



Collecting, Interpreting Risk Data in Africa

SEPTEMBER 2019

Five Steps of Analytics



The Scenario

The Scenario:

- 500-million USD finance for prevention, treatment of HIV/AIDS
- Extensive infrastructure, testing, treatment investment

The Challenges:

- Almost no awareness of program among population
- Services underutilised
- Vulnerable populations at risk

What Data Is Out There, Where?



ARCHIVES

Archives of scanned documents, statements, insurance forms, medical record and customer correspondence, paper archives, and print stream files that contain original systems of record between organizations and their customers



DOCS

XLS, PDF, CSV, email, Word, PPT, HTML, HTML 5, plain text, XML, JSON, etc.



MEDIA

Images, videos, audio, Flash, live streams, podcasts, etc.



DATA STORAGE

SQL, NoSQL, Hadoop, doc repository, file systems, etc.



BUSINESS APPS

Project management, marketing automation, productivity, CRM, ERP content management systems, HR, storage, talent management, procurement, expense management, Google Docs, intranets, portals, etc.



PUBLIC WEB

Government, weather, competitive, traffic, regulatory, compliance, health care services, economic, census, public finance, stock, OSINT, the World Bank, SEC/Edgar, Wikipedia, IMDb, and other Web services



SOCIAL MEDIA

Twitter, LinkedIn, Facebook, Tumblr, Blog, SlideShare, YouTube, Google+, Instagram, Flickr, Pinterest, Vimeo, Wordpress, IM, RSS, Review, Chatter, Jive, Yammer, etc.



MACHINE LOG DATA

Event logs, server data, application logs, business process logs, audit logs, call detail records (CDRs), mobile location, mobile app usage, clickstream data, etc.



SENSOR DATA

Medical devices, smart electric meters, car sensors, road cameras, satellites, traffic recording devices, processors found within vehicles, video games, cable boxes or household appliances, assembly lines, office buildings, cell towers and jet engines, air conditioning units, refrigerators, trucks, farm machinery, etc.

Data compiled by the domain experts at Kapow Software, a Kofax company, and is based on almost a decade of experience helping hundreds of large global enterprises and innovative start-ups across industries leverage critical data from disparate internal and external sources to meet business objectives.

Example Data

1. Archives

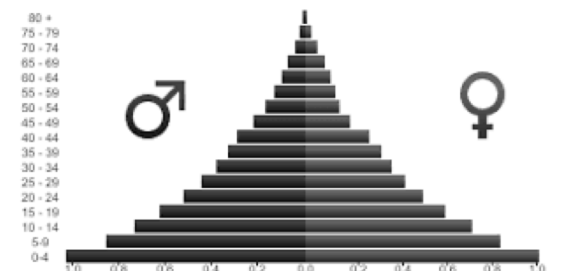
- Historic data (past disasters/events)
- Low cost, valuable for baseline

2. Sensor

- Satellite and/or drone (granular “ground truth”)
- Varied cost, value lies in fidelity

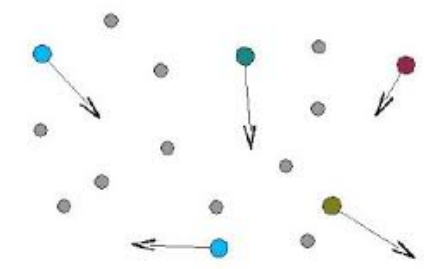
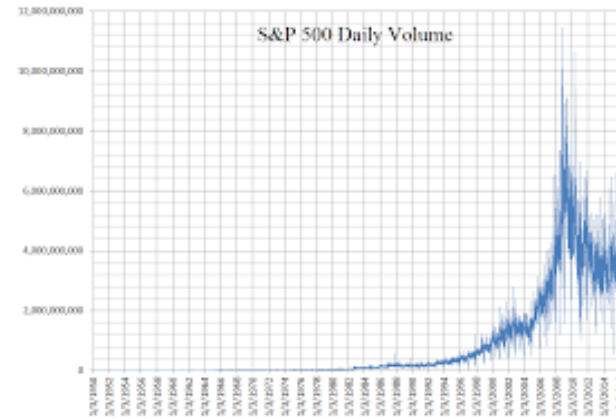
3. Public

- Census data (find high risk populations to direct finances to)
- Usually free, freshness challenge



Characteristics of Data Types – “The V’s”

1. Volume
2. Velocity
3. Variety
4. Veracity
5. And?



Data Challenges in Africa

1. Availability/Access (primary challenge in Africa)
2. Appropriateness/Fit
3. Freshness
4. Cost (free can cost in other ways)
5. Informed Consent (ethics of data)
6. And?

Real-World Development Challenge

The Scenario:

- 500-million USD finance for prevention, treatment of HIV/AIDS
- Extensive infrastructure, testing, treatment investment

The Challenges:

- Almost no awareness of program among population
- Services not used to full capacity
- Vulnerable populations at risk

What Data Could Help, How?



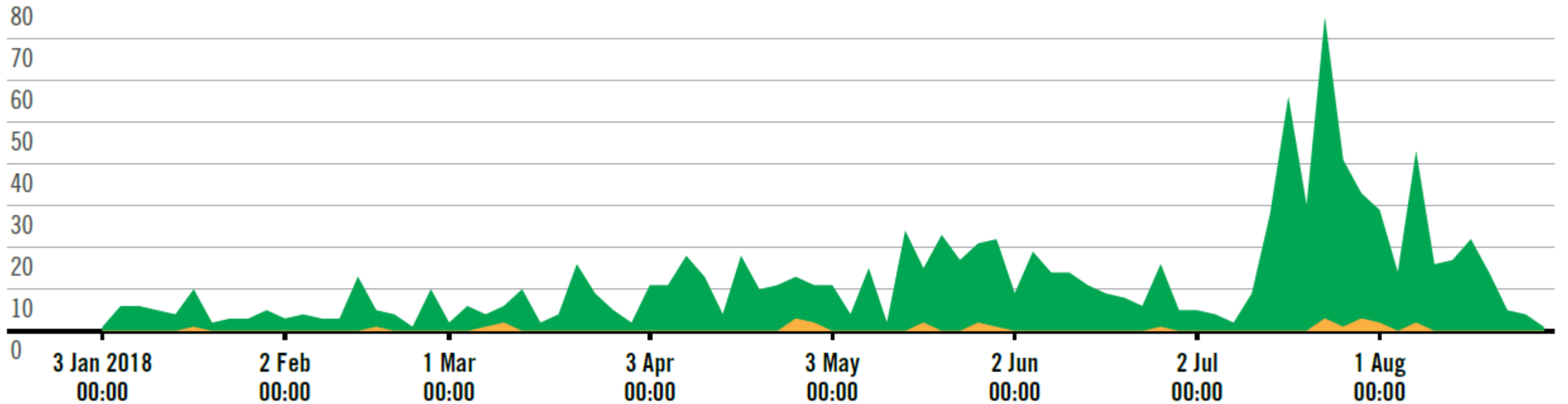
What Data Is Out There, Where?



Questions	Data/Information Required	Potential Sources
How do people find out about our HIV/AIDS testing services?	Media landscape survey.	Media representative bodies, pre-existing surveys.

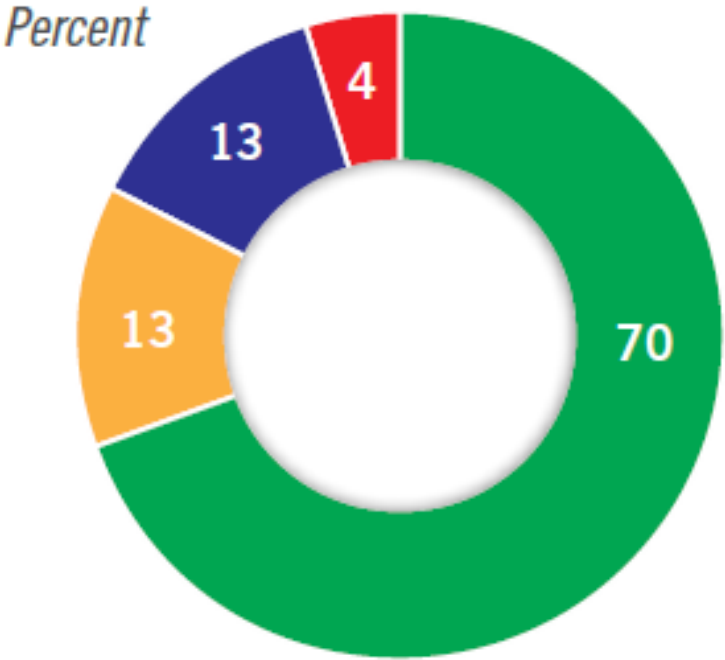
Baseline

Unit of measure

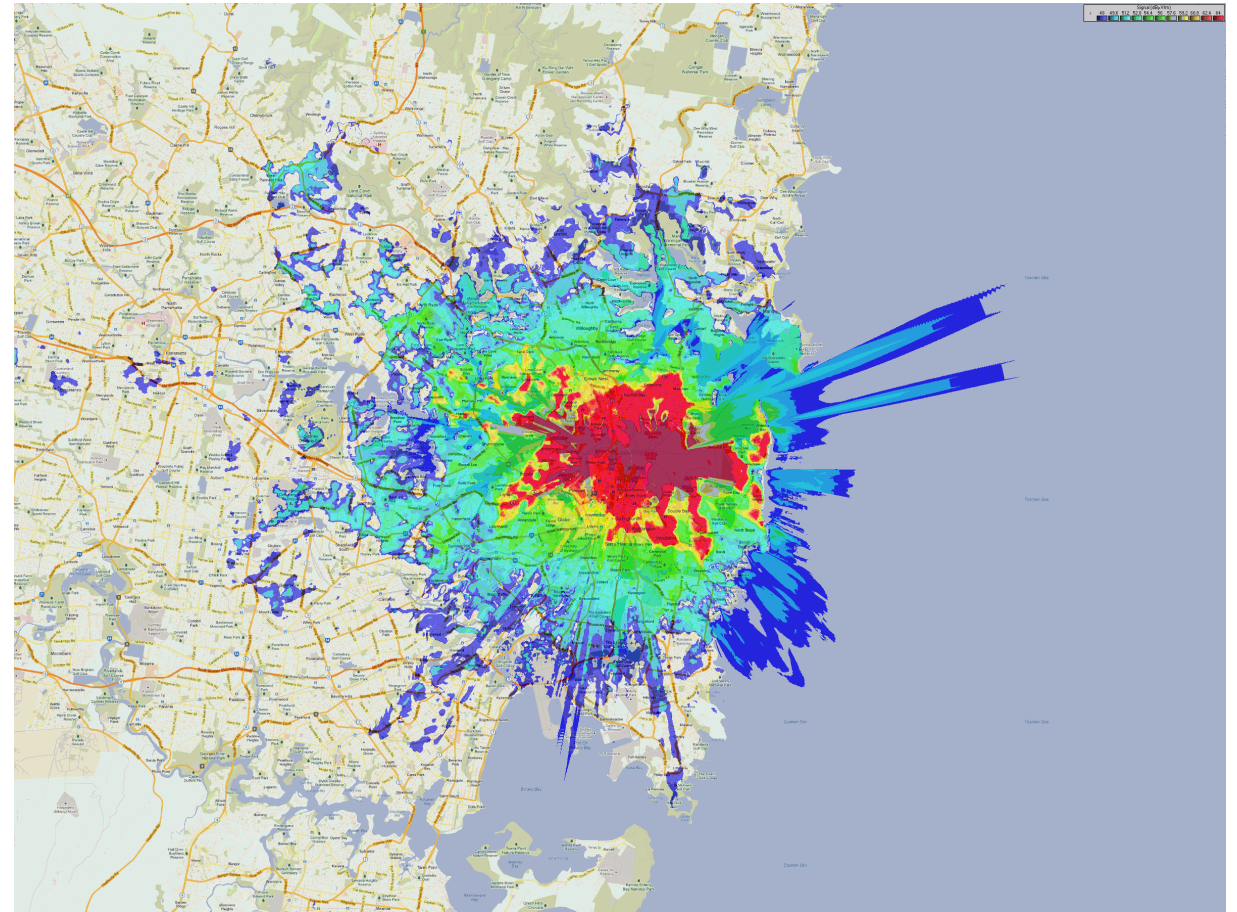
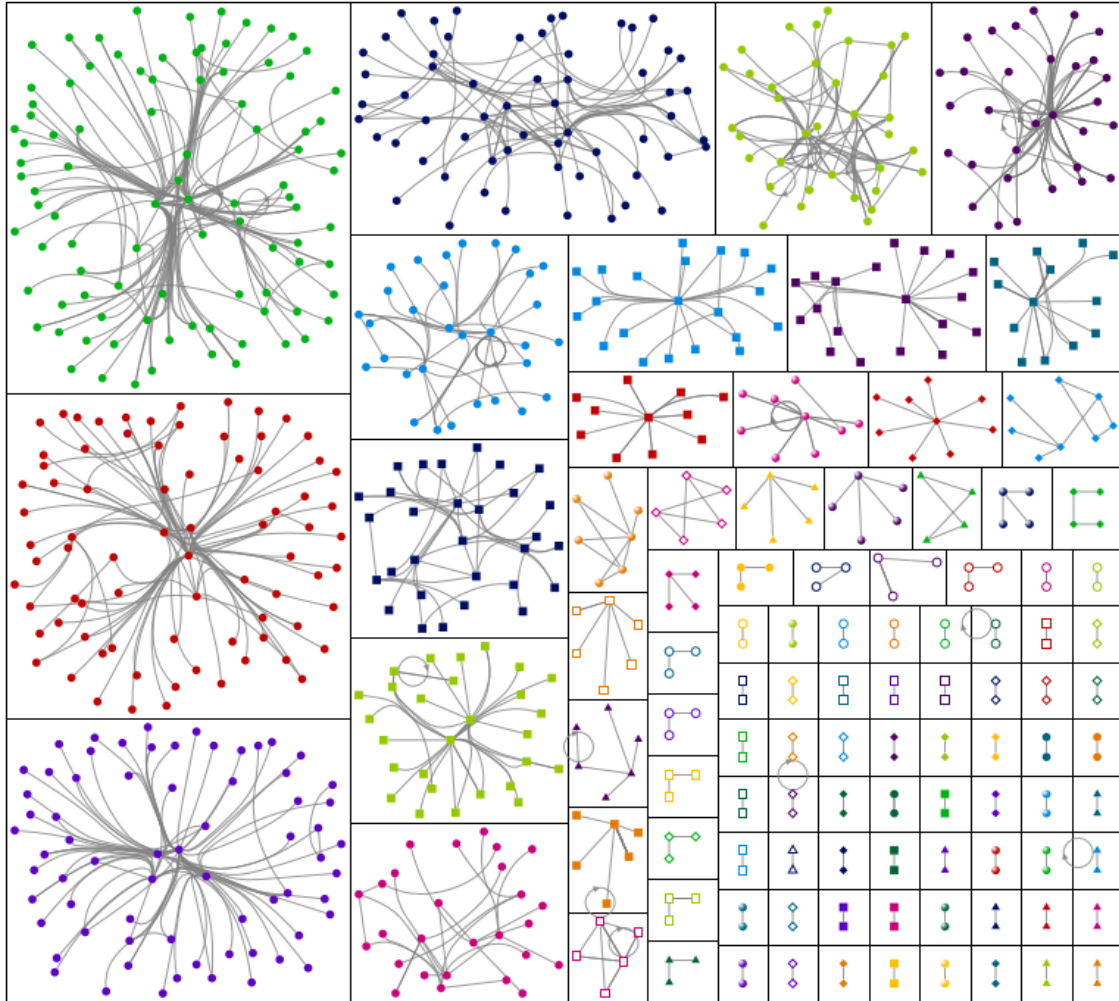


Theme

■ WOMEN ■ GOVERNMENT ■ CELEBRITY ■ TB



Networks



Impact, Bottom Line

- Mention of program doubled in targeted province
- Near doubling of visits to clinics
- Total cost: \$10k (0,002% of total budget)

