Data Visualization (CIS 468)

Introduction

Dr. David Koop
What is Data Visualization?
How is it different from Computer Graphics?
The purpose of computing is about insight, not numbers

- R. W. Hamming
The purpose of visualization is about insight, not pictures

- Card, Mackinlay, Schneiderman
Why do we visualize data? (vs. looking at tables?)
## MTA Fare Data Table

<table>
<thead>
<tr>
<th>REMOTE</th>
<th>STATION</th>
<th>FF</th>
<th>SEN/DIS</th>
<th>7-D AFAS UNL</th>
<th>D AFAS/RR TKT</th>
<th>7-D UNL</th>
<th>30-D UNL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>R011 42ND STREET &amp; 8TH AV</td>
<td>00228985</td>
<td>00008471</td>
<td>00000441</td>
<td>00001455</td>
<td>00000134</td>
<td>00033341</td>
</tr>
<tr>
<td>2</td>
<td>R170 14TH STREET-UNION SQ</td>
<td>00224603</td>
<td>00011051</td>
<td>00000827</td>
<td>00003026</td>
<td>00000660</td>
<td>00089367</td>
</tr>
<tr>
<td>3</td>
<td>R046 42ND STREET &amp; GRAND CENTRAL</td>
<td>00207758</td>
<td>00007908</td>
<td>00000323</td>
<td>00001183</td>
<td>00003001</td>
<td>00040759</td>
</tr>
<tr>
<td>4</td>
<td>R012 34TH STREET &amp; 8TH AV</td>
<td>00188311</td>
<td>00006490</td>
<td>00000498</td>
<td>00001279</td>
<td>00003622</td>
<td>00035527</td>
</tr>
<tr>
<td>5</td>
<td>R293 34TH STREET - PENN STATION</td>
<td>00168768</td>
<td>00006155</td>
<td>00000523</td>
<td>00001065</td>
<td>00005031</td>
<td>00030645</td>
</tr>
<tr>
<td>6</td>
<td>R033 42ND STREET/TIMES SQUARE</td>
<td>00159382</td>
<td>00005945</td>
<td>00000378</td>
<td>00001205</td>
<td>00000690</td>
<td>00058931</td>
</tr>
<tr>
<td>7</td>
<td>R022 34TH STREET &amp; 6TH AV</td>
<td>00156008</td>
<td>00006276</td>
<td>00000487</td>
<td>00001543</td>
<td>00000712</td>
<td>00058910</td>
</tr>
<tr>
<td>8</td>
<td>R084 59TH STREET/COLUMBUS CIRCLE</td>
<td>00155262</td>
<td>00009484</td>
<td>00000589</td>
<td>00002071</td>
<td>00000542</td>
<td>00053397</td>
</tr>
<tr>
<td>9</td>
<td>R020 47-50 STREETS/ROCKEFELLER</td>
<td>00143500</td>
<td>00006402</td>
<td>00000384</td>
<td>00001159</td>
<td>00000723</td>
<td>00037978</td>
</tr>
<tr>
<td>10</td>
<td>R179 86TH STREET-LEXINGTON AVE</td>
<td>00142169</td>
<td>00010367</td>
<td>00000470</td>
<td>00001839</td>
<td>00000271</td>
<td>00050328</td>
</tr>
<tr>
<td>11</td>
<td>R023 34TH STREET &amp; 6TH AV</td>
<td>00134052</td>
<td>00005005</td>
<td>00000348</td>
<td>00001112</td>
<td>00000649</td>
<td>00031531</td>
</tr>
<tr>
<td>12</td>
<td>R029 PARK PLACE</td>
<td>00121614</td>
<td>00004311</td>
<td>00000287</td>
<td>00000931</td>
<td>00000792</td>
<td>00025404</td>
</tr>
<tr>
<td>13</td>
<td>R047 42ND STREET &amp; GRAND CENTRAL</td>
<td>00100742</td>
<td>00004273</td>
<td>00000185</td>
<td>00000704</td>
<td>00001241</td>
<td>00022808</td>
</tr>
<tr>
<td>14</td>
<td>R031 34TH STREET &amp; 7TH AV</td>
<td>00095076</td>
<td>00003990</td>
<td>00000232</td>
<td>00000727</td>
<td>00001459</td>
<td>00024284</td>
</tr>
<tr>
<td>15</td>
<td>R017 LEXINGTON AVENUE</td>
<td>00094655</td>
<td>00004688</td>
<td>00000190</td>
<td>00000833</td>
<td>00000754</td>
<td>00020018</td>
</tr>
<tr>
<td>16</td>
<td>R175 8TH AVENUE-14TH STREET</td>
<td>00094313</td>
<td>00003907</td>
<td>00000286</td>
<td>00001144</td>
<td>00000256</td>
<td>00038272</td>
</tr>
<tr>
<td>17</td>
<td>R057 BARCLAYS CENTER</td>
<td>00093804</td>
<td>00004204</td>
<td>00000454</td>
<td>00001386</td>
<td>00001491</td>
<td>00039113</td>
</tr>
<tr>
<td>18</td>
<td>R138 WEST 4TH ST-WASHINGTON SQ</td>
<td>00093562</td>
<td>00004677</td>
<td>00000251</td>
<td>00000965</td>
<td>00000127</td>
<td>00031628</td>
</tr>
</tbody>
</table>
MTA Fare Data Visualization
Why do we visualize data?

**Figures** are richer; provide more information with less clutter and in less space. Figures provide the gestalt effect: they give an overview; make structure more visible.

Figures are more accessible, easier to understand, faster to grasp, more comprehensible, more memorable, more fun, and less formal.

*List adapted from: Stasko et al. 1998*

[via A. Lex]
What are the purposes for visualization?
Exploration: Subway Ridership Density
Why Peyton Manning's Record Will Be Hard to Beat

By GREGOR AISCH and KEVIN QUEALY  OCT. 19, 2014

The Broncos quarterback set the all-time N.F.L. touchdown passing record — and is still going strong.

[Graph showing historical passing records with Peyton Manning's record highlighted.]
Exploration <-> Communication Spectrum

Consecutive Starts by a Quarterback for a Single Team

[K. Quealy, 2013]
Exploration <-> Communication Spectrum

Consecutive Starts by a Quarterback for a Single Team

[Image: Diagram showing consecutive starts by a quarterback for a single team]

Questions

Exploration

Confirmation

Communication

Answers/Persuasion

[K. Quealy, 2013]
What types of data can we visualize?
Types of Data

- Tables
- Networks (Graphs)
- Spatial Data
  - Geography
  - Physical (e.g. Scientific, Medical)
- Text
- Sets
Where have you seen visualizations?
Books / Posters

[Rock 'N' Roll is Here to Pay, R. Garofalo, 1977]
Books / Posters
Music Timeline

[Music Timeline, Google Research]
What is the advantage of the second version?
Interaction
How do we create modern visualizations?
Tools

• Desktop Applications:
  - Excel (see excelcharts.com)
  - Tableau
  - …

• Programming Frameworks/Languages
  - Processing
  - d3.js
  - deck.gl, MapboxGL
  - vega-lite, …

• Advantages and disadvantages
  - Speed, customization, understanding, dissemination
D3.js is a JavaScript library for manipulating documents based on data. D3 helps you bring data to life using HTML, SVG, and CSS. D3’s emphasis on web standards gives you the full capabilities of modern browsers without tying yourself to a proprietary framework, combining powerful visualization components and a data-driven approach to DOM manipulation.

See more examples.
Why do we care about the design of visualizations?
Design: Focus on only the y-axis

Average Annual Global Temperature in Fahrenheit
1880-2015

[S. Hayward, 2015]
Design: Year on the y-axis

Year

Year (A.D.)
Design: Different y-axis

Average Annual Global Temperature in Fahrenheit
1880-2015

[S. Hayward, 2015]
Visualization exploration is often iterative
Transportation Data - NYC MTA
### MTA Fare Data Exploration

<table>
<thead>
<tr>
<th>REMOTE</th>
<th>STATION</th>
<th>FF</th>
<th>SEN/DIS</th>
<th>7-D AFAS UNL</th>
<th>D AFAS/REML</th>
<th>JOINT RR TKT</th>
<th>7-D UNL</th>
<th>30-D UNL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>R011 42ND STREET &amp; 8TH AVENUE</td>
<td>00228985</td>
<td>00008471</td>
<td>00000441</td>
<td>00001455</td>
<td>00000134</td>
<td>00033341</td>
<td>00071255</td>
</tr>
<tr>
<td>2</td>
<td>R170 14TH STREET-UNION SQUARE</td>
<td>00224603</td>
<td>0011051</td>
<td>00000827</td>
<td>00003026</td>
<td>00000660</td>
<td>00089367</td>
<td>00199841</td>
</tr>
<tr>
<td>3</td>
<td>R046 42ND STREET &amp; GRAND CENTRAL</td>
<td>00207758</td>
<td>0007908</td>
<td>0000323</td>
<td>00001183</td>
<td>00003001</td>
<td>00040759</td>
<td>00096613</td>
</tr>
<tr>
<td>4</td>
<td>R012 34TH STREET &amp; 8TH AVENUE</td>
<td>00188311</td>
<td>0006490</td>
<td>00000498</td>
<td>00001279</td>
<td>00003622</td>
<td>00035527</td>
<td>00067483</td>
</tr>
<tr>
<td>5</td>
<td>R293 34TH STREET - PENN STATION</td>
<td>00168768</td>
<td>0006155</td>
<td>0000523</td>
<td>00001065</td>
<td>00005031</td>
<td>00030645</td>
<td>00054376</td>
</tr>
<tr>
<td>6</td>
<td>R033 42ND STREET/TIMES SQUARE</td>
<td>00159382</td>
<td>0005945</td>
<td>0000378</td>
<td>00001205</td>
<td>00000690</td>
<td>00058931</td>
<td>00078644</td>
</tr>
<tr>
<td>7</td>
<td>R022 34TH STREET &amp; 6TH AVENUE</td>
<td>00156008</td>
<td>0006276</td>
<td>0000487</td>
<td>00001543</td>
<td>00000712</td>
<td>00058910</td>
<td>00110466</td>
</tr>
<tr>
<td>8</td>
<td>R084 59TH STREET/COLUMBUS CIRCLE</td>
<td>00155262</td>
<td>0009484</td>
<td>0000589</td>
<td>00002071</td>
<td>00000542</td>
<td>00053397</td>
<td>00113966</td>
</tr>
<tr>
<td>9</td>
<td>R020 47-50 STREETS/ROCHEFELLER</td>
<td>00143500</td>
<td>0006402</td>
<td>0000384</td>
<td>00001159</td>
<td>00000723</td>
<td>00037978</td>
<td>00090745</td>
</tr>
<tr>
<td>10</td>
<td>R179 86TH STREET-LEXINGTON AVE</td>
<td>00142169</td>
<td>0010367</td>
<td>0000470</td>
<td>00001839</td>
<td>00000271</td>
<td>00050328</td>
<td>00125250</td>
</tr>
<tr>
<td>11</td>
<td>R023 34TH STREET &amp; 6TH AVENUE</td>
<td>00134052</td>
<td>0005005</td>
<td>0000348</td>
<td>00001112</td>
<td>00000649</td>
<td>00031531</td>
<td>00075040</td>
</tr>
<tr>
<td>12</td>
<td>R029 PARK PLACE</td>
<td>00121614</td>
<td>0004311</td>
<td>0000287</td>
<td>00000931</td>
<td>00000792</td>
<td>00025404</td>
<td>00065362</td>
</tr>
<tr>
<td>13</td>
<td>R047 42ND STREET &amp; GRAND CENTRAL</td>
<td>00100742</td>
<td>0004273</td>
<td>0000185</td>
<td>00000704</td>
<td>00001241</td>
<td>00022808</td>
<td>00068216</td>
</tr>
<tr>
<td>14</td>
<td>R031 34TH STREET &amp; 7TH AVENUE</td>
<td>00095076</td>
<td>0003990</td>
<td>0000232</td>
<td>00000727</td>
<td>00001459</td>
<td>00024284</td>
<td>00038671</td>
</tr>
<tr>
<td>15</td>
<td>R017 LEXINGTON AVENUE</td>
<td>00094655</td>
<td>0004688</td>
<td>0000190</td>
<td>00000833</td>
<td>00000754</td>
<td>00020018</td>
<td>00055066</td>
</tr>
<tr>
<td>16</td>
<td>R175 8TH AVENUE-14TH STREET</td>
<td>00094313</td>
<td>0003907</td>
<td>0000286</td>
<td>00001144</td>
<td>00000256</td>
<td>00038272</td>
<td>00074661</td>
</tr>
<tr>
<td>17</td>
<td>R057 BARCLAYS CENTER</td>
<td>00093804</td>
<td>0004204</td>
<td>0000454</td>
<td>00001386</td>
<td>00001491</td>
<td>00039113</td>
<td>00068119</td>
</tr>
<tr>
<td>18</td>
<td>R138 WEST 4TH ST-WASHINGTON SQ</td>
<td>00093562</td>
<td>0004677</td>
<td>0000251</td>
<td>00000965</td>
<td>00000127</td>
<td>00031628</td>
<td>00074458</td>
</tr>
</tbody>
</table>
MTA Fare Data Exploration
MTA Fare Data Exploration
MTA Fare Data Exploration
MTA Fare Data Exploration

East 161st Street and River Avenue

Date

Full Fares Purchased

08-02 08-09 08-16 08-23 08-30 09-06 09-13 09-20 09-27 10-04 10-11 10-18 10-25 11-01
MTA Fare Data Exploration

East 161st Street and River Avenue

New York Yankees

AUGUST

<table>
<thead>
<tr>
<th>Sun</th>
<th>Mon</th>
<th>Tue</th>
<th>Wed</th>
<th>Thu</th>
<th>Fri</th>
<th>Sat</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Date</th>
<th>08-02</th>
<th>08-09</th>
<th>08-16</th>
<th>08-23</th>
<th>08-30</th>
<th>09-06</th>
<th>09-13</th>
<th>09-20</th>
<th>09-27</th>
<th>10-04</th>
<th>10-11</th>
<th>10-18</th>
<th>10-25</th>
<th>11-01</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>80000</td>
<td>100000</td>
<td>120000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

SCHEDULES

- 2013 REGULAR SEASON SCHEDULE

[Game Schedule Details]

**Note:** ALL GAMES ARE EASTERN TIME.

D. Koop, CIS 468, Fall 2018
Administrivia

- **Course Web Site**
- **Syllabus**
  - Plagiarism
  - Accommodations
- **Textbook:**
  - Required: Munzner (VAD)
  - Rec'd: Murray, 2nd ed. (IDV)
- **Assignments**
- **Registration:**
  - Add/Drop is **Wednesday**
  - Make sure you are registered
Important Dates

• Check these now!

• Quizzes:
  - October 2 (in class)
  - November 20 (in class)

• Midterm: October 23 (in class)

• Final: December 13 (8-11am)
Questions?
Programming

• "Programming is blindly manipulating symbols." - B. Victor

• "Code is often the best tool we have because it is the most general tool we have; code has almost unlimited expressiveness" - M. Bostock

• You will write code in this class for assignments

• JavaScript is the language of the Web
  - Somewhat forgiving, not always the easiest to debug
  - Lots of references out there
  - A quickly-changing environment of frameworks
Office Hours & Email

• Scheduled office hours are open to all students
  - M: 3-5pm, TuTh: 11am-12pm
• You do not need an appointment to stop in during scheduled office hours
• If you need an appointment outside of those times, please email me with specific details about what you wish to discuss
• Many questions can be answered via email so try writing your question as an email first
Do not plagiarize

- It is **cheating**. It violates the Academic Honesty Policy at UMassD.
- Do your own work
- Do not copy anyone else's code, text, sentences, …
  - Anyone = another student, an internet source, book, blog, …
- **Cite** sources that you use (two places):
  - in code
  - at the beginning of your assignments
Do not cheat

- Cheating on assignments, quizzes, and exams is not allowed
- You will receive a **zero** on the assignment/quiz/exam
- It will be reported to the department and university
- If it repeats, you will fail the course
- You can be kicked out of the university
Do ask questions!
Do ask questions

• If you are stuck on a specific issue with an assignment:
  - Do email me with **specific** questions
  - Do consult books, online documentation, tutorials
  - Do discuss that specific issue with a classmate

• If you are asked about a question:
  - Do not share your code
  - If the questioner wants to cheat, walk away
  - If you see an obvious mistake, kindly point it out
  - Suggest a specific function or library that may be useful
Questions?
Homework: Reading

• Munzner, Ch. 1
• Murray, Ch. 3-4 (or other HTML/JavaScript background material)