

 **NYC**  
Citywide Administrative Services

Management Academy  
Analytics for the Data-Driven Manager  
Richard Dunks and Julia Marden

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

**Datapolitan**  
Data Solutions for the Modern Metropolis

---

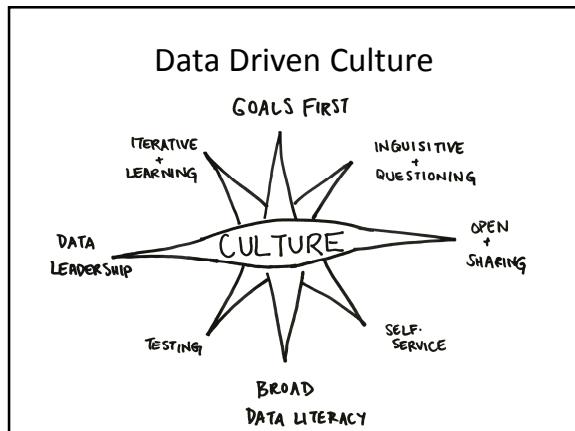
---

---

---

---

---



**Introductions**

Turn to your neighbor and introduce yourself with the following:

- Who you are
- Where you work
- One data-related project you've worked on

You will introduce them to the class

---

---

---

---

---

---

## What to Expect for Today

- 9:15 – Welcome and Introductions
- 9:45 – Data Driven Government
- 10:30 – 15 min break
- 10:45 – Data Analytics and the Data Driven Culture
- 12:00 – Lunch
- 1:00 – Data Analytics Process and Open Data
- 1:55 – 5 min break
- 2:00 – Data Analytics Exercise
- 2:50 – 10 min break
- 3:00 – Data Analytics Exercise (continued)
- 4:30 – Dismissal

*All times approximate*

## Housekeeping

- We'll have one 15 minute break in the morning
  - We'll have an hour for lunch
  - We'll try for two breaks 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 hallway to not disrupt the class

Intel chips could have powered first iPhone, CEO Otellini says

By AppleInsider Staff

Thursday, May 16, 2013, 02:42 pm PT (05:42 pm ET)

As Paul Otellini relinquishes the reins of chipmaker Intel, the outgoing chief executive reveals in an in-depth profile that he ultimately passed on a contract to build the silicon which powered Apple's original iPhone, a regretful decision given the handset's wild success.



Retiring Intel CEO Paul Otellini

<http://appleinsider.com/articles/13/05/16/intel-chips-could-have-powered-first-iphone-ceo-otelini-says>

## The Value of Data

- Data describes phenomena of interest
- Can describe the phenomena directly or indirectly

NYC 311 Noise-Related Service Requests  
1 June - 30 August 2015

Hour of the Day	Number of Requests
0	~5500
1	~4500
2	~2000
3	~1000
4	~500
5	~300
6	~300
7	~350
8	~350
9	~350
10	~350
11	~350
12	~350
13	~350
14	~350
15	~350
16	~350
17	~350
18	~350
19	~350
20	~350
21	~350
22	~350
23	~350
24	~350
25	~350
26	~350
27	~350
28	~350
29	~350
30	~350
31	~350
32	~350
33	~350
34	~350
35	~350
36	~350
37	~350
38	~350
39	~350
40	~350
41	~350
42	~350
43	~350
44	~350
45	~350
46	~350
47	~350
48	~350
49	~350
50	~350
51	~350
52	~350
53	~350
54	~350
55	~350
56	~350
57	~350
58	~350
59	~350
60	~350
61	~350
62	~350
63	~350
64	~350
65	~350
66	~350
67	~350
68	~350
69	~350
70	~350
71	~350
72	~350
73	~350
74	~350
75	~350
76	~350
77	~350
78	~350
79	~350
80	~350
81	~350
82	~350
83	~350
84	~350
85	~350
86	~350
87	~350
88	~350
89	~350
90	~350
91	~350
92	~350
93	~350
94	~350
95	~350
96	~350
97	~350
98	~350
99	~350
100	~350
101	~350
102	~350
103	~350
104	~350
105	~350
106	~350
107	~350
108	~350
109	~350
110	~350
111	~350
112	~350
113	~350
114	~350
115	~350
116	~350
117	~350
118	~350
119	~350
120	~350
121	~350
122	~350
123	~350
124	~350
125	~350
126	~350
127	~350
128	~350
129	~350
130	~350
131	~350
132	~350
133	~350
134	~350
135	~350
136	~350
137	~350
138	~350
139	~350
140	~350
141	~350
142	~350
143	~350
144	~350
145	~350
146	~350
147	~350
148	~350
149	~350
150	~350
151	~350
152	~350
153	~350
154	~350
155	~350
156	~350
157	~350
158	~350
159	~350
160	~350
161	~350
162	~350
163	~350
164	~350
165	~350
166	~350
167	~350
168	~350
169	~350
170	~350
171	~350
172	~350
173	~350
174	~350
175	~350
176	~350
177	~350
178	~350
179	~350
180	~350
181	~350
182	~350
183	~350
184	~350
185	~350
186	~350
187	~350
188	~350
189	~350
190	~350
191	~350
192	~350
193	~350
194	~350
195	~350
196	~350
197	~350
198	~350
199	~350
200	~350
201	~350
202	~350
203	~350
204	~350
205	~350
206	~350
207	~350
208	~350
209	~350
210	~350
211	~350
212	~350
213	~350
214	~350
215	~350
216	~350
217	~350
218	~350
219	~350
220	~350
221	~350
222	~350
223	~350
224	~350
225	~350
226	~350
227	~350
228	~350
229	~350
230	~350
231	~350
232	~350
233	~350
234	~350
235	~350
236	~350
237	~350
238	~350
239	~350
240	~350
241	~350
242	~350
243	~350
244	~350
245	~350
246	~350
247	~350
248	~350
249	~350
250	~350
251	~350
252	~350
253	~350
254	~350
255	~350
256	~350
257	~350
258	~350
259	~350
260	~350
261	~350
262	~350
263	~350
264	~350
265	~350
266	~350
267	~350
268	~350
269	~350
270	~350
271	~350
272	~350
273	~350
274	~350
275	~350
276	~350
277	~350
278	~350
279	~350
280	~350
281	~350
282	~350
283	~350
284	~350
285	~350
286	~350
287	~350
288	~350
289	~350
290	~350
291	~350
292	~350
293	~350
294	~350
295	~350
296	~350
297	~350
298	~350
299	~350
300	~350
301	~350
302	~350
303	~350
304	~350
305	~350
306	~350
307	~350
308	~350
309	~350
310	~350
311	~350
312	~350
313	~350
314	~350
315	~350
316	~350
317	~350
318	~350
319	~350
320	~350
321	~350
322	~350
323	~350
324	~350
325	~350
326	~350
327	~350
328	~350
329	~350
330	~350
331	~350
332	~350
333	~350
334	~350
335	~350
336	~350
337	~350
338	~350
339	~350
340	~350
341	~350
342	~350
343	~350
344	~350
345	~350
346	~350
347	~350
348	~350
349	~350
350	~350
351	~350
352	~350
353	~350
354	~350
355	~350
356	~350
357	~350
358	~350
359	~350
360	~350
361	~350
362	~350
363	~350
364	~350
365	~350
366	~350
367	~350
368	~350
369	~350
370	~350
371	~350
372	~350
373	~350
374	~350
375	~350
376	~350
377	~350
378	~350
379	~350
380	~350
381	~350
382	~350
383	~350
384	~350
385	~350
386	~350
387	~350
388	~350
389	~350
390	~350
391	~350
392	~350
393	~350
394	~350
395	~350
396	~350
397	~350
398	~350
399	~350
400	~350
401	~350
402	~350
403	~350
404	~350
405	~350
406	~350
407	~350
408	~350
409	~350
410	~350
411	~350
412	~350
413	~350
414	~350
415	~350
416	~350
417	~350
418	~350
419	~350
420	~350
421	~350
422	~350
423	~350
424	~350
425	~350
426	~350
427	~350
428	~350
429	~350
430	~350
431	~350
432	~350
433	~350
434	~350
435	~350
436	~350
437	~350
438	~350
439	~350
440	~350
441	~350
442	~350
443	~350
444	~350
445	~350
446	~350
447	~350
448	~350
449	~350
450	~350
451	~350
452	~350
453	~350
454	~350
455	~350
456	~350
457	~350
458	~350
459	~350
460	~350
461	~350
462	~350
463	~350
464	~350
465	~350
466	~350
467	~350
468	~350
469	~350
470	~350
471	~350
472	~350
473	~350
474	~350
475	~350
476	~350
477	~350
478	~350
479	~350
480	~350
481	~350
482	~350
483	~350
484	~350
485	~350
486	~350
487	~350
488	~350
489	~350
490	~350
491	~350
492	~350
493	~350
494	~350
495	~350
496	~350
497	~350
498	~350
499	~350
500	~350

Being data-driven doesn't mean  
blindly following data.

Augment decision makers with objective, trustworthy, and relevant data.

### What is Analysis?

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

- Carter Hewgley, Director of Analytics,  
Center for Government Excellence

---

---

---

---

---

---



Having lots of reports does not make you data-driven.

---

---

---

---

---

---



Having lots of dashboards does not make you data-driven.

---

---

---

---

---

---



Having lots of alerts does not make you data-driven.

---

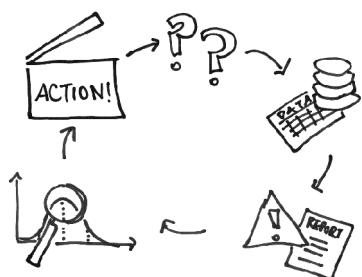
---

---

---

---

### It's Putting Them All Together



---

---

---

---

---

"If you do not know  
how to ask the  
right question,  
you discover  
nothing."  
W. Edward Deming



(c) iStockphoto / Thinkstock

---

---

---

---

---

## TYPES OF ANALYSIS

---



---



---



---



---



---

### 1. Quantifying Needs

- How much of X do I need?
  - Analyzing inputs (resources, people, etc.)
- How much does my need change given a different set of conditions?
  - What are the conditions that influence X?
- Important Considerations:
  - How does X play into my organization's mission and goals?
  - What's the most meaningful way of quantifying X?

---



---



---



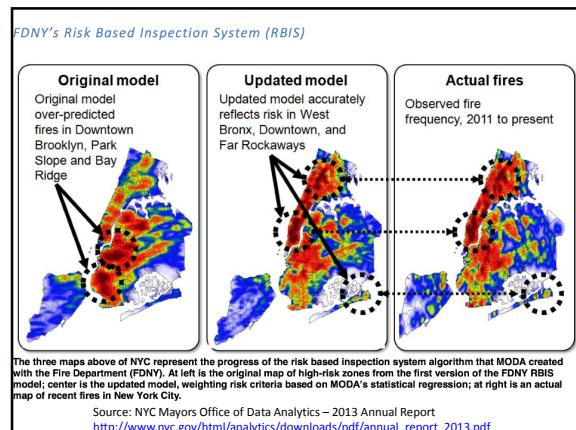
---



---



---




---



---



---



---



---



---

## 2. Operational Analysis

- What is my organization doing?
  - Assessment
- How might my organization do things better?
- Important Considerations:
  - What are your organization's mission and goals?
  - How do your employees do their work?
  - What's the best way to measure this work?

---



---



---



---



---



---



---



---

 64°  
Sign In | Subscribe

**The Charlotte Observer**

**BUSINESS** AUGUST 28, 2015

## Cincinnati racks up more than \$130,000 in late fees to Duke Energy

To address the issue, workers in Cincinnati's new Office of Performance and Data Analytics identified the impact the fees were having on the city and spent three days this summer with city department heads to come up with a solution, the station reported.

City leaders spent time in what's known as the Innovation Lab, where they figured out a "new way" to pay bills on time and avoid such fees. It wasn't immediately clear what their solution was.

Source: <http://www.charlotteobserver.com/news/business/article32617293.html>

---



---



---



---



---



---



---



---

## 3. Performance Metrics

- How is my organization doing?
  - Monitoring and evaluation
- How do we make this data visible to the people who need it?
- Important Considerations:
  - What is most important to measure (think mission and goals)?
  - How do we best measure performance?
  - "When a measure becomes a target, it ceases to be a good measure" - Charles Goodhart

---



---



---



---



---



---

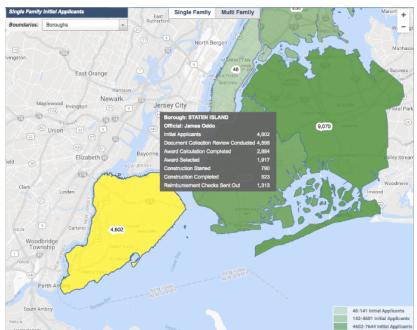


---



---

### Tracking Hurricane Sandy Relief Funds



Source: [http://www.slideshare.net/Leanalytics/startup-metrics-toronto-march-19](http://www.slideshare.net/Leananalytics/startup-metrics-toronto-march-19)

### Goodhart's Law

"When a measure becomes a target, it ceases to be a good measure."

- Charles Goodhart,  
Economist



<https://philebersole.wordpress.com/2013/01/10/goodharts-law-on-not-going-by-the-numbers/>

#### 4. Prioritization

- How do I meet optimal outcomes with limited resources?
  - Optimizing allocation
- Important considerations:
  - Minimize disruption
  - Work within current workflow
  - Support existing business practices

---



---



---



---

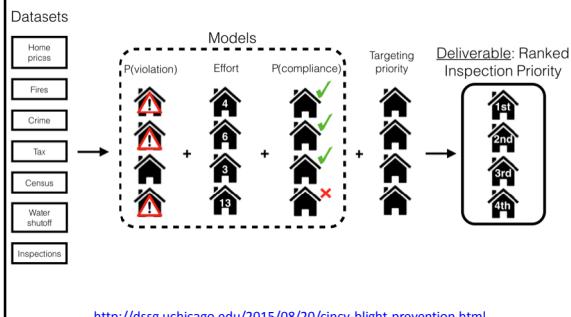


---



---

#### Cincinnati Blight Prevention Model




---



---



---



---

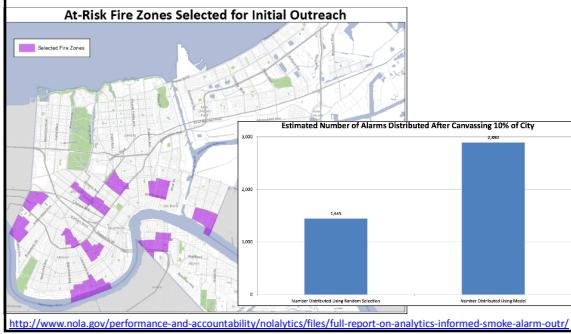


---



---

#### New Orleans Smoke Alarm Targeted Outreach




---



---



---



---



---



---

## NOFD saves lives by targeted smoke alarm outreach

New Orleans Fire Department working hard to prevent fires: Chief Tim McConnell  
[www.nola.com](http://www.nola.com/politics/index.ssf?/base/politics&coll=1#4338111111111111)

Just two weeks ago, a family of 11 members, including a 2-month-old baby, escaped a disastrous fire in their Central City home. They were alerted by a smoke alarm at 3 a.m., fleeing before the smoke was bad enough to impede their escape. New Orleans firefighters had installed that smoke alarm in the home as part of this year's campaign. The truth is that I have never

<http://www.nola.com/politics/index.ssf?/base/politics&coll=1#4338111111111111>

---

---

---

---

---

---

---

## WHAT KIND OF ANALYSIS DOES YOUR OFFICE DO?

---

---

---

---

---

---

---

## 4 TYPES OF CONCERN TO BE MINDFUL OF

---

---

---

---

---

---

---

## 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*

---



---



---



---



---



---

The New York Times

HEALTH

**Hipaa's Use as Code of Silence Often Misinterprets the Law**

JULY 17, 2015

**Paula Span**  
THE NEW OLD AGE



How do people use, misuse or abuse Hipaa, the federal regulations protecting patients' confidential health information? Let us count the ways:

■ Last month, in a continuing care retirement community in Ithaca, N.Y., Helen Wyvill, 72, noticed that a friend hadn't shown up for their regular swim. She wasn't in her apartment, either.

Email Share Tweet Save

[http://www.nytimes.com/2015/07/21/health/hipaas-use-as-code-of-silence-often-misinterprets-the-law.html?\\_r=0](http://www.nytimes.com/2015/07/21/health/hipaas-use-as-code-of-silence-often-misinterprets-the-law.html?_r=0)

---



---



---



---



---



---

### 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 agency and ultimately to the city as a whole*

---

---

---

---

---

### Benefits

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

---

---

---

---

---

## Group Exercise – Universal Pre-K

- Define the analytical problem and data needs in one of these key areas
  - Capacity
  - Outreach
  - Enrollment (CBOs/Students)
  - Monitoring/Evaluation
- Situation
  - ~104,000 4-year olds in NYC
  - 58,528 current seats
  - 26,364 in public schools
  - 32,164 in community based organizations
- Goals
  - Increase enrollment by 30,000 for 2014-2015
  - Increase enrollment by 20,000 for 2015-2016

---



---



---



---



---



---



---

## BUILDING A 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

---

---

---

---

---

## 6 ANALYTICAL STEPS

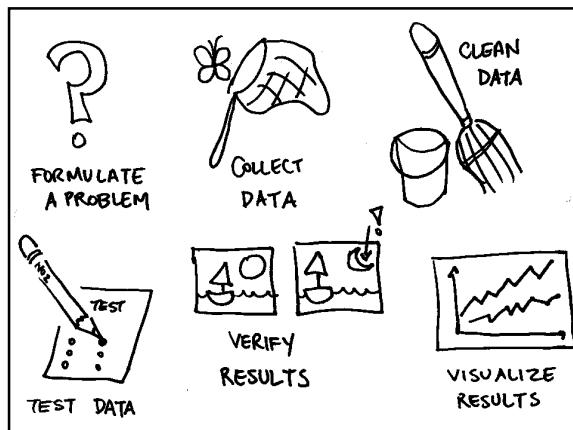
---

---

---

---

---



---

---

---

---

---

### 1. Problem Formulation

- What question or need am I trying to answer?
- What's my organization's mission and goals?
- How can I best apply data to this task?

---

---

---

---

---

### 2. Data Gathering/Preliminary Analysis

- What data do I think I'm going to need?
- What condition is it in?
- Does it tell me what I need?
- What other data might I need?
- How much work do I need to put into the data?

---

---

---

---

---

### 3. Data Cleaning/Data Munging

- Make the data usable and compatible
- Takes up the most amount of time
- May require more sophisticated tools depending on the state and size of the data

---

---

---

---

---

## 4. Hypothesis Testing

- Am I getting the results I'd hoped for?
- What other questions come up?

---



---



---



---



---



---

## 5. Verification

- Do my results make sense?
- Did I make a simple mistake?
- Check twice and you'll sleep easier

---



---



---



---



---



---

**MarketWatch**    

NEWS VIEWER MARKETS INVESTING TRADING DECK PERSONAL FINANCE RETIREMENT ECONOMY REAL ESTATE

Home > Personal Finance

## 88% of spreadsheets have errors

Published: Apr 20, 2013 7:39 a.m. ET

    139

*Research: Bad math rampant in family budgets and Harvard studies*

 By JEREMY OLSHAN EDITOR  

Microsoft Excel makes it easy for anyone to do the kind of number crunching once reserved for accountants and statisticians. But the world's best-selling spreadsheet software has also contributed to the proliferation of bad math.

Close to 90% of spreadsheet documents contain errors, a 2008 analysis of multiple studies suggests. "Spreadsheets, even after careful development, contain errors in 1% or more of all formula cells," writes Ray Panko, a professor of IT management at the University of Hawaii and an authority on bad spreadsheet practices. "In large spreadsheets with thousands of formulas, there will be dozens of undetected errors."

<http://www.marketwatch.com/story/88-of-spreadsheets-have-errors-2013-04-17>

---



---



---



---



---



---

## 6. Visualization

- “A picture is worth a thousand words”
- Communicate results clearly and concisely
- Help to better understand your data
- The eyes have a much higher bandwidth into the brain

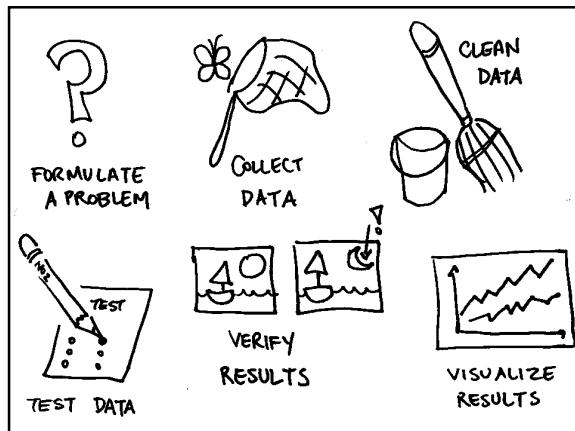
---

---

---

---

---



---

---

---

---

---

HOW DOES YOUR OFFICE ANALYZE DATA?

---

---

---

---

---

## Definition of Open Data

- Definition:
  - Open data is data that can be freely used, shared and built-on by anyone, anywhere, for any purpose
- Key Features
  - Availability and access
  - Reuse and redistribution
  - Universal participation

<http://blog.okfn.org/2013/10/03/defining-open-data/>

---

---

---

---

---

---

## Open Data Benefits

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

---

---

---

---

---

---

## Open Data Concerns

- Privacy
  - Personally identifiable information (PII)
  - Mosaic Effect
- Confidentiality
- Security

---

---

---

---

---

---

## LINKS TO OPEN DATA PORTALS

NYC Open Data Portal - <https://data.cityofnewyork.us/>  
 NYS Open Data Portal - <https://data.ny.gov>  
 US Federal Government Open Data Portal - <http://www.data.gov/>

### Data for Exercise

Borough-Block-Lot:  
1001210001

Assessed Value:  
\$70,771,500

Community District:  
Manhattan 1

School District: 2

Latitude: 40.71306  
Longitude: -74.0037

Address:  
1 Centre Street  
New York, NY 10007

Census Tract:  
36061002900

Census Blocks:  
360610029004012  
360610029004019

X: 983210  
Y: 199016

## Exploratory Data Analysis

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

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

---



---



---



---



---



---



---



---



---

## Question-Driven Analysis

- Goal -> Answer a specific problem or concern
- Approach
  - Have a question or problem in mind when analyzing data
  - “I need to know X”
  - Problem-focused discovery with the data

---



---



---



---



---



---



---



---



---

## Question Driven Analysis

### Vision Zero (dB)

#### Tasks:

- Given 311 noise complaint data, assist enforcement efforts by identifying community districts that have a high volume of noise complaints and the time frame enforcement resources should be deployed to combat the noise issue at its peak
- Identify the prevalent types of noise complaints in these areas to guide enforcement in each community district

---



---



---



---



---



---



---



---



---

## Analytical Resources

- Internal Agency Teams
  - Offices within your organization charged with performing analysis for internal or external stakeholders
- Mayor's Office Task Forces
  - Special inter-agency efforts around critical policy areas bringing together critical skills and experience in a subject area
- NYC Center for Innovation through Data Intelligence (CIDI)
  - Conducts inter-agency research to identify areas of service need in the City of New York
  - Collaborates with all Health and Human Service (HHS) agencies and other City partners to improve services
  - CIDI values the contextual interpretation of data and respects persons' confidentiality in its research activities
  - <http://www.nyc.gov/html/cidi/html/home/home.shtml>

---

---

---

---

---

---

## Analytical Resources

- Mayor's Office of Data Analytics (MODA)
  - New York City's civic intelligence center
  - Aggregating and analyzing data from across City agencies
  - More effectively address crime, public safety, and quality of life issues
  - Uses analytic tools to:
    - Prioritize risk more strategically
    - Deliver services more efficiently
    - Enforce laws more effectively
    - Increase transparency

---

---

---

---

---

---

## What We've Covered

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

---

---

---

---

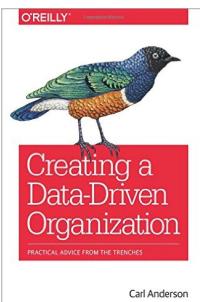
---

---

## Technical Resources

- Stack Overflow
  - <http://stackoverflow.com/>
  - One of the best Q&A sites for technical questions of all kinds
- Microsoft Office Support
  - <http://office.microsoft.com/en-us/support/>
  - Documentation on various MS Office products
- Excel Tips
  - <http://excel.tips.net/>
  - Various tips and tricks for using Excel
- Data Science Cheatsheet
  - <https://github.com/govex/Data-Science>

## Resources



## Contact Information

### Instructors

- Richard Dunks
  - Email: [richard@datapoltan.com](mailto:richard@datapoltan.com)
  - Blog: <http://blog.datapolitan.com>
  - Twitter: @datapoltan
- Julia Marden
  - Email: [julia@tinypanther.com](mailto:julia@tinypanther.com)
  - Website: <http://tinypanther.com>
  - Twitter: @juliaem