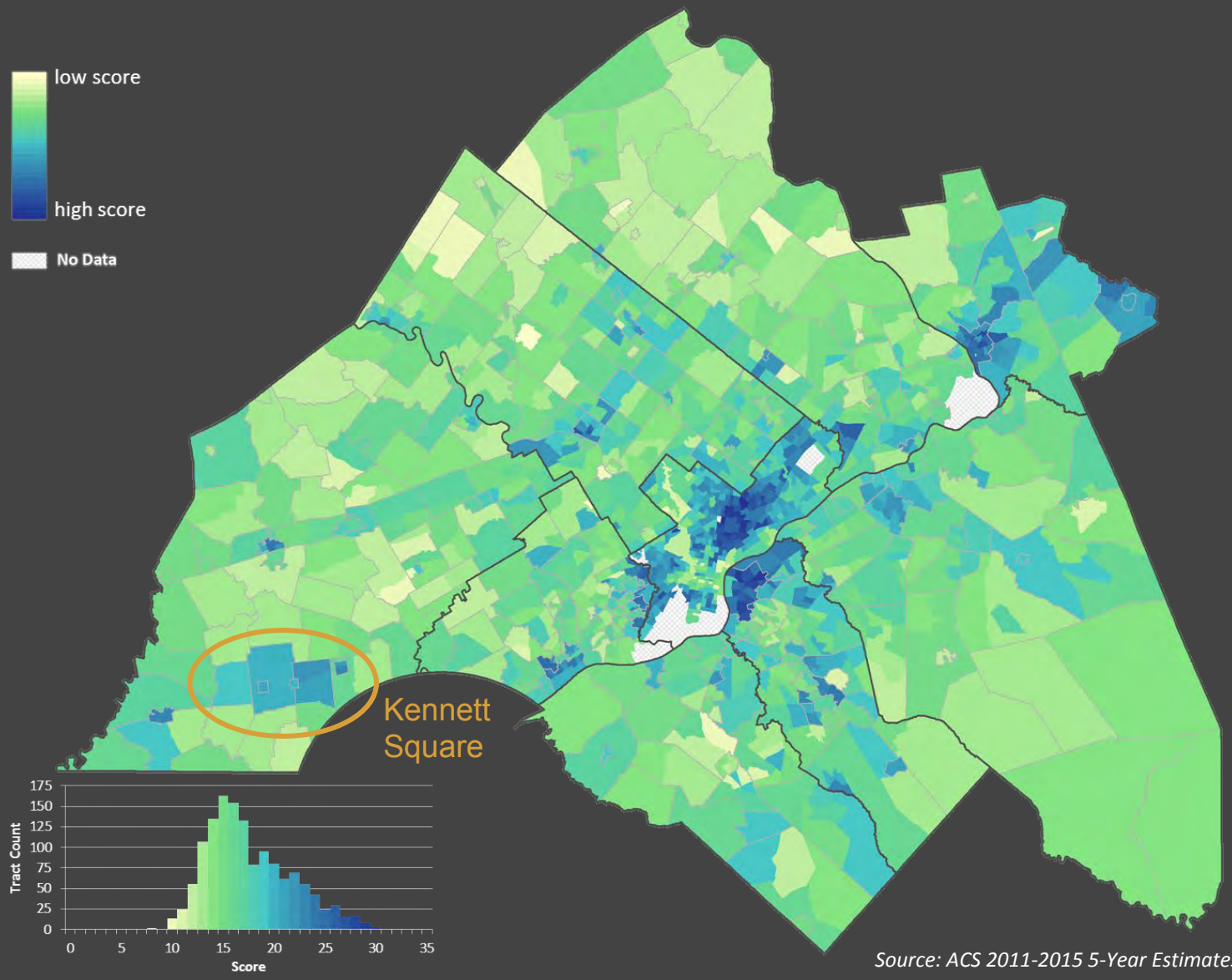
IPD 2.0 METHODOLOGY:
THE END RESULT OF BINNING THE DATA



Source: ACS 2011-2015 5-Year Estimates

dvrpc

IPD 2.0 METHODOLOGY: THE END RESULT OF BINNING THE DATA



Source: ACS 2011-2015 5-Year Estimates

dvrpc

Equity Analysis for the Greater Philadelphia Region (Beta)

Home | About

Census Tract: 3117

IPD Indicators

The information below provides the estimated percentages, Margin of Error (MOE), and IPD Score Classification of our nine indicators in reference to U.S Census Tract **3117**. The entire set of indicators data and scores may be downloaded from our Open Data Portal for further analysis.

Youth
30.5% +/- 3.1%
of residents are 18 years or younger
above average

Older Adults
10.3% +/- 1.5%
of residents are 65 years or older
below average

Female
49.7% +/- 2.6%
residents are female
average

Racial Minority
10.9% +/- 3.5%
of residents identify as one or more races considered under Title VI and EJ
below average

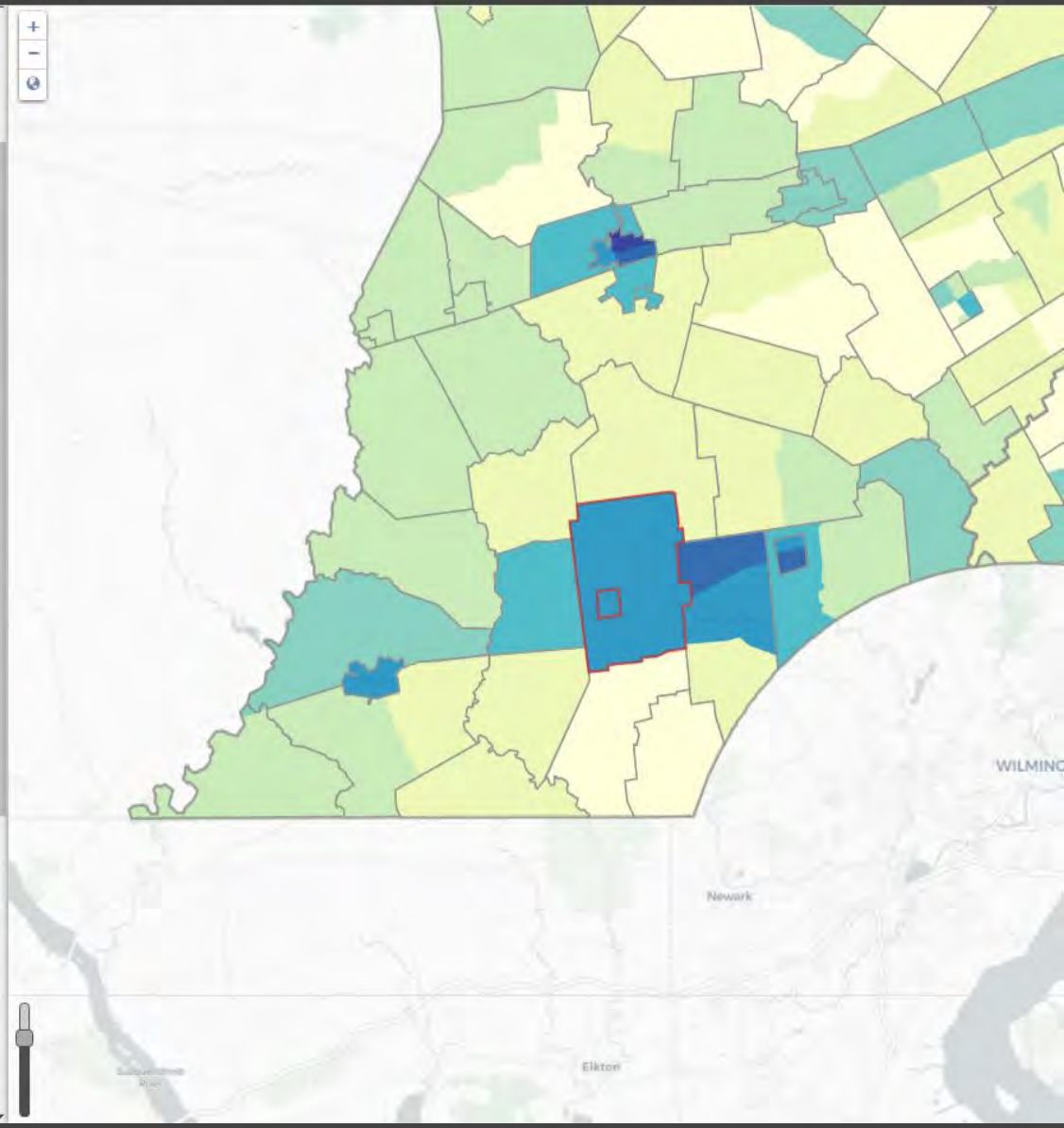
Ethnic Minority
34.0% +/- 8.0%
of residents identified themselves as being of Hispanic or Spanish origin.
well above average

Foreign Born
20.9% +/- 5.6%
of residents were born outside of the United States
above average

Limited English Proficiency
16.7% +/- 4.0%
of residents report having English proficiency below "very well"
above average

Disabled
7.6% +/- 3.4%
of residents with one or more physical and/or mental disabilities
below average

Low-Income
31.8 % +/- 9.8%
of residents live in households with an income below 200% of the national poverty level
average



Source: ACS 2011-2015 5-Year Estimates

dvrpc

Equity Analysis for the Greater Philadelphia Region (Beta)

Home | About

below average

Ethnic Minority

34.0% +/- 8.0%

of residents identified themselves as being of Hispanic or Spanish origin.

well above average

Foreign Born

20.9% +/- 5.6%

of residents were born outside of the United States

above average

Limited English Proficiency

16.7% +/- 4.0%

of residents report having English proficiency below "very well"

above average

Disabled

7.6% +/- 3.4%

of residents with one or more physical and/or mental disabilities

below average

Low-Income

31.8 % +/- 9.8%

of residents live in households with an income below 200% of the national poverty level

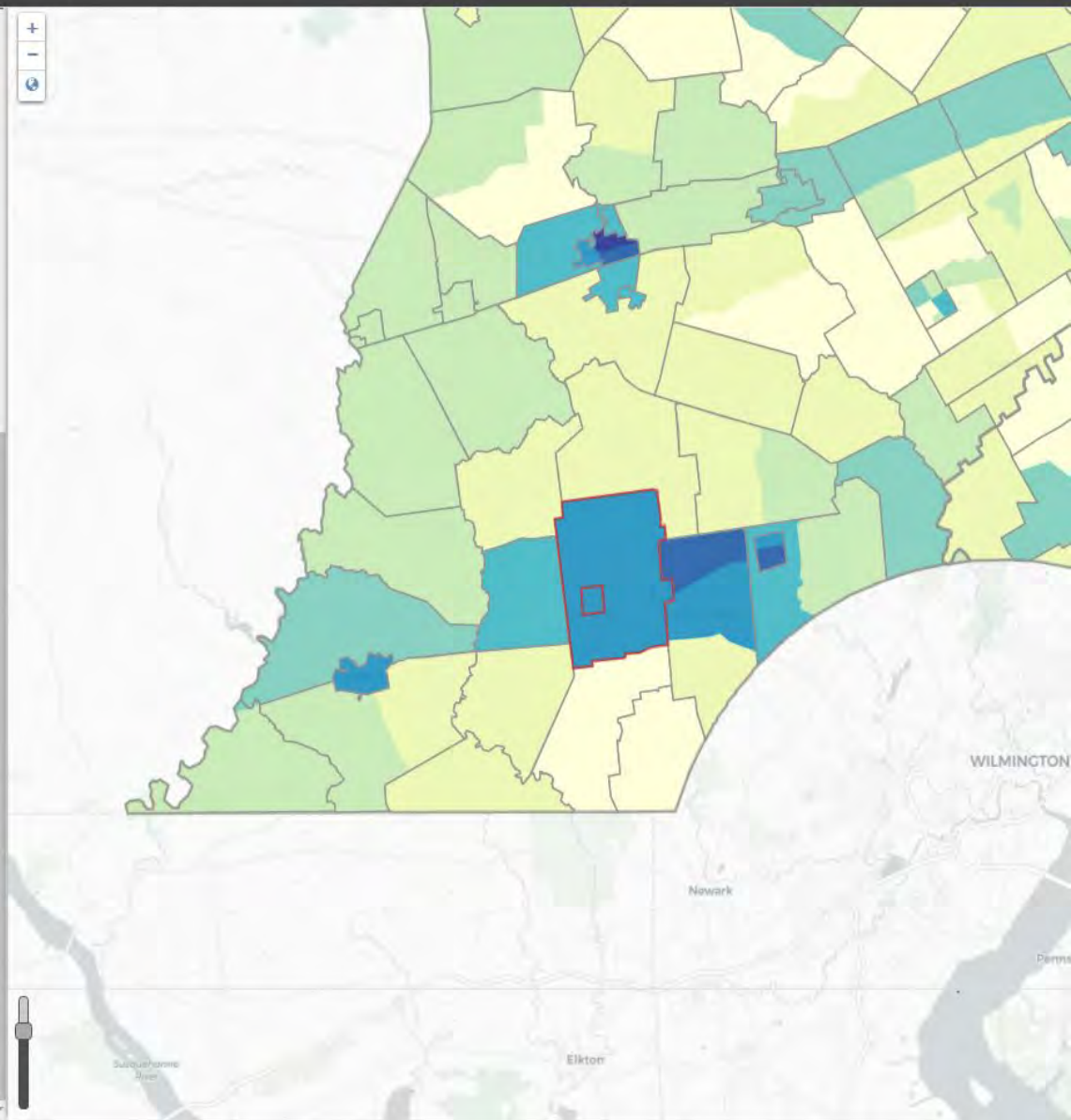
average

IPD Scoring Analysis

The information below provides tract-level IPD scores and percentages for our nine indicators. The entire set of indicators may be downloaded from our Open Data Portal for further analysis.

IPD Score

22



Source: ACS 2011-2015 5-Year Estimates

dvrpc

LESSONS LEARNED: BEYOND THE LEGISLATION

dvrpc

BEYOND THE LEGISLATION

Air quality

Carless household

Cost of living

Asthma rates

Access to jobs

Obesity rates

Purpose:
Identify and address adverse human health or environmental effects

Health insurance

Workplace and JTW

Crime

Auto crash rates

Vacant properties

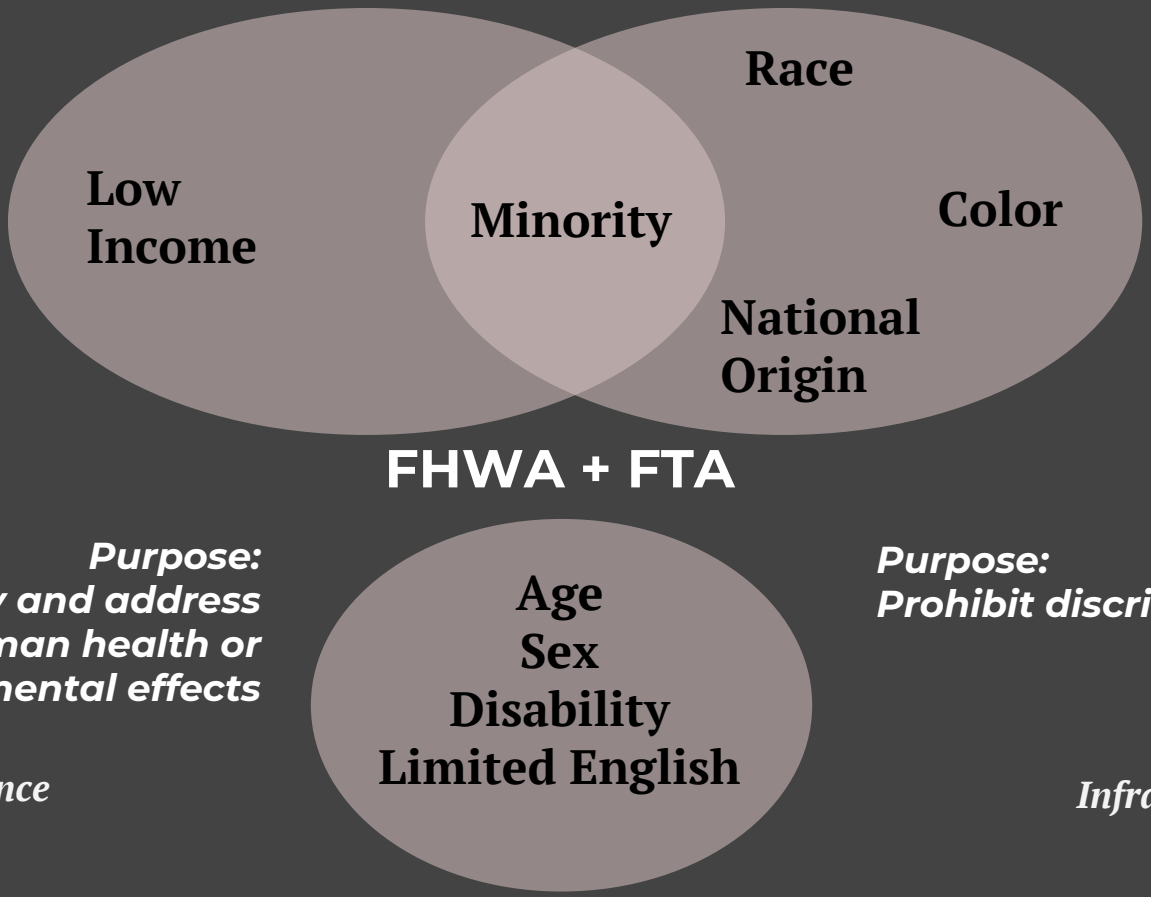
Bike lanes

Parks and open space

Commute length

Purpose:
Prohibit discrimination

Infrastructure quality



dvrpc

BEYOND THE LEGISLATION

Air quality

Carless household

Cost of living

Asthma rates

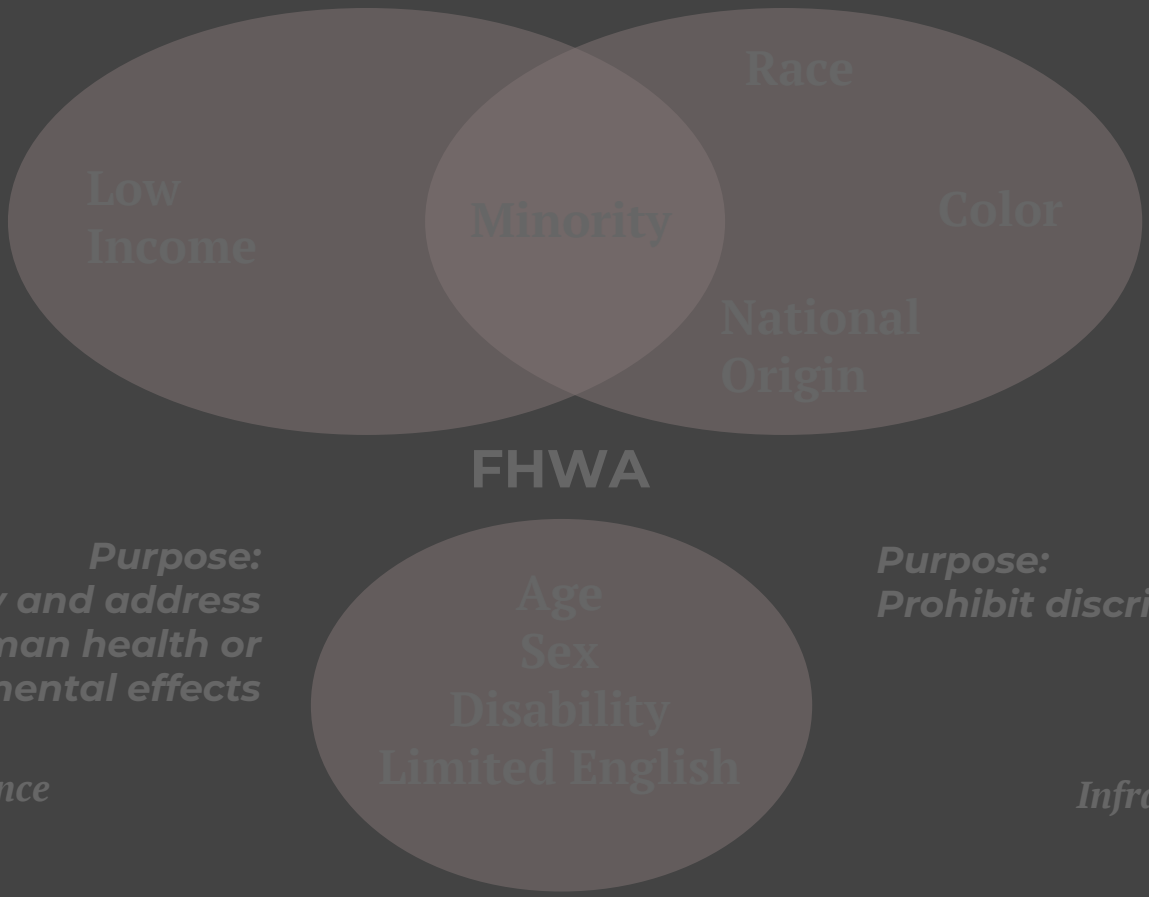
Access to jobs

Obesity rates

Purpose:
Identify and address adverse human health or environmental effects

Health insurance

Workplace and JTW



Crime

Auto crash rates

Vacant properties

Bike lanes

Parks and open space

Commute length

Purpose:
Prohibit discrimination

Infrastructure quality

dvrpc

IPD 2.1 EXPERIMENT #1

GO BEYOND RESIDENCE

dvrpc

CENSUS TRANSPORTATION PLANNING PRODUCTS (CTPP) EXPERIMENTATION

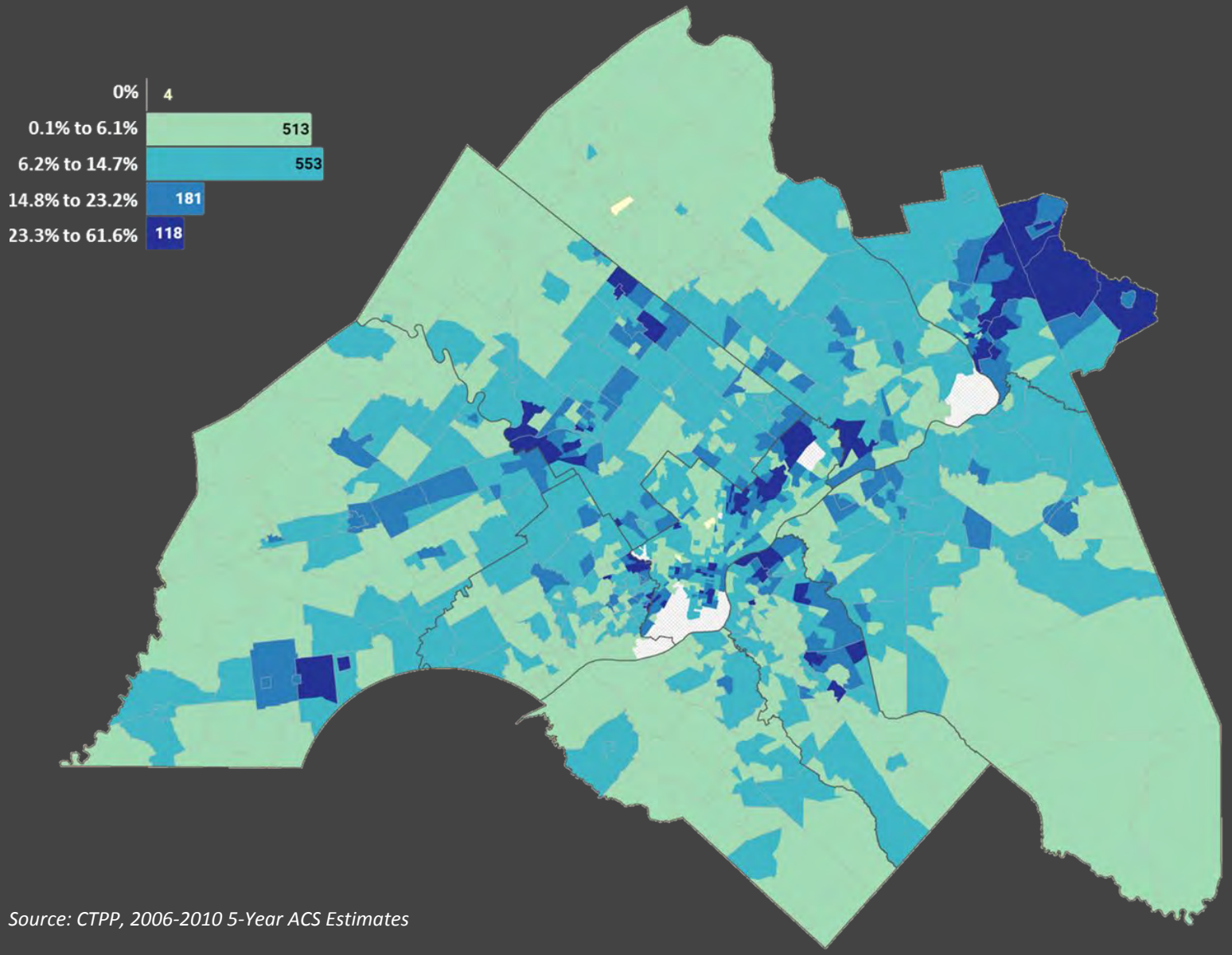
Workplace

- Minority (Race & Ethnicity)
- Foreign Born
- Limited English Proficiency
- Low Income (150% Poverty Rate)
- Carless Households

Journey-to-Work Flows

- Minority (Race & Ethnicity)
- Low Income (150% Poverty Rate)

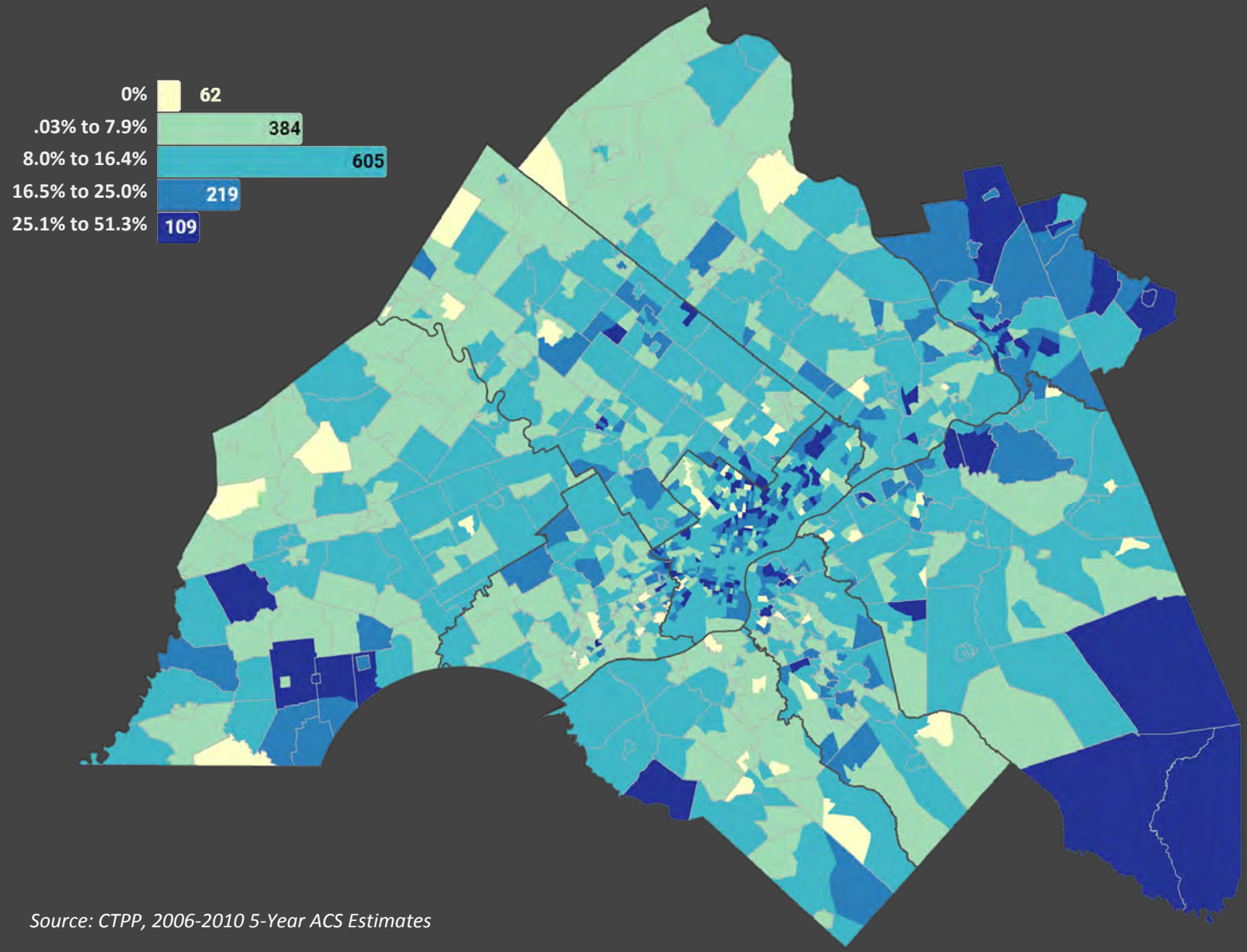
FOREIGN BORN - RESIDENCE



Source: CTPP, 2006-2010 5-Year ACS Estimates

dvrpc

FOREIGN BORN - WORKPLACE



Source: CTPP, 2006-2010 5-Year ACS Estimates

dvrpc

THANK YOU!

SHOSHANA AKINS
Public Participation Planner
sakins@dvrpc.org

KIM KOREJKO
Manager of Data Coordination
kkorejko@dvrpc.org

BEN GRUSWITZ
Senior Planner
bgruswitz@dvrpc.org

RESOURCES

DVRPC's IPD Interactive Map: www.dvrpc.org/webmaps/IPD/

FHWA Title VI guidance: https://www.fhwa.dot.gov/civilrights/programs/tvi.cfm

FTA EJ guidance:
https://www.transit.dot.gov/regulations-and-guidance/fta-circulars/environmental-justice-policy-guidance-federal-transit

CTPP Data: http://ctpp.transportation.org/Pages/5-Year-Data.aspx

List of CTPP EJ Tables: CTPP Status Report - April 2017
www.fhwa.dot.gov/planning/census_issues/ctpp/status_report/sr0417/index.cfm

dvrpc

EXTRA SLIDES

INDICATORS AND CENSUS TABLES

Indicator in IPD analysis update	ACS data table for indicator in IPD analysis	Protected class indicator represents
Youth	S0101: Age and Sex	Age
Older Adults	S0101: Age and Sex	Age
Female	S0101: Age and Sex	Sex
Racial Minority	B02001: Race	Race and Minority
Ethnic Minority	B03002: Hispanic or Latino Origin by Race	Minority and National Origin
Foreign Born	B05012: Nativity in the United States	National Origin
Limited English Proficiency	S1601: Language Spoken at Home	Limited English Proficiency, and National Origin
Disabled	S1810: Disability Characteristics	Disability
Low-Income	S1701: Poverty Status in the Past 12 Months	Low-Income

CTPP EXPERIMENTATION

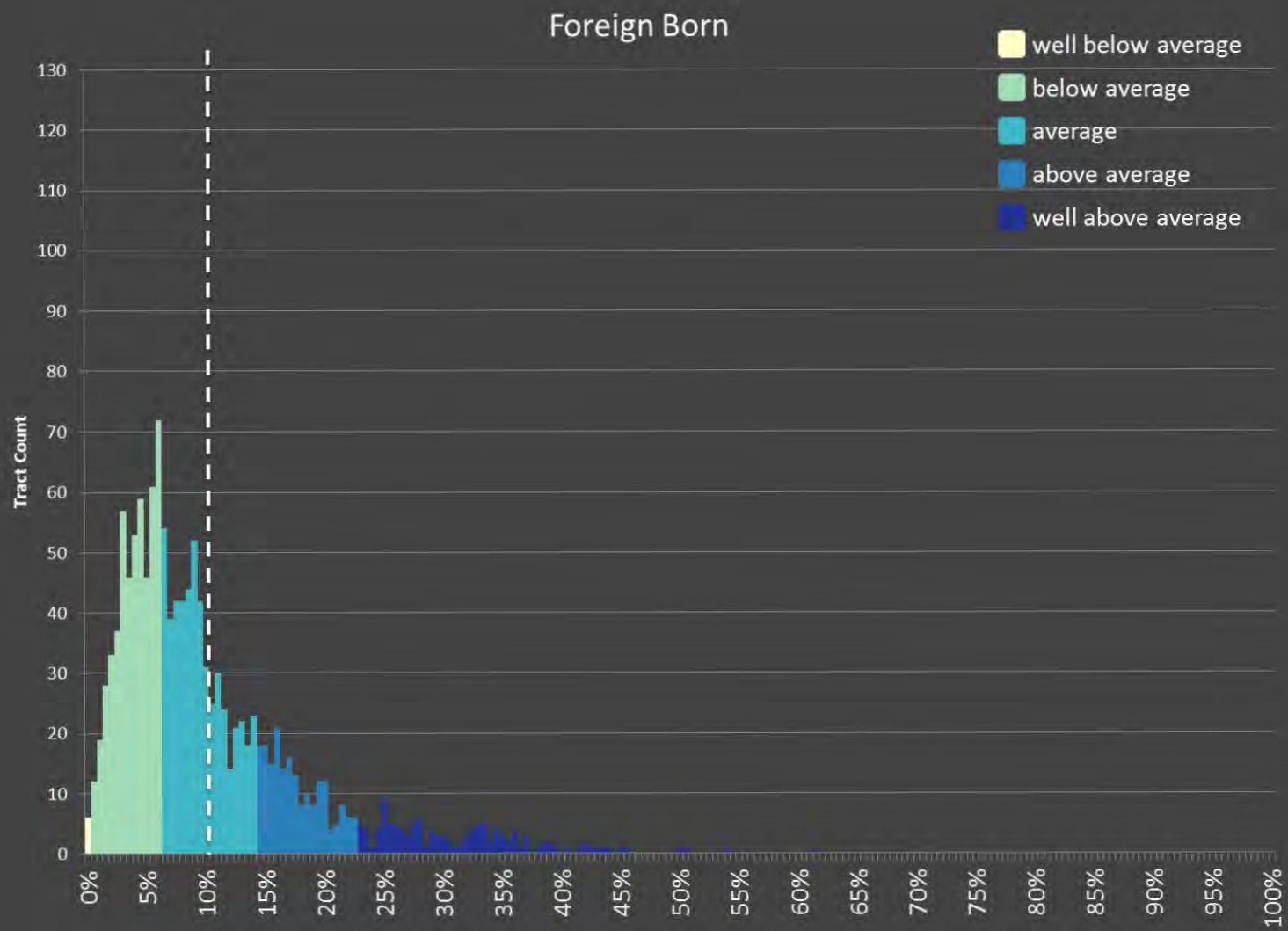
Workplace

- Minority (Race & Ethnicity)
- Foreign Born
- Limited English Proficiency
- Low Income (150% Poverty Rate)
- Carless Households

Journey-to-Work Flows

- Minority (Race & Ethnicity)
- Low Income (150% Poverty Rate)

FOREIGN BORN - RESIDENCE



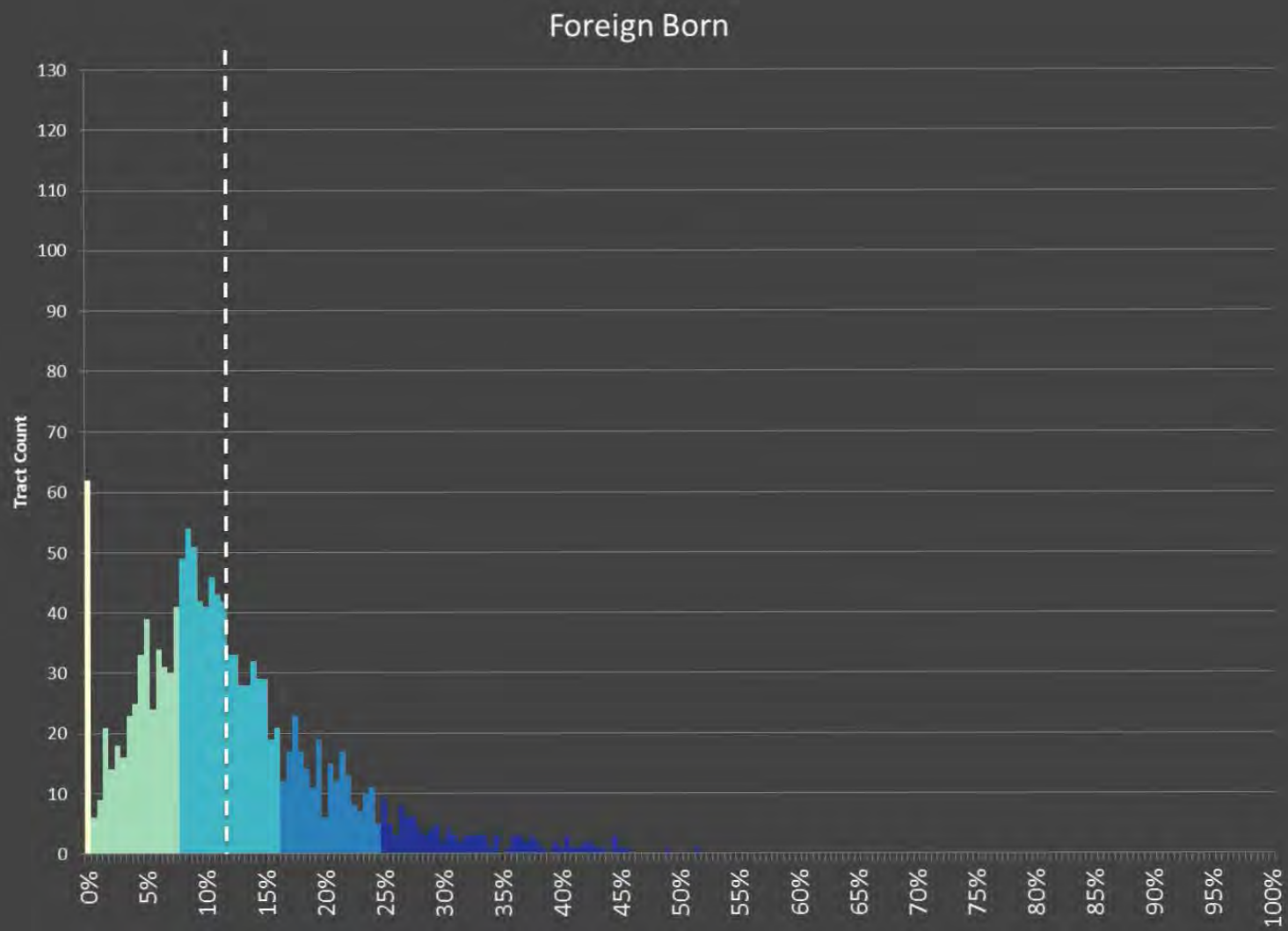
dvrpc

IPD 2.1 EXPERIMENT #1

GO BEYOND RESIDENCE

dvrpc

FOREIGN BORN - WORKPLACE



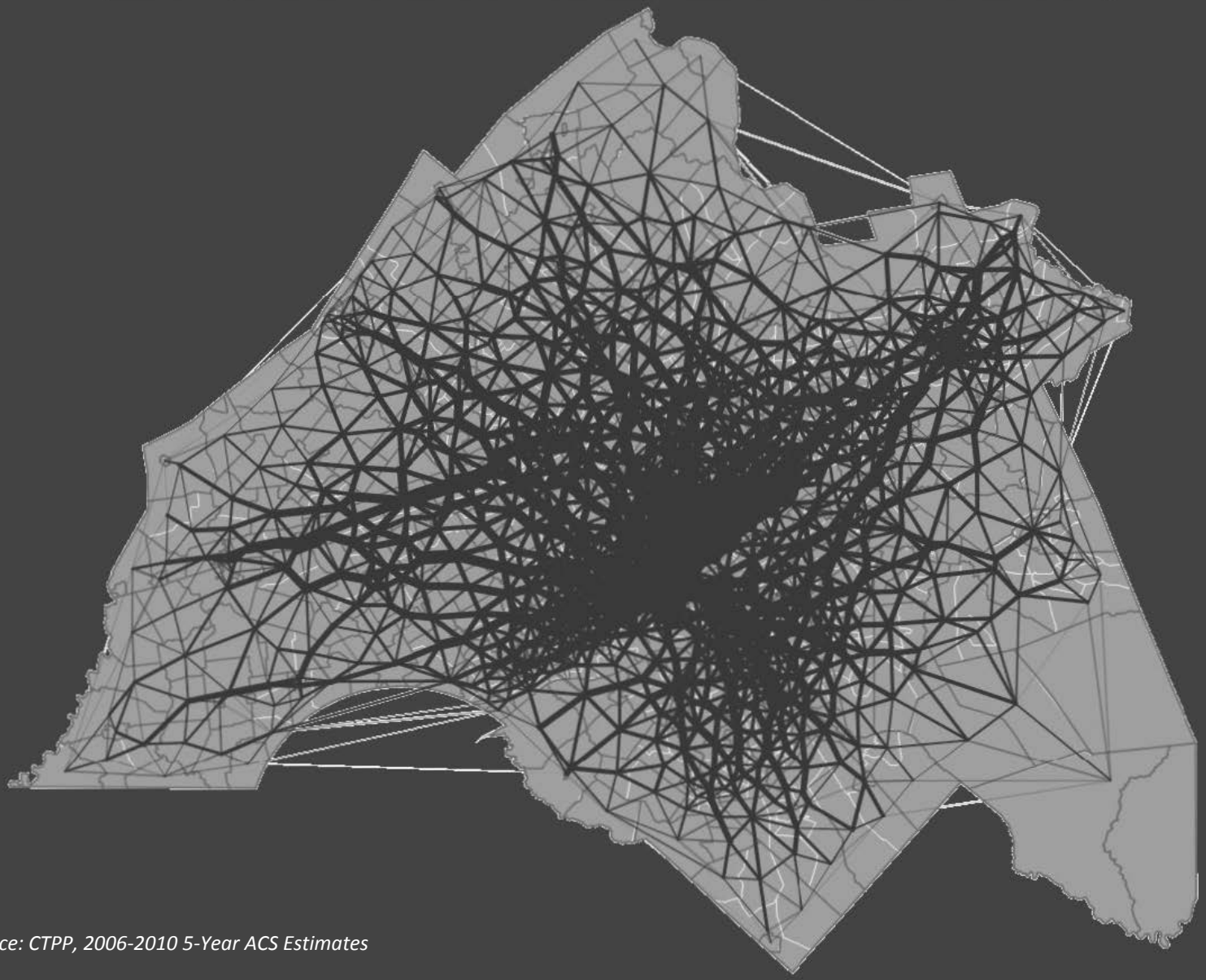
dvrpc

DESIRE LINES - TOTAL FLOW (20+ WORKERS)

Source: CTPP, 2006-2010 5-Year ACS Estimates

dvrpc

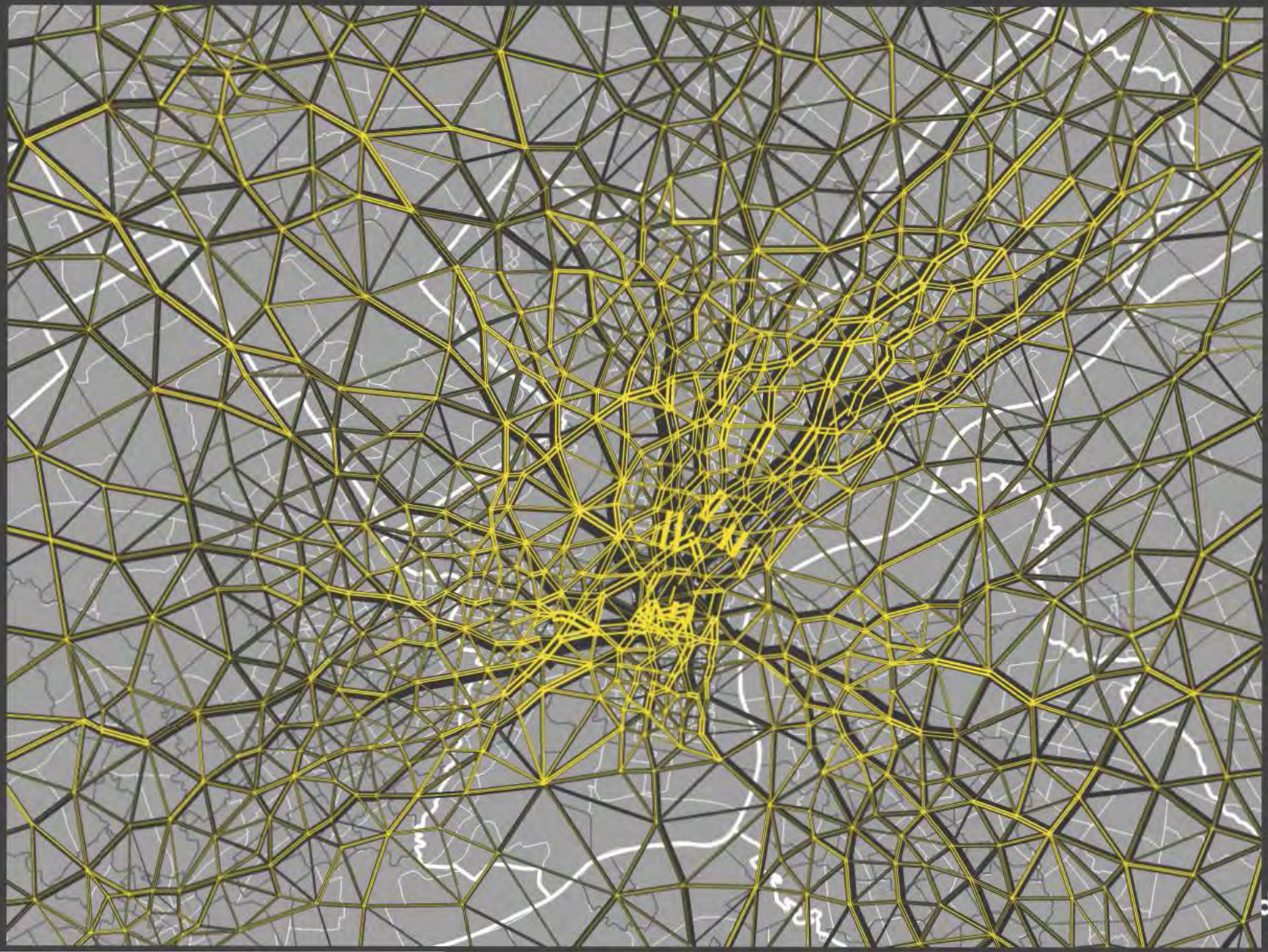
DELAUNAY LINES - TOTAL FLOW



Source: CTPP, 2006-2010 5-Year ACS Estimates

dvrpc

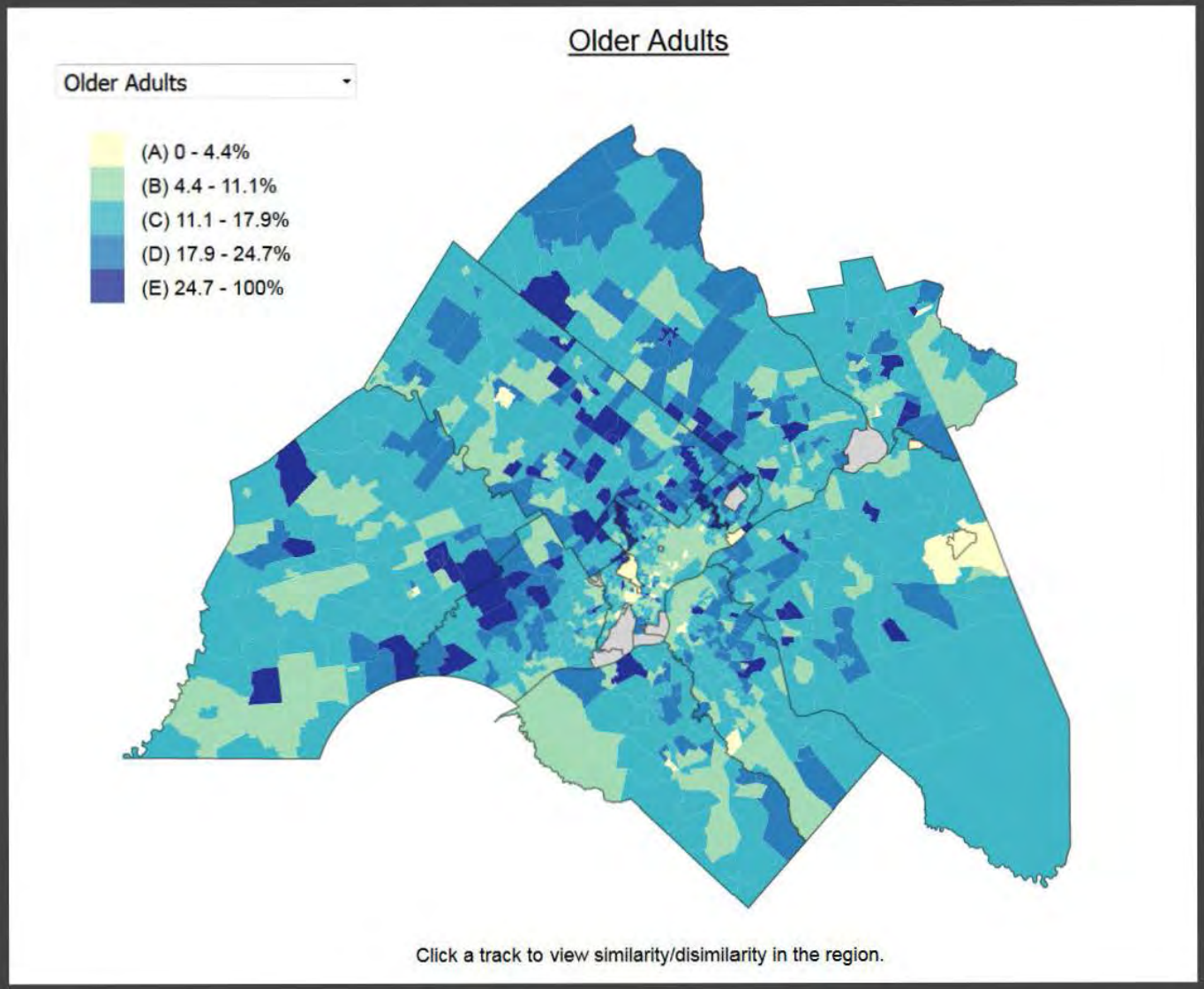
DELAUNAY LINES - LOW INCOME FLOW



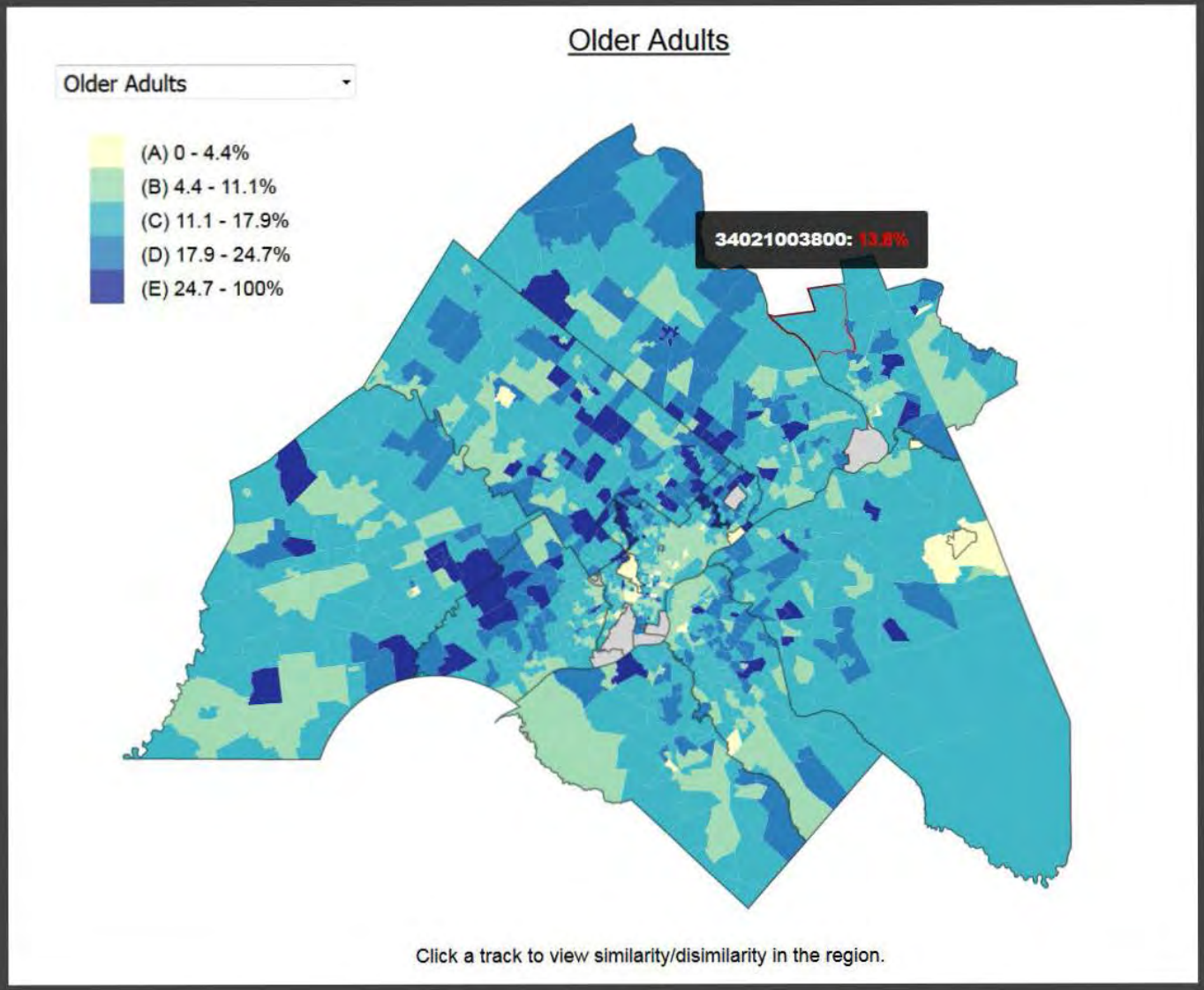
IPD 2.1 EXPERIMENT #2
COMMUNICATE STATISTICAL SIGNIFICANCE

dvrpc

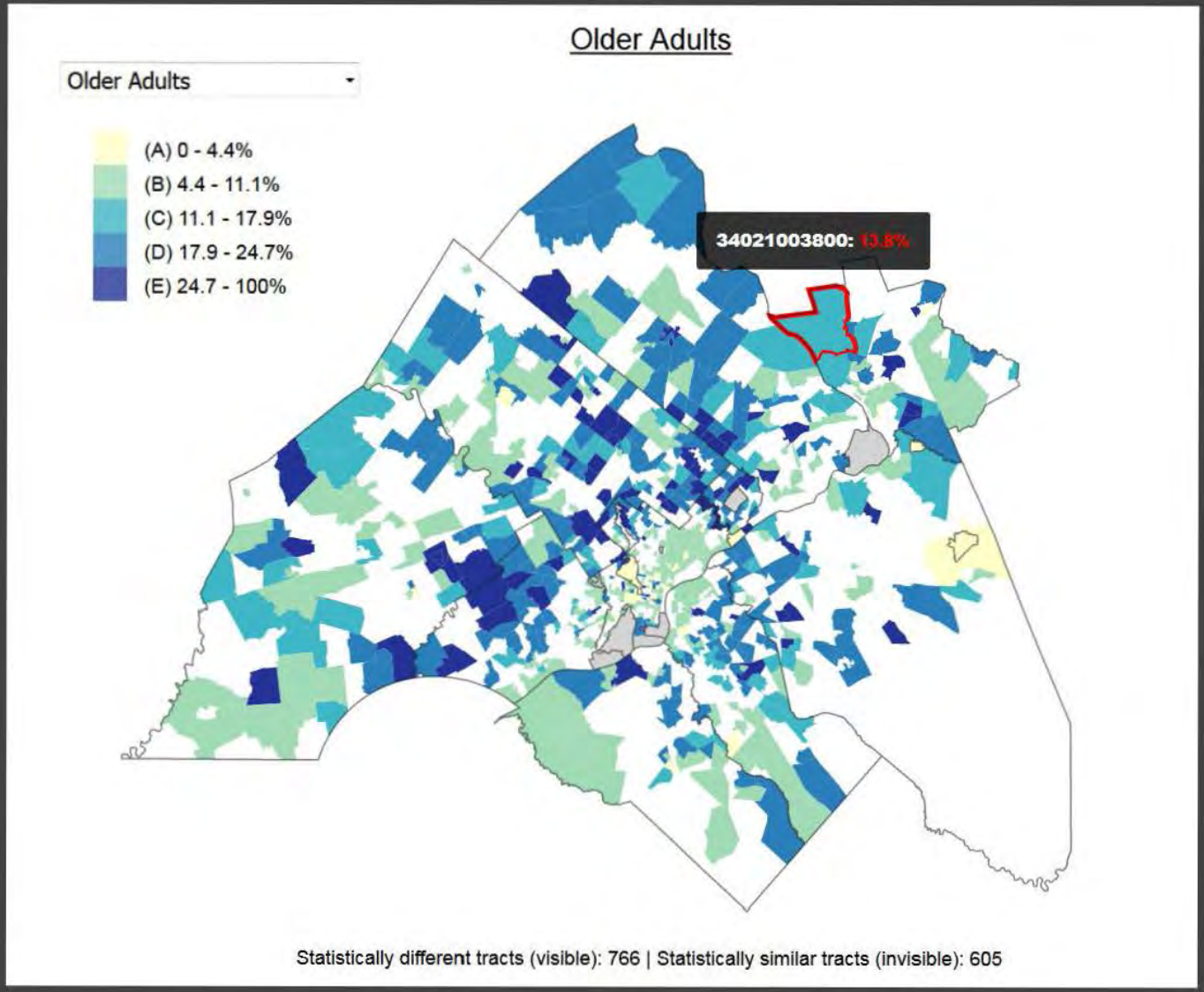
COMMUNICATE STATISTICAL SIGNIFICANCE



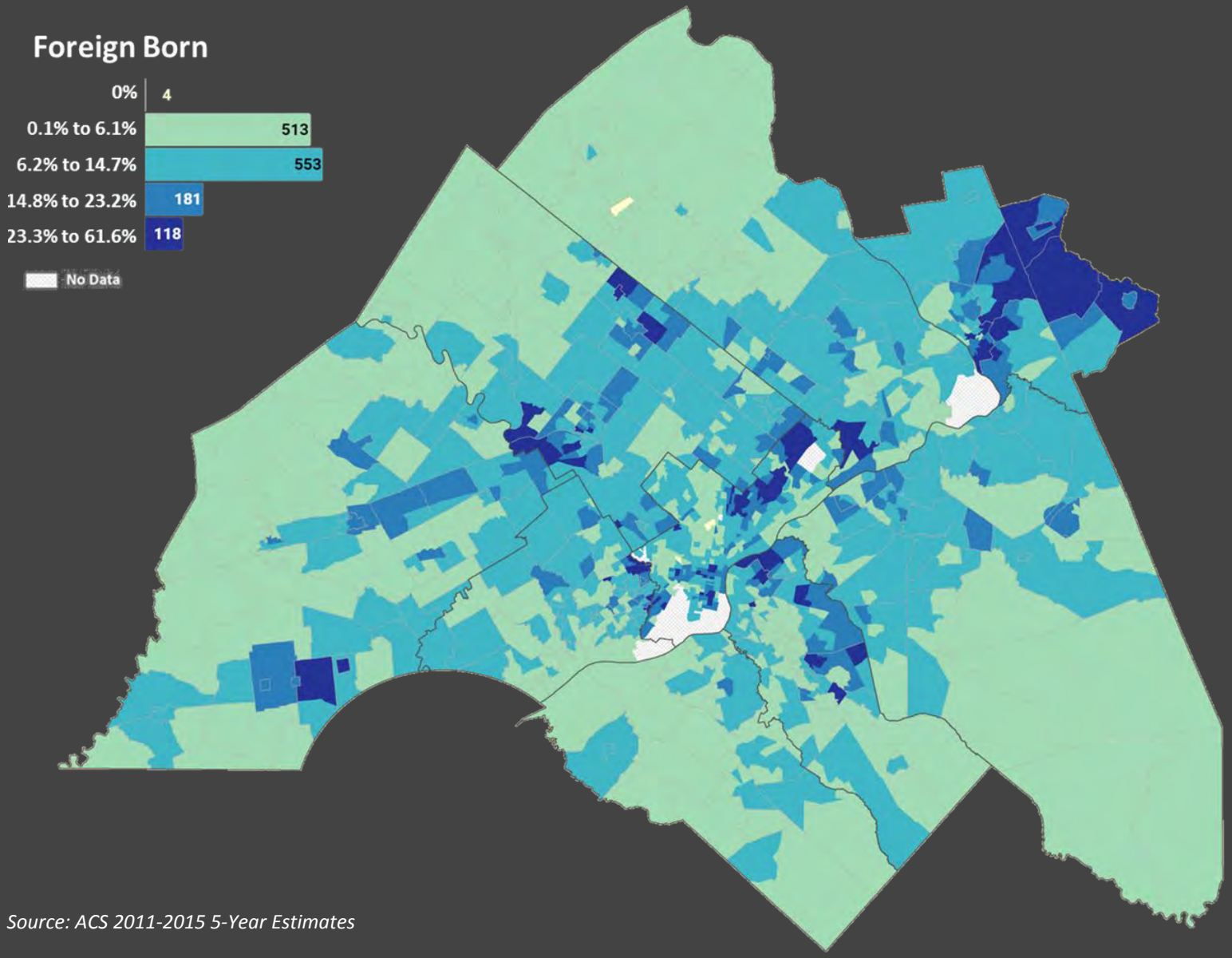
COMMUNICATE STATISTICAL SIGNIFICANCE



COMMUNICATE STATISTICAL SIGNIFICANCE



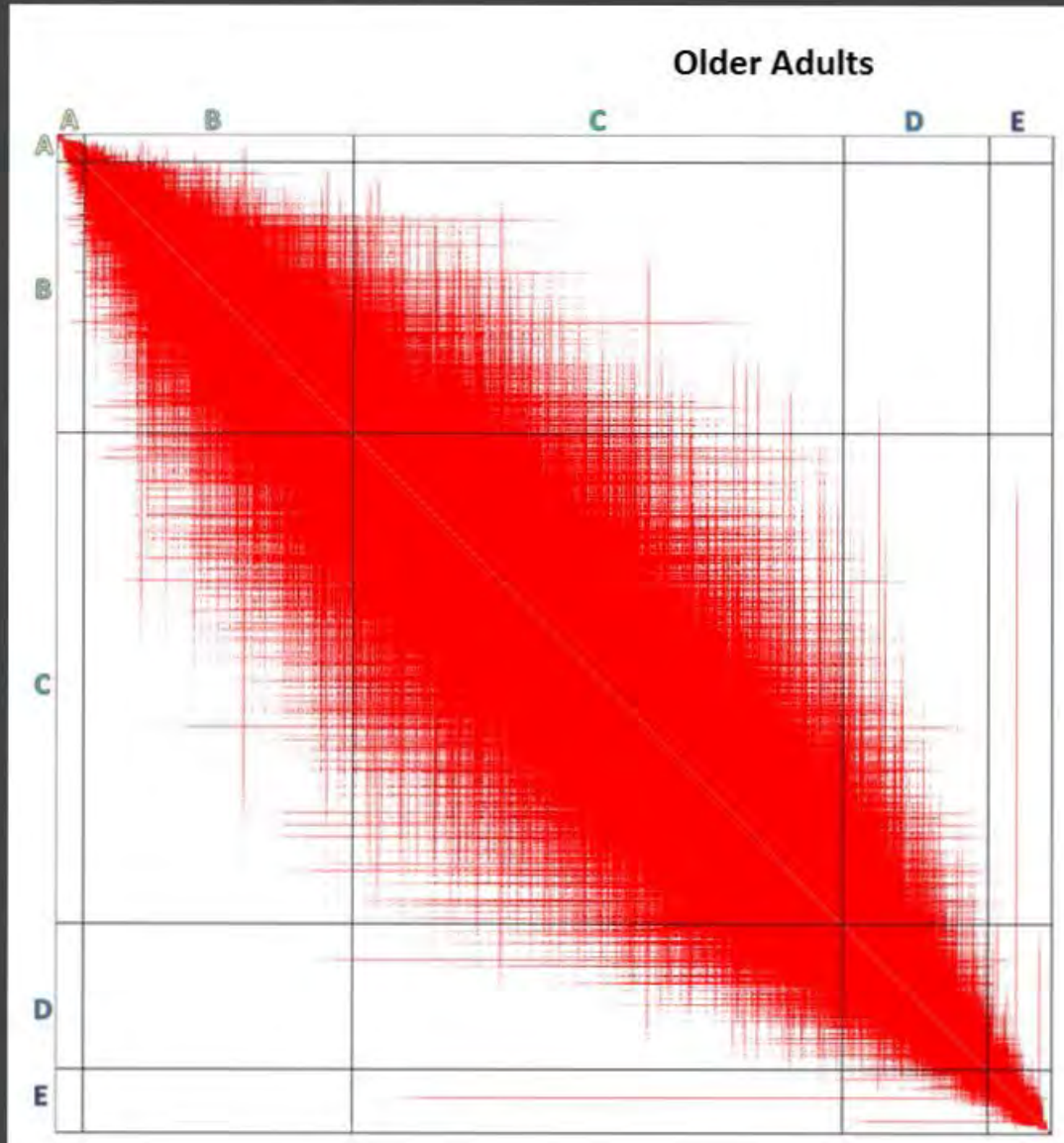
IPD 2.0 INDICATORS + METHODOLOGY MAPPED



Source: ACS 2011-2015 5-Year Estimates

dvrpc

COMMUNICATE STATISTICAL SIGNIFICANCE



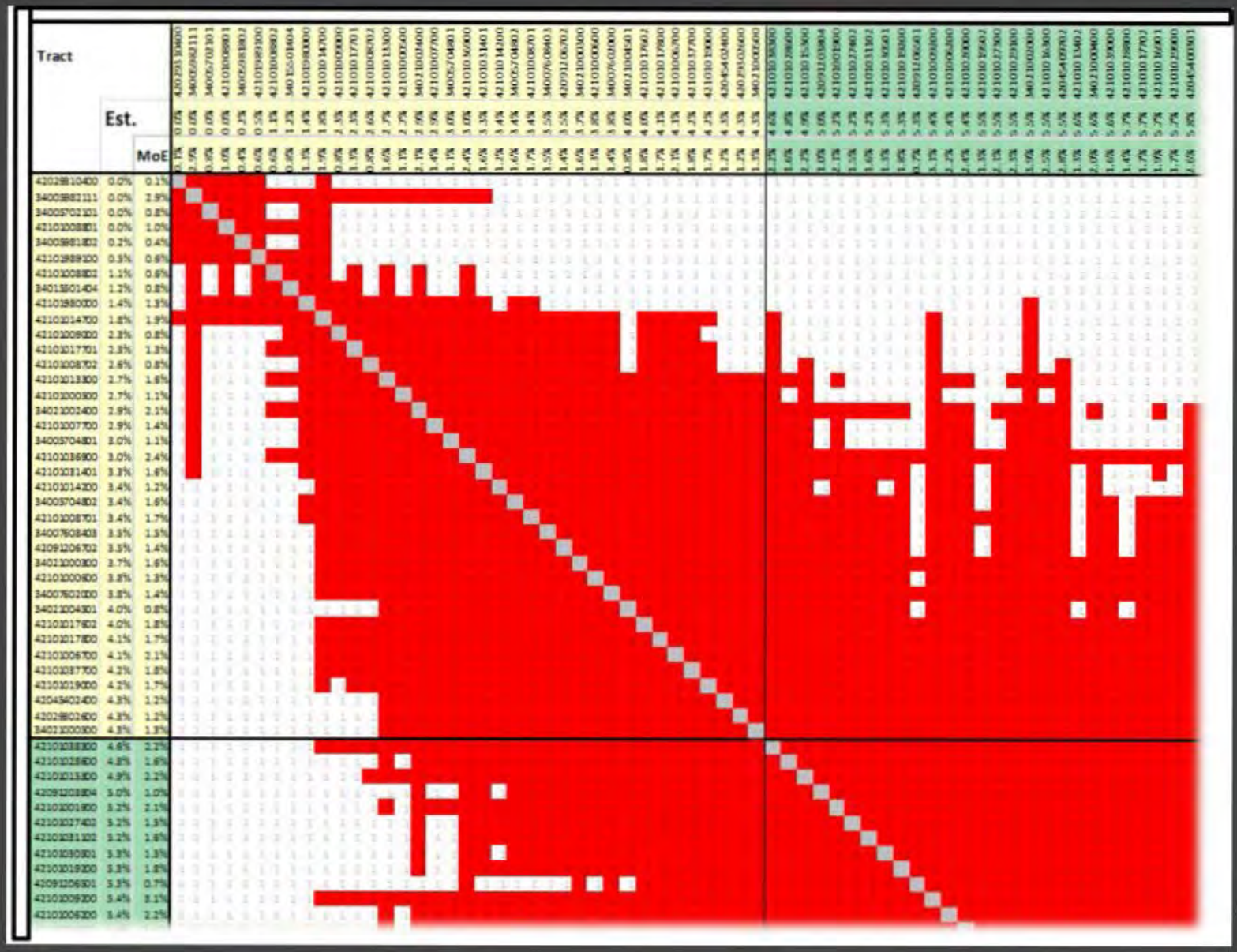
Percent Not Distinct

	A	B	C	D	E
A	65%	7%	0%	0%	0%
B	7%	76%	22%	0%	0%
C	0%	22%	78%	23%	1%
D	0%	0%	23%	88%	18%
E	0%	0%	1%	18%	61%

A – well below average
B – below average
C – average
D – above average
E – well above average

dvrpc

COMMUNICATE STATISTICAL SIGNIFICANCE



dvrpc

ALIGNING INDICATORS WITH LEGISLATION

Indicator (ACS 5-year estimates)	Executive Order 12898	Title VI of the Civil Rights Act of 1964	FHWA's Title VI and EJ documents	FTA's Title VI and EJ documents
Youth			✔	
Older Adults			✔	
Female			✔	
Racial Minority	✔	✔	✔	✔
Ethnic Minority	✔	✔	✔	✔
Foreign Born		✔	✔	✔
Limited English Proficiency		✔	✔	✔
Disabled			✔	
Low-income	✔		✔	✔

ALIGNING INDICATORS WITH LEGISLATION

Indicator (ACS 5-year estimates)	Executive Order 12898	Title VI of the Civil Rights Act of 1964	FHWA's Title VI and EJ documents	FTA's Title VI and EJ documents
Youth			✔	
Older			✔	
Female			✔	
Racial Minority	✔	✔	✔	✔
Ethnic Minority	✔	✔	✔	✔
Foreign-Born		✔	✔	✔
Limited English Proficiency		✔	✔	✔
Disabled			✔	
Low-income	✔		✔	✔
Carless Households	X	X	X	X
Female Head of Household	?	X	?	X