Transforming Geospatial Data for Visualization with D3

FOSS4G — Boston
August 17, 2017

Beatrice Jin and Benjamin Krepp

Boston Region Metropolitan Planning Organization

Agenda

- Who we are
- Project context
- Implementation vehicle
- Overview of dashboard data
- Presentation of linear data
- Conclusion
- Q&A

Who We Are

- Metropolitan Planning Organizations (MPOs)
- Boston Region MPO
- Central Transportation Planning Staff (CTPS)

Metropolitan Planning Organizations

- Federally mandated and federally funded transportation planning agencies
- Introduced by Federal-Aid Highway Act of 1962
- Required for all U.S. urbanized areas with populations greater than 50,000

BOSTON REGION METROPOLITAN PLANNING ORGANIZATION

BOSTON REGION
MPO
METROPOLITAN PLANNING ORGANIZATION

Boston Region MPO

- Geographic scope
  - 101 municipalities in the Boston metro area
  - Boundary roughly defined by I-495

- Membership
  - MassDOT, MBTA, MBTA Advisory Board, Massport, MAPC, RTAC, 6 cities, 7 towns
  - FHWA, FTA (non-voting)

Central Transportation Planning Staff

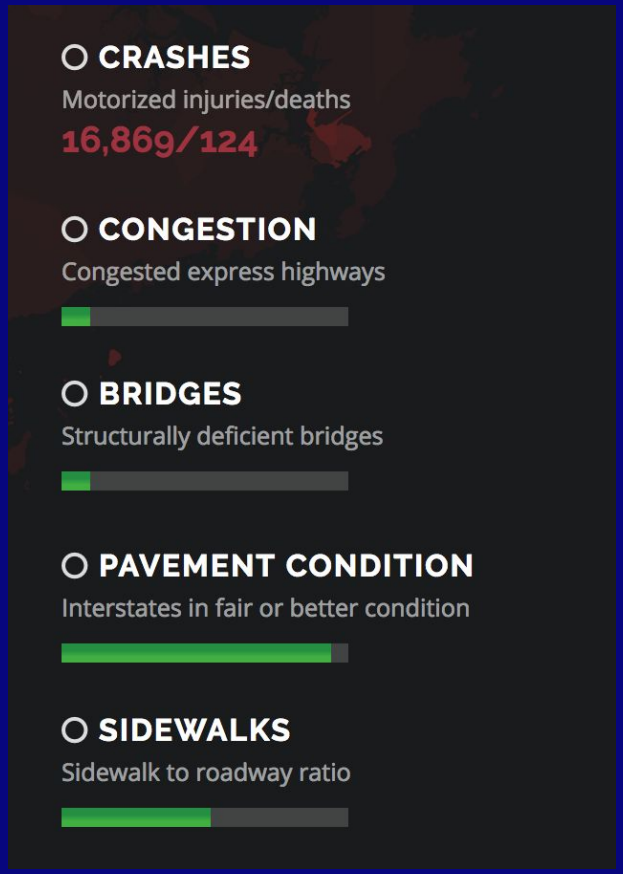
- Staff to the Boston Region MPO
- Expertise in comprehensive, multimodal transportation planning and analysis
- In-house dedicated GIS, data-development and software group

CTPS
CENTRAL TRANSPORTATION PLANNING STAFF

BOSTON REGION MPO
METROPOLITAN PLANNING ORGANIZATION

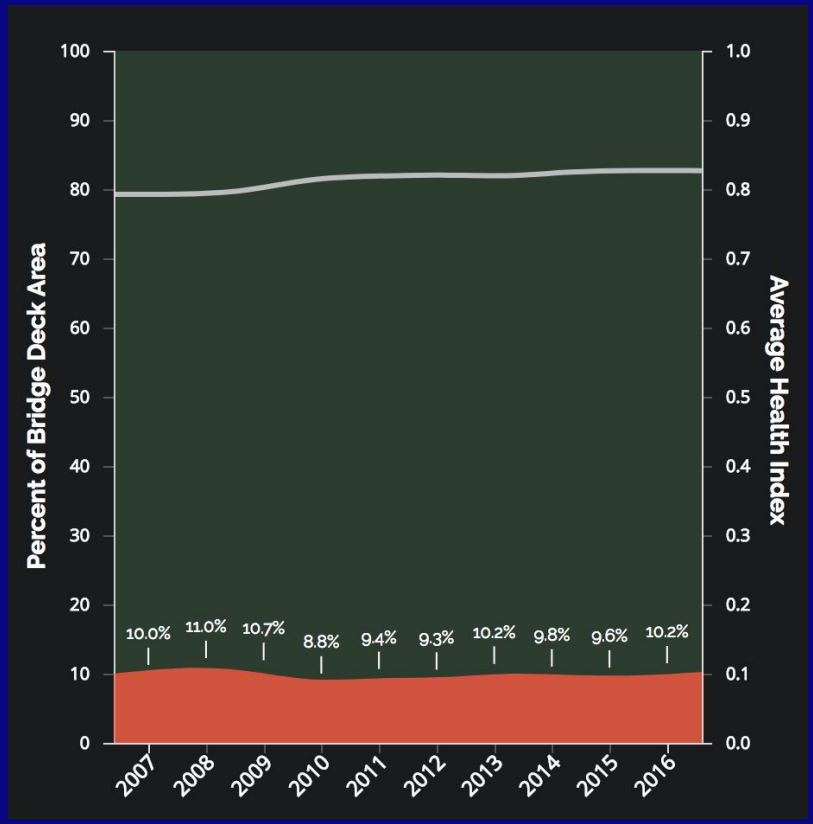
Project Context

- Federal mandate for performance-based planning
- Define performance metrics
- Set goals
- Collect data
- Track and report progress



The Performance Dashboard

- Fulfill federal requirements
- Present data to the public on the web
- Accurate, navigable, interactive presentation



BOSTON REGION MPO
METROPOLITAN PLANNING ORGANIZATION

Implementation Vehicle: D3.js

- D3 = Data-driven documents
- JavaScript library
- Developed by Mike Bostock at Stanford visualization lab
- First released in 2011
- Free / open-source

Transformation, Not Representation

- A web page <u>is</u> a data structure
- Transform it, based on user data
  - Structure
  - Appearance
  - Behavior
- Foundation: *The Grammar of Graphics*

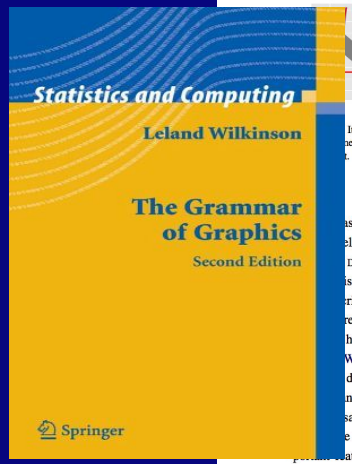
A LAYERED GRAMMAR OF GRAPHICS 19

Iteratively reproducing the depiction of Napoleon's March by Minard. (Top) Displaying the key ent data. (Center) Adding town locations as reference points. (Bottom) Tweaking scales to produce t.

st rudimentary tools for importing, restructuring, transforming, and aggregating lying on other tools in R, `ggplot2` does not need three elements of Wilkinson's DATA, TRANS, and the algebra.

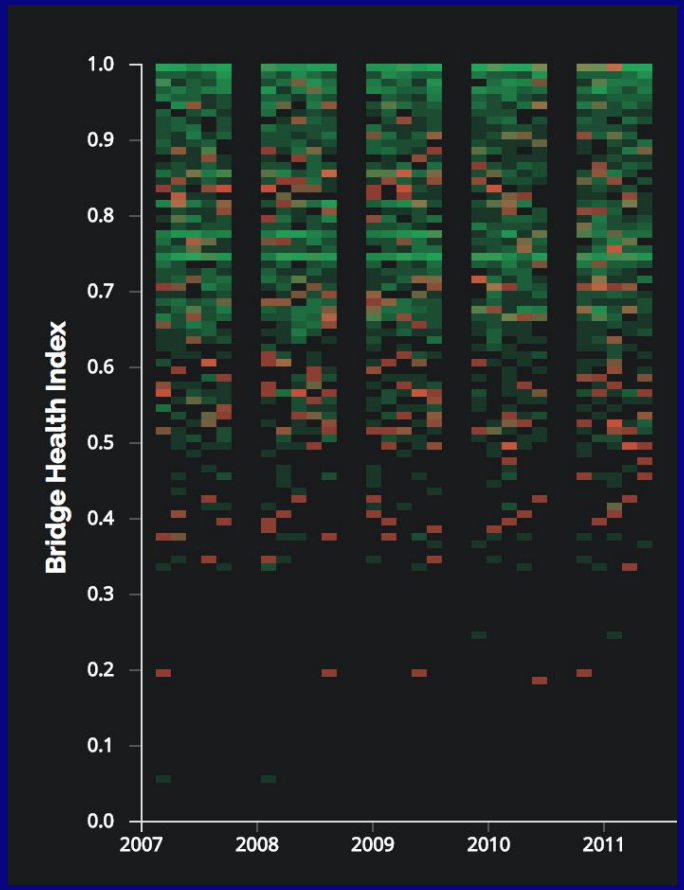
is no longer needed because data are stored as R data frames; they do not need ribed as part of the graphic. TRANS can be dropped because variable transfor- e already so easy in R; they do not need to be part of the grammar. The algebra how to reshape data for display and, in R, can be replaced by the `reshape` Wickham 2005). Separating data manipulation from visualization allows you to data, and the same restructuring can be used in multiple plots. Additionally, the nnot easily perform aggregation or subsetting, but `reshape` can.

sadvantages of embedding the grammar are somewhat more subtle, and center e grammatical restrictions applied by the host language. One of the most im- atures of the grammar is its declarative nature. To preserve this nature in R, `ggplot2` uses + to create a plot by adding pieces of the definition together. The `ggplot` function creates a base object, to which everything else is added. This base object is not necessary in a stand-alone grammar.



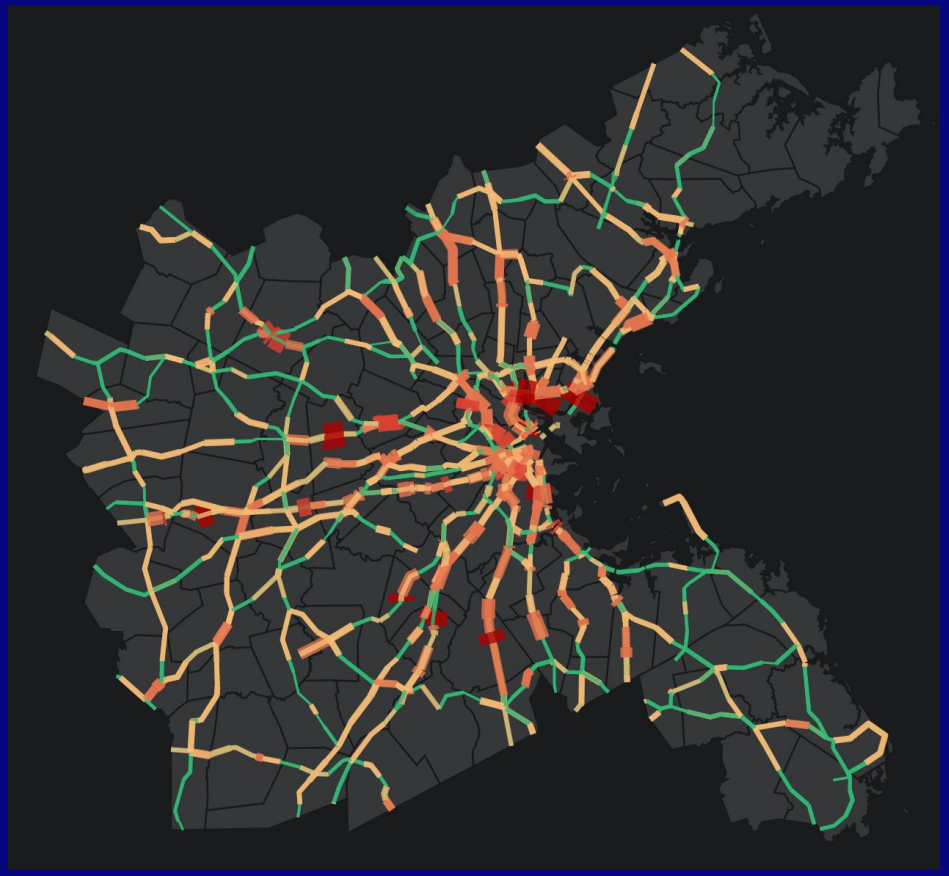
Dashboard Data

- Tabular
  - Census demographics
  - Crash count per town
- Geospatial
  - Point – crash location
    - Data is confidential
  - Line – roads
  - Polygon – town- and census-boundaries



Linear Geospatial Data

- Pavement condition
- Congestion
- Sidewalk coverage
- Bicycle facilities



BOSTON REGION MPO
METROPOLITAN PLANNING ORGANIZATION

Data Sources

- Pavement condition, sidewalk coverage
  - Annual MassDOT Road Inventory
- Congestion
  - INRIX speed-run data processed by CTPS
- Bicycle facilities
  - MassDOT Bicycle Accommodation Inventory
  - MAPC Bicycle and Pedestrian Mapping Index

Source Data Formats

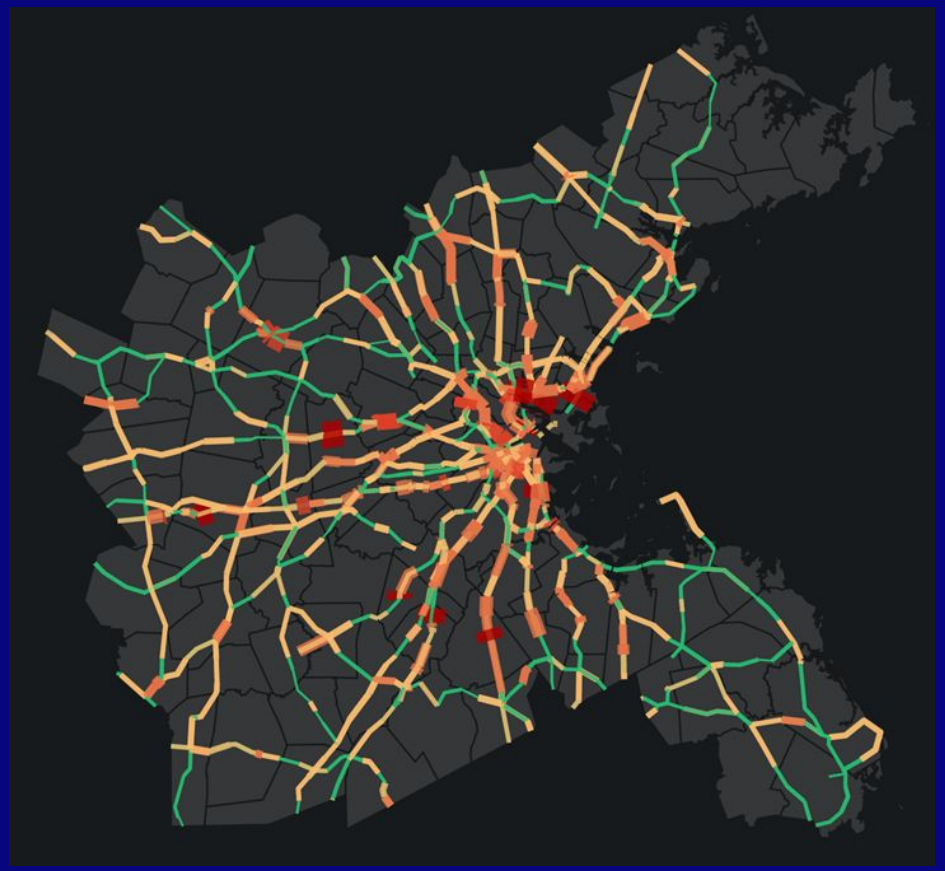
- Pavement, sidewalks, congestion
  - ArcSDE feature classes
- Bicycle facilities
  - Shapefile
  - ArcSDE feature class
- Source data exported to GeoJSON format
  - ogr2ogr, esri-to-open

Linear Data

- Linear transformation
- *Non*-linear transformation

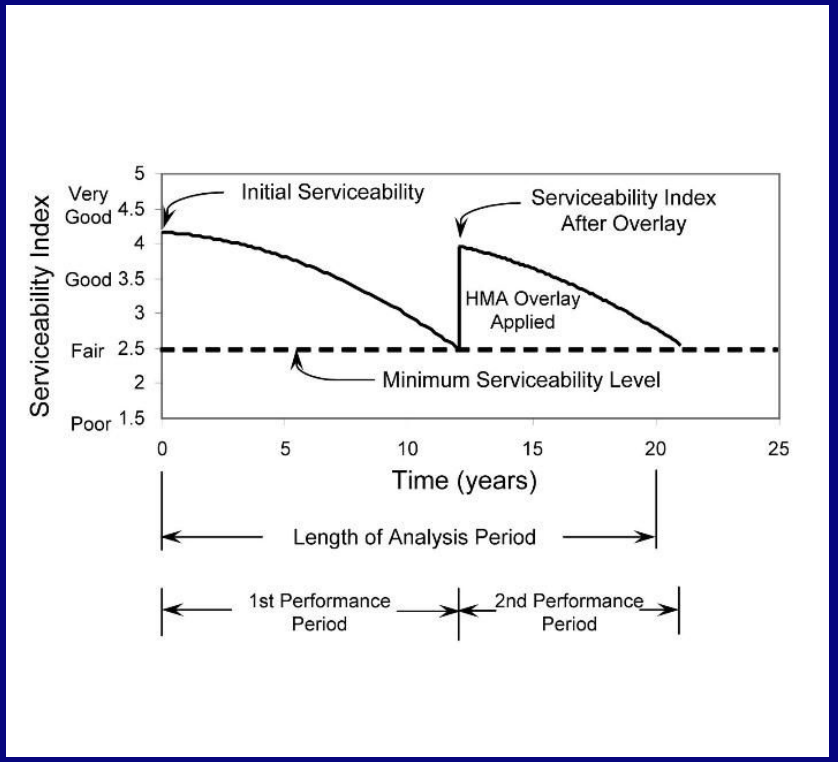
Linear Transformation

- Pavement condition
- Congestion



Pavement Condition

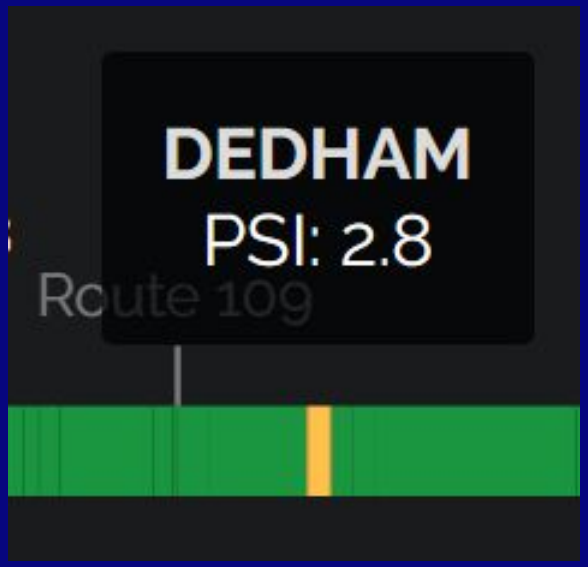
- Metric: **Present Serviceability Index** (PSI)
- PSI is a measure of a road's roughness
  - 5 → Excellent
  - 0 → Poor



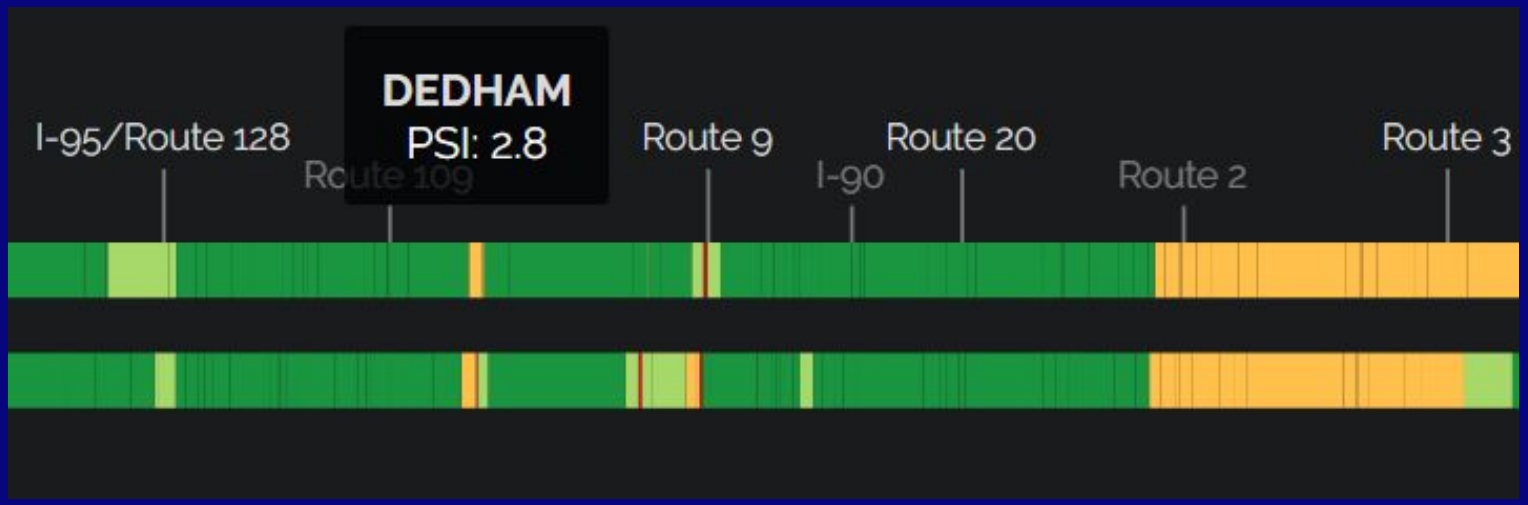
Sample Road Inventory Record (GeoJSON Format)


```
({""coordinates": [[-71.22469420065103,
42.30296805460002], [-71.22474753776642,
42.30307874118052]]},
"type": "Feature",
"id": 1,
"properties": {
        "ROUTESYSTEM": "I",
        "MILEAGECOUNTED": 1,
        "SURFACETYPE": 6,
        "IRIYEAR": 2013,
        "TOMEASURE": 0.02,
        "FEDERALFUNCTIONALCLASS": 1,
        "ROUTETO": 35.555800000000005,
        "NUMBEROFPEAKHOURLANES": 4,
        "TRUCKNETWORK": 1,
        "RPA": "MAPC",
        "ROUTEDIRECTION": "NB",
        "ASSIGNEDLENGTH": 0.02,
        "FUNCTIONALCLASSIFICATION": 1,
        "ROUTEKEY": "I95 NB",
        "ROUTETYPE": 1,
        "ROADINVENTORY_ID": 26246800,
        "PSIYEAR": 2013,
        "ADTDERIVATION": 1,
        "ADT": 169398,
        "JURISDICTION": "1",
        "COUNTY": "K",
        "IRI": 124,
        "ROUTENUMBER": "95",
        "FACILITYTYPE": 1,
        "MHS": 0,
        "CITY": 199,
        "FROMMEASURE": 0.0,
        "PSI": 2.88707553,
        "ROUTEFROM": 35.5477,
        "ROADSEGMENT_ID": 262468,
        "NUMBEROFTRAVELLANES": 3,
        "STRUCTURALCONDITION": 2,
        "TRUCKROUTE": 1,
        "FEDERALAIDROUTENUMBER": "I-95",
        "ADTYEAR": 2013,
        "MPO": "Boston Region",
        "SPEEDLIMIT": 55,
        "NHSSTATUS": 1,
        "STREETNAME": "YANKEE DIVISION
        HIGHWAY"}}
```

One Road Segment

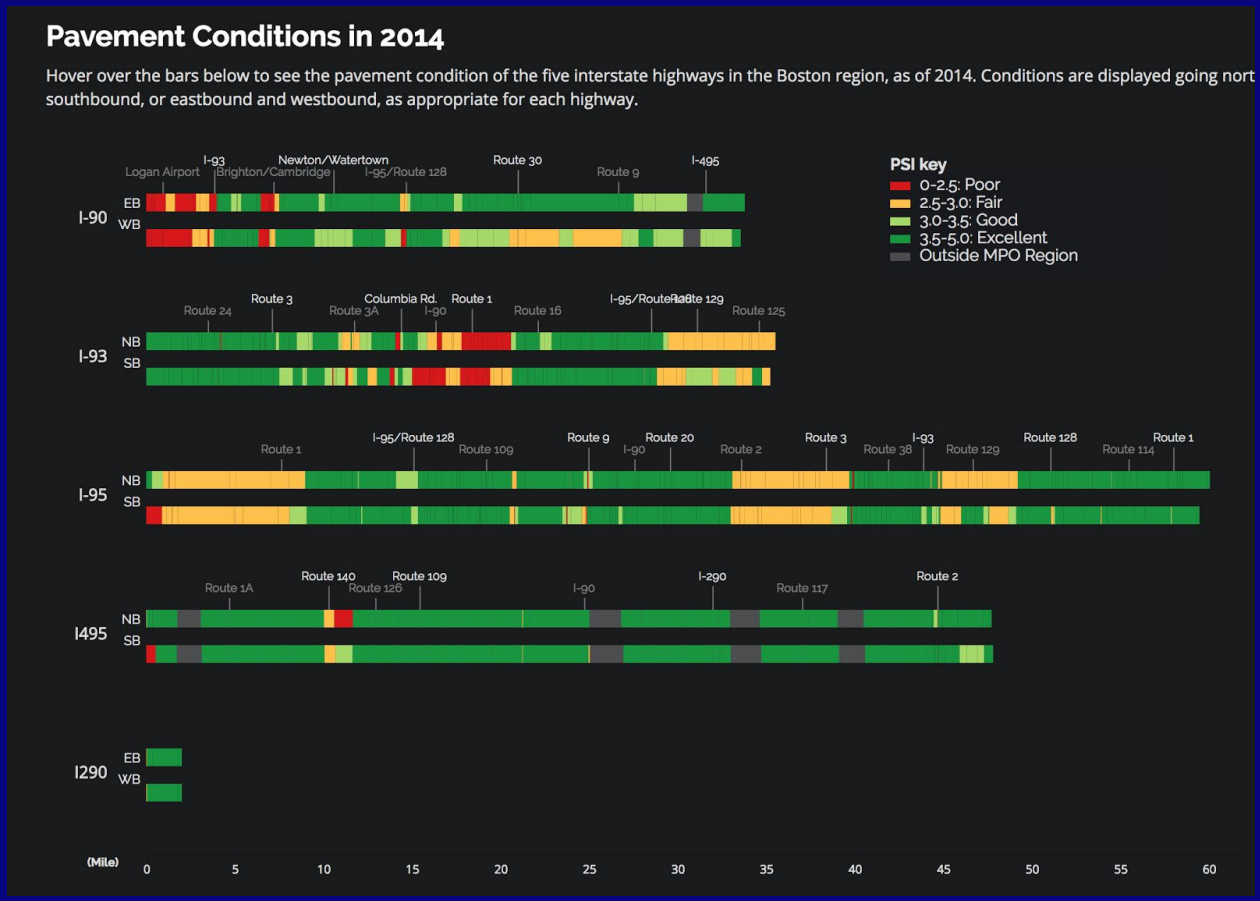


One Road Segment in Local Context



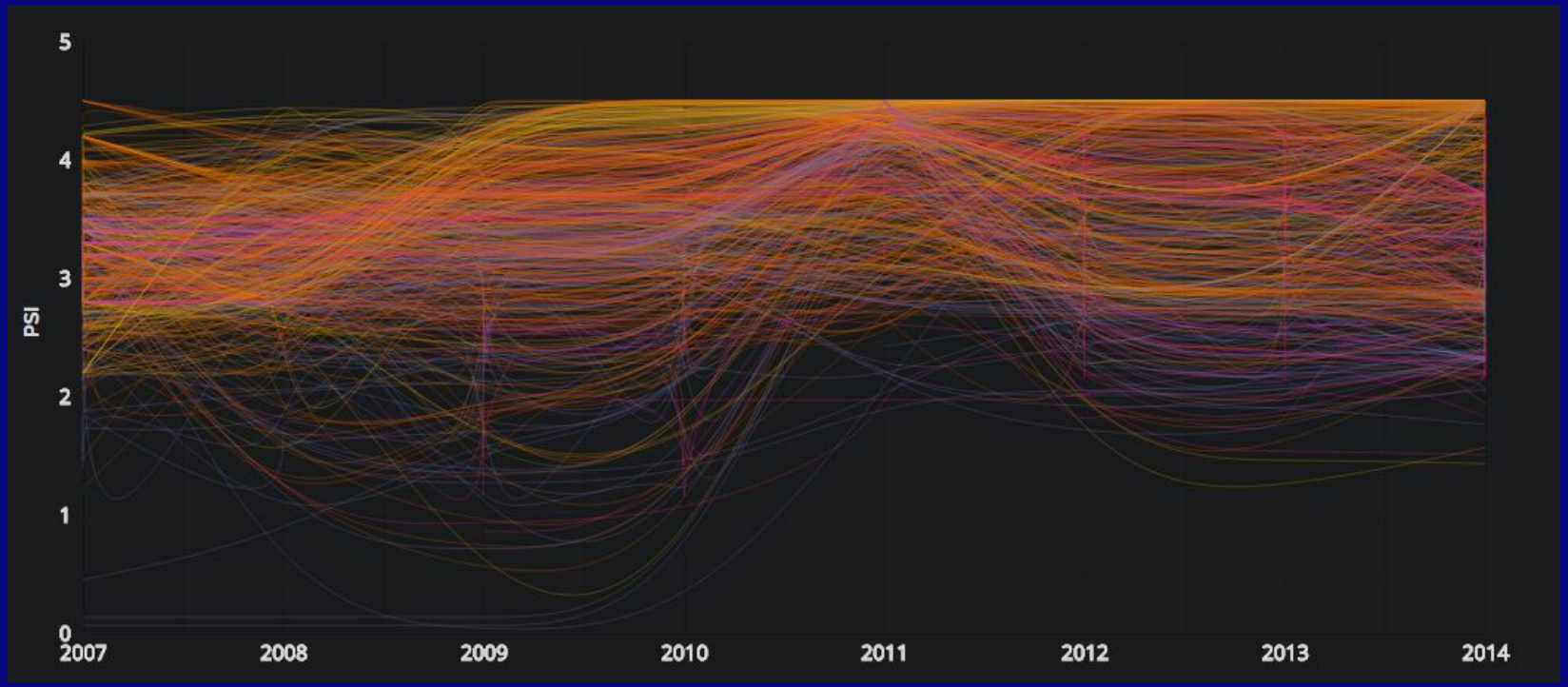
In the Context of an Entire Route

In the Context of the Regional Express Highway System

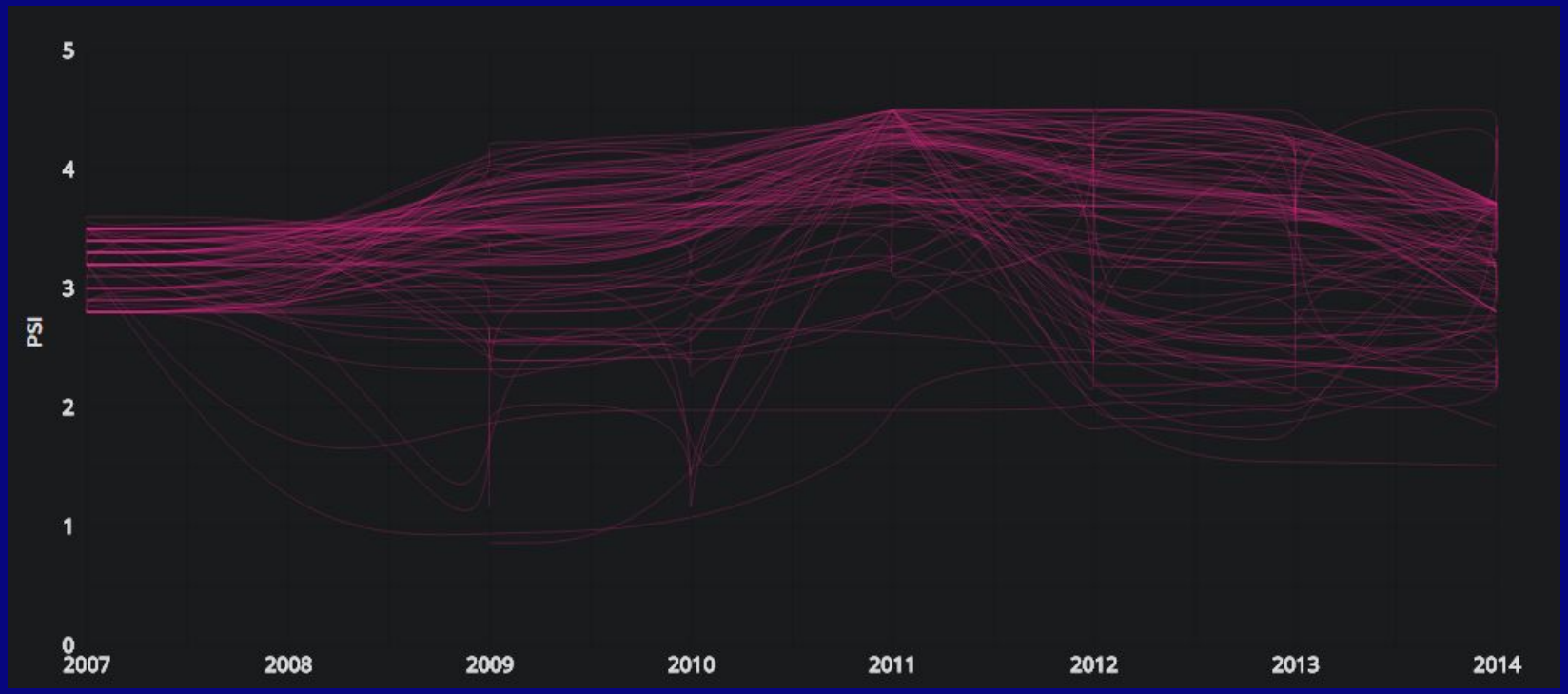


BOSTON REGION MPO METROPOLITAN PLANNING ORGANIZATION

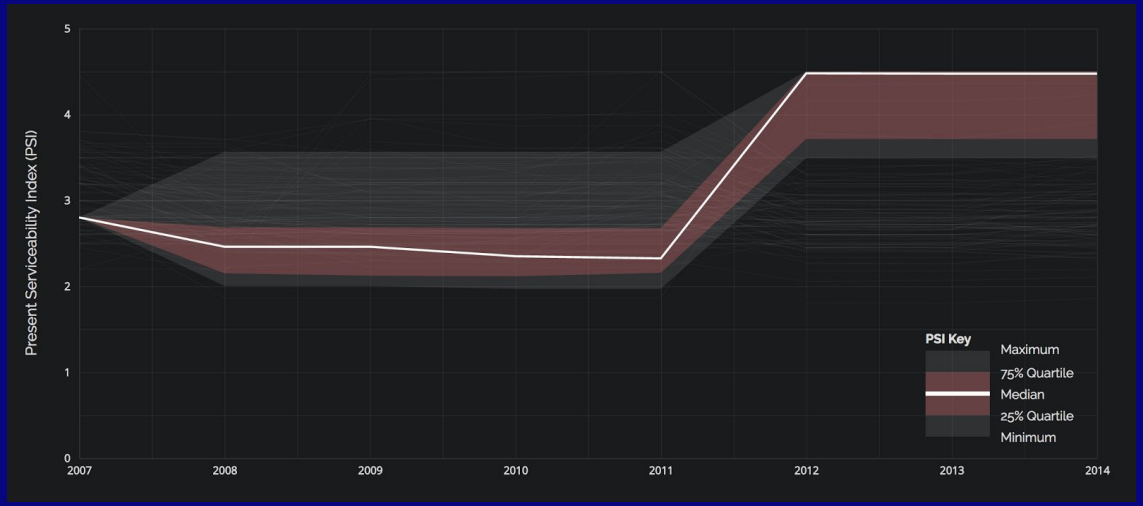
Temporal PSI Overview – All Interstates

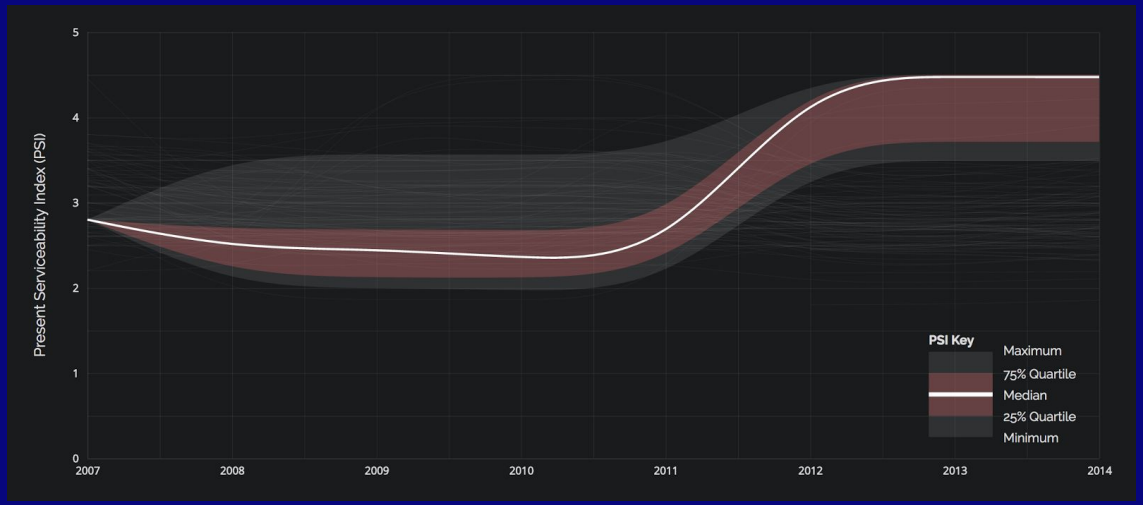


Temporal PSI Overview – I-90



D3's curveBasis smooths discrete data points





Congestion

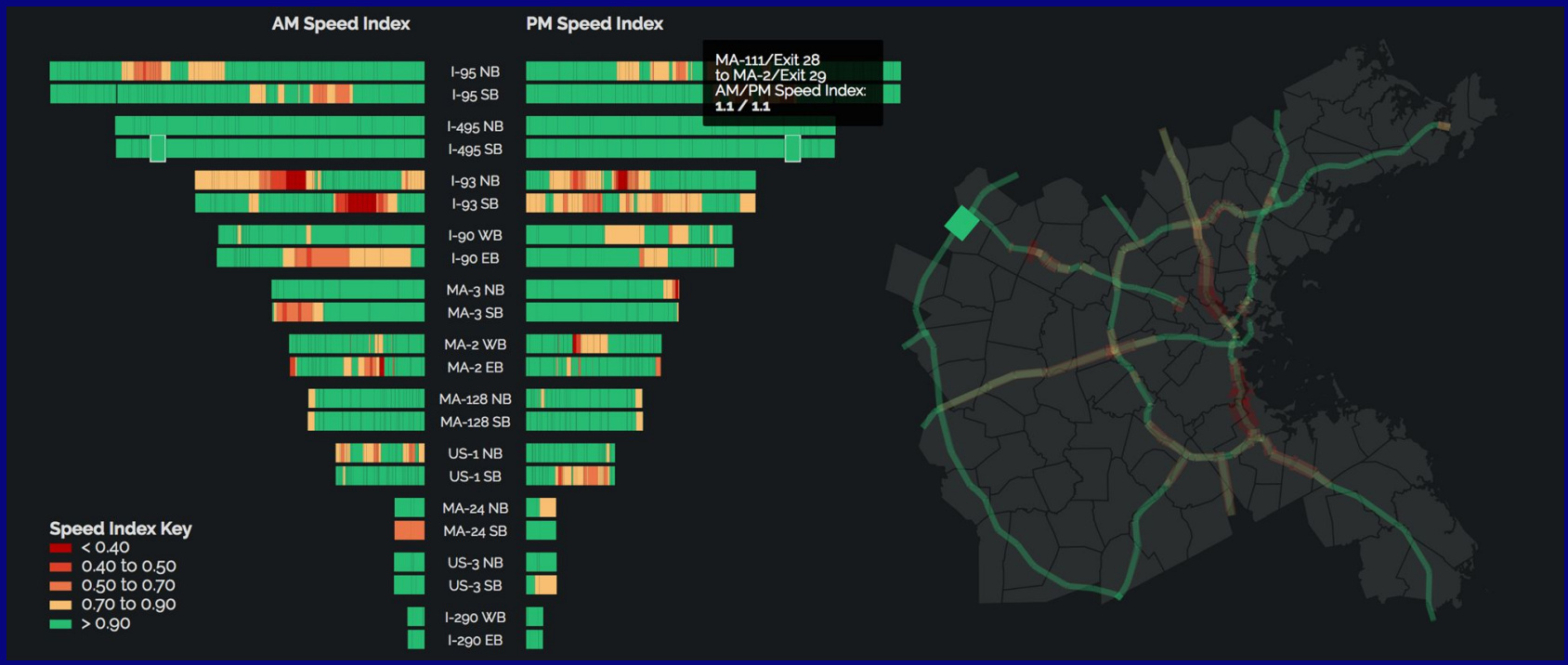
Metric: Speed Index (SI)

$$SI = \frac{observed\ speed}{posted\ speed\ limit}$$

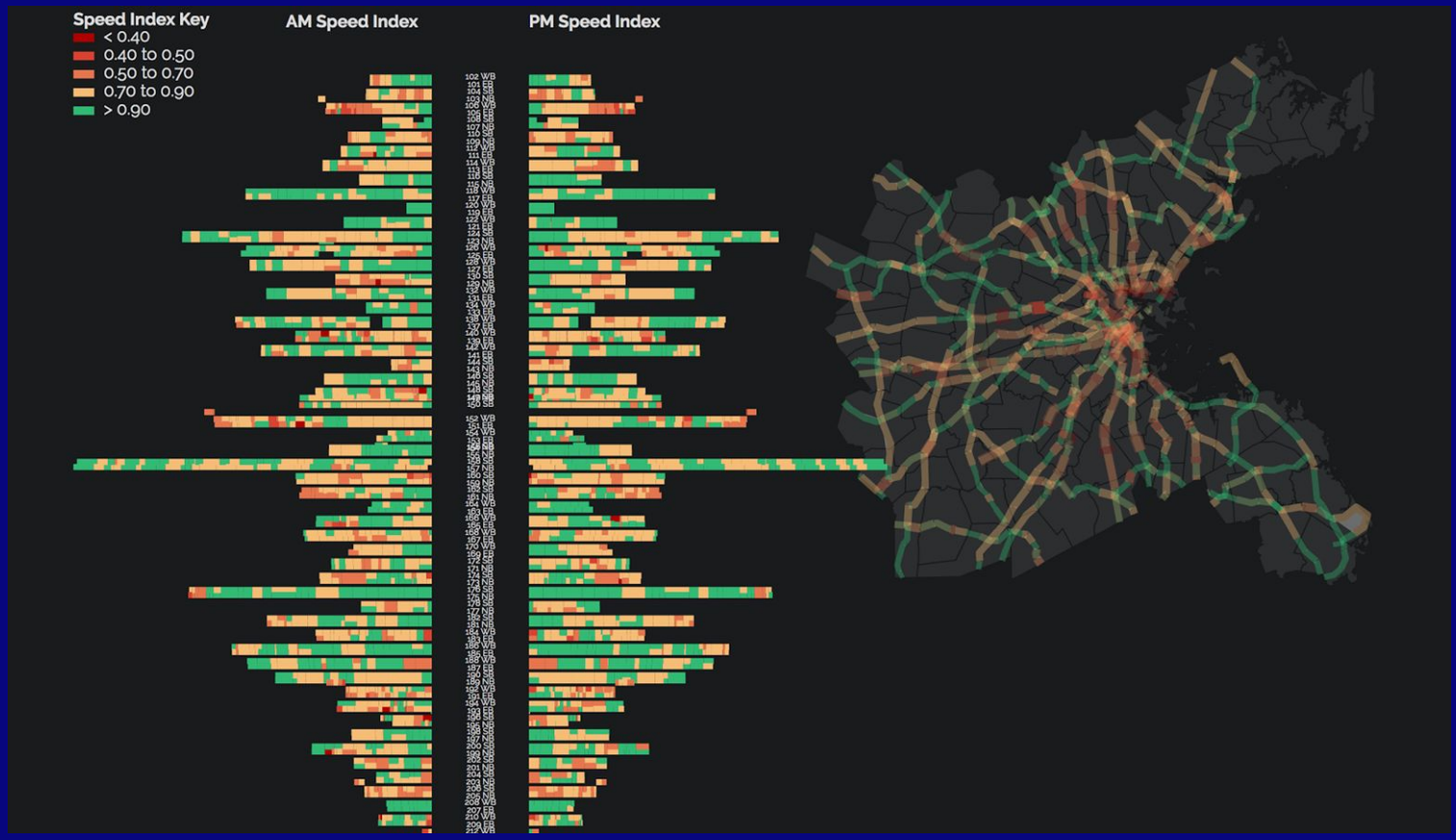
Sample Congestion Data Record (GeoJSON Format)


```
{"type":"LineString",
"id":8,
"properties":{
        "SPD_LIMIT":30,
        "TO_MEAS":11478.2641,
        "FROM_MEAS":10704.506300000001,
        "ROAD_NAME":"NULL",
        "DIRECTION":"Westbound",
        "AM_SPD_IX":1.2600923,
        "SEG_END":"Toll Plaza/Exit 18",
        "PM_SPD_IX":0.98233948,
        "RID":6,
        "ROUTE_NUM":"I-90",
        "SEG_BEGIN":"Toll Plaza/Exit 20"},
        "arcs":[1]}
```

Speed Index: Express Highways



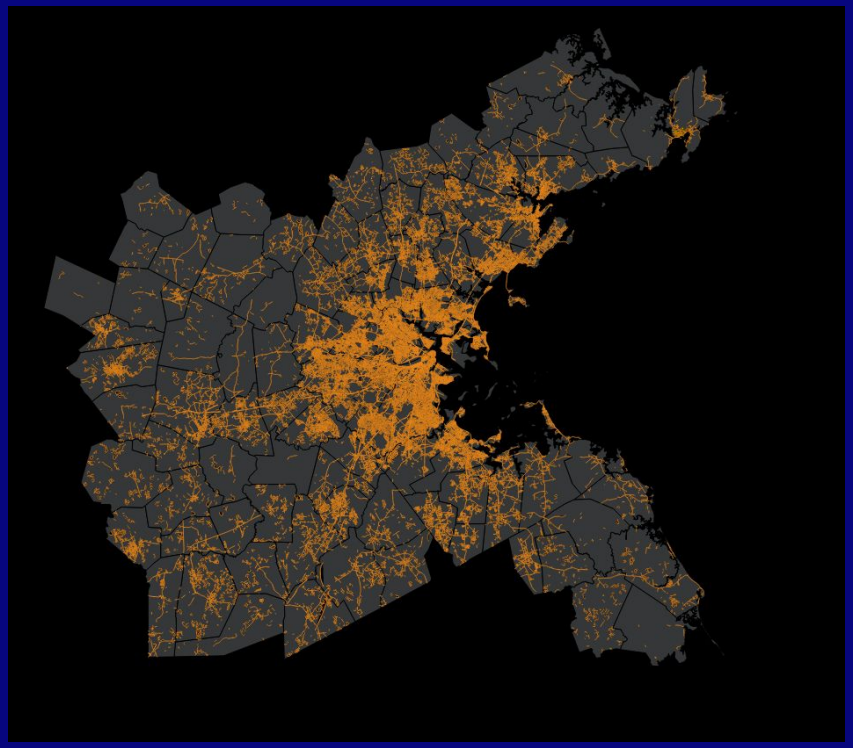
Speed Index: Arterial Routes



Speed Index: One Arterial Route

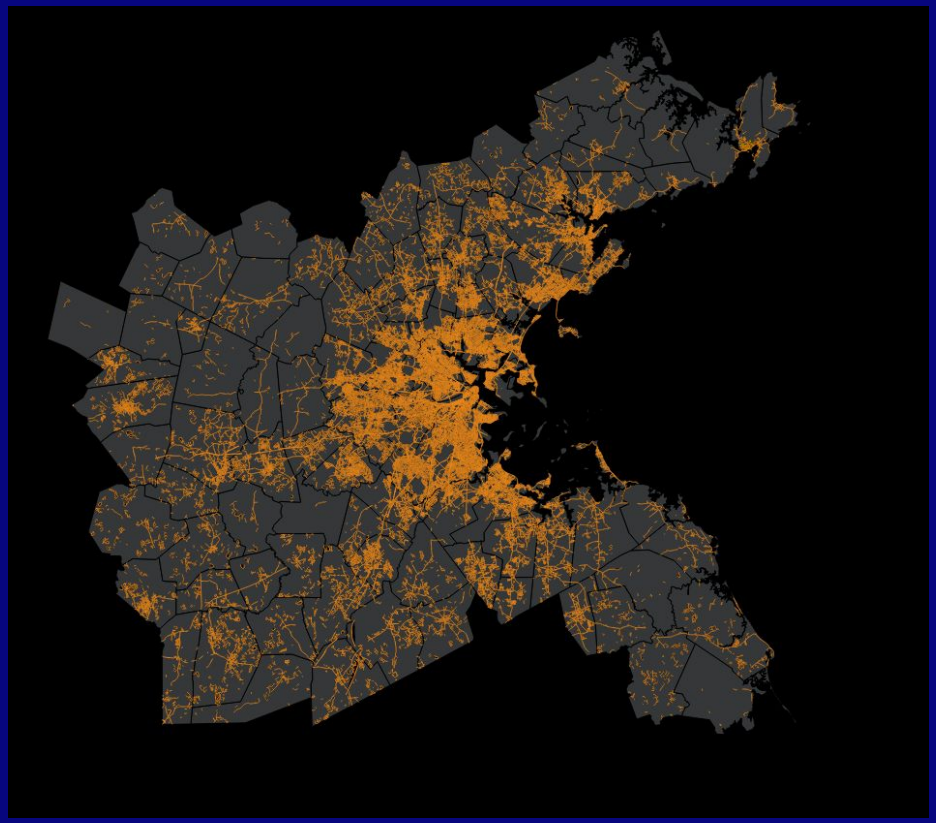
Non-linear Transformation

Linear data isn't always best presented as a line!



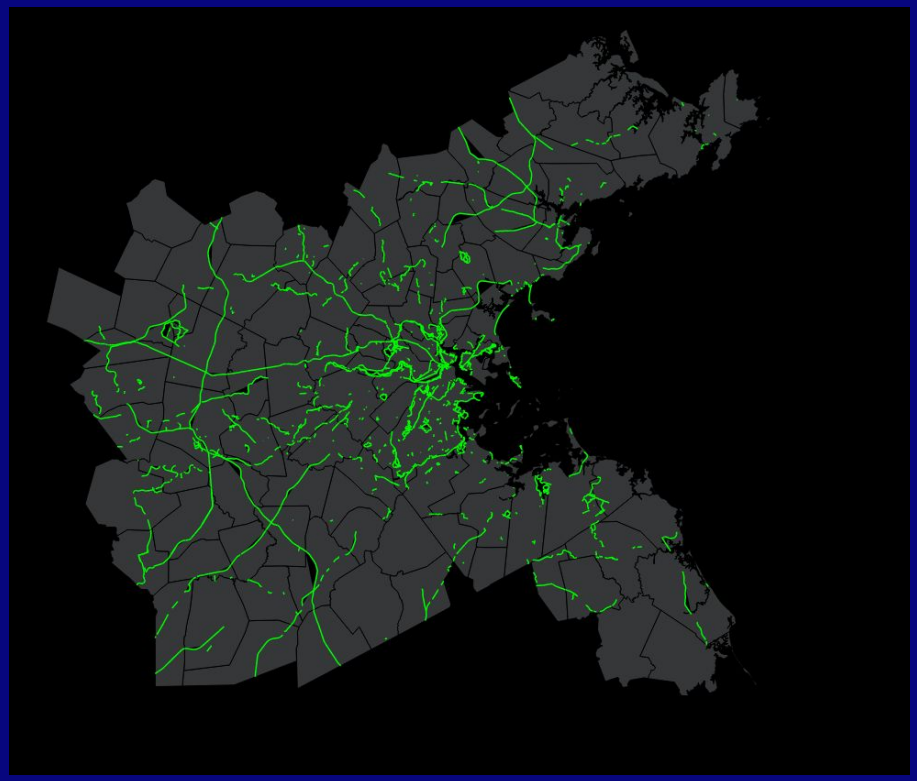
Dense Linear Data

Sidewalk coverage



Sparse Linear Data

Off-road bicycle facilities



Solution

- Aggregate data by geographic unit
  - Municipality
  - Census tract
  - Traffic analysis zone (TAZ)
- Display aggregated data
  - Choropleth map
  - Bar chart
  - Enhanced x-y scatter plot

Metrics

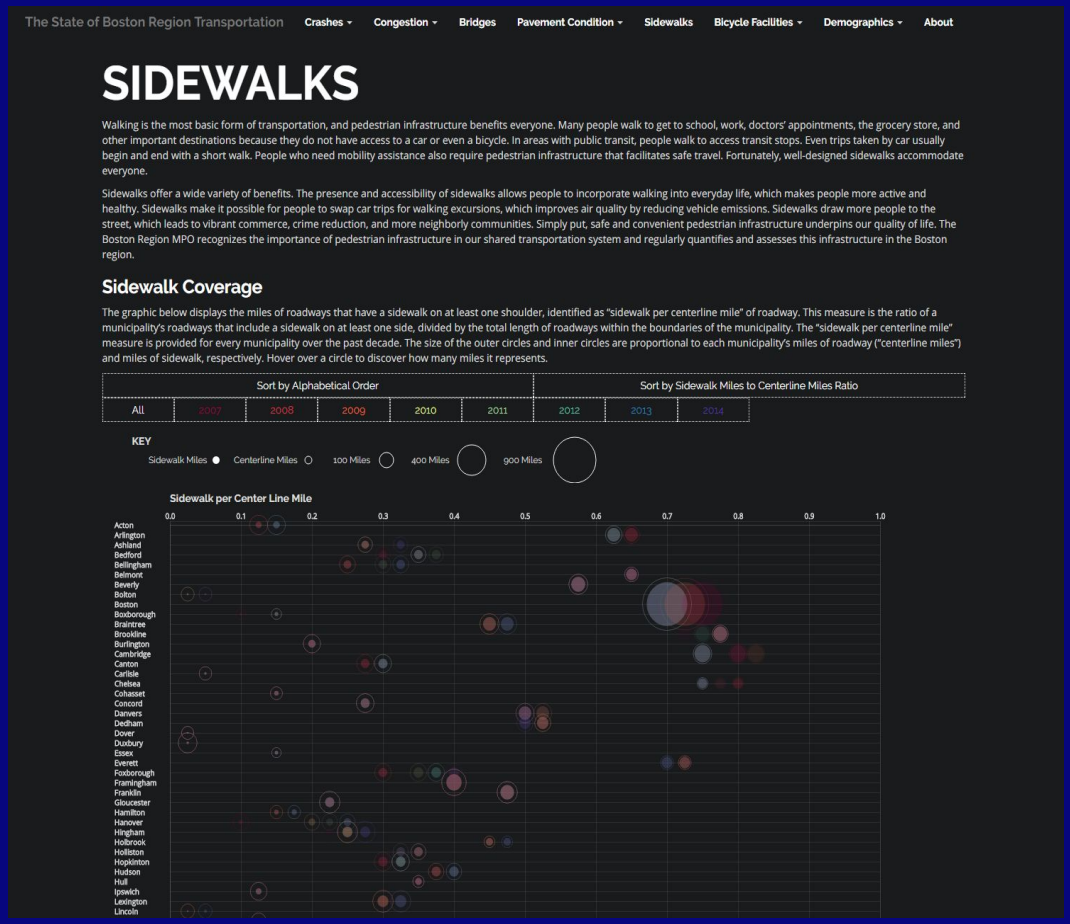
- **Sidewalk coverage**

$$\frac{miles\ of\ non-limited\ access\ roads\ with\ sidewalks}{centerline\ miles\ of\ non-limited-access\ roads}$$

- **Off-road bicycle facilities**
  - **Miles per municipality**

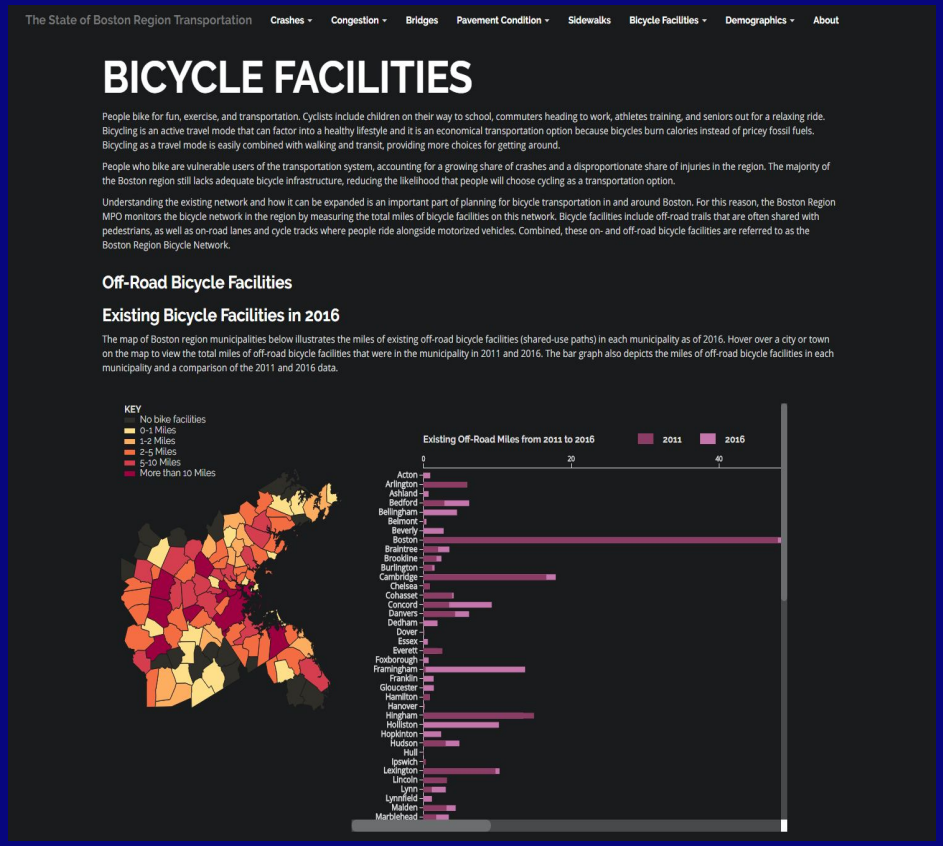
Sidewalk Coverage

- Enhanced x-y scatter plot



Off-road Bicycle Facilities

- Choropleth map
- Bar chart



Conclusion

- Single visualization framework
  - Spatial data
  - Non-spatial data
- Web server *is* spatial data server
- Open format data
  - CSV, TSV, GeoJSON, TopoJSON
- Rich presentation

Check Out the Dashboard

www.ctps.org/dv/lrtp_dashboard

Thank You

- David Knudsen – CTPS GIS Analyst
- Kathy Jacob – CTPS GIS Analyst
- Paul Reim – CTPS GIS Analyst
- Matt Archer – CTPS Intern
- Jennifer Rowe – CTPS Public Participation Manager
- Anne McGahan – CTPS LRTP Project Manager
- Lourenço Dantas – CTPS Certification Activities Manager
- Scott Peterson – CTPS Director of Technical Services
- Robin Mannion – CTPS Deputy Executive Director
- Karl Quackenbush – CTPS Executive Director

Q&A

- Beatrice Jin — bjin@ctps.org
- Ben Krepp — bkrepp@ctps.org