

Data**politan**

Data Solutions for the Modern Metropolis

Data Analytics for Managers

Instructor: Richard Dunks

Follow along at: <http://bit.ly/data-driven-gov>

See the code at: <http://bit.ly/data-driven-gov-code>

1 / 121

Data**politan**

Data Solutions for the Modern Metropolis

Data Analytics for Managers

Welcome

A Few Ground Rules

- Step up, step back
- Be curious and ask questions!
- Assume noble regard and positive intent
- Respect multiple perspectives
- Listen deeply
- Be present (phone, email, social media, etc.)

Introduce Yourself to Your Neighbor

- Who are you?
- Where do you work?
- What has been the proudest moment in your job?

What to Expect Today

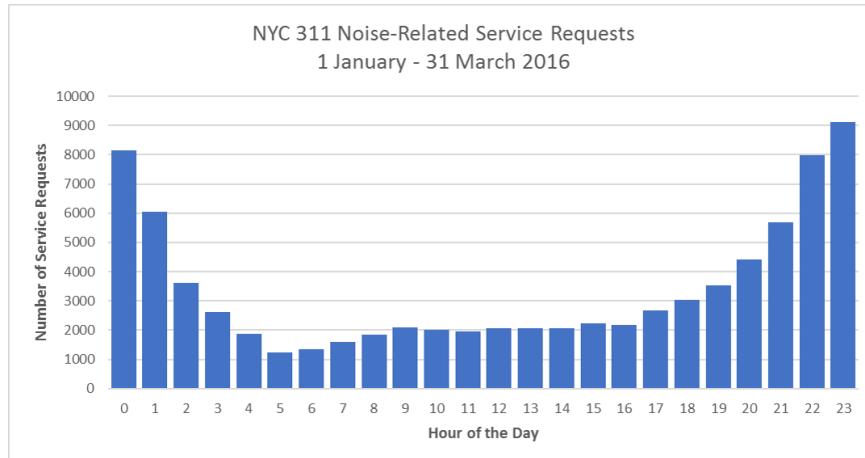
- 9:40 – Data Analytics 101
- 10:00 - Introduction to Problem Ideation
- 10:30 – 15 min break
- 10:45 – Process Mapping and the Types of Analysis
- 12:00 – Lunch
- 1:00 – Overview of Open Data
- 2:00 – Data Analytics Exercise
- 2:30 – 15 min break
- 2:45 – Data Analytics Exercise (continued)
- 4:30 – Dismissal

Housekeeping

- We'll have one 15 minute break in the morning
- We'll have an hour for lunch
- We'll have a 15 minute break in the afternoon
- Class will start promptly after breaks
- Feel free to use the bathroom if you need during class
- Please take any phone conversations into the hall to not disrupt the class

The Value of Data

- Data tells a story about something that's happened
- Can describe what happened directly or indirectly



Are All Data Points Created Equal?

“ Facts do not “speak for themselves.” They speak for or against competing theories. Facts divorced from theory or visions are mere isolated curiosities.

-Thomas Sowell *A Conflict of Visions*

Data Driven Decisions Require Humans



Image Credit: 100 lion, CC BY-SA 4.0, via Wikimedia Commons

How Intel lost out on the contract of a lifetime

‘ Data is only as valuable as the decisions it enables.

-[Ion Stoica](#)

‘ If data is enabling important decisions, then the data is important too

-[Richard Dunks](#)

What is Analysis?

‘ “Analysis is simply the pursuit of understanding, usually through detailed inspection or comparison”

-[Carter Hewgley](#), (Former) Director of Analytics, [Center for Government Excellence](#)

What Analysis Isn't

Data Analytics for Managers by Richard Dunks and Julia Mardon is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](#).



13 / 121

Writing a report



Image Credit: RRZEicons, [CC BY-SA 3.0](#), via [Wikimedia Commons](#)

Data Analytics for Managers by Richard Dunks and Julia Mardon is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](#).



14 / 121

Creating a dashboard



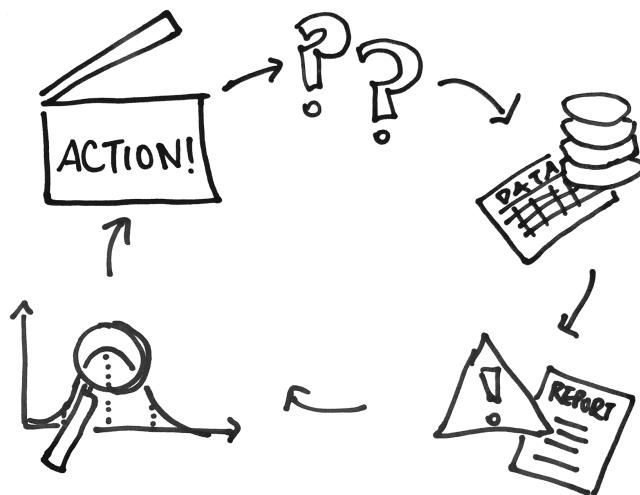
Image Credit: Chabe01, CC BY-SA 4.0, via Wikimedia Commons

Generating an alert



Image Credit: Tokyoship, CC BY-SA 3.0, via Wikimedia Commons

It's Putting Them All Together

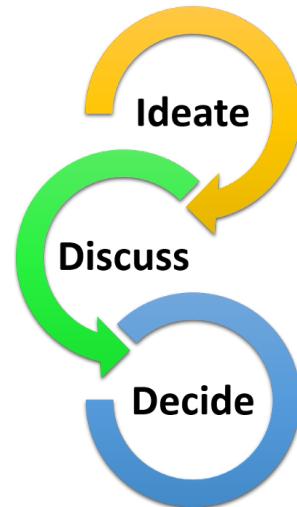


“If you do not know how to ask the right question, you discover nothing.”

- W. Edward Deming

Our Method For Generating Ideas

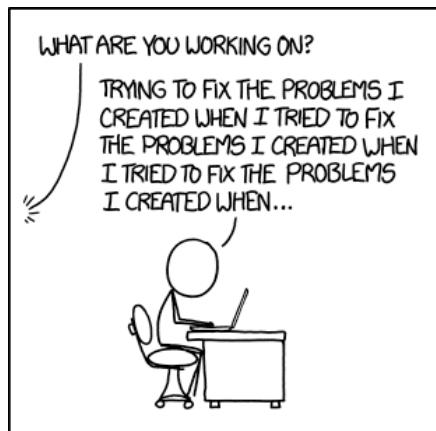
- **Ideate** - On your own, generate at least 3 ideas (ideally more), each on their own Post-It Note
- **Discuss** - Review the ideas generated
- **Decide** - Come to a consensus as a group



Exercise - Reducing Noise Complaints in NYC

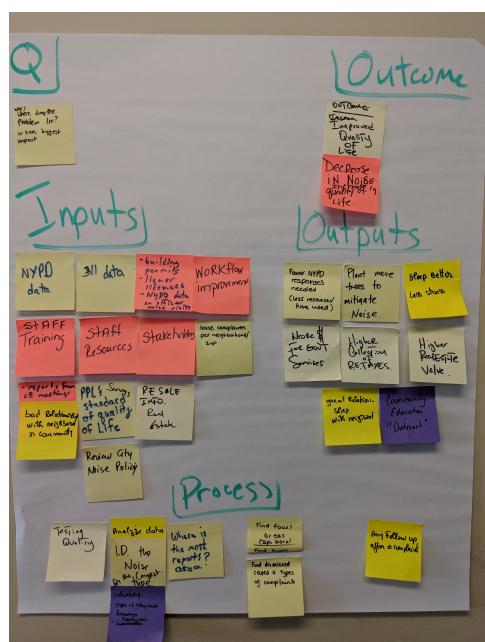
- Between 1 Jan and 30 Sept 2017, there were an average of 1,271 noise-related 311 service requests a day
- The same period in 2016 had an average of 1,180 noise-related 311 service requests a day
- You've been tasked with decreasing noise complaints in the city
- **What questions would you ask to kick off a data analysis?**

15 MIN BREAK



[Source](#)

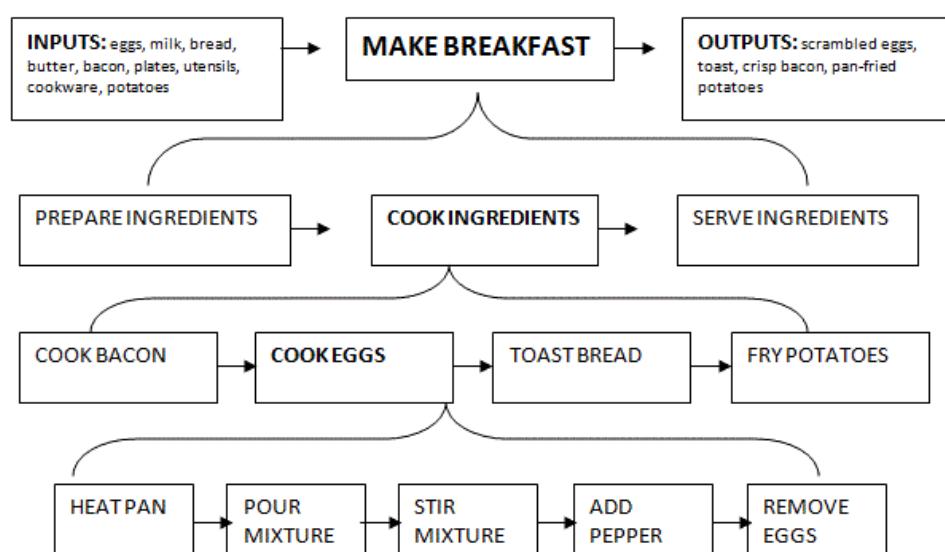
Process Mapping (Our method)



Process Mapping

- Allows you to identify and strategize for key steps in your analysis
- Helps sequence tasks and identify gaps in understanding
- Provides a basis for documenting work

Process Mapping

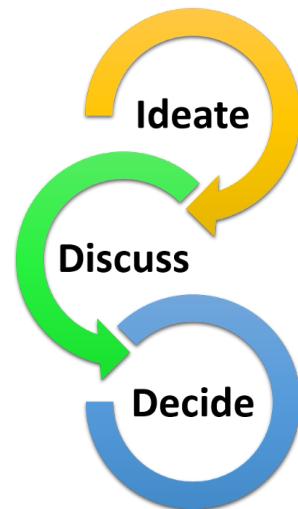


By Scottsm1991 (Own work) CC BY-SA 3.0, via Wikimedia Commons

How to Create a Process Map

- Identify the key challenge
- Identify the outcome
- Identify key way to validate the outcome (outputs)

How do we know we've got it right?

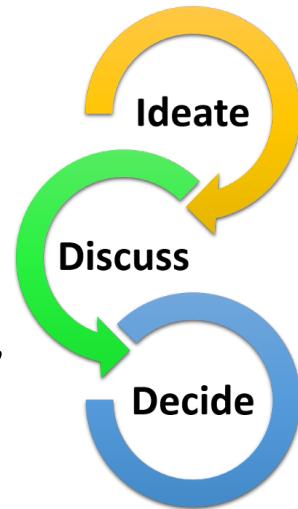


Outcomes vs Outputs

- Outcomes are the larger benefits and/or achievements you're trying to realize (happiness, health, well-being, etc.)
- Outputs are the tangible parts of your outcome (survey responses, measured results, etc.)
- Outputs enable us to find outcomes
- Without outcomes, there is no need for outputs

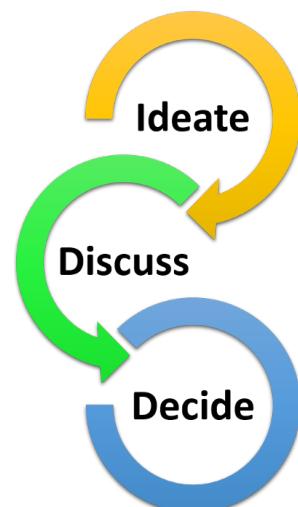
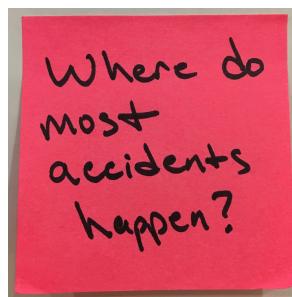
How to Create a Process Map

- Identify the key challenge
- Identify the outcome
- Identify key way to validate the outcome (outputs)
How do we know we've got it right?
- Identify the key inputs (data, partners, etc.)
- Sequence the key questions to turn inputs into outputs



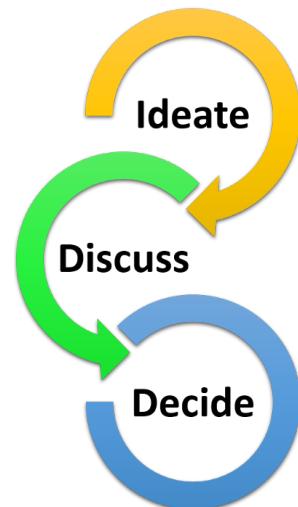
Key Tips

- Place each step on a Post-It Note



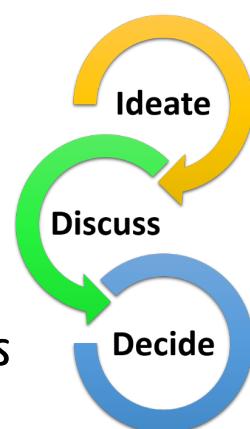
Key Tips

- Place each step on a Post-It Note
- Order and reorder as necessary
- Some steps will need to be broken down



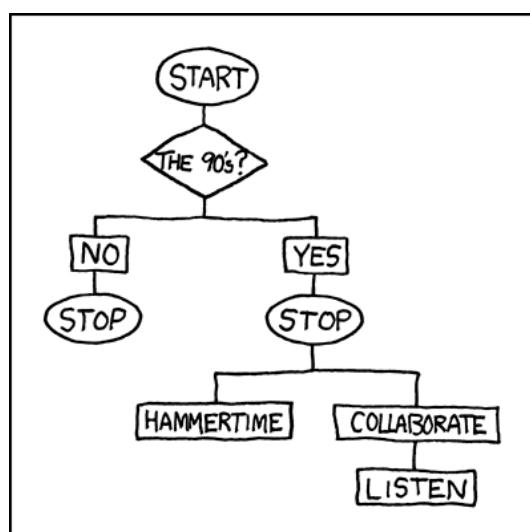
Reducing Noise Complaints

- Identify the key question/challenge
- Identify the key outcome
- Identify the outcome measures
- Identify the key inputs (data, partners, etc.)
- Sequence the key questions to turn inputs into outputs



WRAP-UP

LUNCH



[Source](#)

WELCOME BACK!

Link to NYC Open Data Portal for Exercise

[NYC Open Data Portal](#)

Data for Exercise

The screenshot shows a data visualization interface with two main sections: 'Filter' and 'Download'.
The 'Filter' section allows users to refine a dataset based on specific criteria. It includes dropdown menus for 'Created Date' (set to 'is between' 01/01/2016 12:00:00 AM and 04/01/2016 12:00:00 AM) and 'Complaint Type' (set to 'contains' 'noise'). Both of these filter criteria are highlighted with red boxes.
The 'Download' section provides options to export the filtered dataset in various formats. The 'Download As' dropdown menu lists 'CSV', 'CSV for Excel' (which is also highlighted with a red box), 'JSON', 'RDF', 'RSS', 'TSV for Excel', and 'XML'.
Below the interface, a blue link reads 'Click to download if you have problems'.

5 Data Analytics Tasks

1. Sorting
2. Filtering
3. Aggregating (PivotTable)
4. Manipulating
5. Visualizing

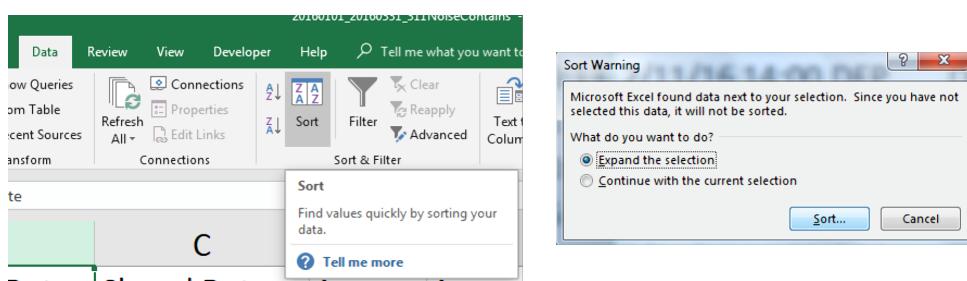
1. Sorting

- Reorganize rows in a dataset based on the values in a column
- Can sort on multiple columns

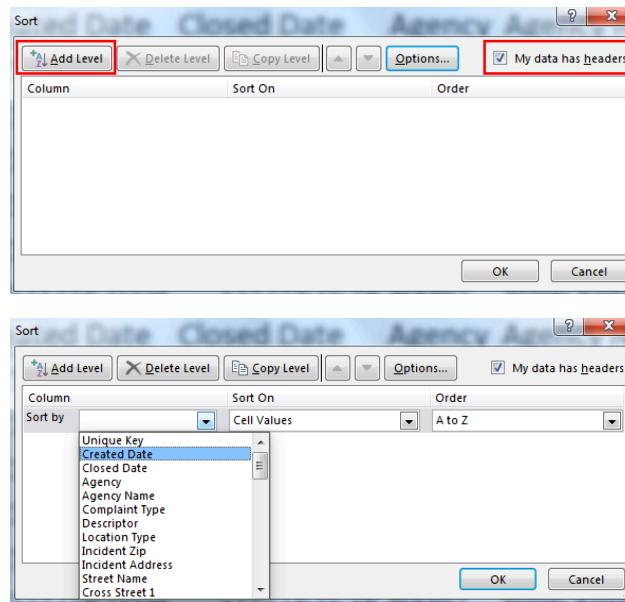
Sorting by Date

	A	B	C
1	Unique Key	Created Date	Closed Date
2	32576232	2/1/16 10:13	2/11/16 14:00
3	32884992	3/11/16 13:48	3/12/16 8:00
4	32445685	1/18/16 10:56	1/20/16 8:00
5	32816604	3/2/16 15:17	3/3/16 8:00
6	32480740	1/22/16 10:19	1/24/16 20:00
7	32518960	1/25/16 7:35	1/28/16 15:45

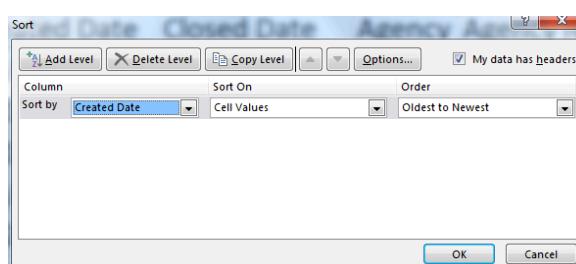
	A	B	C
1	Unique Key	Created Date	Closed Date
2	32305299	1/1/16 0:00	1/1/16 1:57
3	32310343	1/1/16 0:00	1/1/16 3:12
4	32308578	1/1/16 0:02	1/1/16 23:35
5	32305983	1/1/16 0:03	1/1/16 3:24
6	32305208	1/1/16 0:03	1/1/16 2:43
7	32309484	1/1/16 0:04	1/1/16 0:28



Sorting by Date



Sorting by Date



	A	B	C
1	Unique Key	Created Date	Closed Date
2	32305299	1/16 0:00	1/16 1:57
3	32310343	1/16 0:00	1/16 3:12
4	32308578	1/16 0:02	1/16 23:35
5	32305983	1/16 0:03	1/16 3:24
6	32305208	1/16 0:03	1/16 2:43
7	32309484	1/16 0:04	1/16 0:28

2. Filtering

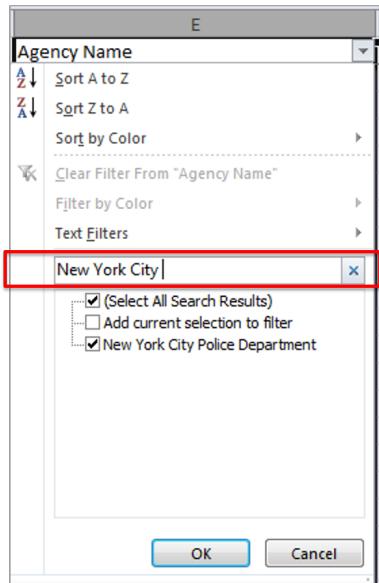
- Only show rows that contain some value
- Can filter by multiple values
- Can filter by values in multiple columns

Filtering by Agency Name



	A	B	C	D	E	F
1	1	Created Date	Closed Date	Agency Name	Complaint Type	

Filtering by Agency Name

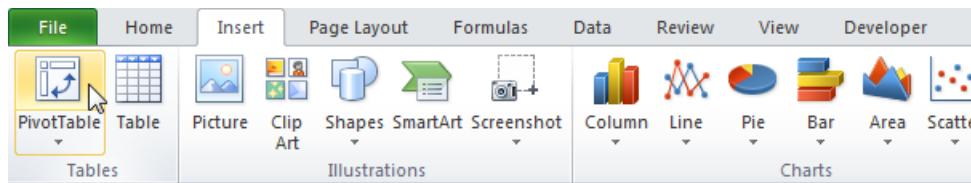


3. Aggregating Data

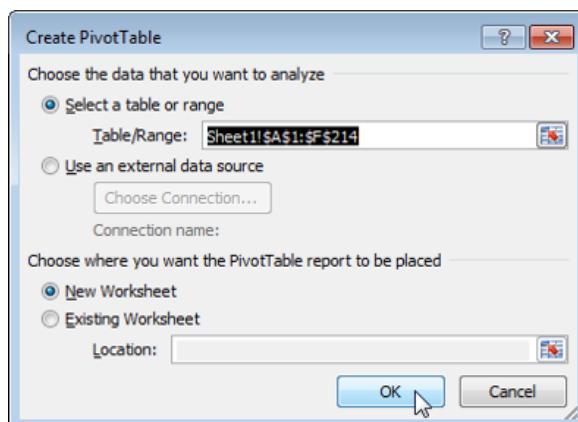
- Trends only become clear in aggregate
- Often where you discover the "so what"
- Aggregating data meaningfully can be tricky

PivotTables

- A data summarization tool
- Useful to quickly understand data
- Can use to graph data totals



Creating a PivotTable

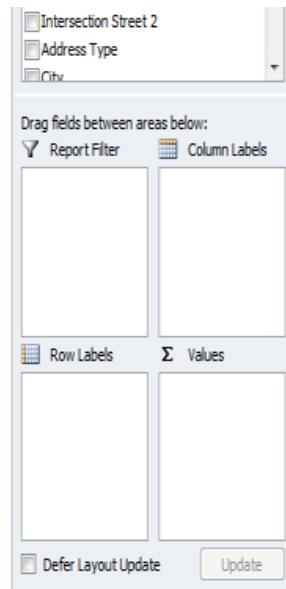


- Should default to all your data unless you have any cells selected
- Should default to a new worksheet

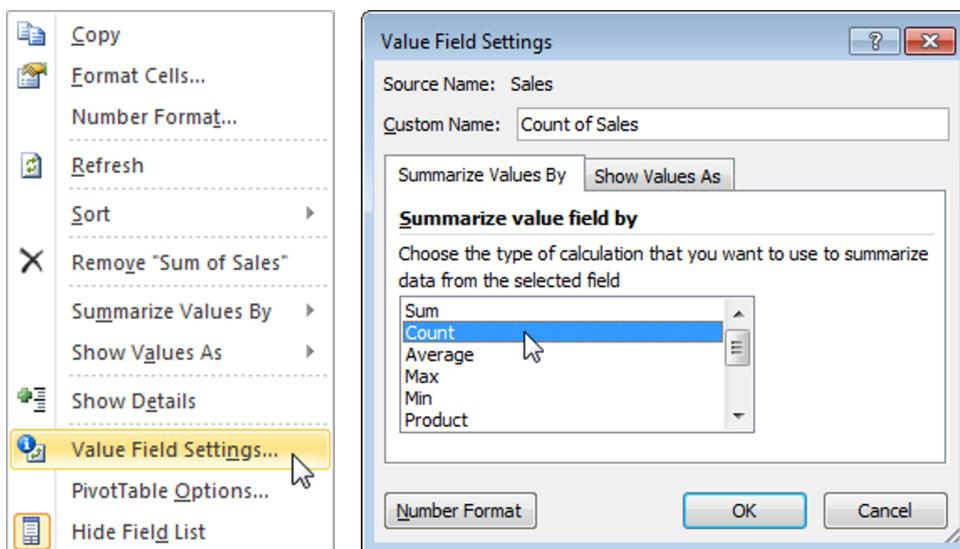
Creating a PivotTable

Drag and drop fields to visualize

- Row labels
- Values
- Filter
- Column Labels



Creating a PivotTable



4. Manipulating Data (In a good way...)

- Sometimes available categories don't make sense
- Values may not be in the format you need (or have mistakes)
- You always want to have a clean copy of the data to go back to
- Best to keep track of what you've done

Extracting Hour From Timestamp

A	B	C
Unique Key	Created Date	hour
30195273	3/18/15 2:12	=HOUR(B2)
30203057	3/18/15 2:00	
30197320	3/18/15 1:58	
30194112	3/18/15 1:37	
30202379	3/18/15 1:36	
30199506	3/18/15 1:28	

HOUR function

This article describes the formula syntax and usage of the HOURfunction in Microsoft Excel.

Description

Returns the hour of a time value. The hour is given as an integer, ranging from 0 (12:00 A.M.) to 23 (11:00 P.M.).

Syntax

`HOUR(serial_number)`

The HOUR function syntax has the following arguments:

- **Serial_number** Required. The time that contains the hour you want to find. Times may be entered as text strings within quotation marks (for example, "6:45 PM"), as decimal numbers (for example, 0.78125, which represents 6:45 PM), or as results of other formulas or functions (for example, `TIMEVALUE("6:45 PM")`).

Extracting Hour From Timestamp

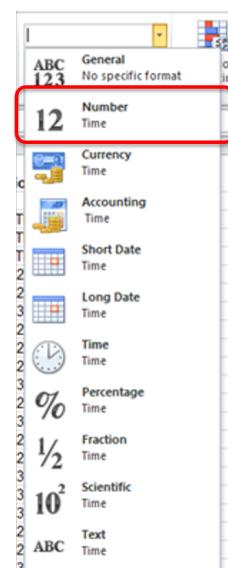
Unique Key	Created Date	hour
30195273	3/18/15 2:12	1/3/04 0:00
30203057	3/18/15 2:00	
30197320	3/18/15 1:58	
30194112	3/18/15 1:37	
30202379	3/18/15 1:36	
30199506	3/18/15 1:28	

Extracting Hour From Timestamp

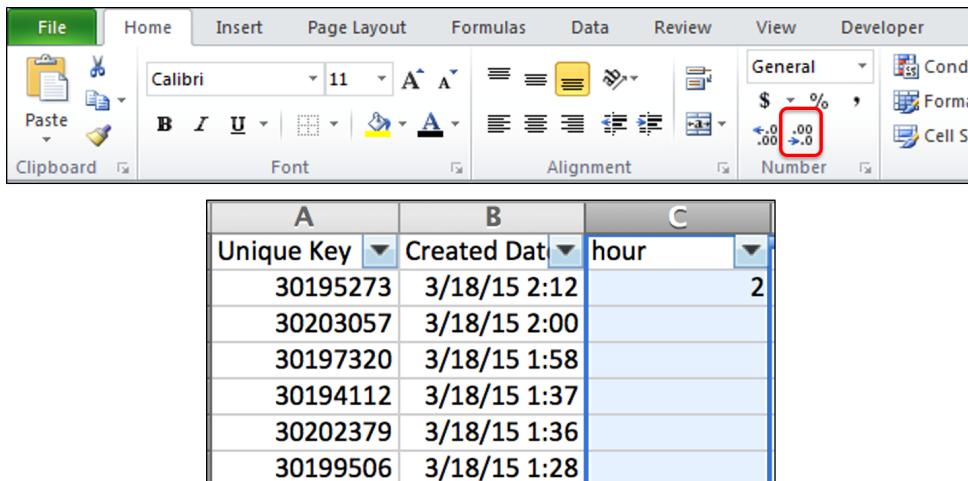
A	B	C
Unique Key	Created Date	hour
30195273	3/18/15 2:12	1/3/04 0:00
30203057	3/18/15 2:00	
30197320	3/18/15 1:58	
30194112	3/18/15 1:37	
30202379	3/18/15 1:36	
30199506	3/18/15 1:28	

Extracting Hour From Timestamp

A	B	C
Unique Key	Created Date	hour
30195273	3/18/15 2:12	2.00
30203057	3/18/15 2:00	
30197320	3/18/15 1:58	
30194112	3/18/15 1:37	
30202379	3/18/15 1:36	
30199506	3/18/15 1:28	



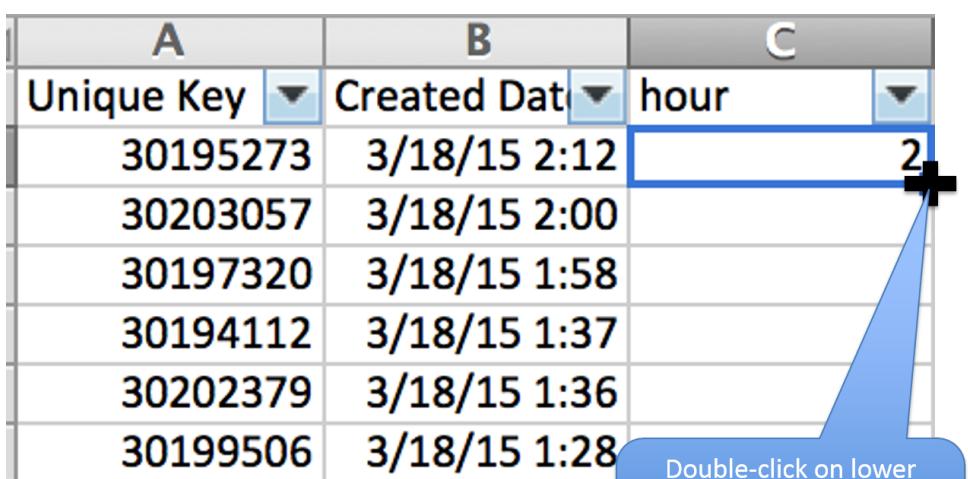
Extracting Hour From Timestamp



A screenshot of Microsoft Excel showing a table with three columns: Unique Key, Created Date, and hour. The Unique Key column contains values like 30195273, 30203057, etc. The Created Date column contains timestamp values like 3/18/15 2:12, 3/18/15 2:00, etc. The hour column contains the value 2. The formula bar shows the formula =HOUR(B2). The Excel ribbon is visible at the top, and the status bar at the bottom indicates "Data Analytics for Managers by Richard Dunkle and Julia Mardon is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License".

A	B	C
Unique Key	Created Date	hour
30195273	3/18/15 2:12	2
30203057	3/18/15 2:00	
30197320	3/18/15 1:58	
30194112	3/18/15 1:37	
30202379	3/18/15 1:36	
30199506	3/18/15 1:28	

Extracting Hour From Timestamp



A screenshot of Microsoft Excel showing a table with three columns: Unique Key, Created Date, and hour. The Unique Key column contains values like 30195273, 30203057, etc. The Created Date column contains timestamp values like 3/18/15 2:12, 3/18/15 2:00, etc. The hour column contains the value 2. The formula bar shows the formula =HOUR(B2). A blue callout bubble points to the bottom-right corner of the cell containing the value 2, with the text "Double-click on lower right hand corner to expand formula to bottom of column". The Excel ribbon is visible at the top, and the status bar at the bottom indicates "Data Analytics for Managers by Richard Dunkle and Julia Mardon is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License".

A	B	C
Unique Key	Created Date	hour
30195273	3/18/15 2:12	2
30203057	3/18/15 2:00	
30197320	3/18/15 1:58	
30194112	3/18/15 1:37	
30202379	3/18/15 1:36	
30199506	3/18/15 1:28	

Double-click on lower
right hand corner to
expand formula to bottom
of column

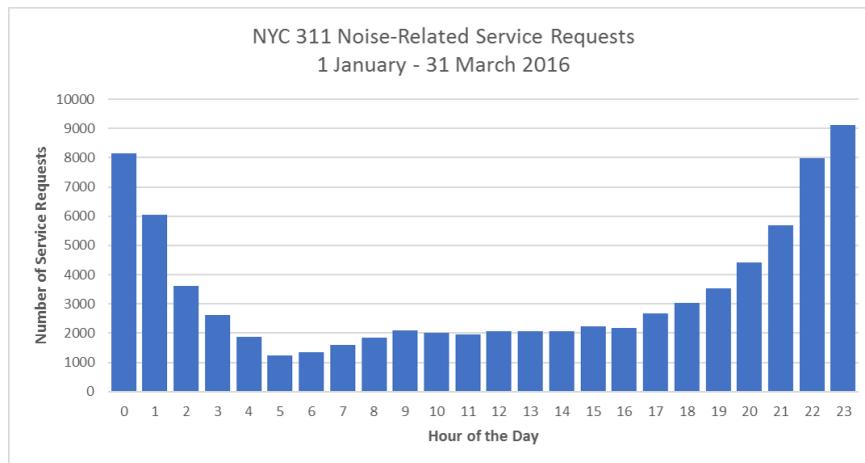
Extracting Hour From Timestamp

A	B	C
Unique Key	Created Date	hour
30195273	3/18/15 2:00	2
30203057	3/18/15 2:00	2
30197320	3/18/15 1:58	1
30194112	3/18/15 1:37	1
30202379	3/18/15 1:36	1
30199506	3/18/15 1:28	1

5. Visualizing Data

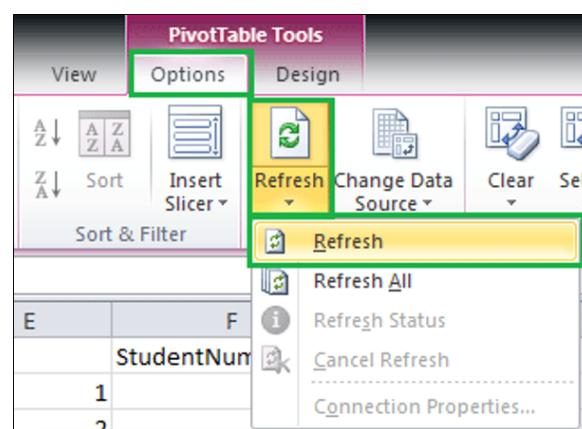
- Quickly communicate information
- Tell a clearer story
- A picture is worth a thousands words

When Are Noise Complaints Received?



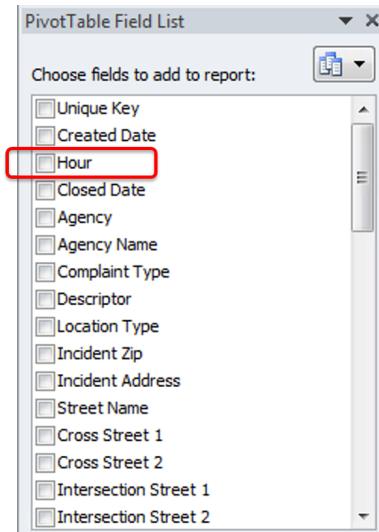
Charting Noise Complaints by Hour

- Refresh PivotTable



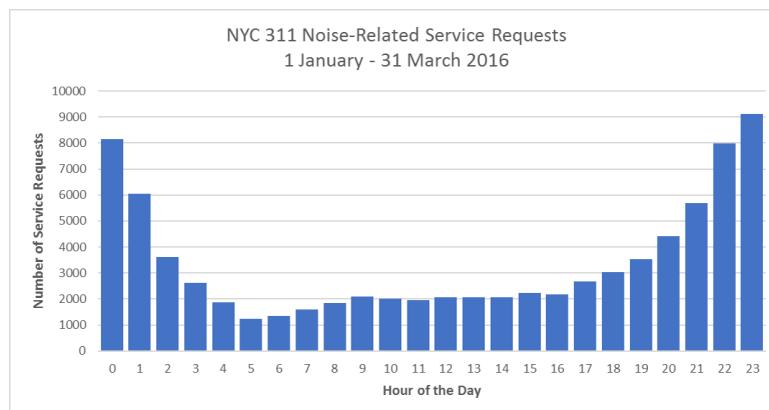
Charting Noise Complaints by Hour

- Refresh PivotTable
- Find **Hour** and add it to "Rows"



Charting Noise Complaints by Hour

- Refresh PivotTable
- Find **Hour** and add it to "Rows"
- Style the chart



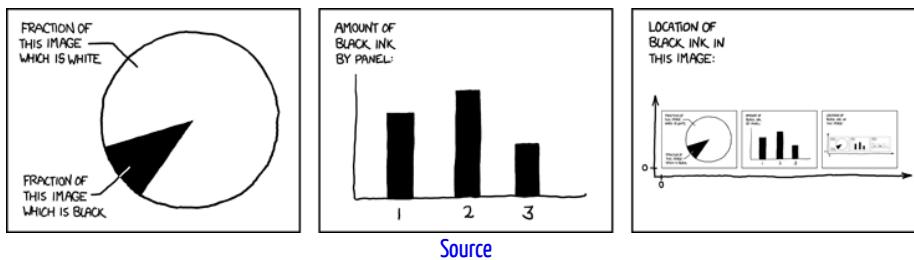
Exploratory Data Analysis

- Goal -> Discover patterns in the data
- Understand the context
- Summarize fields
- Use graphical representations of the data
- Explore outliers

Tukey, J.W. (1977). Exploratory data analysis. Reading, MA: Addison-Wesley

Exploring the Data

15 MIN BREAK



Source

Let's get back to our Process Map

Outline the Process

- What are the steps to creating our outputs?
- What is the best order of steps?
- How granular do we need to break this down for clarity?

Collection
Verification
Description
Sensemaking
Communication/Visualization

<p class="footer">

67 / 121

4 Concerns

1. Technical

- Having the right tools
- Having the people who can use them
- Making everything work together
- *Potential trap: having a solution in search of a problem*

2. Legal

- Laws
- Regulations
- Practices
- *Potential trap: not doing something because of mistaken assumptions*

When HIPAA Gets In The Way of Health Care



Image Credit: Hipaa Nurse Shredding Papers, by [Atlantic Training](#). CC BY-SA 3.0

[Learn more](#)

3. Cultural

- “We’ve always done it this way”
- “I’m not sure I understand how this works”
- *Potential trap: being afraid of rocking the boat*

4. Political

- Inter-departmental
- Intra-departmental
- Public relations
- *Potential trap: not putting the necessary effort into something that will pay dividends to your office and ultimately to the organization as a whole*

Benefits

- Time, money, lives saved
- Better delivery of services to stakeholders
- More transparency
- More accountability

What could be some of the benefits we'd realize from what we've mapped?

What did you notice about this process?

- Knowing the problem and sequencing the steps can be harder than working with data
- Having these answers can make the analysis much easier
- This is a process you can do with any challenge, no matter how big
- Everyone has something to add, no matter how technical (or non-technical) they are

Let's Work Our Plan

NYC Community Districts

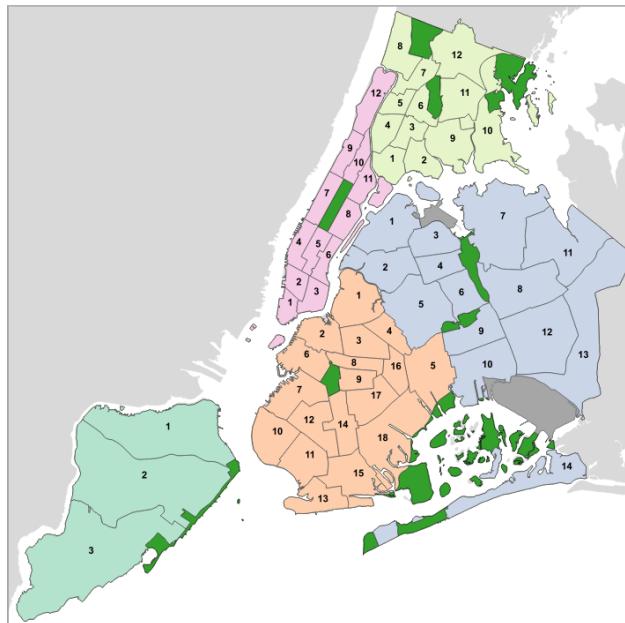


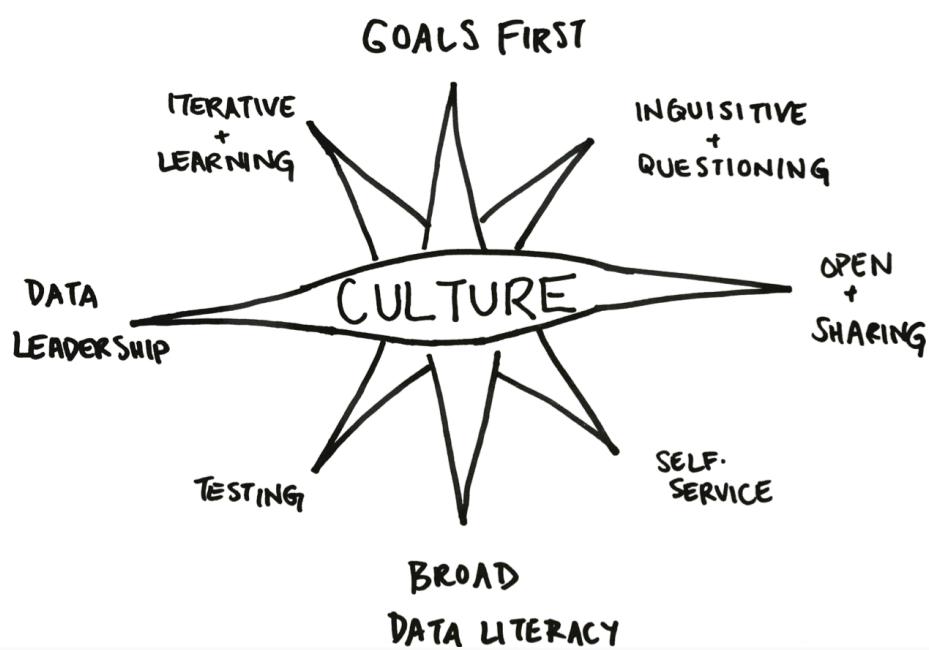
Image Credit: Data**politan** CC BY-SA 4.0

Links to Other Open Data Portals

- New York State - <https://data.ny.gov>
- US Federal Government - <https://www.data.gov/>

BUILDING A DATA-DRIVEN CULTURE

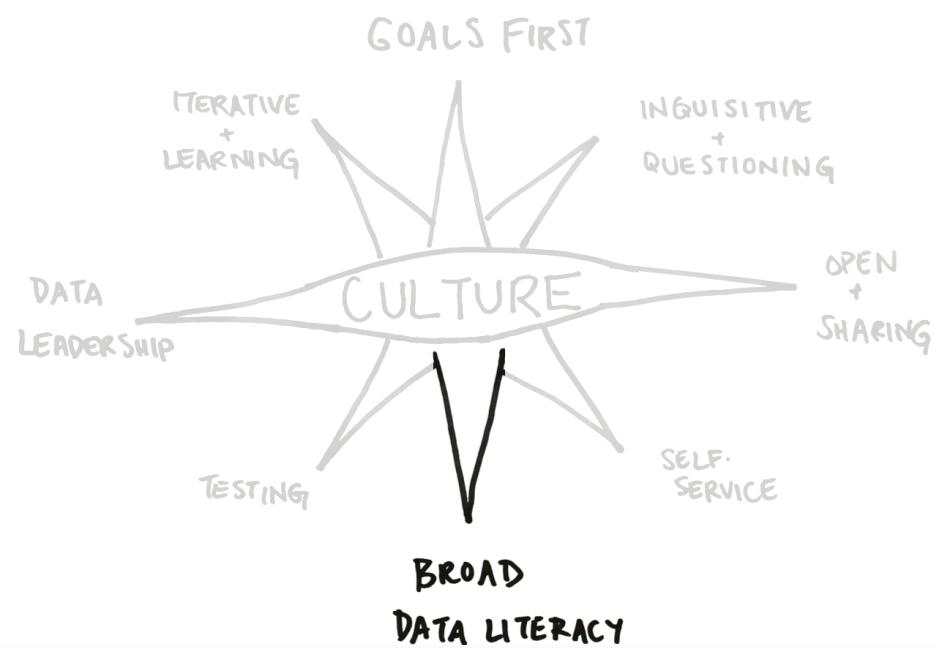
Data Driven Culture



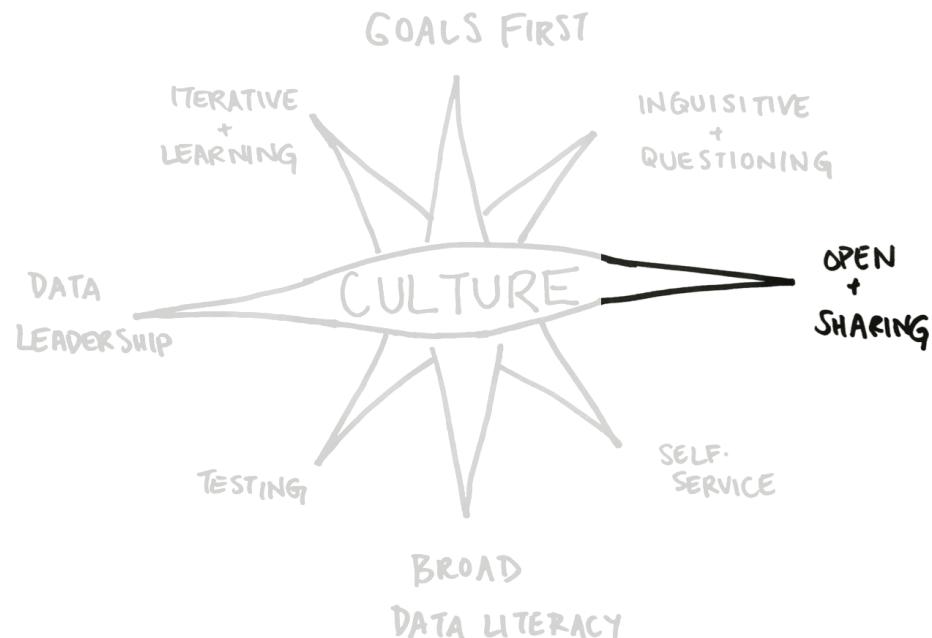
Data Driven Culture



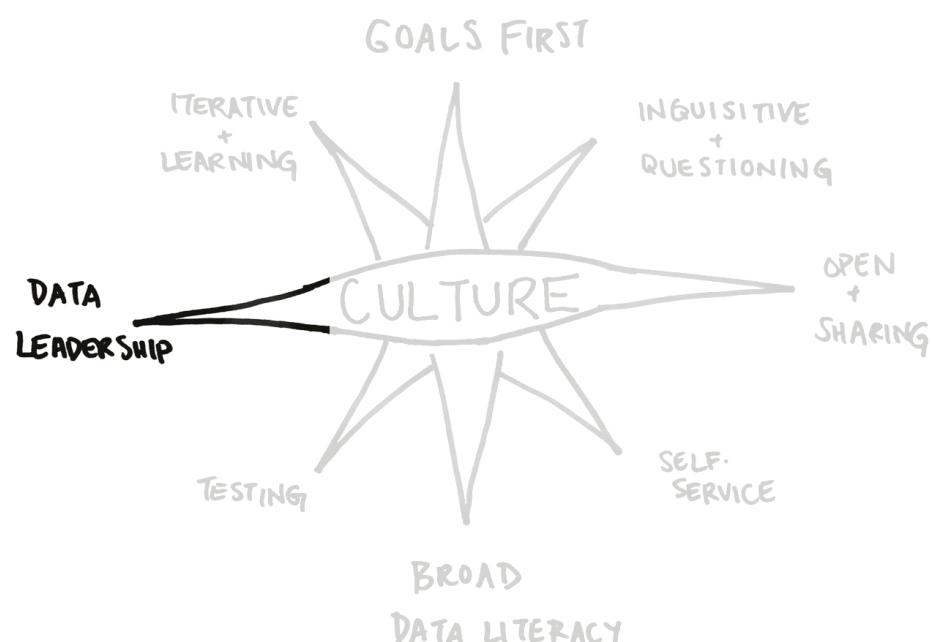
Data Driven Culture



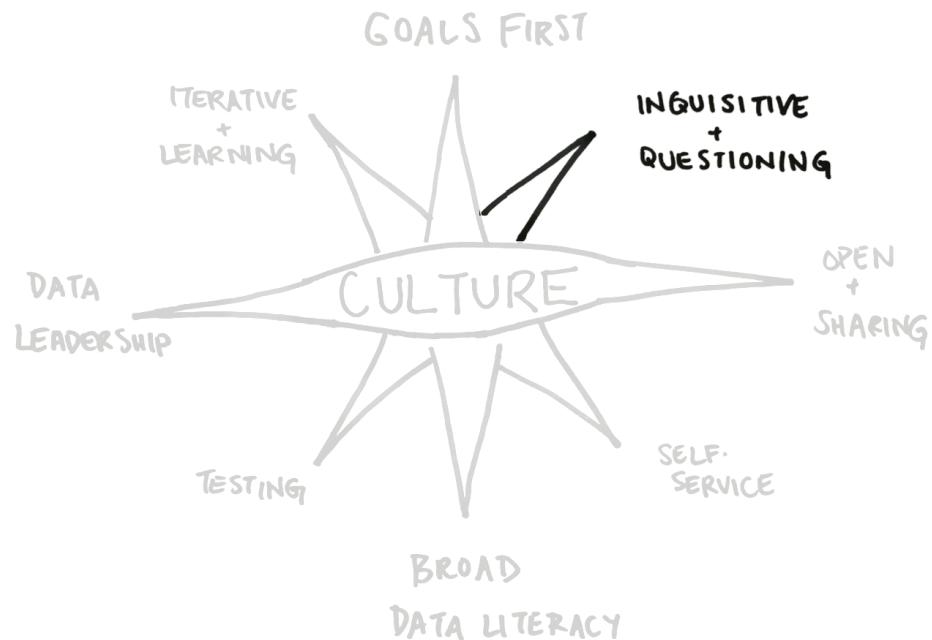
Data Driven Culture



Data Driven Culture



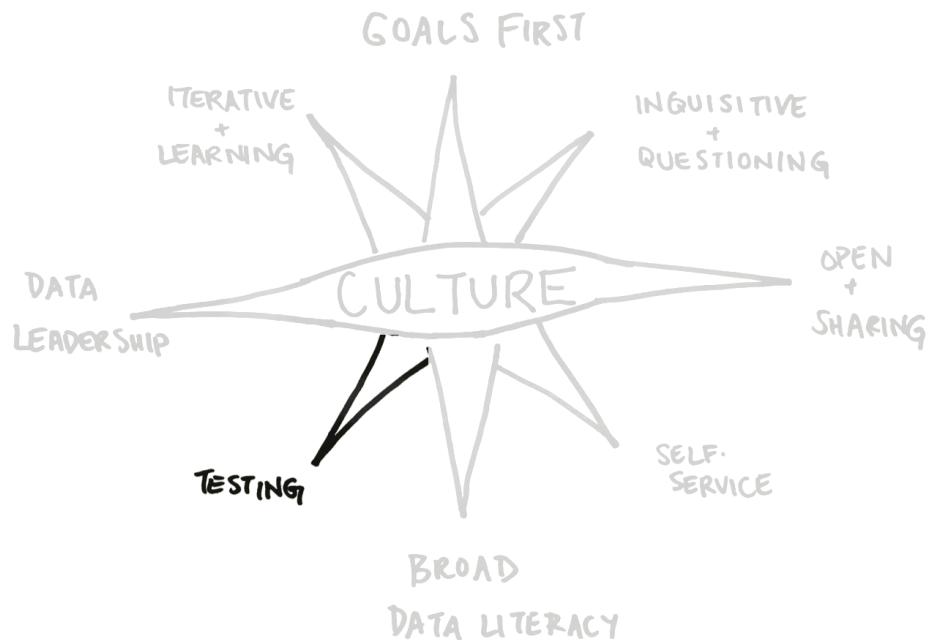
Data Driven Culture



Data Driven Culture



Data Driven Culture



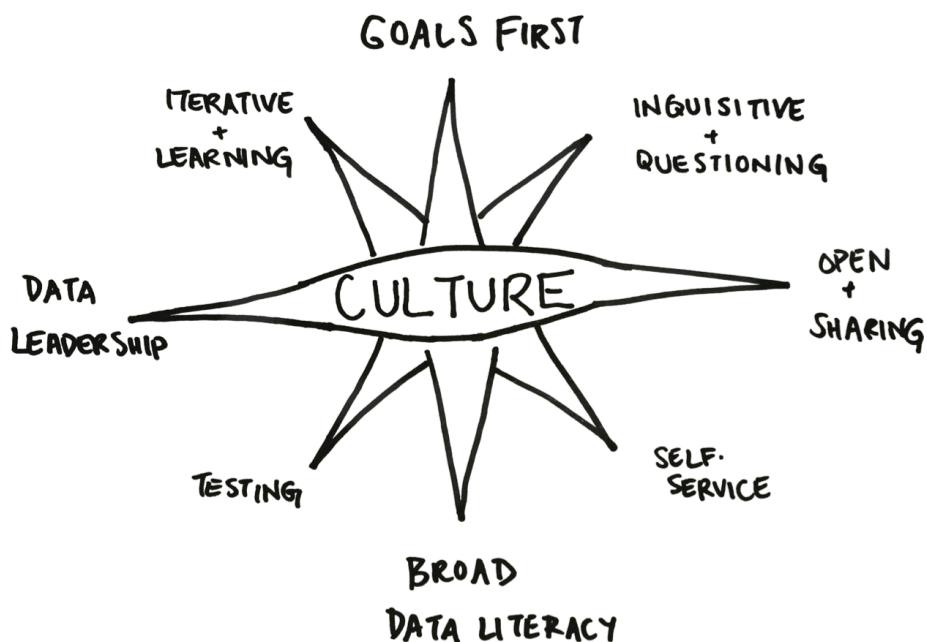
Data Driven Culture



Data Driven Culture



Data Driven Culture



Data Driven Culture

“Do you have data to back that up?” should be a question that no one is afraid to ask (and everyone is prepared to answer)

- Julie Arsenault

[Source](#)

WRAP-UP

What We've Covered

- Elements of a data-driven culture
- Types of analysis
- The analytics process
- What else?

What might you do different with this information when you go back to your offices?

Final Thoughts

- Data can tell a story, but doesn't speak for itself
- Analysis is the search for understanding and where we learn to tell that story
- Be good to your data and it will be good to you

Technical Support

- [Microsoft Office Support](#) - Documentation on various MS Office products
- [Data Science Cheatsheet](#) - Includes various terms and concepts related to data science
- [Open Data Handbook](#) - Guides, case studies and resources for government & civil society on the "what, why & how" of open data
- [Copy of today's handout](#)

Resources

- Harvard Government Performance Lab Results-Driven Contracting
- Carl Anderson *Creating a Data-Driven Organization*
- DJ Patil & Hilary Mason *Data Driven: Creating a Data Culture*
- IDEO Design Kit - Resource for design thinking techniques
- Datapolitan training classes

Contact Information

Richard Dunks

- Email: richard[at]datapolitan[dot]com
- Website: <http://www.datapolitan.com>
- Twitter: [@datapolitan](#)

THANK YOU!

Types of Analysis

THE NOLALYTICS
CIVIC ANALYTICS TYPOLOGY

A TYPOLOGY OF PROJECTS FORMS OUR TEMPLATE FOR DEVELOPING NEW USE CASES WITH DEPARTMENTS

CITY OF NEW ORLEANS

Civic Analytics Network

See more at <http://datadriven.nola.gov/>

Based on the work of the [City of New Orleans, Office of Performance and Accountability](#)

Graphics: Copyright © Harvard University Ash Center (Used with Permission)



Finding the
needle in a
haystack



Prioritizing
work for impact



Early warning
tools



Better, quicker
decisions



Optimizing
resource
allocation



Experimenting
for what works

Finding a Needle in Haystack

- Challenge: Targets are difficult to identify or locate within a broader population
- Opportunity: Data analysis and predictive modeling to identify targets based on existing data



Finding the
needle in a
haystack

New Orleans Distributes Smoke Alarms



Image Credit: Michael Barnett [CC BY-SA 2.5](#), via Wikimedia Commons

Targeted Outreach Saves Lives

At-Risk Fire Zones Selected for Initial Outreach

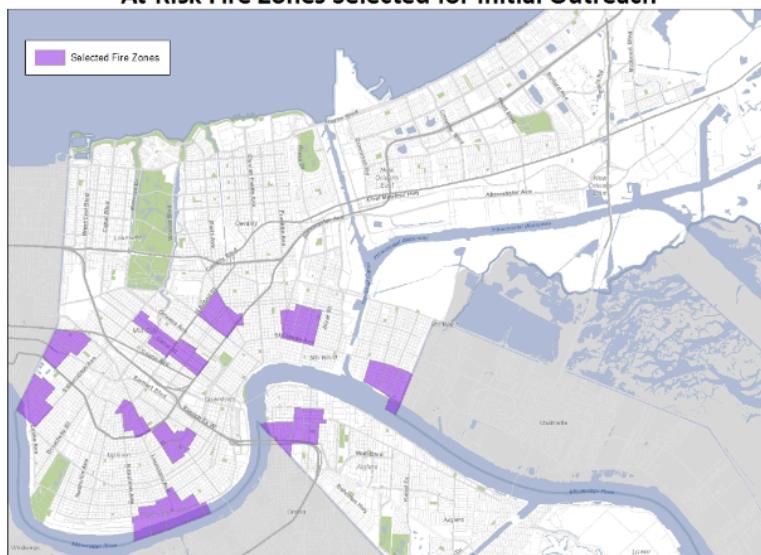


Image Credit: City of New Orleans, via [nola.gov](#)

Targeted Outreach Saves Lives

Predictive power of model

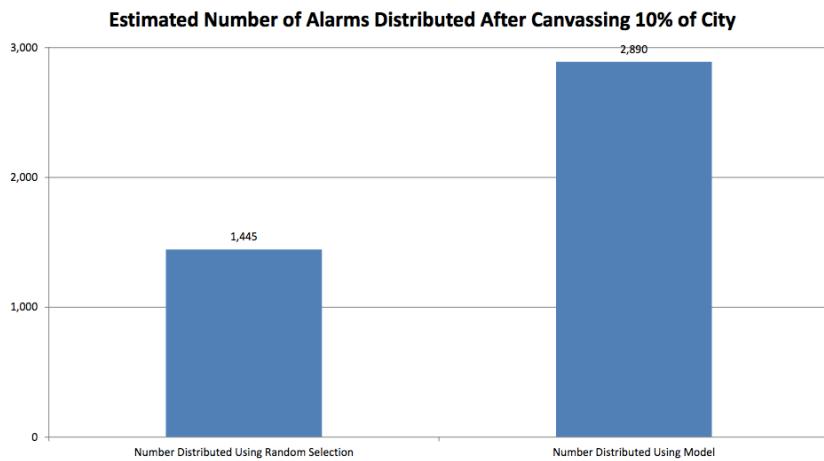


Image Credit: City of New Orleans, via [nola.gov](#)

Prioritizing Work for Impact

- Challenge: Services do not categorize high-priority cases early
- Opportunity: Data analysis and predictive modeling to prioritize cases



Prioritizing
work for impact

NYC Restaurant Inspectors Save Time



Image Credit: Grease Trap Cover Asphalt by Christopher Sessums, CC BY 2.0

Source: http://www.nyc.gov/html/dep/html/press_releases/12-71pr.shtml#.W6v2nRNKhZo

Early Warning Tools

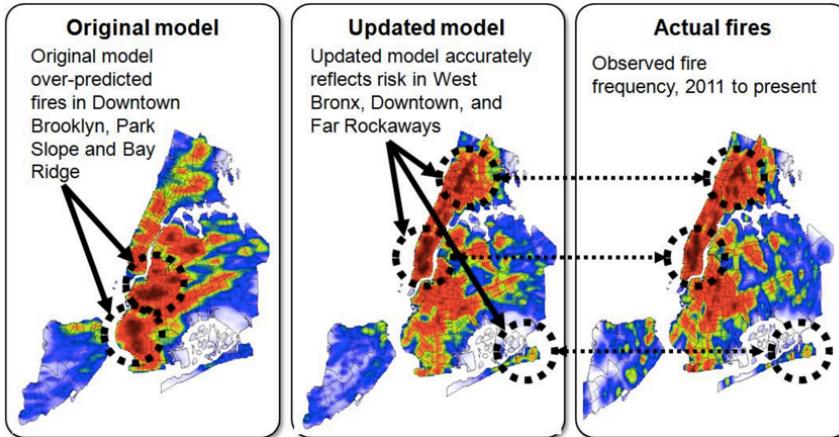
- Challenge: Resources overly focused on reactive services
- Opportunity: Developing tools to predict need based on historic patterns



Early warning tools

Using Data to Fight Fires

FDNY's Risk Based Inspection System (RBIS)



The three maps above of NYC represent the progress of the risk based inspection system algorithm that MODA created with the Fire Department (FDNY). At left is the original map of high-risk zones from the first version of the FDNY RBIS model; center is the updated model, weighting risk criteria based on MODA's statistical regression; at right is an actual map of recent fires in New York City.

Image Credit: NYC Mayor's Office of Data Analytics, 2013 Annual Report

Better, Quicker Decisions

- **Challenge:** Repeated decisions are made without access to all relevant information
- **Opportunity:** Developing recommendation tools for operational decisions



Better, quicker decisions

Cincinnati Targets Urban Blight



Image Credit: Wholtone, via [Wikimedia Commons](#)

[Learn more](#)

Optimizing Resource Allocation

- **Challenge:** Assets are scheduled or deployed without input of latest service data
- **Opportunity:** Use data to drive decisions on the deployment of resources



Optimizing
resource
allocation

Shortening Public School Bus Rides



Image Credit: [Patrick Hudepohl](#), used under [Creative Commons BY-NC-SA 2.0 license](#).

[Source](#)

Experimenting for What Works

- Challenge: Services have not been assessed for impact
- Opportunity: Experimental testing and improvement of service options



Experimenting
for what works

Redesign of NYC Summons

CRC-326 (512)
Criminal Court Appearance Ticket

Name (Last, First, MI) _____ Date of Birth (mm/dd/yy) _____

Offense Number (where court may contact you) _____ Name from Muster (where court may contact you) _____

Show up to court: _____ Court Appearance Date (mm/dd/yy): _____ at: 9:30 a.m.

Your court appearance location: _____ (where you appear)
 Bronx Criminal Court Brooklyn Community Justice Center Bronx Community Justice Center Kings Criminal Court
 Bronx Borough Courthouse Bronx County Courthouse Bronx Family Court Bronx Juvenile Court
 Bronx Supreme Court Bronx Trial Court Bronx Traffic Court Bronx Appellate Division Court

To avoid a warrant for your arrest, you must show up to court.*
At court, you may plead guilty or not guilty.

Please note that in the execution of Public Conveyance of Alcohol and Public Urination offenses

Court Locations: You must appear at the court location indicated below.
Bronx Criminal Court - 215 E. 167th Street, Bronx, NY 10451
Kings & Bronx Criminal Court - 345 Broadway, New York, NY 10013
Bronx Community Justice Center - 184-34 Myrtle Place, Brooklyn, NY 11223
Bronx Community Justice Center - 314 W. 181st Street, Bronx, NY 10451
Queens Criminal Court - 120-55 Queens Boulevard, Kew Gardens, NY 11415
Richmond Criminal Court - 26 Centre Ave, Staten Island, NY 10301

Title of Offense:
Title 21 (Hour Law) _____ Date of Offense (mm/dd/yy) _____ County _____
Place of Occurrence _____ District _____

Violation of: Subsection _____ Part _____ Subpart _____ Article _____

For Additional Information and Questions:
Visit the website or call the number below for additional information about your court appearance: www.nycsummons.nyc
OR
Call 646-360-3010

Defendant stated in my presence (in substance): _____

I solemnly declare that the contents of the offense charged above, fully statement made before me, punishable as a Class A Misdemeanor pursuant to section 210.03 of the Penal Law. Affirmed under penalty of law.

Complainant's (SI) Name Printed _____ Rank/Title Signature of Complainant _____ Date Affirmed _____

Tax Registry # _____ Agency _____ Command Code _____

DEFENDANT'S COPY

CRC-326 (512)
Complaint/Information
The People of the State of New York vs. _____

Name (Last, First, MI) _____ Age, _____ Agt. No. _____

Street Address _____ City _____ State _____ Zip Code _____

ID/Assume Number _____ State _____ Type/Class _____ Expires (mm/yy) _____ Sex _____

Date of Birth (mm/yy) _____ M _____ W _____ Y _____ H _____ M _____ P _____ Reg _____

Reg. State _____ Expires (mm/yy) _____ Plw Type _____ Vw Type _____ Make _____ Year _____ Color _____

The Person Described Above Is Charged as Follows:

Time 24 Hour (hh:mm) _____ Date of Offense (mm/yy) _____ County _____

Place of Occurrence _____

In Violation of Section _____ Subsection _____ VTL _____ Admin _____ Prob _____ Park _____ Other _____

Title of Offense:

Bronx Criminal Court - 215 E. 167th Street, Bronx, NY 10451
Kings Criminal Court - 346 Broadway, New York, NY 10013
Bronx Community Justice Center - 184-34 Myrtle Place, Brooklyn, NY 11223
New York Criminal Court - 314 W. 181st Street, Bronx, NY 10451
Midtown Criminal Court - 314 W. 42nd Street, New York, NY 10018
Queens Criminal Court - 120-55 Queens Boulevard, Kew Gardens, NY 11415
Richmond Criminal Court - 67 Targis Street, Staten Island, NY 10301

Defendant stated in my presence (in substance): _____

I personally observed the commission of the offense charged herein. False statements made before me are punishable as a Class A Misdemeanor pursuant to section 210.03 of the Penal Law. Affirmed under penalty of law.

Complainant's Full Name Printed _____ Rank/Title Signature of Complainant _____ Date Affirmed _____

Agency _____ Tax Registry # _____ Command Code _____

Complainant's Name Printed _____ Rank/Title Signature of Complainant _____ Date Affirmed _____

Complainant's Name Printed _____ Rank/Title Signature of Complainant _____ Date Affirmed _____

The person described above is summoned to appear at NYC Criminal Court _____ Summoned Per: _____ County _____

Date of Appearance (mm/yy) _____ At 9:30 a.m. _____

Image credit [City of New York](#). Click for more information

Definition of Open Data

Open data is data that can be freely used, shared and built-on by anyone, anywhere, for any purpose

- Open Knowledge International

Key Features of Open Data

- Availability and access
- Reuse and redistribution
- Universal participation

Open Data Benefits

- Transparency
- Releasing social and commercial value
- Participation and engagement

Keeping NYC Accountable on Parking Tickets



Image Credit: Parking Violator by [Atomische ✘ Tom Giebel](#), CC BY-NC-ND 2.0

Source: <http://iquantny.tumblr.com/post/87573867759/success-how-nyc-open-data-and-reddit-saved-new>

Open Data Concerns

- Privacy ([PII and the Mosaic Effect](#))
- Accuracy
- Security

When Good Data Goes Bad



Image Credit: Kenny Louie, CC BY 2.0, via [Wikimedia Commons](#).jpg)

Gawker matches Taxi and Limousine Data with Paparazzi Photos