BETTER DATA FOR BETTER HEALTH
Enhancing Community Data in New York City

Massachusetts Health Data Consortium, October 9, 2019
Presented by Dan Benevento and Cambria Brown
AGENDA

1. Who are we and a little about HealthDataViz
2. Community health data in New York City: the problem and goals for the solution
3. The approach (tips and ideas you can apply to any data visualization project)
4. Demo of the new site
5. Q&A
WHO ARE WE?

Dan Benevento, AB
Principal, HealthDataViz

- Data Visualization Expert
- 9+ years Tableau and Tableau Server Experience
- 9+ years meaningfully communicating health and healthcare data
- Author of The Best Boring Book of Tableau for Healthcare
- Tableau Advanced Course Instructor

Cambria Brown, MPH
Consultant, HealthDataViz

- Data Visualization Expert
- 5+ years Tableau and Tableau Server Experience
- 10 years meaningfully communicating health and public health data
- Public health and biostatistics expert
- Tableau Beginner/Intermediate Course Instructor
We are data visualization experts who help health & healthcare organizations see the opportunities to improve:

CLINICALLY • FINANCIALLY • OPERATIONALLY

And raise public awareness about health & healthcare issues
WHY?

BECAUSE YOU CAN'T IMPROVE WHAT YOU CAN'T SEE
A COMMON PROBLEM

Health & healthcare is swimming in data, but it’s locked away.

Decision makers don’t see the data or what it means.
THAT’S WHERE **HEALTDATAVIZ** COMES IN

- Founded in 2013 by and for health and healthcare professionals
- Recognized experts in the science of how people see and understand data—visual intelligence
- Deep health and healthcare expertise
- Data visualization software expertise
- Lots of credentials: MBA, MD, MEd, MHA, MS, MPH, RPh, RN
- We help health and healthcare organizations see how they’re doing
ENHANCING COMMUNITY DATA IN NEW YORK CITY

THE WHAT

• NYC’s website for disseminating public health data
• Houses survey, vital records and disease surveillance data
• Includes Epiquery (Epidemiology + Query) and Community Health Profiles
ENHANCING COMMUNITY DATA IN NEW YORK CITY

THE PROBLEM

• The website was outdated, aesthetically unappealing, fragmented and hard to use

• Data users were unable to access the data they need to make decisions

• Health department staff were burdened with custom data requests
THE OLD EPIQUERY

EpiQuery: NYC Interactive Health Data

EpiQuery provides data on the health of New Yorkers from a variety of sources, including surveys, surveillance data, and vital records (births and deaths).

- To begin using EpiQuery: Select a module below. Click the "Information" link next to the module name to view detailed information about the data source.
- Not sure where to find the data you’re interested in? Select the EpiQuery Summary tab at the top of this page to view a brief summary of all EpiQuery modules, as well as a list of topics and the modules within EpiQuery where related data may be obtained.
- Tutorial and more information: Select the FAQ tab at the top of this page for frequently asked questions and a step-by-step guide to using EpiQuery.

Survey Modules

- NEW  Community Health Survey (Information)
  NYC CHS 2002-2017
- Child Health Data (Information)
  NYC Child Data 2009 and 2015
- NYC Health and Nutrition Examination Survey (Information)
  NYC NHANES 2004 and 2013-2014
- New York City Youth Risk Behavior Survey (Information)
  NYC YRBS 1997-2017 (odd years)
- World Trade Center Health Registry (Information)
  WTCHR Baseline Survey 2001-2004

Surveillance Modules

- Communicable Disease Surveillance Data (Information)
  CDSS 2000-2017
- HIV/AIDS Surveillance Data (Information)

Navigation by data source only – only those familiar with the survey content know where to go to get data on topics of interest
THE OLD EPIQUERY

Data presented in piecemeal fashion, providing a fragmented view of New Yorkers’ health needs.

Community Health Survey 2017

Please select a single year or trend to view a list of available survey measures. Then select a measure and click SUBMIT at the bottom of the page.

Year: 2017 ▼  Trend (All available years)

Chronic Conditions
- Diabetes ever
- High blood pressure ever
- Blood pressure medication ever
- Current blood pressure medication
- Overweight and obesity

Cigarette Smoking
- Smoking status
- Type of smoker
- Smoke electronic cigarette
- Second-hand smoke from outside

Alcohol
- Alcohol drinker
- Binge drinking
- Heavy drinking

Physical Activity
- Exercise in past 30 days
- Rides bicycle in past year

The Community Health Survey (CHS) includes self-reported data from adults. CHS has included adults with landline phones since 2002 and, starting in 2009, also has included adults starting in 2011. CHS weighting methods were updated to use Census 2010 and additional demographic information.

Copyright HealthDataViz, LLC 2019
No part of this presentation may be replicated without permission. All information is confidential.
THE OLD EPIQUERY

Controls and results on different web pages results in frustrating back clicks.

[Chart showing smoking status by sex, with details on smoking status and prevalence]

[Table showing smoking status by sex and age-adjusted prevalence]
THE OLD EPIQUERY

Outdated user interface and visuals

Smoking status by Sex and age group, 2017 (Unadjusted for age)

Smoking status: Smoking status is defined as being a current or former smoker or never smoker (having smoked less than 100 cigarettes ever).
THE OLD EPIQUERY

HIV/AIDS Surveillance Data

New York City

Year: ● 2015 ● Trend (available for the most recent 5 years of data)

HIV/AIDS Outcome Measure
- Diagnoses
  - HIV
  - AIDS
- People Living with HIV/AIDS
- Deaths among people with HIV/AIDS

Geographic Unit
- Citywide
- Borough
- High-Risk Neighborhoods
  - Neighborhood Health Action Center areas
- Area-Based Poverty
- Neighborhood Map
  - All 42 NYC neighborhoods with maps by 42 United Hospital Fund areas when no additional factors selected

Demographic/Risk Factors
- Gender (data 2011 and later)
- Sex (data 2010 and earlier)
- Race/Ethnicity
- Age Group
- HIV Transmission Risk
- Place of Birth

Submit

For additional reports and information, visit New York City HIV/AIDS Surveillance Data

Confusing interactivity and instructions

HIV diagnoses by Neighborhood, New York City, 2015

Rate per 100,000 population

Copyright HealthDataViz, LLC 2019
No part of this presentation may be replicated without permission. All information is confidential.
THE OLD COMMUNITY HEALTH PROFILES

New York City Community Health Profiles

The New York City Community Health Profiles capture the health of 59 community districts across the city. They contain over 50 measures of neighborhood health, emphasizing that our health starts where we live, work and play. These reports highlight the disparities among neighborhoods and can be used by policymakers, community groups, health professionals, researchers and residents to encourage community engagement and action.

Profiles by Borough

Browse through the New York City Community Health Profiles currently available:

- Bronx
- Brooklyn
- Manhattan
- Queens
- Staten Island

All the below Community Health Profiles publications are in PDF format. You need to have Adobe Reader to read PDF files.

For more data on these community health profiles, check out Additional Resources.

Bronx

- Mott Haven and Melrose (PDF)
- Hunts Point and Longwood (PDF)
- Morrisania and Crotona (PDF)
THE OLD COMMUNITY HEALTH PROFILES

Who We Are

POPULATION BY RACE AND ETHNICITY*%

<table>
<thead>
<tr>
<th></th>
<th>New York City</th>
<th>Mott Haven and Melrose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asian</td>
<td>1%</td>
<td>1%</td>
</tr>
<tr>
<td>Black</td>
<td>22%</td>
<td>21%</td>
</tr>
<tr>
<td>Latino</td>
<td>23%</td>
<td>23%</td>
</tr>
<tr>
<td>White</td>
<td>12%</td>
<td>12%</td>
</tr>
<tr>
<td>Other</td>
<td>4%</td>
<td>4%</td>
</tr>
</tbody>
</table>

TOTAL POPULATION

8,537,673

98,403

POPULATION BY AGE

<table>
<thead>
<tr>
<th></th>
<th>New York City</th>
<th>Mott Haven and Melrose</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-4</td>
<td>21%</td>
<td>21%</td>
</tr>
<tr>
<td>5-9</td>
<td>8%</td>
<td>8%</td>
</tr>
<tr>
<td>10-14</td>
<td>11%</td>
<td>10%</td>
</tr>
<tr>
<td>15-24</td>
<td>16%</td>
<td>15%</td>
</tr>
<tr>
<td>25-44</td>
<td>23%</td>
<td>23%</td>
</tr>
<tr>
<td>45-64</td>
<td>23%</td>
<td>23%</td>
</tr>
<tr>
<td>65+</td>
<td>22%</td>
<td>22%</td>
</tr>
</tbody>
</table>

BORN OUTSIDE THE US

37%

30%

HAVE LIMITED ENGLISH PROFICIENCY

23%

33%

Social and Economic Conditions

Economic stress
Living in high-poverty neighborhoods limits healthy options and makes it difficult to access quality health care and resources that promote health. In Mott Haven and Melrose, 29% of residents live in poverty, compared with 20% of NYC residents. Access to affordable housing and employment opportunities with fair wages and benefits are also closely associated with good health. Mott Haven and Melrose’s unemployment rate is higher than the citywide average of 9%. Rent burdened households pay more than 30% of their income for housing and may have difficulty affording food, clothing, transportation and health care. Fifty-eight percent of Mott Haven and Melrose residents are rent burdened, a higher rate than residents citywide. One way to consider the effect of income on health is by comparing death rates among neighborhoods. "Avoidable deaths" are those that could have been avoided if each neighborhood had the same death rate as the five wealthiest neighborhoods. Using this measure, 40% of deaths could have been averted in Mott Haven and Melrose.

ECONOMIC STRESS

<table>
<thead>
<tr>
<th></th>
<th>Mott Haven and Melrose</th>
<th>Bronx</th>
<th>NYC</th>
<th>Lowest %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poverty</td>
<td>29%</td>
<td>25%</td>
<td>20%</td>
<td>7%</td>
</tr>
<tr>
<td></td>
<td>(percent of residents)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unemployment (percent of people ages 16 and older)</td>
<td>12%</td>
<td>13%</td>
<td>9%</td>
<td>5%</td>
</tr>
<tr>
<td>Rent Burden (percent of renter-occupied households)</td>
<td>58%</td>
<td>58%</td>
<td>51%</td>
<td>37%</td>
</tr>
</tbody>
</table>

Note: Unemployment data may differ from rates presented in other published sources. See technical notes at the Public Use Data disk for more details.

Is your neighborhood gentrifying?
Gentrification transforms a low-income area into a high-income area through neighborhood redevelopment. It is often defined as changes in the racial and ethnic makeup, education level and average income of a neighborhood’s residents, as well as changes in housing and commercial businesses. While development may be beneficial, it is often inequitable, and can lead to displacement of long-time residents and businesses. Gentrification can be measured in many ways. One measure that is used in NYC is to determine if a low-income neighborhood (those with the lowest 40% of average household income in 1990) saw higher than median rent growth over the past 20 years. Based on this definition, 24 neighborhoods were considered low income, and Mott Haven and Melrose is one of 17 neighborhoods that is gentrifying.

Source: NYU Furman Center, 2015

Copyright HealthDataViz
No part of this presentation may be replicated without permission. All information is confidential.
ENHANCING COMMUNITY DATA IN NEW YORK CITY

THE GOAL

• Make the website intuitive, informative and beautiful

• Tailor reports to different audiences

• Enable researchers, policy makers, program developers and the general public to easily access data that tells a story and get the information they need to make decisions and improve health

• Decrease the burden on health department staff for maintaining and updating the website, and responding to custom requests
ENHANCING COMMUNITY DATA IN NEW YORK CITY

THE PROCESS

1. Persona development
2. Guided analytics
3. Best practices in data visualization
4. Designing for the modern user
STEP 1: PERSONA DEVELOPMENT

A technique that employs fictitious users (based on research and exploration of very real users’ needs) to guide decision making regarding the design of products, services, communication, presentations

- Steers us away from self-referential thinking and moves us to a point of view more aligned with our audiences’
- Should be based on research and interviews, not assumptions or imagination
- Be able to list real, tangible goals and problems to solve, and how the data presentation will help move people toward solutions
- Good to leverage existing mental models in our data presentations and interactivity
STEP 1: PERSONA DEVELOPMENT

- Epiquery User Personas
  1. Community Based Organizers
  2. Researchers at the Health Department
  3. Researchers external to the Health Department
  4. Educators, Parents and the General Public
  5. Press/Journalists
  6. Public Relations Liaison at the Health Department
  7. Health Providers/Insurance Companies/Health Navigators
STEP 1: PERSONA DEVELOPMENT

Nancy – Community Based Organizer
Level of education: College/Masters MPA, MSW, MPH

Scope of Responsibility: Needs assessment, planning, identifying target populations, grant writing, presentation development, program planning and evaluation, resource allocation

Need: Summary statistics and to use data as a management indicator

Familiarity and Experience with Data and Statistics: Has experience using summary statistics only

Frustrations: Wants to see data by zip codes and catchment areas, and use it for evaluation purposes but couldn’t (with old website)
STEP 2: GUIDED ANALYTICS

THE FRAMEWORK

Summary Overview

Focused Report
  Detailed Lists
  Analytic Tools

Focused Report
  Detailed Lists

Focused Report
  Detailed Lists
STEP 2: GUIDED ANALYTICS

Views tailored for different audiences and purposes

List of Available Indicators/Data

Indicator Overview

Analyze by Neighborhood

Analyze by Demographics

Compare Two Indicators

Test Statistical Significance

More complex for more advanced users
STEP 3: BEST PRACTICES IN DATA VISUALIZATION

1. Bar charts are NOT boring

   – Bars are a pillar of data visualization
   – Humans are great at distinguishing differences in length
     • We are very bad at determining differences in areas or angles, unless the differences are large
   – Great for comparing values for discrete categories of data
   – Can order data such as by rank, or see distribution of the data
   – Can be horizontal or vertical (unlike its labels which must be horizontal)
STEP 3: BEST PRACTICES IN DATA VISUALIZATION

Percentage Cases by Country

- Spain
- Italy
- Sweden
- Switzerland
- Norway
- France
- Germany
- UK
STEP 3: BEST PRACTICES IN DATA VISUALIZATION

Percent Cases by Country

- Germany: 32.0%
- UK: 23.0%
- France: 12.0%
- Norway: 11.5%
- Switzerland: 10.5%
- Sweden: 6.0%
- Italy: 5.0%
- Spain: 2.0%
STEP 3: BEST PRACTICES IN DATA VISUALIZATION

2. Labels and colors are NOT arbitrary
   – The indiscriminate use of color can be confusing and at best adds nothing to your visualization
   – Absolutely use color to impart meaning, but choose colors wisely
     • Don’t overwhelm
     • Colors can have powerful emotional associations
     • 7% of males and 0.5% of females are colorblind
   – Clear legible labels go a long way to making data meaningful
     • Acronyms and vertical text are often unnecessary impediments to understanding the data
3. Chart junk and data-ink-ratio

Edward Tufte invented the concept of Chart Junk (the excessive and unnecessary use of graphical effects in graphs) and Data-Ink-Ratio

• Data Ink is the ink on a graph that represents data.
• Tufte claims that good graphical representations maximize data-ink and minimize non-data-ink
• 5 principles:
  1. Above all else show data
  2. Maximize the data-ink ratio
  3. Erase non-data-ink
  4. Erase redundant data-ink
  5. Revise and edit
STEP 3: BEST PRACTICES IN DATA VISUALIZATION

Chart junk

Excess non-data ink

<table>
<thead>
<tr>
<th>Priority Area</th>
<th>CY18</th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
<th>Q4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Falls</td>
<td>9.4%</td>
<td>8.46%</td>
<td>5.41%</td>
<td>6.6%</td>
<td></td>
</tr>
<tr>
<td>Serious Reportable Events</td>
<td>0.4%</td>
<td>0.4%</td>
<td>0.4%</td>
<td>0.4%</td>
<td></td>
</tr>
<tr>
<td>Medical Records</td>
<td>7.2%</td>
<td>6.6%</td>
<td>6.5%</td>
<td>6.7%</td>
<td></td>
</tr>
<tr>
<td>MD Notes Composite</td>
<td>100%</td>
<td>90%</td>
<td>50%</td>
<td>50%</td>
<td></td>
</tr>
<tr>
<td>H&amp;P Compliance</td>
<td>60%</td>
<td>60%</td>
<td>60%</td>
<td>60%</td>
<td></td>
</tr>
<tr>
<td>Medication Management/Inpatient</td>
<td>63%</td>
<td>63%</td>
<td>63%</td>
<td>63%</td>
<td></td>
</tr>
<tr>
<td>Medication Secured Proper</td>
<td>63%</td>
<td>63%</td>
<td>63%</td>
<td>63%</td>
<td></td>
</tr>
<tr>
<td>Expired Medic (Days)</td>
<td>60%</td>
<td>60%</td>
<td>60%</td>
<td>60%</td>
<td></td>
</tr>
<tr>
<td>Patients’ Own Media/Labeling</td>
<td>60%</td>
<td>60%</td>
<td>60%</td>
<td>60%</td>
<td></td>
</tr>
<tr>
<td>Recording/Reporting Fridge Temp</td>
<td>61%</td>
<td>61%</td>
<td>61%</td>
<td>61%</td>
<td></td>
</tr>
<tr>
<td>Patient Education on AntiCoag</td>
<td>60%</td>
<td>60%</td>
<td>60%</td>
<td>60%</td>
<td></td>
</tr>
<tr>
<td>Med Rec Admission</td>
<td>60%</td>
<td>60%</td>
<td>60%</td>
<td>60%</td>
<td></td>
</tr>
<tr>
<td>Med Rec-Discharge</td>
<td>60%</td>
<td>60%</td>
<td>60%</td>
<td>60%</td>
<td></td>
</tr>
</tbody>
</table>

Percent of readmissions by primary payer

- Medicare Traditional/Indemnity: 44%
- Medicare/Managed Care: 13%
- Medicaid Traditional/Indemnity: 11%
- Commercial/Private PPO: 10%
- Medicaid/Managed Care: 9%
STEP 3: BEST PRACTICES IN DATA VISUALIZATION

Top 10 Highest Grossing Music Tours of all Time, Adjusted for Inflation

Artist Tour

<table>
<thead>
<tr>
<th>Artist/Organization</th>
<th>Tour Description</th>
<th>Gross (adjusted for inflation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC/DC</td>
<td>Black Ice World Tour</td>
<td>$506,821,238.00</td>
</tr>
<tr>
<td>Coldplay</td>
<td>A Head Full of Dreams Tour</td>
<td>$546,024,326.00</td>
</tr>
<tr>
<td>Ed Sheeran</td>
<td>+ Tour</td>
<td>$775,046,937.00</td>
</tr>
<tr>
<td>Guns N’ Roses</td>
<td>Not In This Lifetime Tour</td>
<td>$553,360,000.00</td>
</tr>
<tr>
<td>Madonna</td>
<td>Sticky &amp; Sweet Tour</td>
<td>$476,674,691.00</td>
</tr>
<tr>
<td>Roger Waters</td>
<td>The Wall Live</td>
<td>$493,364,420.00</td>
</tr>
<tr>
<td>The Rolling Stones</td>
<td>Licks Tour</td>
<td>$423,768,493.00</td>
</tr>
<tr>
<td>U2</td>
<td>360° Tour</td>
<td>$525,158,550.00</td>
</tr>
<tr>
<td>U2</td>
<td>Vertigo Tour</td>
<td>$820,194,886.00</td>
</tr>
<tr>
<td>U2</td>
<td>360° Tour</td>
<td>$820,194,886.00</td>
</tr>
<tr>
<td>U2</td>
<td>Vertigo Tour</td>
<td>$483,351,954.00</td>
</tr>
</tbody>
</table>
**STEP 3: BEST PRACTICES IN DATA VISUALIZATION**

U2 and The Rolling Stones had 4 of the top 10 highest grossing music tours of all time (adjusted for inflation)

- **U2: 360°** - $820M
- **Ed Sheeran:** - $775M
- **Guns N’ Roses: Not in This Lifetime...** - $563M
- **Coldplay: A Head Full of Dreams** - $546M
- **The Rolling Stones: Voodoo Lounge** - $526M
- **AC/DC: Black Ice World** - $507M
- **Roger Waters: The Wall Live** - $493M
- **U2: Vertigo** - $484M
- **Madonna: Sticky & Sweet** - $476M
- **The Rolling Stones: Licks** - $424M
ENHANCING COMMUNITY DATA IN NEW YORK CITY

STEP 4: DESIGNING FOR THE MODERN USER

• Adapts for use on different devices (desktop, tablet, mobile)
• Seamless links to related resources – integration between data and website
• Ability to share insights with colleagues and on social media (deep linking)
• Accessibility/WCAG compliance