

FINDING STORIES IN DATA

Local voices. Global change.

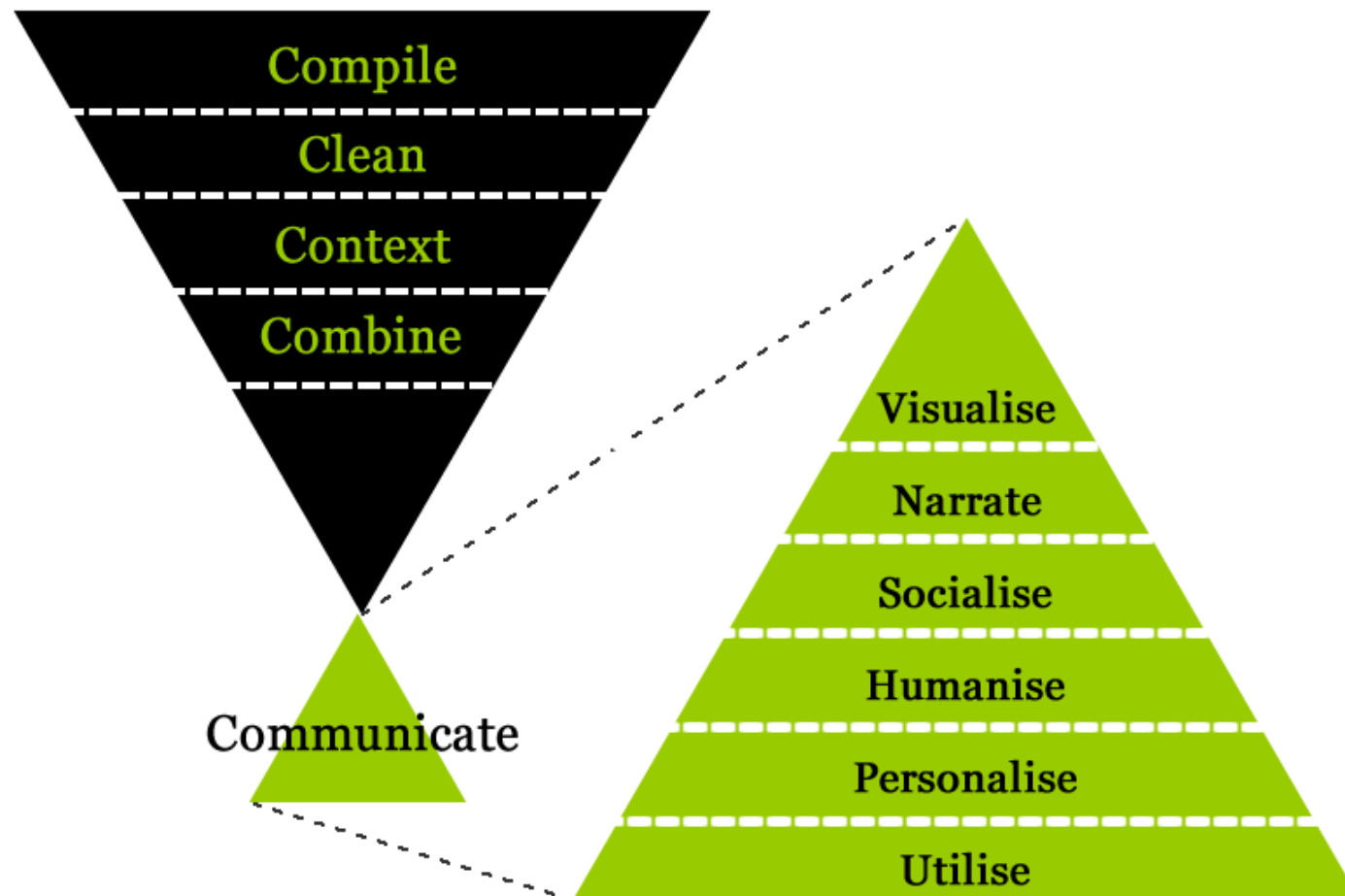
Contents

- The Data Journalism Process
- Data sources: recurring events
- Exercise
- Data sources: sent or leaked data
- Data sources: breaking news
- Data sources: theories to be explored
- Exercise
- Summary & References

The Data Journalism Process

The Inverted Pyramid of Data Journalism (complete)

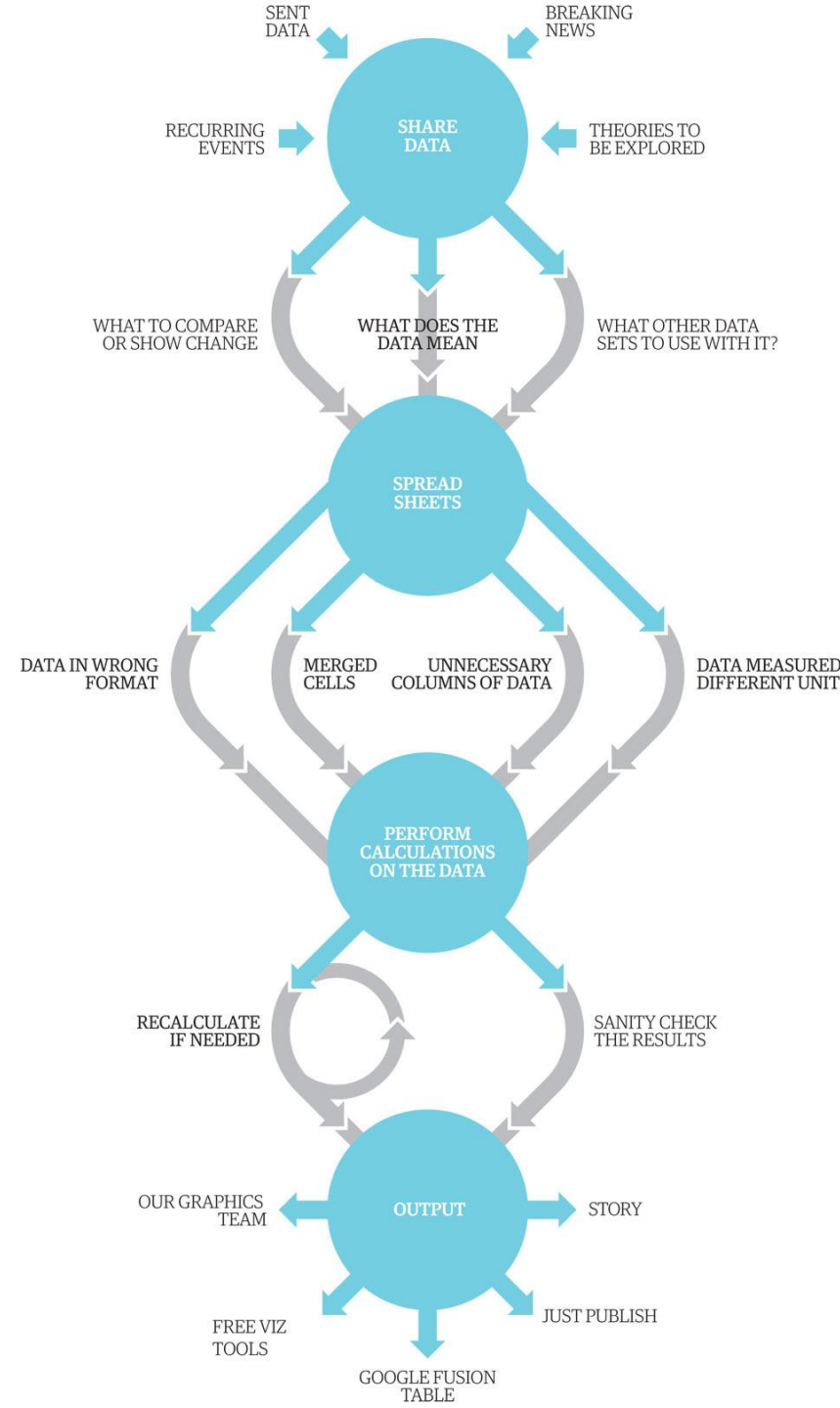
Paul Bradshaw, OnlineJournalismBlog.com



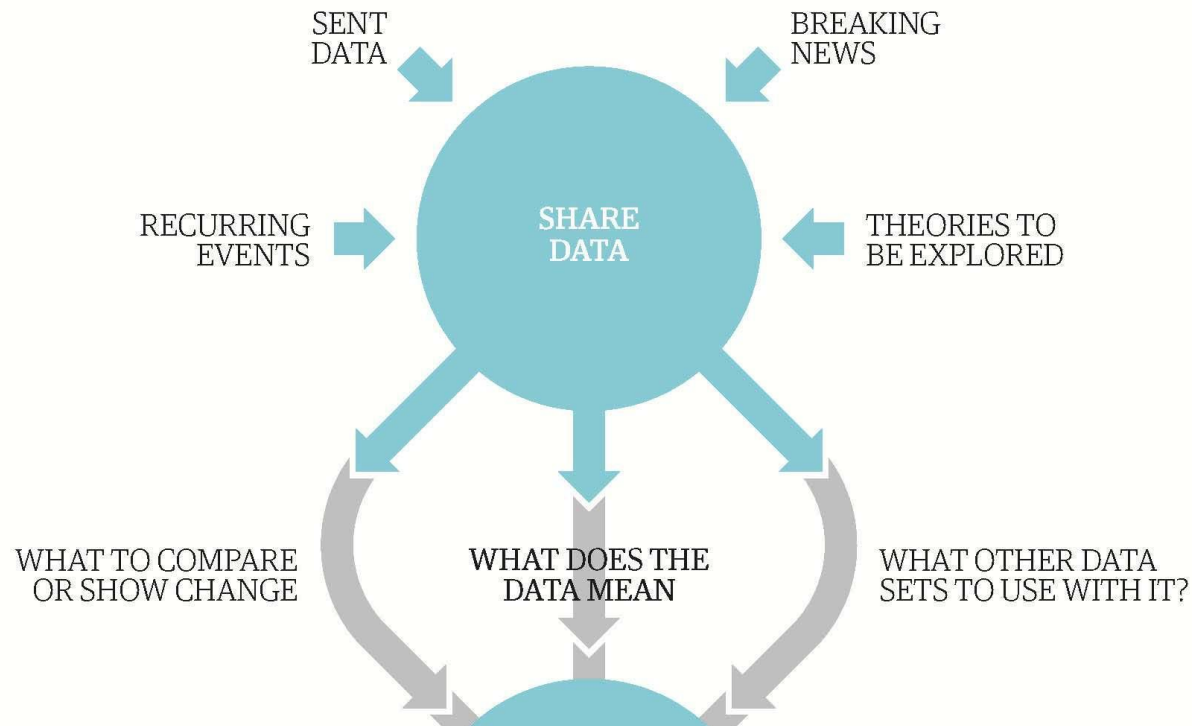
The Data Journalism Process

The Guardian:

1. Share Data
2. Spread Sheets
3. Perform Calculations
4. Output



The Data Journalism Process



Data sources

- **Data sources: recurring events**
- Data sources: sent or leaked data
- Data sources: breaking news
- Data sources: theories to be explored

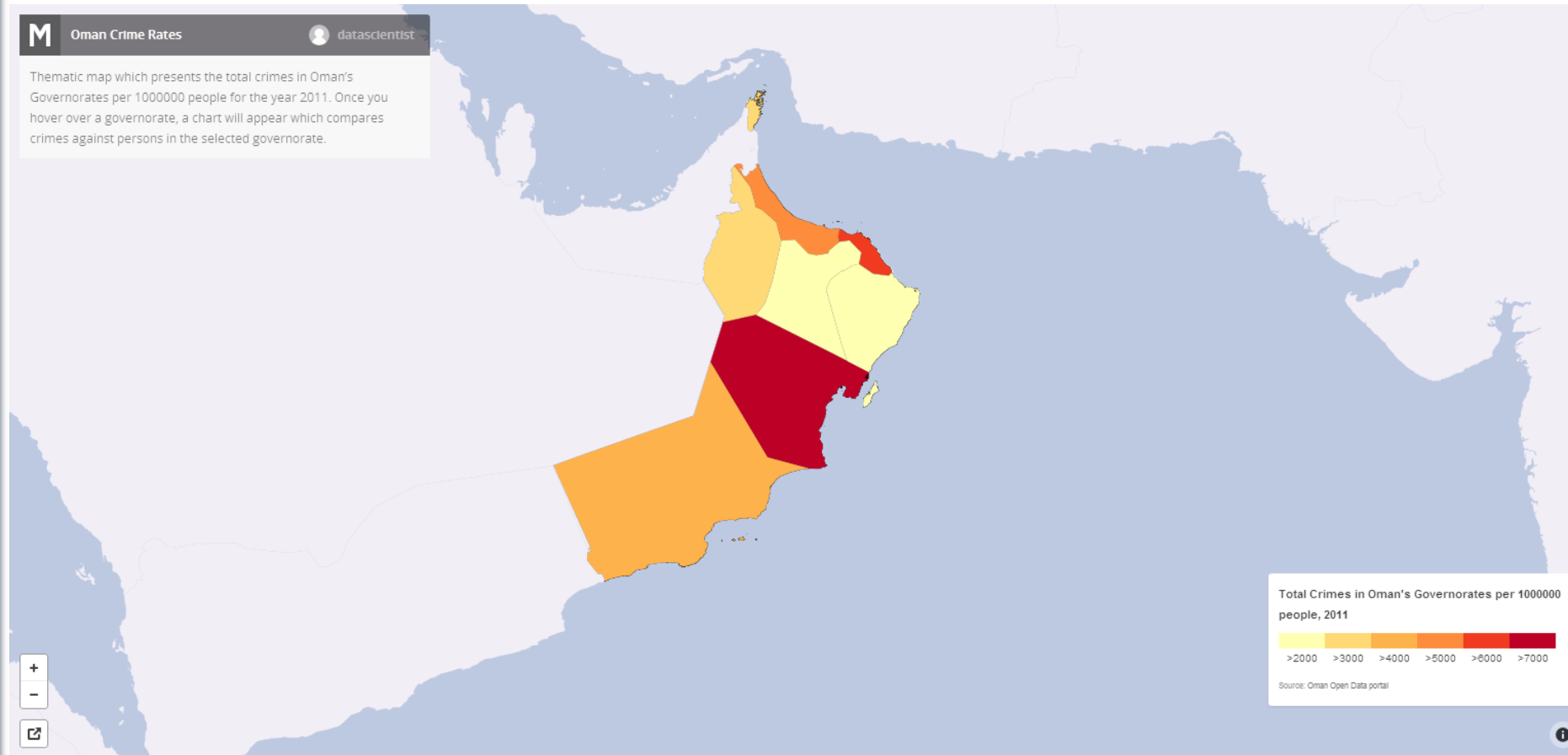
Data sources

- **Data sources: recurring events**
 - Governance & Security Indicators
 - Economic Indicators & Budgets
 - Education Indicators
 - Human rights & Equality Indicators
 - Environmental Indicators
 - Health Indicators

Data sources

- Data sources: recurring events
 - **Governance & Security Indicators**
 - Economic Indicators & Budgets
 - Education Indicators
 - Human rights & Equality Indicators
 - Environmental Indicators
 - Health Indicators

Data Sources: Recurring Events



Crime In Oman Visualized: For the first time, the government of Oman released crime data, which was visualized by a local blogger



Internews
Local voices. Global change.

Data Sources

- Data Sources: recurring events
 - Governance & Security Indicators
 - **Economic Indicators & Budgets**
 - Education Indicators
 - Human rights & Equality Indicators
 - Environmental Indicators
 - Health Indicators

Data Sources: Recurring Events

Ukraine 2013

budget: A Ukrainian newspaper created an interactive data visualization when the government released its annual budget so that citizens could evaluate how the government is spending its tax dollars

facebook.com/sumynews
twitter.com/sumynews
youtube.com/thesumynews

SUMY NEWS

Інтернет-телебачення

Медіа-платформа та прес-центр неурядової аналітичної організації "Центр досліджень регіональної політики" [м. Суми]

Засновано в січні 2009 року

sumy Медіа-платформа Колонка автора Тексти Data journalism Відео Громадський простір Історичний портал ЦДРП

Новини Політика Журналістика цифрового майбутнього

news **BREAKING NEWS** Жесть відео: губернатора Волині прикували наручниками до сцени

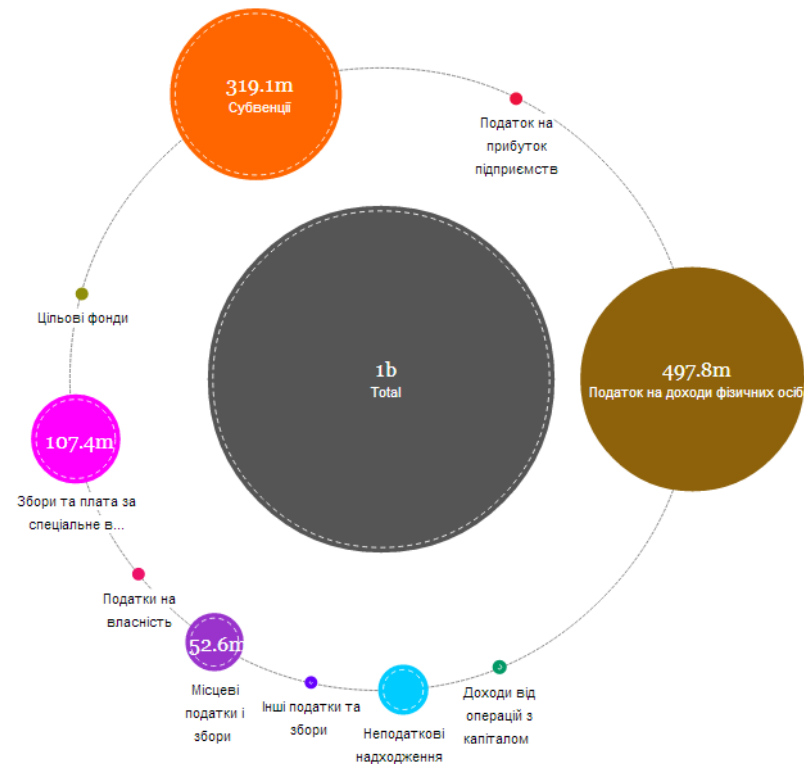
02 січ. 2013 Бюджет м. Суми на 2013 рік [OpenSpending]

 Tweet 0  Подобається! Зареєструйтеся, щоб побачити, що подобається Вашим друзям.  +1

26 грудня 2012 року Сумська міська рада схвалила міський бюджет на 2013 рік, але при цьому велике занепокоєння викликає перспектива його виконання з причин:

- обсяг доходів, які враховуються при визначенні міжбюджетних трансфертів, затверджений в сумі 498651,1 тис. грн. (додаток 6 Закону України "Про Державний бюджет на 2013 рік"), що на 67696,6 тис. грн., або на 15,7% більше очікуваного надходження в 2012 році. При цьому, констатуємо, що невиконання затвердженого показника прогнозується в сумі 32183,3 тис. грн.

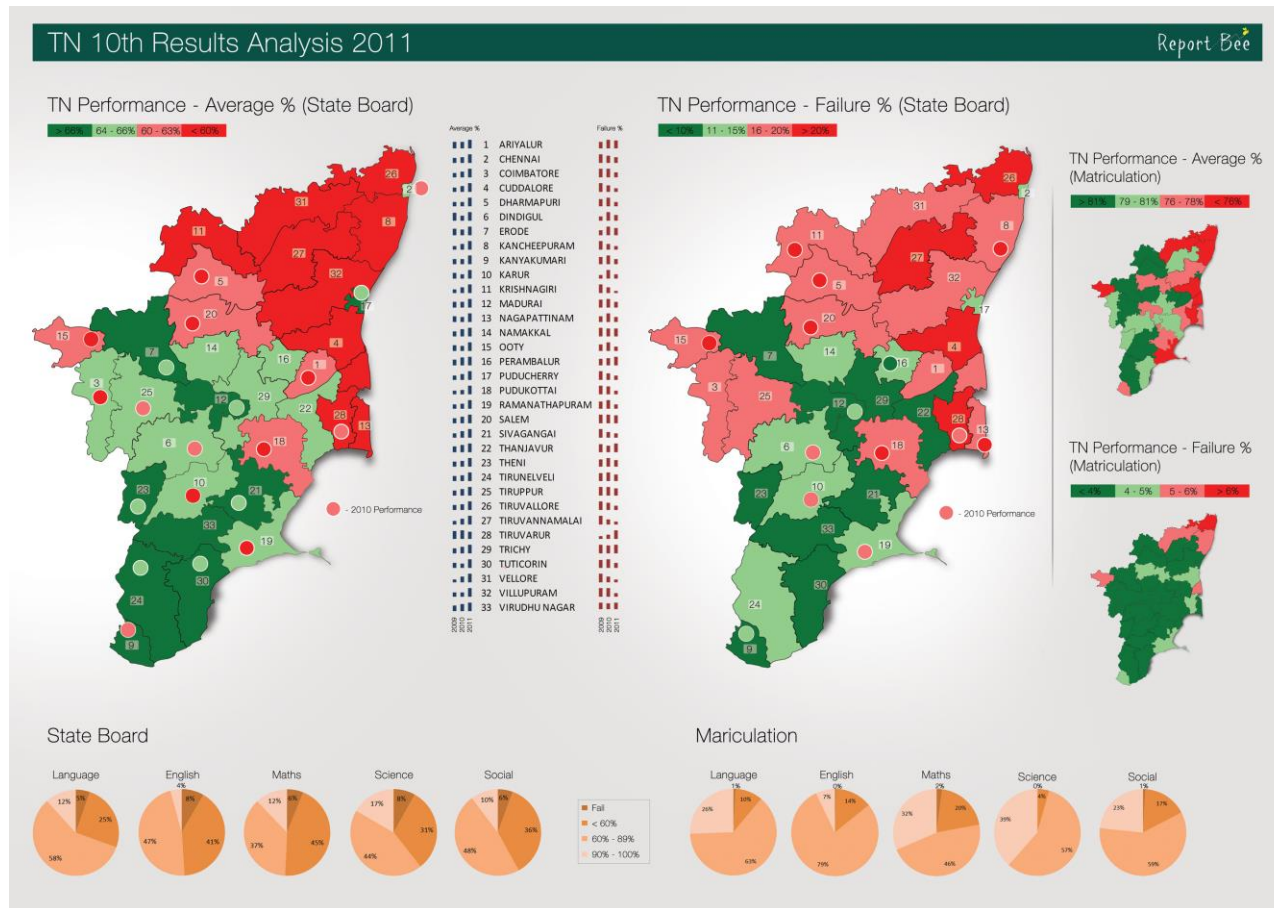
Доходи бюджету м. Суми в 2013 році



Data Sources

- Data Sources: recurring events
 - Governance & Security Indicators
 - Economic Indicators & Budgets
 - **Education Indicators**
 - Human rights & Equality Indicators
 - Environmental Indicators
 - Health Indicators

Data sources: Recurring Events



2011 Tamilnadu School Performance: India's Reporter Bee visualized the performance of different school districts based on annual performance reports.

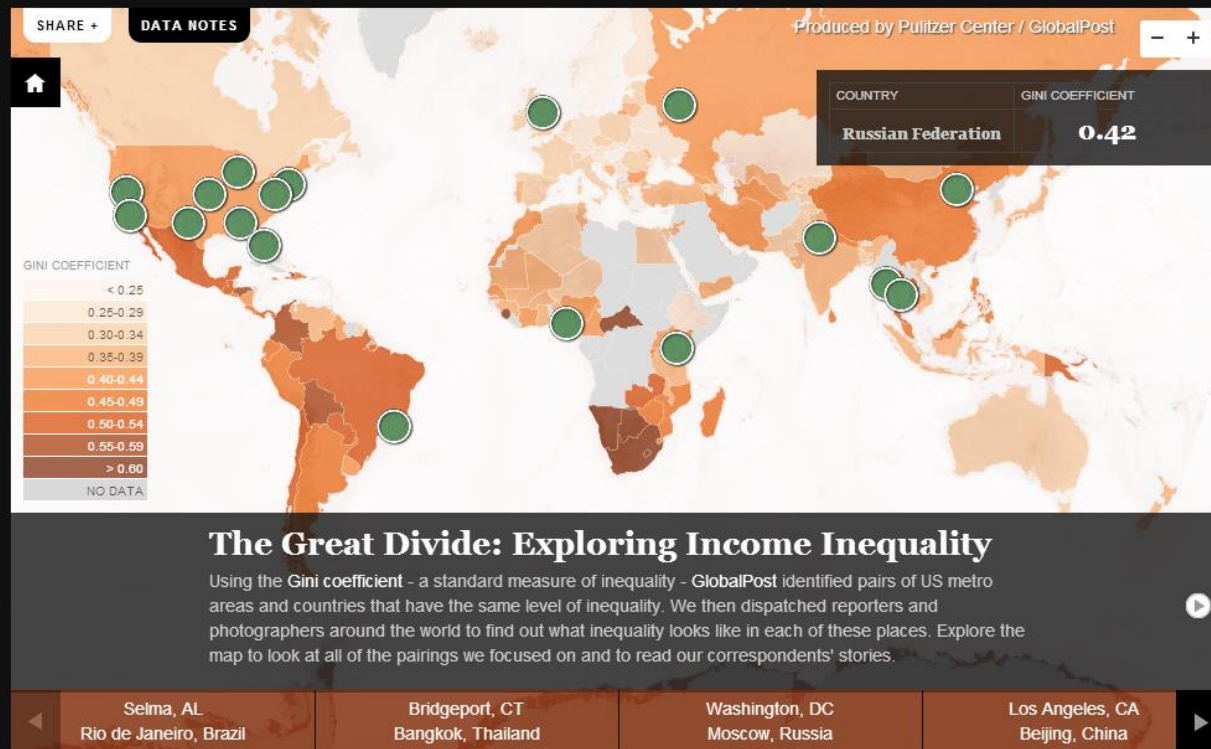
Data Sources

- Data Sources: recurring events
 - Governance & Security Indicators
 - Economic Indicators & Budgets
 - Education Indicators
 - **Human rights & Equality Indicators**
 - Environmental Indicators
 - Health Indicators

Data sources: Recurring Events

Mapping the Divide

IN PARTNERSHIP WITH:
Pulitzer Center



[The Great Divide Global Income Inequality](#): The Global Post maps annual data on global inequality by country and compares cities with similar levels of inequality.



Internews
Local voices. Global change.

Data Sources

- Data Sources: recurring events
 - Governance & Security Indicators
 - Economic Indicators & Budgets
 - Education Indicators
 - Human rights & Equality Indicators
 - **Environmental Indicators**
 - Health Indicators

Data sources: Recurring Events



PROJECT POTICO

<http://wri.org/project/potico>

INDONESIA HAS THE WORLD'S 3rd LARGEST TROPICAL FOREST (2010)

FWI/GFW (2002). The State of the Forest Indonesia. Bogor, Indonesia. Washington D.C.: World Resources Institute, Forest Watch Indonesia, Global Forest Watch.



and yet Indonesia ranks **2nd** in the world for tropical deforestation

498.000 Ha/year

2000-2010 (FAO State of the Forest, 2011)

HALF

of Sumatra's forest cover has disappeared from 1985 to 2008. Forest cover has gone from 50% to 25%.



THE FACTS ABOUT DEFORESTATION AND PALM OIL

Peatland and Land Use/Land Use Change & Forestry it's responsible for

79%
OF INDONESIA'S TOTAL GREENHOUSE GAS EMISSIONS

Almost **90%** of oil palm plantation in Kalimantan from 1990-2010 came at the expense of forest cover.

Carlson, K.M., et al. 2013. "Carbon Emissions from Forest Conversion by Kalimantan Oil Palm Plantations", Nature Climate Change, Nature Publishing Group.

Biodiversity Loss



BUT PALM OIL BRINGS MANY BENEFITS TO INDONESIA

GOVERNMENT

Palm oil is Indonesia's Top Exported Commodity, accounting for 18.97% of 2011's total export (Ministry of Trade, Ministry of Industry) with competitiveness value constantly increasing from 24.1 to 41.05 (2000-2008).

BUSINESS

Global demand for palm oil is expected to increase by 32% to 60 MMT in 2020. This number is 10 times greater than China's annual palm oil consumption in 2012. (Oil World Data Base, 2012)

LOCAL

Almost all consumer products, from soaps to chocolates, depend on crude palm oil. Supporting palm oil production means supporting affordable food prices. (RSPO, 2012)

GLOBALLY

Indonesia is the world's #1 producer and exporter of palm oil, accounting for 50% of the world's total export. Indonesia has set a target for 40 million tonnes of CPO production in 2020, including expansion of oil palm plantation to 10 million Ha in 2014.

THE DILEMMA



THE POTICO CONCEPT: SHIFT TO DEGRADED LAND



POTICO's goal is to prevent deforestation by diverting oil palm plantations away from forests and on to already degraded areas.

Project Potico: World Resource Institute explores deforestation in Indonesia



Internews
Local voices. Global change.

Data Sources

- Data Sources: recurring events
 - Governance & Security Indicators
 - Economic Indicators & Budgets
 - Education Indicators
 - Human rights & Equality Indicators
 - Environmental Indicators
 - **Health Indicators**

Sources of Data: Recurring Events

Child Injury: The Center for Disease Control in the United States created this infographic for journalists to draw attention to a major health risk for children

INJURY

The #1 killer of children in the US



For every **1** child that dies there are...

25

hospitalizations

925

treated in ER

Many

more treated in doctors' offices



In 2005, injuries that resulted in death, hospitalization or an ER visit cost nearly \$11.5 billion in medical expenses.

SOURCES: Web-based Injury Statistics Query and Reporting System (WISQARS), CDC, 2009.

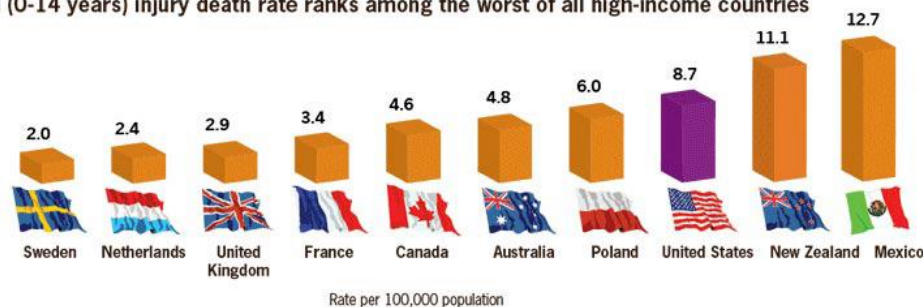
National Health Interview Survey, 2009 data release, CDC, National Center for Health Statistics.

Global view

The US child (0-14 years) injury death rate ranks among the worst of all high-income countries

The US child injury death rate is 4 times that of countries with the lowest rates.

SOURCE: Global Health Observatory Data Repository. Mortality and burden of disease: disease and injury country estimates, 2008, by sex and age.



Rate per 100,000 population



Internews
Local voices. Global change.

Contents

- The Data Journalism Process
- Data sources: recurring events
- **Exercise**
- Data sources: sent or leaked data
- Data sources: breaking news
- Data sources: theories to be explored
- Exercise
- Summary & References

Exercise

How could these recurring reports be used in your health reporting?

- Disease burden by county
- Economic growth data by sector
- Maternal and child mortality reports
- Other reports, press releases and data sent to your media?

Contents

- The Data Journalism Process
- Data sources: recurring events
- Exercise
- **Data sources: sent or leaked data**
- Data sources: breaking news
- Data sources: theories to be explored
- Exercise
- Summary & References

Data Sources: Sent or Leaked Data

Wikileaks in Kenya

- The [Looting of Kenya](#) published by *The Guardian* in 2007.
- [Kenya: The Cry of Blood - Report on Extra-Judicial Killings and Disappearances](#) - the [Kenya National Commission on Human Rights](#) (KNCHR) 2008.
- Bringing international attention to corruption issues too dangerous to print in-country.

[Project Potico](#): World Resource Institute explores deforestation in Indonesia



Internews
Local voices. Global change.

Data sources: sent or leaked data



Offshore Leaks: A leaked hard drive of 2.5 million documents about offshore bank accounts to the International Consortium of Investigative Journalists

Contents

- The Data Journalism Process
- Data sources: recurring events
- Exercise
- Data sources: sent or leaked data
- **Data sources: breaking news**
- Data sources: theories to be explored
- Exercise
- Summary & References

Data Sources: Breaking News

Data collection:

- Crowd-sourcing via Ushahidi-type platforms

- Citizen media contributions

- Sensor data for environmental catstrophes

Risks:

- Selection bias

- Verification process

- Context



Data sources: breaking news

Webpage Screenshot

THE WESTGATE ATTACK A SOCIAL MEDIA PERSPECTIVE

EVENT

GLOBAL CONTEXT

ATTACKERS

***●

KEY ACTORS

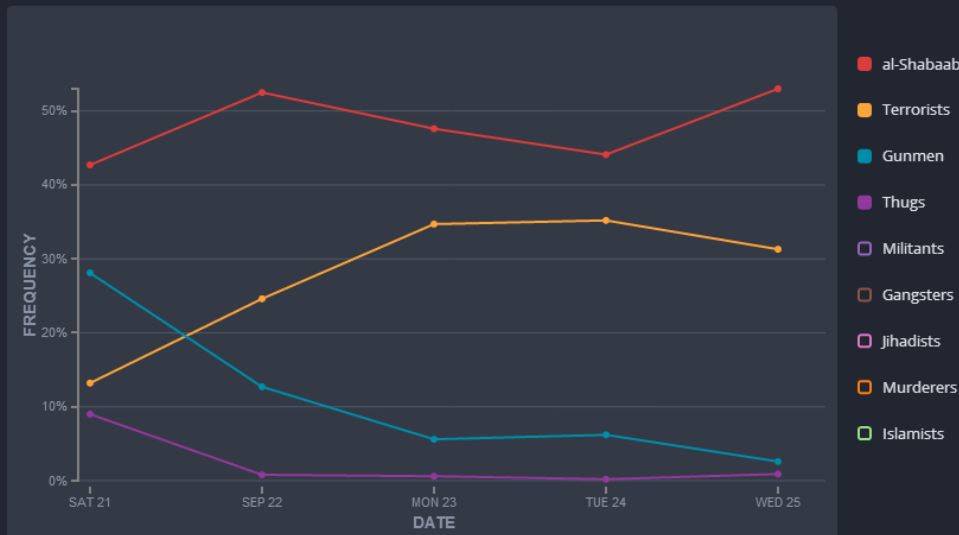
WITNESSES

**

WORD FREQUENCY

This multiple line graph depicts the usage dynamics of the most common words used to describe the attackers by Twitter users in the online conversation.

Check the box for the words you're interested in to visualize their frequency throughout the four days of the attack. For easier comparison among words, check multiple boxes at the same time. Additionally, you can hover over a word to hide unrelated line segments.



<http://lstout.github.io/westgate/html/> Tue May 20 2014 10:42:37 GMT+0300 (E. Africa Standard Time)

The Westgate Attack: A Social Media Perspective: An analysis of the events of Westgate driven by analysis of social media content.

Data Sources: Theories to be Explored

Global and open data catalogs

[World Health Organization](#)

[United Nations](#)

[World Bank](#)

[DataCatalogs.org](#)

[The Guardian's world government data portal](#)

[Google's public data directory](#)

[The data hub](#)

[DBPedia Datasets](#)

[Factual](#)

[Free GIS data](#)

[List of open data resources](#)

[Energy data repositories](#)

[World Research Institute](#)

[Data wrangling](#)

[Quora thread: "Where can I find large datasets open to the public?"](#)

[Directory of APIs](#)

[Infochimps](#)

[Datamarket](#)

[Offshore Leaks](#)

[Investigative Dashboard](#)

[Open Corporates](#)

[Natural Earth data](#)

[UNEP Data](#)

[Transparency International](#)

[Corruption Index](#)

[Land Ownership Database](#)

[Gapminder World](#)



Data Sources: Theories to be explored

Kenyan Data Sources

[Mars Group](#)

[Africa Open Data](#)

[Kenya National Bureau of Statistics](#)

[E-promis Kenya](#)

[International Budget Partnership County Budgets](#)

[Kenya Open Data Initiative](#)

Kenya Health Data

[World Bank Boost](#)

[Kenya Demographic Health Survey](#)

[Internews Data Dredger](#)

[Kenya Service Delivery Indicators](#)



Internews
Local voices. Global change.

Exercise

Divide into two groups and read:

- [Tourism Players Want State to Wave Taxes](#)
- [Kenya Bids to Revive Tourism as al-Shabaab Vows Broader War](#)

Exercise

Answer

1. What is the journalist's hypothesis?

- What topics does the journalist cover in the story?
- What are the most important impact in changing tourism revenue according to the story?
- What does he or she want the reader to understand after reading the story?

Exercise

Answer

2. What is the journalist's evidence for the story?

- What data does he or she use to prove his hypothesis?
- What experts does he or she quote?
- Does the author use information from the national budgets or data on Kenya's revenue sources to enrich the story?

Exercise

Answer

3. What information and data is missing?

- What questions does the story not answer?
- What other data about the budget, taxes, economics, etc. would add to the story?
- What would your data angle for the story be?

Contents

- The Data Journalism Process
- Data sources: recurring events
- Exercise
- Data sources: sent or leaked data
- Data sources: breaking news
- Data sources: theories to be explored
- Exercise
- **Summary & References**

Summary & Resources

- Data sources include recurring events, sent or leaked data, breaking news and theories to be explored.
- Data from these stories can be used to enrich traditional reporting, find new angles on old stories, uncover corruption or engage citizens in the reporting process.
- Data sources are all around us: NGOs, research institutions, government and ordinary people.

Summary & Resources

