Data Journalism

Michael Friendly
Psych 6135

https://friendly.github.io/6135/
Topics

• What is DDJ?
• Some early examples
• Modern examples
  ▪ election reporting
  ▪ health statistics
  ▪ social media
• Best examples??
  ▪ Panama Papers
• Satellite journalism
What is data journalism?

- **Data journalism** reflects the increased role of numerical data for reporting in the digital era. **Data driven** or DDJ more useful
  - Reflects increased interaction between journalists and visualization design, computer science and statistics.

- Aspects:
  - **Database analysis**: FOI requests on topics like crime, racial discrimination, judicial sentencing, toxic pollution, ...
  - **Visualization**: infographic→data vis., emphasis on the story
  - Increasing use & development of data vis. **software tools**
    - analysis: R, python, ...
    - visualization libraries: R, D3, javascript, ...
  - Increasing use of **dynamic updating** of data sources
    - election polls / prediction
    - disease tracking
What is data journalism?

To an outsider, it is often hard to tell concretely what #DDJ is, or is not

Two motivational infographics of #DDJ as a process
❖ How does it actually work?
❖ What makes it successful?
Bill Dedman of *The Atlanta Journal-Constitution* analyzed racial discrimination by banks and other lenders in middle-income black neighborhoods of Atlanta, GA.

In a series of articles, May 1-4, 1988, *The Color of Money*, he showed that although banks made loans in even the poorest white neighborhoods, they did not lend in middle-class or more affluent black neighborhoods.

This work won the Pulitzer prize in investigative reporting in 1989.

It helped establish the idea of computer-assisted reporting and the importance of compelling visualizations in telling the story.

See: [http://powerreporting.com/color/](http://powerreporting.com/color/) for all the articles and images
Hurricane Andrew

- In Aug. 1992, a category 5 hurricane hit Miami-Dade county
- Huge numbers of houses were destroyed, but this varied by neighborhood
- Investigative reporters at the *Miami Herald* discovered why:
  - Building codes and inspection became lax after 1980.
  - Older homes survived.
- This report by Liz Balmaseda et al. won the Pulitzer Prize in 1993.

Source: https://www.flickr.com/photos/juggernautco/sets/72157607210036175/
The report in the *Miami Herald* contained many detailed analysis maps.

There was overwhelming evidence for the cause of destruction.

The combination of investigative reporting, analysis and graphics was seen as powerful & compelling.

Data-driven journalism had arrived.
In 1962, John Tukey and a statistical team joined with NBC to attempt to predict the congressional election.

Over time, they developed new methods to combine past results with exit polls and current returns to give up-to-date estimates and measures of uncertainty as the numbers came in.

Key idea: “borrowing strength”

Now known as “hierarchical Bayes”

NBC execs were freaked out when the Tukey team began to call elections quite early with 95%, then 97.5% certainty. They waited until it reached 99%, and missed the scoop.
CNN Magic Wall

Direct, interactive visualization of election news on TV took a big jump in 2008 with the development of the Magic Wall, a large touch-screen display, tied to databases of election results & with on-screen controls to zoom in/out, compare, highlight, etc.

Video: https://www.youtube.com/watch?v=rpLHnK-MSy4
In the special election for the Alabama U.S. Senate, Dec. 12, 2017, *The New York Times* used dynamic, interactive maps to compare the vote to the 2016 Presidential election.

Hovering the mouse over a county showed the details.

A map-based display is natural & easy.

Q: What other graphic might show the comparison better?
In the 2016 US election, the NYT used dynamically updated dashboard displays of the election returns and model-based predictions of the winner of the Alabama Senate race, predicting a win by the Democrat, Doug Jones, over rival Republican, Roy Moore.

Doug Jones has won the election, according to A.P.

Two dials:
- Predicted outcome
- Certainty

Jones +1.5

>95% Jones
The jittery icon of liberal trauma

But in the Hilary Clinton – Donald Trump contest the needle earned a place in media infamy. “For much of the night, ‘the needle’ — a quivering indicator of which candidate is more likely to win — showed Hillary Clinton as the favorite. And then, suddenly, it didn’t.”

“As the needle swung in Trump’s direction around 9 p.m., it became much more than a data visualization. For Clinton voters, it became an encapsulation of everything that was going terribly, unfathomably wrong: The fate of the country and democracy distilled to a lone meter that was suddenly pointing to ‘red alert.’”

Another dynamically updated chart showed the projected vote margin, together with two levels (50%, 95%) of uncertainty intervals for the current estimate.

**Projected Vote Margin**

Once all the votes have been counted, our estimated margin and the reported margin will match. As a rule, when our estimated margin is steady, our forecast is more trustworthy.
Novel graphic designs

This graph, from fivethirtyeight.com was designed to show how some presidential candidates had shifted positions before the 2016 election.

The axes are a score on social and economic policy, but they rotated the axes by 45° to create zones related to political thought.

This info graphic is eye-catching and self-explanatory:
• colored/labeled zones
• interpretive labels on axes
• arrows showing movement to extremes
Measles and vaccines

Visualizing the impact of health policy interventions

In 2015 Tynan DeBold & Dov Friedman in the Wall Street Journal tried to show the effect of the introduction of vaccination programs in the US states on disease incidence, using color-coded heat maps for a variety of diseases.

The long time series ~70 years made this work.

The heat map color scale is not exemplary, but the message is still clear: disease incidence declined after vaccines were introduced.

The images are “interactive,” in the weak sense that tool-tips are shown on mouse movement.

Making the message more explicit

What you should try to show here in an Info graphic is the overall impact of vaccination on measles.

Ed Tufte did this by adding a histogram at the top showing total # of cases by year.

With the overall impact made clear, details about individual states can now be explored.

“Overview first; details on demand”
Effect ordering & color scales

Ordering the states by an overall measure of measles incidence makes it easier to study details.
A uni-dimensional scale of color is more effective than the ad-hoc scale used in the original.

This chart, by Xan Gregg, also shows missing data explicitly.

Color scales for data graphs is a growing topic.

What’s wrong in Wis. & VT?

From: https://community.jmp.com/t5/JMP-Blog/Graph-makeover-Measles-heat-map/ba-p/30550
Before – After Circle Charts

What Happened When Trump Was Banned on Social Media?

❖ Circles are statements by Trump.
❖ Circle size ~ likes & shares on social media

“First Lady and I tested positive for COVID-19 ...”
Both President Trump’s own posts and those
directly quoting him were liked and shared
4.6 million times on Twitter,
Facebook and Instagram.

“I WON THIS ELECTION,
BY A LOT!”
4.7 million

Donald Trump
banned from social media

“This has been yet another phase of the greatest
witch hunt in the history of our Country...”
Posts quoting this Trump statement on his second impeachment
got 2 million likes and shares across social media, even without
Mr. Trump being able to post it himself.

“To all of those who have asked, I will not be
going to the Inauguration on January 20th.”
2.5 million

Before – After: Who pushed the lies?

How/Who Spread Trump statements before and after the ban

Posts referencing each Trump statement, sized by total number of likes and shares

<table>
<thead>
<tr>
<th>Facebook</th>
<th>Donald J. Trump post 250,000 likes and shares on 57 posts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Twitter</td>
<td>@realDonaldTrump post 248,000 likes and retweets on 2 posts</td>
</tr>
</tbody>
</table>

An Oct. 8, 2020 statement with about half a million total likes and shares

<table>
<thead>
<tr>
<th>Instagram</th>
<th>170,000 likes and reposts on 14 posts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facebook</td>
<td>390,000 likes and shares on 100 posts</td>
</tr>
</tbody>
</table>

A March 21, 2021 statement with about half a million total likes and shares

<table>
<thead>
<tr>
<th>Instagram</th>
<th>Donald Trump Fan Page 31,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facebook</td>
<td>Breitbart 83,000</td>
</tr>
<tr>
<td>Twitter</td>
<td>Breitbart 77,000</td>
</tr>
<tr>
<td>Twitter</td>
<td>Fox News 43,000</td>
</tr>
<tr>
<td>Twitter</td>
<td>President Donald Trump Fan Club 48,000</td>
</tr>
<tr>
<td>Twitter</td>
<td>I Love My Freedom 47,000</td>
</tr>
<tr>
<td>Twitter</td>
<td>@JennaEllisEsq 37,000</td>
</tr>
</tbody>
</table>

CNN News18
This powerful NYT chart compares cases of diseases and firearm injuries to the number of N.I.H. research awards.

How well do you like this design?

Is it a fair depiction of the data?

Can you suggest anything different?
Votes for vs. COVID vaccinations

% of state pop. with at least one dose, vs. who won (Biden, Trump) in 2020

circle size ~ state pop.

What makes this chart effective?

From: [https://twitter.com/kjhealy/status/1402691186826039301](https://twitter.com/kjhealy/status/1402691186826039301)
NYT Infographic: Gun deaths

“a single death is a tragedy, a million deaths are a statistic” – Joseph Stalin

Infographics use simple icons to give emotional impact to raw numbers.

But journalists always use the ‘single death’ in the lede.
NYT: Gun deaths & policy

Among developed countries, the United States has by far the highest rate of firearms ownership.

<table>
<thead>
<tr>
<th>Guns per 100 people</th>
</tr>
</thead>
<tbody>
<tr>
<td>UNITED STATES</td>
</tr>
<tr>
<td>CANADA</td>
</tr>
<tr>
<td>SWITZERLAND</td>
</tr>
<tr>
<td>SWEDEN</td>
</tr>
<tr>
<td>FRANCE</td>
</tr>
<tr>
<td>GERMANY</td>
</tr>
<tr>
<td>AUSTRALIA</td>
</tr>
<tr>
<td>ITALY</td>
</tr>
<tr>
<td>SPAIN</td>
</tr>
<tr>
<td>ENGLAND, WALES</td>
</tr>
<tr>
<td>JAPAN</td>
</tr>
</tbody>
</table>

How does the US compare to other Western countries in gun ownership & murders?

Gun murders per 100,000 people

America's private arsenal is five times as lethal as Canada's, and 30 times worse than Australia's.

<table>
<thead>
<tr>
<th>Gun murders per 100,000 people</th>
</tr>
</thead>
<tbody>
<tr>
<td>UNITED STATES</td>
</tr>
<tr>
<td>CANADA</td>
</tr>
<tr>
<td>FRANCE</td>
</tr>
<tr>
<td>SWEDEN</td>
</tr>
<tr>
<td>ITALY</td>
</tr>
<tr>
<td>SWITZERLAND</td>
</tr>
<tr>
<td>AUSTRALIA</td>
</tr>
<tr>
<td>GERMANY</td>
</tr>
<tr>
<td>SPAIN</td>
</tr>
<tr>
<td>ENGLAND, WALES</td>
</tr>
<tr>
<td>JAPAN</td>
</tr>
</tbody>
</table>

Sources: Small Arms Survey, United Nations Office on Drugs and Crime. * Ownership rates are for 2017. Murder rates for the U.S., Canada, Sweden, Switzerland, Australia and Spain are from 2016; otherwise, the latest available rates are used.

These simple barcharts have a grim conclusion.
Seatbelts: Story telling

New York Times: An annotated chart of vehicle deaths in relation to seatbelts

Telling the story of a public safety policy with explanatory milestones

Deaths per 100 million vehicle miles traveled

Source: National Highway Traffic Safety Administration
This graph directly linked gun deaths to the strength of states’ laws on gun ownership.
Best examples of DDJ? Ask Twitter

Michael Friendly @datavisFriendly · 22h
1/ Thinking about an upcoming class in my #psyc6135 #datavis course on the topic #datajournalism, with a guest star @JeeRankin. It got me thinking of the idea of "Great moments in Data Journalism": investigative reporting, incisive analyses, compelling graphic displays.

Michael Friendly @datavisFriendly · 22h
2/ I would love ❤️ for nominations on this topic. @albertocairo and others have done some of this, but perhaps this is a topic on which others could contribute their nominations for "Great moments..."

Alberto Cairo @albertocairo · 22h
Twitter #ddj and news #dataviz, contribute with nominations.

I'd do @FiveThirtyEight's: projects.fivethirtyeight.com/redistricting-

Anything by @ProPublica, like @lenagroeger's appropriately creepy: projects.propublica.org/graphics/worke...

Michael Friendly @datavisFriendly · 20h
3/
Reciprocity: How has #DDJ contributed to or advanced #datavis itself? And, vice-versa?

Team work: How do teams of journalists, data people, graphic designers and web developers work together on projects? Anything special/different from other team projects?
Now trending ...

❖ Digital publishing and online, interactive media
   ▪ Many news organizations have interactive graphic developers or entire online departments
   ▪ A revenue stream responding to less from print & ads

❖ Story telling frames a wider enterprise
   ▪ A new focus and techniques for data visualization
   ▪ New online forms: scrollytelling

❖ Opening data
   ▪ Maps-GIS (Google, Open Street maps)
   ▪ Gov’t data: census, health, crime, ...
   ▪ NGO initiatives: Our World in Data
   ▪ Unsealed: WikiLeaks, Panama Papers

❖ Software infrastructure
   ▪ Javascript as base for online, interactive methods
   ▪ Libraries, frameworks building on this (D3.js, Processing, ...)
There’s a lot of complaining about gerrymandering, but what should US Congressional districts look like? What would they look like to serve various goals?

This map shows the current boundaries, with expected wins by Rep and Dem.

Why do I find this app appealing?

From:
https://projects.fivethirtyeight.com/redistricting-maps/
How could district boundaries be re-drawn to promote more competitive elections?

Election district boundaries are consequential & highly political. They can be changed by state legislatures.

Different scenarios show the changes in representation, w/o a single voter switching parties!

From: https://projects.fivethirtyeight.com/redistricting-maps/
What if the Democrats could control election districts to favor themselves?

Note that the overall impression of the map is still Rep. overall.

Changes in representation take place largely on the east/west coasts.

From: https://projects.fivethirtyeight.com/redistricting-maps/
Aaron Williams & Armand Emamdjomeh at the *Washington Post* studied changes in racial composition and diversity in U.S. cities over 30+ years. They relate these to demographic, policy, and urban/suburban issues using well-designed thematic maps.

[Images of thematic maps showing racial composition and diversity in U.S. cities.]

An impressive range of social & environmental issues presented by Our World in Data.
This series of charts from the Financial Times was hailed for best-practices in the early stages of the pandemic.

❖ Log scale make it easy to compare rates of increase
❖ Visual meaning for “leveling the curve”
❖ Labels for interventions (lockdown)
❖ Sensible (relative) origin – days since 10th death
Data journalists at Reuters created live trackers, pulling data from different sources to generate real-time graphics with text, telling the story of the pandemic. An interactive globe serves as a novel way to explore vaccination rates.
Other interactive displays provide the possibility to compare countries by geographic region and income level.

“Reuters also took on the hard task of tracking the vaccination policies that determined who had access to COVID-19 vaccines in countries around the world.

That meant creating a unifying structure to catalogue the different phases of each country's vaccination rollout plan.”

-- Prasanta Kumar Dutta

Daniel Dale & Ed Tubb at the *Toronto Star* maintained a database of every false claim made by Donald Trump since the Inauguration on Jan. 20, 2017

Donald Trump has said **4557** false things as U.S. president

(as of 3-13-2019)

Every false claim is documented: place, claim, truth, “slug”, categories

<table>
<thead>
<tr>
<th>number</th>
<th>date</th>
<th>place</th>
<th>claim</th>
<th>truth</th>
<th>slug</th>
<th>repeat</th>
<th>source</th>
<th>category1</th>
<th>category2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2017-01-20</td>
<td>Post-inaugura</td>
<td>&quot;I looked at the rain, which was coming...&quot;</td>
<td>The rain began right at the beginning...</td>
<td>Inauguration</td>
<td></td>
<td><a href="https://www.thestar.com/ne">https://www.thestar.com/ne</a>...</td>
<td>Inauguration</td>
<td>Weather</td>
</tr>
<tr>
<td>2</td>
<td>2017-01-20</td>
<td>Post-inaugura</td>
<td>&quot;Even the media said the c...&quot;</td>
<td>The major media reported that...</td>
<td>Inauguration</td>
<td></td>
<td><a href="https://www.thestar.com/ne">https://www.thestar.com/ne</a>...</td>
<td>Inauguration</td>
<td>Exaggerated</td>
</tr>
<tr>
<td>3</td>
<td>2017-01-21</td>
<td>Speech at Cent</td>
<td>&quot;Honestly, it looked like a...&quot;</td>
<td>The crowd, which may not have...</td>
<td>Inauguration</td>
<td></td>
<td><a href="https://www.youtube.com/w">https://www.youtube.com/w</a>...</td>
<td>Inauguration</td>
<td>Exaggerated</td>
</tr>
<tr>
<td>4</td>
<td>2017-01-21</td>
<td>Speech at Cent</td>
<td>&quot;It was almost raining, the...&quot;</td>
<td>Neither of these claims is true...</td>
<td>Inauguration</td>
<td></td>
<td><a href="https://www.youtube.com/w">https://www.youtube.com/w</a>...</td>
<td>Inauguration</td>
<td>NA</td>
</tr>
<tr>
<td>5</td>
<td>2017-01-21</td>
<td>Speech at Cent</td>
<td>&quot;So a reporter for TIME ma...&quot;</td>
<td>Trump's numbers are well off...</td>
<td>TIME co...</td>
<td></td>
<td><a href="https://www.youtube.com/w">https://www.youtube.com/w</a>...</td>
<td>Media</td>
<td>Exaggerated</td>
</tr>
<tr>
<td>6</td>
<td>2017-01-23</td>
<td>Private meeting</td>
<td>Trump told Congressional...</td>
<td>This claim, also reported by...</td>
<td>Mass fr...</td>
<td></td>
<td><a href="http://www.foxnews.com/pol">http://www.foxnews.com/pol</a>...</td>
<td>Voter fraud</td>
<td>Election</td>
</tr>
<tr>
<td>7</td>
<td>2017-01-25</td>
<td>Interview with A...</td>
<td>&quot;They can mean anything, but n...&quot;</td>
<td>They cannot mean anything, but not...</td>
<td>Inauguration</td>
<td></td>
<td><a href="http://abcnews.go.com/Polit">http://abcnews.go.com/Polit</a>...</td>
<td>Inauguration</td>
<td>Exaggerated</td>
</tr>
<tr>
<td>8</td>
<td>2017-01-25</td>
<td>Interview with A...</td>
<td>&quot;No, no, you have to unde...&quot;</td>
<td>Trump's victory was not close...</td>
<td>Election</td>
<td></td>
<td><a href="http://abcnews.go.com/Polit">http://abcnews.go.com/Polit</a>...</td>
<td>Election</td>
<td>Exaggerated</td>
</tr>
<tr>
<td>9</td>
<td>2017-01-25</td>
<td>Interview with A...</td>
<td>&quot;I think you're demeaning...&quot;</td>
<td>Saying &quot;you&quot; here, Trump won...</td>
<td>Media</td>
<td></td>
<td><a href="http://abcnews.go.com/Polit">http://abcnews.go.com/Polit</a>...</td>
<td>Media</td>
<td>Exaggerated</td>
</tr>
<tr>
<td>10</td>
<td>2017-01-25</td>
<td>Interview with A...</td>
<td>With regard to his speech...</td>
<td>Most of the audience was inside...</td>
<td>CIA sp...</td>
<td></td>
<td><a href="http://abcnews.go.com/Polit">http://abcnews.go.com/Polit</a>...</td>
<td>Trump White House</td>
<td>US intelligence</td>
</tr>
<tr>
<td>11</td>
<td>2017-01-25</td>
<td>Interview with A...</td>
<td>Regarding Healthcare.gov...</td>
<td>Healthcare.gov did not cost $5...</td>
<td>Health...</td>
<td></td>
<td><a href="http://abcnews.go.com/Polit">http://abcnews.go.com/Polit</a>...</td>
<td>Exaggerated</td>
<td>NA</td>
</tr>
<tr>
<td>12</td>
<td>2017-01-25</td>
<td>Interview with A...</td>
<td>&quot;I called the author...&quot;</td>
<td>Groveling means &quot;to draw back...&quot;</td>
<td>Grovel...</td>
<td></td>
<td><a href="http://abcnews.go.com/polit">http://abcnews.go.com/polit</a>...</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>13</td>
<td>2017-01-25</td>
<td>Interview with A...</td>
<td>&quot;Now you're telling me...&quot;</td>
<td>Trump was trying to use a 2012...</td>
<td>Few fr...</td>
<td></td>
<td><a href="http://abcnews.go.com/polit">http://abcnews.go.com/polit</a>...</td>
<td>Election</td>
<td>NA</td>
</tr>
<tr>
<td>14</td>
<td>2017-01-25</td>
<td>Interview with A...</td>
<td>Regarding his false claim...</td>
<td>These large numbers of illegal...</td>
<td>Mass fr...</td>
<td></td>
<td><a href="http://abcnews.go.com/polit">http://abcnews.go.com/polit</a>...</td>
<td>Election</td>
<td>NA</td>
</tr>
<tr>
<td>15</td>
<td>2017-01-25</td>
<td>Interview with A...</td>
<td>Regarding voting fraud:...</td>
<td>There is no evidence of a significant...</td>
<td>Mass fr...</td>
<td></td>
<td><a href="http://abcnews.go.com/polit">http://abcnews.go.com/polit</a>...</td>
<td>Voter fraud</td>
<td>Chicago</td>
</tr>
<tr>
<td>16</td>
<td>2017-01-25</td>
<td>Interview with A...</td>
<td>Regarding voting fraud:...</td>
<td>There is no evidence of a significant...</td>
<td>Mass fr...</td>
<td></td>
<td><a href="http://abcnews.go.com/polit">http://abcnews.go.com/polit</a>...</td>
<td>Voter fraud</td>
<td>NA</td>
</tr>
<tr>
<td>17</td>
<td>2017-01-25</td>
<td>Interview with A...</td>
<td>&quot;Look, Barack Obama -- if...&quot;</td>
<td>This is a gross mischaracterization...</td>
<td>Obama...</td>
<td></td>
<td><a href="http://abcnews.go.com/polit">http://abcnews.go.com/polit</a>...</td>
<td>Barack Obama</td>
<td>Crime</td>
</tr>
<tr>
<td>18</td>
<td>2017-01-25</td>
<td>Interview with A...</td>
<td>&quot;So, look, when President...&quot;</td>
<td>There were not only no homicides...</td>
<td>Obama...</td>
<td></td>
<td><a href="http://abcnews.go.com/polit">http://abcnews.go.com/polit</a>...</td>
<td>Barack Obama</td>
<td>Crime</td>
</tr>
</tbody>
</table>

- place: Inauguration
- media
- voter fraud
Graphic artist Jelle Geertsma designed this animated graphic showing the top 5 claim “slugs” & the top 20 categories over time, by month.

Bars at right show cumulative totals

This animation uses the gganimate package.


Data source: Daniel Dale/Toronto Star; Graphic design: Jelle Geertsma
Scrollytelling: The Impatient List


Waitlist length / Donors & blood type, by state

““The Impatient List”, an example of Scrollytelling by Gupta et al."

From https://xujunjiejack.github.io/
Scrollytelling: The Impatient List

Since Kidney Transplant can be time sensitive, it is best for you to choose the ones adjacent to where you live. Please enter your zip code to view centers around you:

![Zip Code Input]

You’ll also need to compare between centers around you for their death rates, survival rates, and waitlist waiting time, in order to find the optimal center to register as a candidate. How would you balance these different factors, when each choice you make [might influence your chance of getting a transplant to being cured?]

*Click on each pin on the map to view the data for each center compared to other centers around you.*
Since Kidney Transplant can be time sensitive, it is best for you to choose the ones adjacent to where you live. Please enter your zip code to view centers around you:

30318

You’ll also need to compare between centers around you for their death rates, survival rates, and waitlist waiting time, in order to find the optimal center to register as a candidate. How would you balance these different factors, when each choice you make might influence your chance of getting a transplant to being cured?

Click on each pin on the map to view the data for each center compared to other centers around you.
Drill-down to tell a story

The *Wall Street Journal* published this interactive story of potential conflicts of interest in the business holdings of Donald Trump and his family.

“The most difficult part of many visualization projects, and certainly of this one, was distilling a complex, messy dataset into a clean visualization with a clear takeaway,” says Eastwood.

In each part of the story, hovering with the mouse reveals the details of a particular investment and links to others.

**EVERY HOLDING IS A POTENTIAL CONFLICT OF INTEREST**

Mr. Trump is at the center of a complex network of business holdings that, according to the May 2016 financial disclosures he filed with the federal government, was made up of more than 500 businesses including hotels, condo towers, golf courses and foreign and domestic licensing deals. His personal connections to such a vast empire eclipse the business interests of any other prominent politician in modern times.

**TRUMP MARKS PANAMA LLC**

*New York, N.Y. USA*

License deal with K GROUP DEVELOPERS INC.

**ASSET**
Type: license deal | Value: $15,001 to $50,000

**INCOME**
Type: royalties | Amount: $1,000,001 to $5,000,000

_Trump Financial Disclosure Report (May 2016)_{*}
In May 2009, an unknown government employee copied a terabyte of documents on MP expenses and released these to *The Daily Telegraph*

**MPs' expenses: how scoop came to light and why journalists fear a 'knock on the door'**

*The Daily Telegraph's* exclusive was offered to at least three other papers by shadowy figure described as a 'rank amateur'
In June 2009, The British House of Commons released 700,000 documents detailing every expense claim over for 4 years by all 644 MPs. *The Guardian* launched a crowd sourced experiment inviting readers to investigate these claims.
Crowd sourcing: The App

• Using our system, you will be able to find your MP - or any member - and look at their records directly. For every page for every MP, you will be able to:
  ▪ comment on individual expenses
  ▪ highlight ones of interest
  ▪ tell us how interesting that receipt is
  ▪ Help by entering the numbers on the page

• Using that information, we will begin to be able to piece together a unique picture of how MPs claims have changed over time.
MPs' expenses: how we visualised the data

How do you display a complicated set of numbers on the story of the week? See how we did it

- Who claimed what?
- The Legg report’s full list of repayments

MP expenses visualised. Click on image for the full graphic
In January 2017, The *International Consortium of Investigative Journalists* released the “Panama Papers”, exposing the tax avoidance schemes of the Panama-based firm Mossack Fonseca. Two terabytes of data, over 40 years, 210,000 companies in 21 offshore jurisdictions

IJIC: 280 top investigative journalists, >100 media partners, >100 countries

https://www.icij.org/investigations/panama-papers/
Panama Papers

Key findings:
❖ Files reveal offshore holdings of 140 politicians & public officials around the world
❖ Current & former world leaders include: PM of Iceland, the king of Saudi Arabia, ... 
❖ >214,000 offshore entities, connected to people in more than 200 countries
❖ Major banks have driven the creation of hard-to-trace companies in offshore havens
Satellite Journalism

- An exponential growth in the number of small, privately owned satellites has opened wide new vistas for investigative reporting.
- Media access to satellite company images data bases has resulted in journalists gaining a near real-time reporting source for breaking stories.
Satellite Journalism

• Satellite imagery has provided ‘smoking-gun’ evidence for:
  ▪ investigative reporting of human rights abuses (Htusan et al., 2015)
  ▪ exposure of environmental destruction in remote locations (e.g., the Amazon: Weir, 2017)
  ▪ before & after imagery of natural disasters (Murphy, 2018)
  ▪ analyses of military build-ups (e.g.: China’s construction of artificial islands in the South China Sea Wu et al., 2018)
  ▪ missile launches and nuclear testing in North Korea (Fifield and Taylor, 2018)

Source: Mark Corcoran (2018), *Satellite Journalism – The Big Picture*,
https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2018-10/Mark%20Corcoran%2C%20Satellite%20Journalism_1.pdf
The La Pampa gold mine in Peru

In 2016, mining operations were extended S. of the Malinowski River, illegally entering the Tambopata National Reserve—a protected forest.

The Amazon Conservation Association used Planet Labs data to track 100s of hectares of illegal expansion and mapped alterations to the course of the Malinowski River.

The Peruvian government intervened and actively targeted illegal mining operations inside the reserve.

Images: Planet Labs
Summary

• DDJ reflects increased interaction between journalists and {visualization design, computer science and statistics} in the digital age.

• New data sources:
  ▪ FOI requests
  ▪ Open Data initiatives
  ▪ Satellite journalism
  ▪ Crowd sourcing

• New graphic forms & reporting
  ▪ Novel graphic displays
  ▪ Interactive story telling

• Powerful impact