Dalberg Data Insights

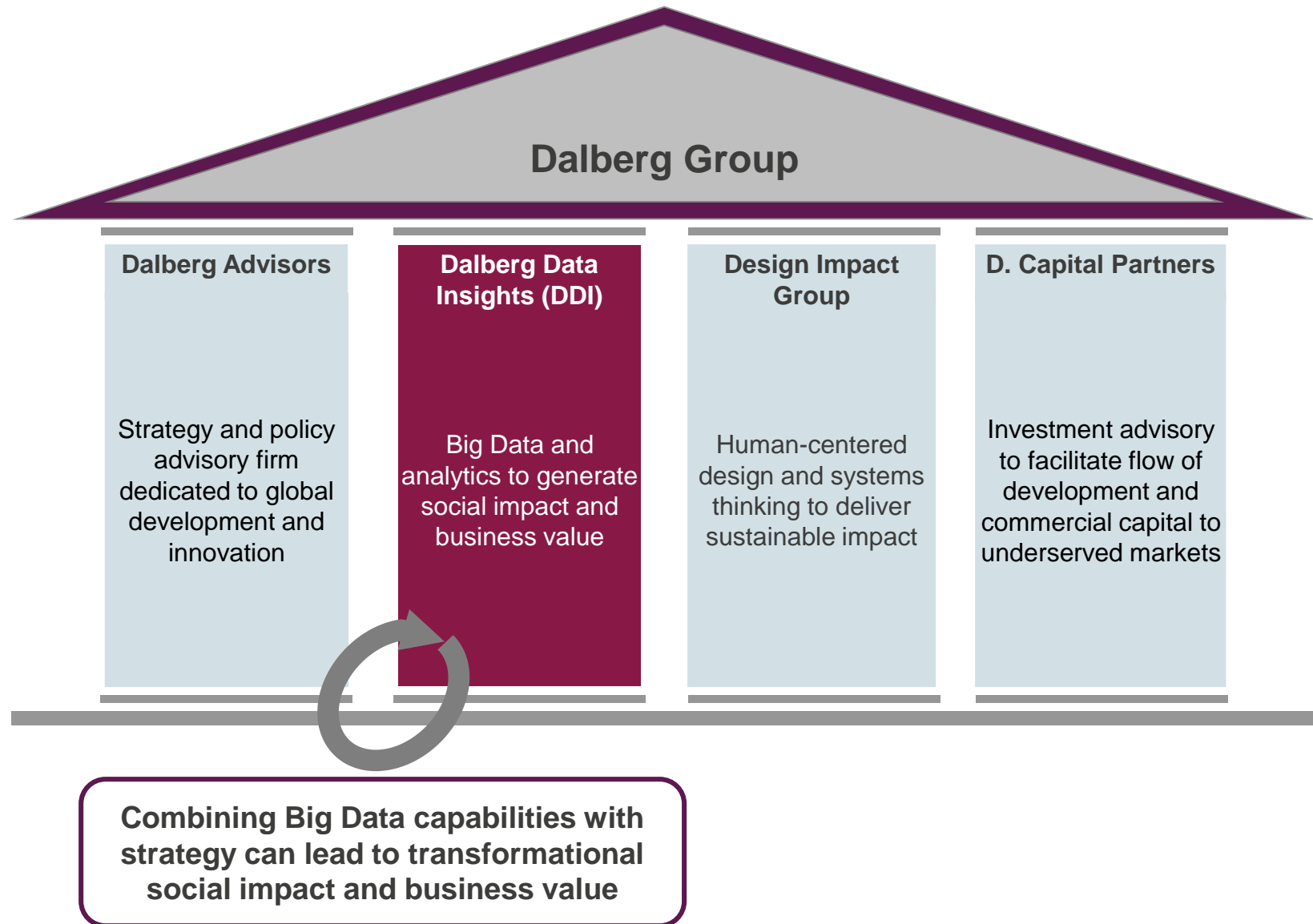# Big Data, Social Impact and Migration

Ispra

November 2017

Agenda

- **Dalberg Data Insights – Our Big Data approach**

- Social Impact & Migration

- Challenges – Data Biases, Partnerships and Sustainability

Dalberg is a platform combining skills and assets

Dalberg Group

Dalberg Advisors

Strategy and policy advisory firm dedicated to global development and innovation

Dalberg Data Insights (DDI)

Big Data and analytics to generate social impact and business value

Design Impact Group

Human-centered design and systems thinking to deliver sustainable impact

D. Capital Partners

Investment advisory to facilitate flow of development and commercial capital to underserved markets

Combining Big Data capabilities with strategy can lead to transformational social impact and business value

Dalberg

The Dalberg Group covers 4 continents and employs 400+ professionals

SAN FRANCISCO
NEW YORK
WASHINGTON DC
DAKAR
LAGOS
BRUSSELS
LONDON
COPENHAGEN
GENEVA
PARIS
ABU DHABI
NEW DELHI
MUMBAI
HONG KONG
ADDIS ABABA
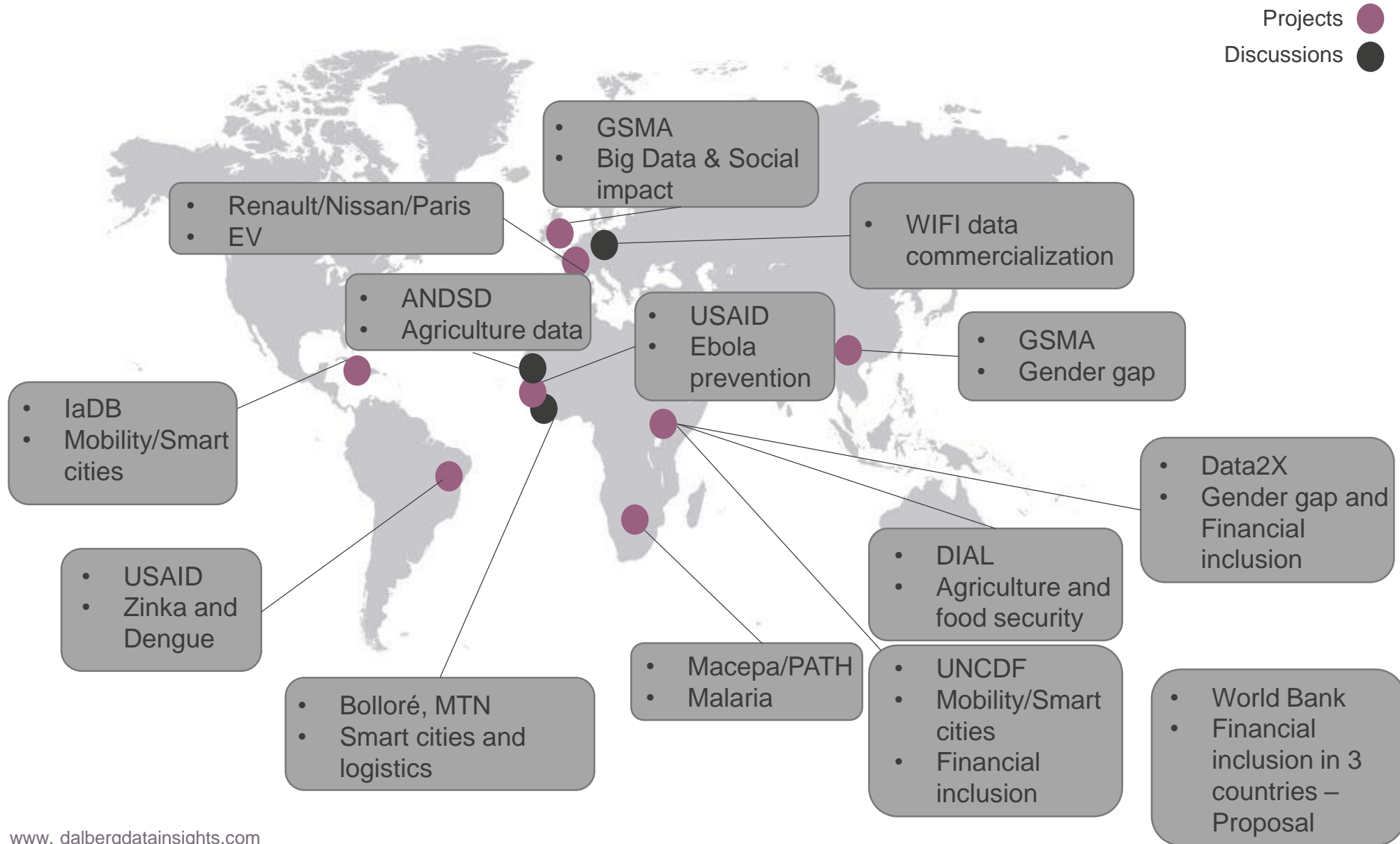SINGAPORE
NAIROBI
DAR ES SALAAM
JOHANNESBURG

Dalberg Data Insights

- **ITU 2017 Prize for Usage of telecom data to improve mobility**

- Part of the Technical Advisory Group leading the Global Partnership for Sustainable Development Data (UN)

GOBAL FACTS

- Founded in 2001
- Experience in 90+ countries

- 400+ professionals, 40+ nationalities
- 50+ languages spoken

- More than 400 clients
- >90% repeat clients

Dalberg

Examples of selected projects over multiple geographies and topics

Projects ●
Discussions ●

- GSMA
- Big Data & Social impact

- Renault/Nissan/Paris
- EV

- WIFI data commercialization

- ANDSD
- Agriculture data

- USAID
- Ebola prevention

- GSMA
- Gender gap

- IaDB
- Mobility/Smart cities

- Data2X
- Gender gap and Financial inclusion

- USAID
- Zinka and Dengue

- DIAL
- Agriculture and food security

- Bolloré, MTN
- Smart cities and logistics

- Macepa/PATH
- Malaria

- UNCDF
- Mobility/Smart cities
- Financial inclusion

- World Bank
- Financial inclusion in 3 countries – Proposal

More specifically, we currently work on 5 topics

1 **Financial Inclusion**

- Where and how to push digital payments?
- Where to further develop distribution network?

2 **Smart cities**

- What are the traffic patterns?
- Where to further develop urban infrastructures?
- How to optimize public transport?
- How to increase usage of EV?

3 **Public health**

- Where to prioritize disease control and eradication?
- Are the quarantine zones enforced?

4 **Agriculture and Gender**

- How to detect early signs of food crises?
- Where do people work and live?
- Where are the poor communities?
- What are the female communities?

Creating platforms of algorithms and tools, while securing privacy

- **All individual data remain within the premises of the data providers**

- **All individual data are anonymized**

- **All individual data are aggregated**

- **All algorithms are open and available**

- **Pushing algorithms to the data**

Big Data smart city platform

Module 1 – Telecom data module	**Module 2 –** Survey data module	**Module 3 –** Administrative data module	**Module 4 –** Retailers' data module
Module 5 – Satellite data module	**Module 6 –** Public transport data module	**Module 7 –** Social media data module	**Module 8 –** Basic technical layers
Module 9 – Mobility monitoring module	**Module 10 –** Public transport module	**Module 11 –** Road network module	…

…

Dalberg

Agenda

- Dalberg Data Insights – Our Big Data approach

- **Social Impact & Migration**

- Challenges – Data Biases, Partnerships and Sustainability

Private and public data sources to offer different insights

Telecom data

CDRs
- Where do people live?
- Where do people work?
- What are the mobility patterns?
- Where to put roads?
- What are the speed?

Signalling
- What are the different modes of transport?

DPI
- What are different socio-economic categories?

Satellite data

Free data
- What are different crops?
- What are the crop yields?

High Res. data
- How to allocate fields?

Public and other private data

Open govern-ment data
- Where do people live?
- What are the demographic trends?

Car data
- How to increase usage of electric cars?
- Where to optimize the power distribution network?

Banking data
- Where to expand the retail network?

Dalberg

Multiple topics around migration share algorithms or data sources

Domestic migrations

- Where do people live? Where do they work? What road do they take?
- What is the impact of changing urban and road infrastructure?
- How to optimize public transport?

- How do people migrate within a country and impact the probability for a disease to spread?
- How to control migrations to reduce public health risks?

- How does the food supply impact migration of populations?
- Is there any correlation between migrations and early alert of food crisis?

Cross-border migrations

- Is it possible to identify international migrations by mapping sources of domestic migration flows at international borders?
- How to map cross-border movements by identifying settlements?

- How border population movements affect the domestic spread of contagious diseases?

- How to manage the distribution of food for international settlements?

Dalberg

We integrate telecom data in intuitive tools for end-users

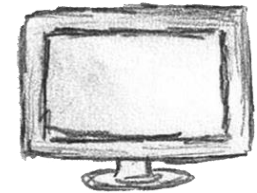| **We integrate data from telco's**	**We anonymize the data**	**We aggregate & enrich data**	**We build intuitive apps for end-users**
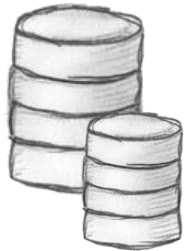
▸ We collect, filter and reconcile CDRs and RAN logs from telco data sources
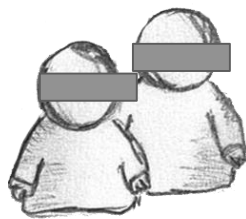
▸ We anonymize location data & subscriber info with hashing algorithms

▸ Our algorithms further process and add intelligence to the raw data

▸ Our intuitive apps provide direct insights and actions to the end-users

Dalberg

Identifying and estimating all traffic flows in and around Kampala

Identify and prioritize the commutes in terms of flow of people and travel time

Prioritization Matrices of Commute Routes
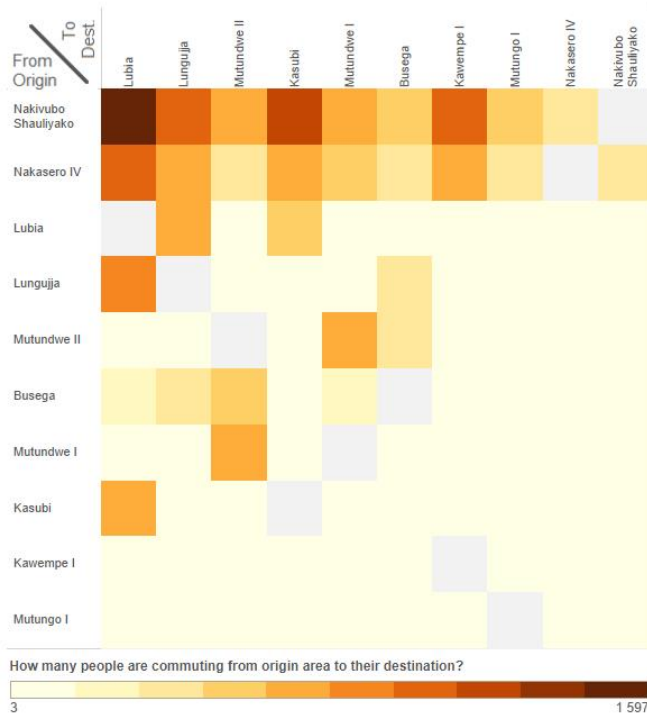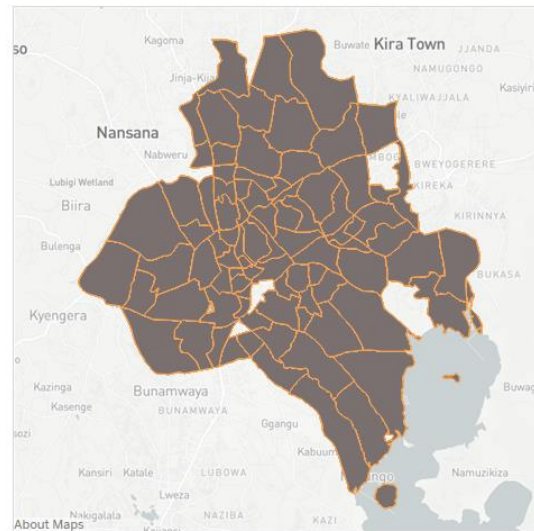
- Identify which commutes are the most popular

- Understand which commutes are the most **time-consuming**

- Obtain more insights about a **specific commute** by selecting a square

- Choose the **number** of commutes displayed in the matrices based on **traffic volume**

Dalberg

Identifying and estimating all traffic flows in and around Port-au-Prince
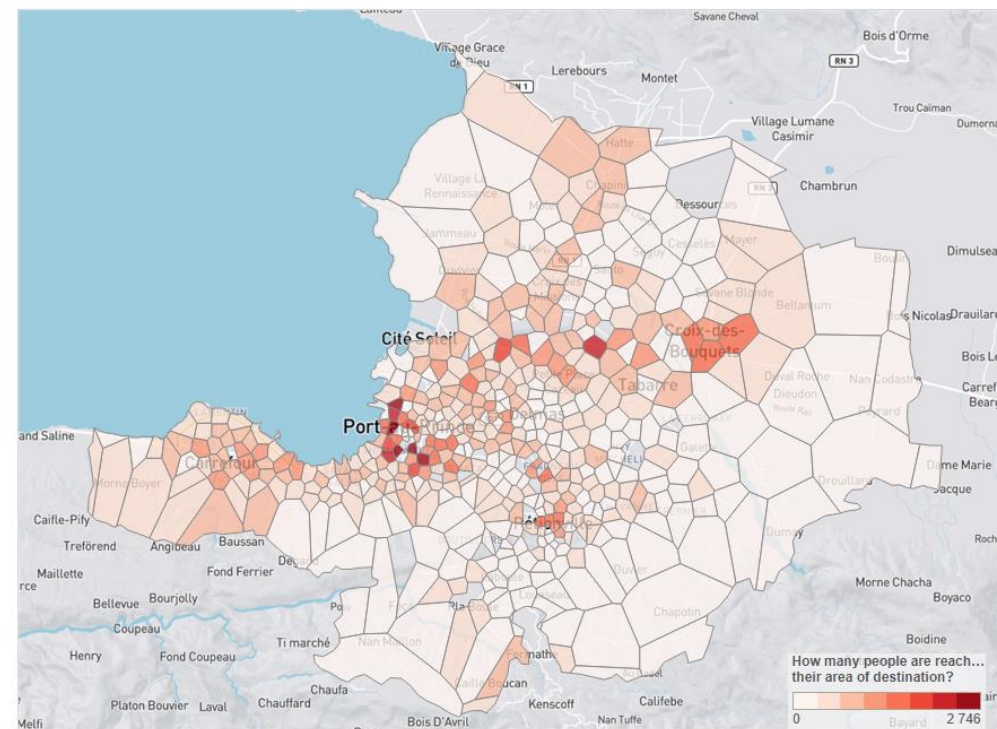
Overview of Commuting Patterns from Origin to Destination

FROM ORIGIN...

⊙ Population leaving their area of origin
◯ Travel time when leaving their area of origin

How many people are leaving their area of origin?

0   Bayard   1 747
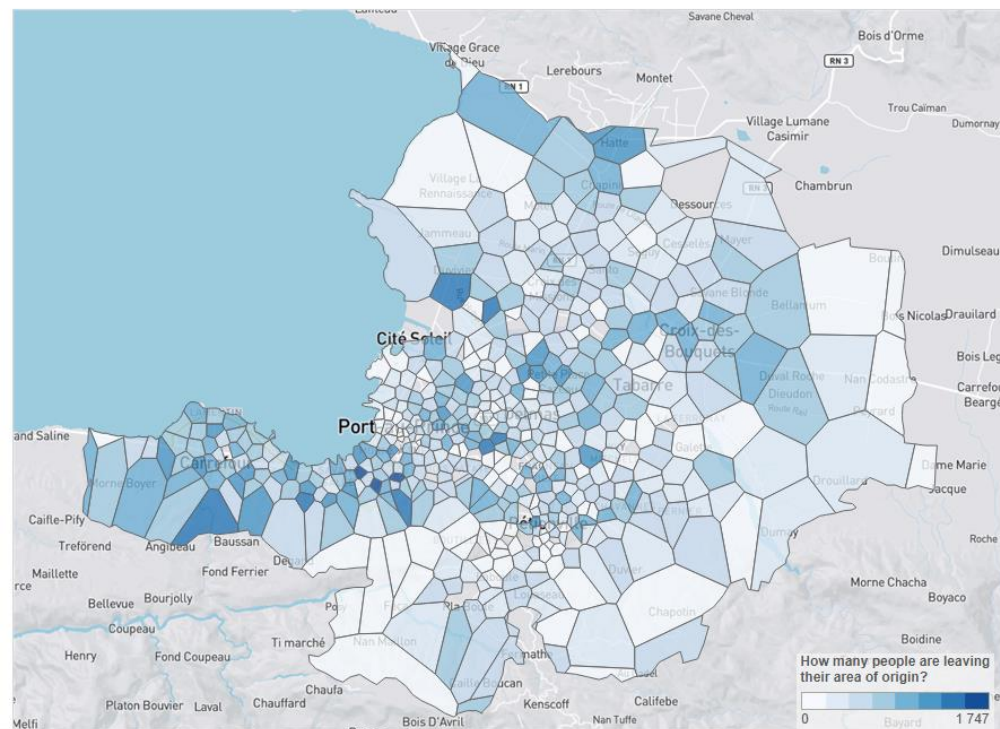
... TO DESTINATION

⊙ Population reaching their area of destination
◯ Travel time for reaching their area of destination

How many people are reach... their area of destination?

0   Bayard   2 746

Dalberg

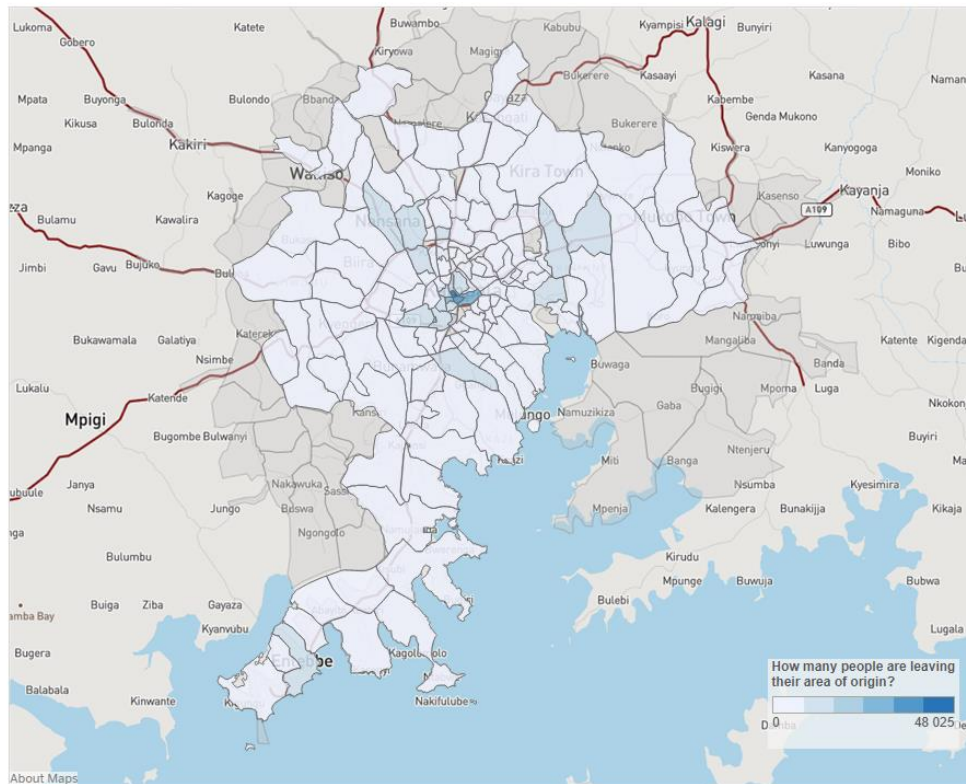# Estimating travel times between any locations

Measure the time spent for commuting from and to any neighborhoods

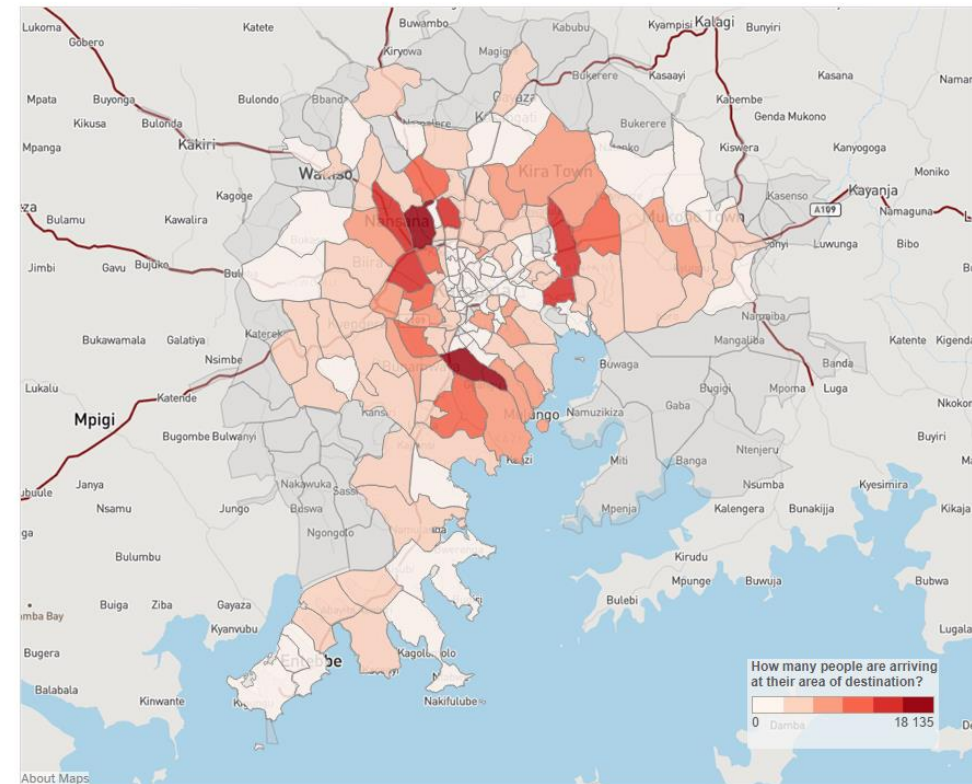- For each neighborhood, understand the travel time **from this neighborhood to every other part of the city**

- For each neighborhood, understand the travel time **to this neighborhood from every other part of the city**

- For every pair of neighborhoods, see the **average commuting time at specific times**

Dalberg

Comparing impact from changes on urban infrastructure
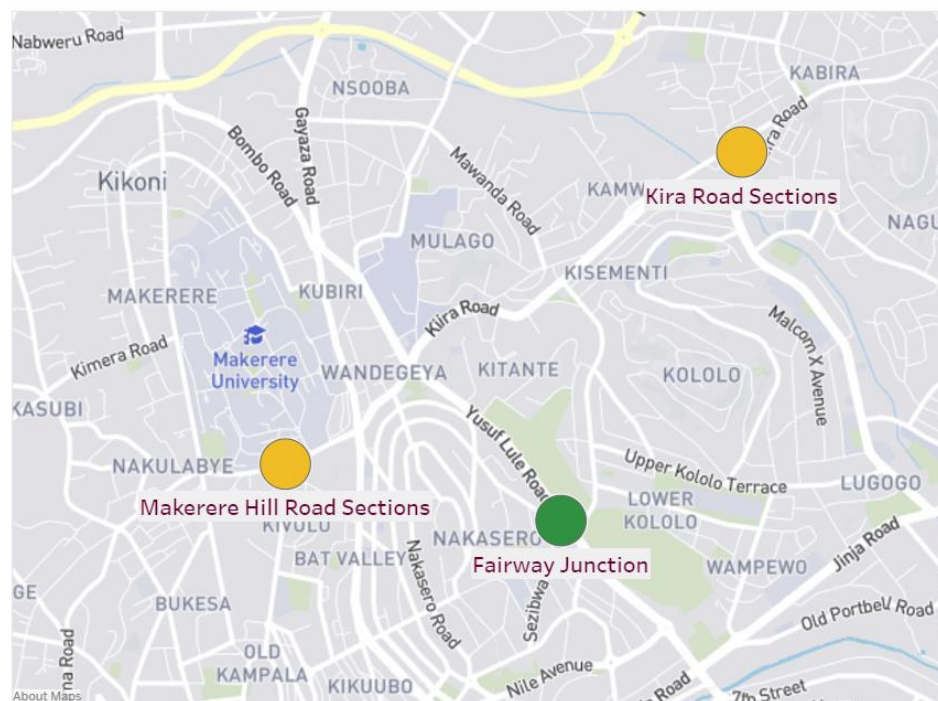
TRAFFIC DIRECTION
- ⊙ Only Traffic entering Junction or Road Sections
- ○ Only Traffic exiting Junction or Road Sections

DAY TYPE

| Business Day | Saturday | Sunday |

TIME PERIOD

| AM Peak | Inter Peak | PM Peak |

Junction / Road Section	Hourly Traffic (ppl/h)	Average Speed (km/h)	Total Time Lost (hours)
Kira Road Sections	+ 68%	- 13%	+ 151%
Fairway Junction	+ 12%	+ 16%	- 15%
Makerere Hill Road Sections	+ 23%	- 13%	+ 147%

Dalberg

Conducting pilots and evaluating granular performance

Using telecom data, we measured mobility performance before and after the building of a new junction in Kampala to understand its impact.

Impact Assessment for Fairway Junction between 2015 and 2017

SHOW ME: Impact Evaluation | Situation in May 2015 | Situation in Feb 2017

TRAFFIC DIRECTION: ○ Only Traffic entering Fairway Junction ● Only Traffic exiting Fairway Junction

DAY TYPE: Business Day | Saturday | Sunday

TIME PERIOD: 04:49 A49/P49 — 10:32 A32/P32 — AM Peak | Inter Peak | PM Peak

Wandegeya
Hourly traffic change: + 19%
Average speed change : + 19%
Total time lost change: - 34%

Kololo
Hourly traffic change: - 6%
Average speed change : + 4%
Total time lost change: - 23%

Nakasero
Hourly traffic change: + 52%
Average speed change : - 8%
Total time lost change: + 95%

Kibuli
Hourly traffic change: + 19%
Average speed change : + 30%
Total time lost change: - 43%

Traffic Zones Legend — Hourly Traffic change (%): -50% … +50%
Road Legend — Average Speed change (...): -50% … +50%

Hourly Traffic (ppl/h)
+ 15%
May 2015: 1 030 — Feb 2017: 1 185

Impact Analysis per Zone
Zones	Change (%)	2017	2015
Kibuli	+ 19%	179	150
Kololo	- 6%	422	449
Nakasero	+ 52%	337	222
Wandegeya	+ 19%	248	209

Traffic over time (ppl)

Average Speed (km/h)
+ 12%
May 2015: 12,9 — Feb 2017: 14,4

Impact Analysis per Zone
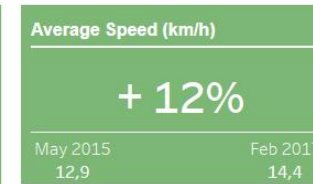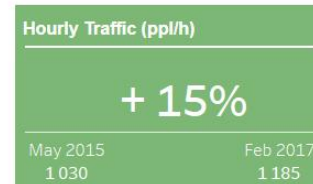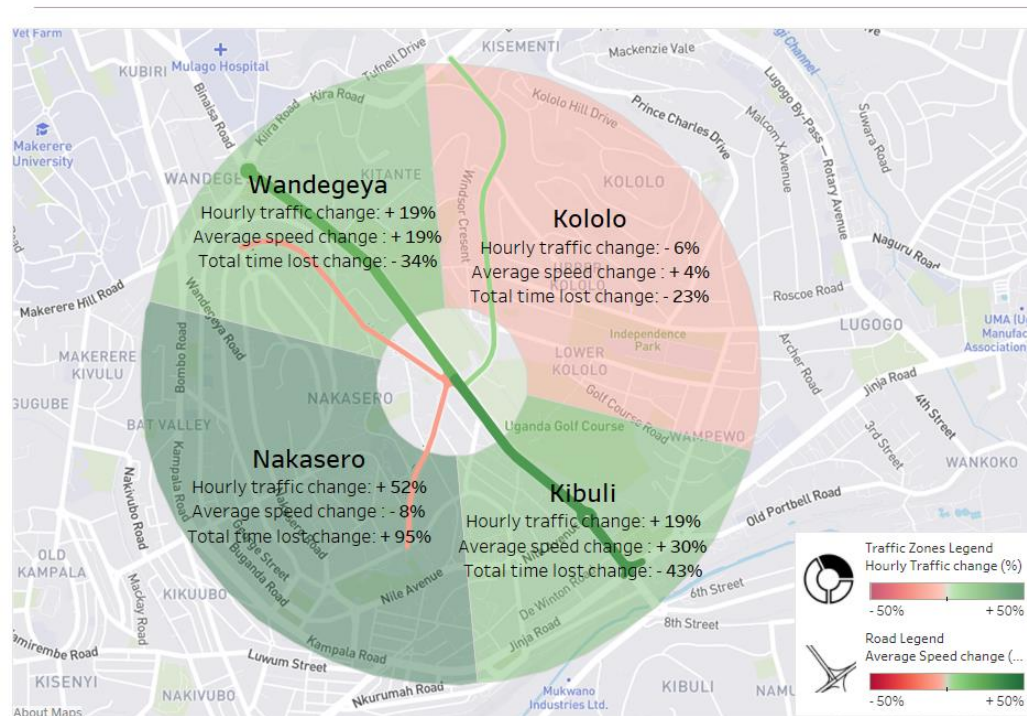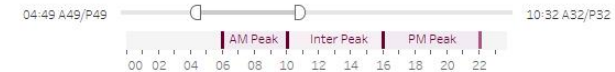| Zones	Change (%)	2017	2015
Kibuli	+ 30%	17,1	13,1
Kololo	+ 4%	13,0	12,4
Nakasero	- 8%	11,4	12,4
Wandegeya	+ 19%	16,1	13,5

Average Speed over time (km/h)

Total Time Lost (hours)
- 6%
May 2015: 379 — Feb 2017: 355

Impact Analysis per Zone
Zones	Change (%)	2017	2015
Kibuli	- 43%	42	74
Kololo	- 23%	114	148
Nakasero	+ 95%	144	74
Wandegeya	- 34%	55	83

Total Time Lost over time (hours)

- Using our algorithms we mapped the origin/destination of daily commuters, estimating the flow of people and their travel time over various time periods.

These insights can be used for:
- Infrastructure planning and decision-making
- Assessing future investments in infrastructure projects
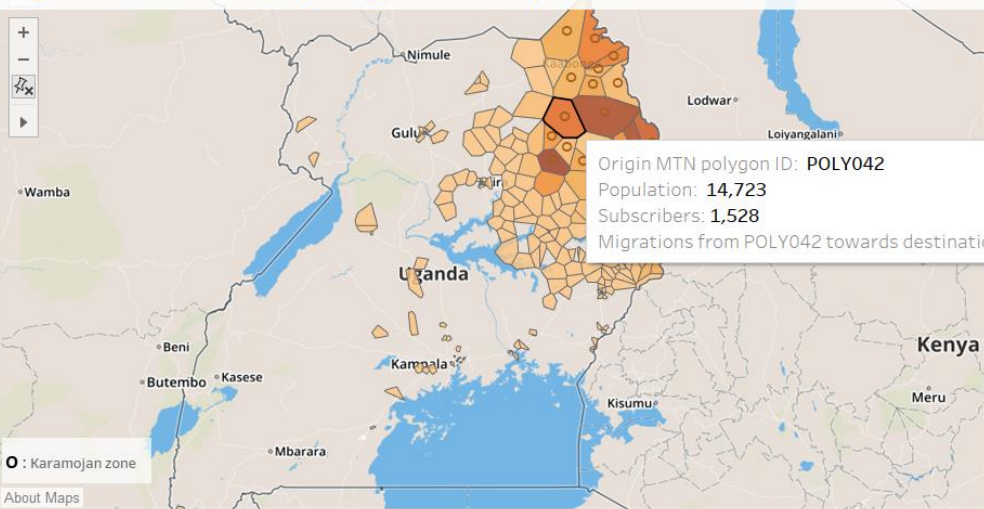
Dalberg

People migrations from (Origin zones) and towards (Destination zones) Karamoja during the defined period of interest (months selection: (All) ▼)

Methodology 💡

Origin zones from where people migrate to Karamoja

Nimule
Gulu
Lodwar
Loiyangalani
Uganda
Kenya
Beni
Butembo · Kasese
Kampala
Kisumu
Meru
Mbarara

Origin MTN polygon ID: POLY042
Population: 14,723
Subscribers: 1,528
Migrations from POLY042 towards destination polygons in Karamoja: 4,397

O : Karamojan zone
About Maps

Destination zones where the people (coming from the selected oirgin zones) settle

Nimule
Gulu
Lodwar
Loiyangalani
Marsabit
Uganda
Kenya
Beni
Butembo · Kasese
Kampala
Kisumu
Meru

O : Karamojan zone
About Maps

Migrations flows from the Origin(s) towards the impacted Destination(s)

Origin	Destination	
POLY015	POLY01380	2,054
	POLY05	582
	POLY01082	581
	POLY083C	519
	POLY01403	340
	POLY016	262
	POLY0818	228
	POLY01412	213
	POLY013	202

Origin of the total migrations towards Karamoja

90%

359,446
internal migrations

■ Migrations from Karamoja
■ Migrations from outside Karamoja

Migration flows towards the destination zone(s) selected over the time

60,829	83,597	83,301	84,768	87,898
201705	201706	201707	201708	201709

People migrations from (Origin zones) and towards (Destination zones) Karamoja during the defined period of interest (months selection: (All) ▼)

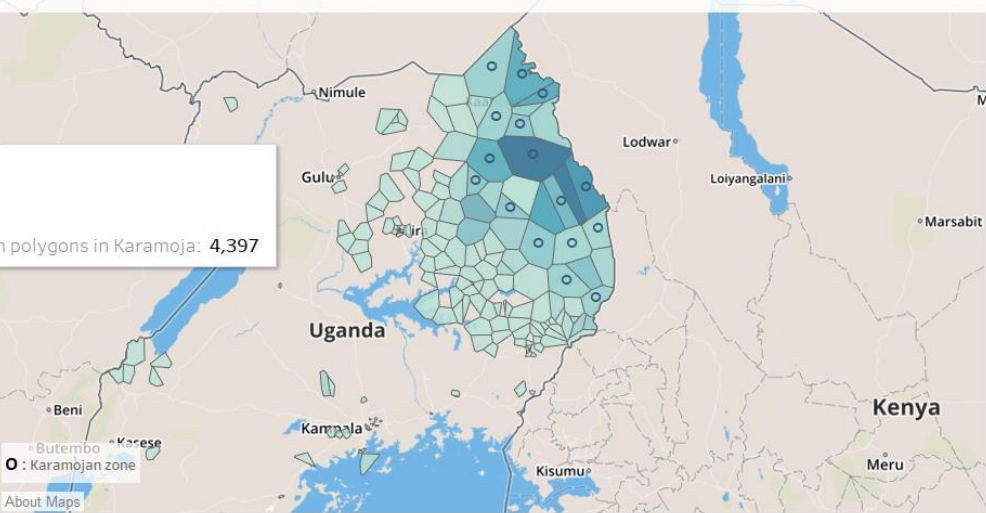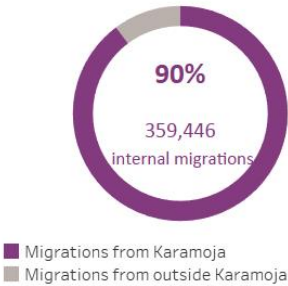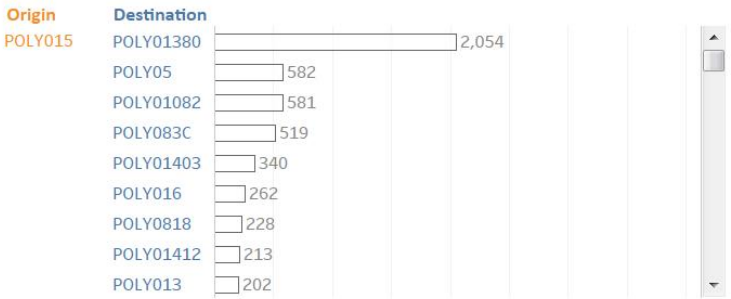Methodology 💡
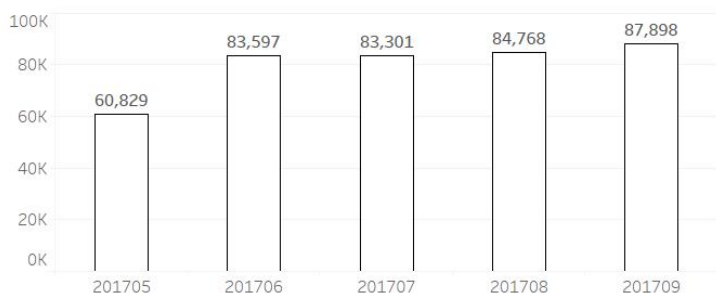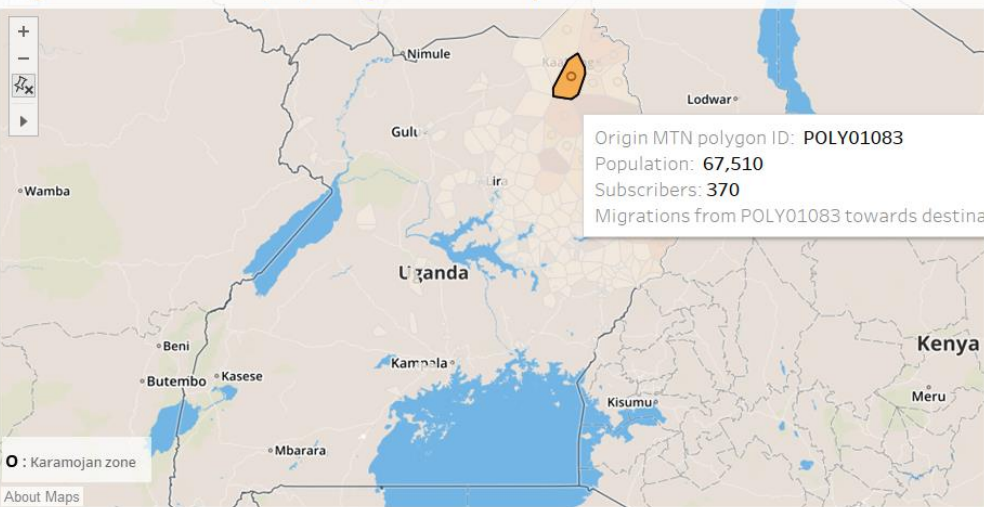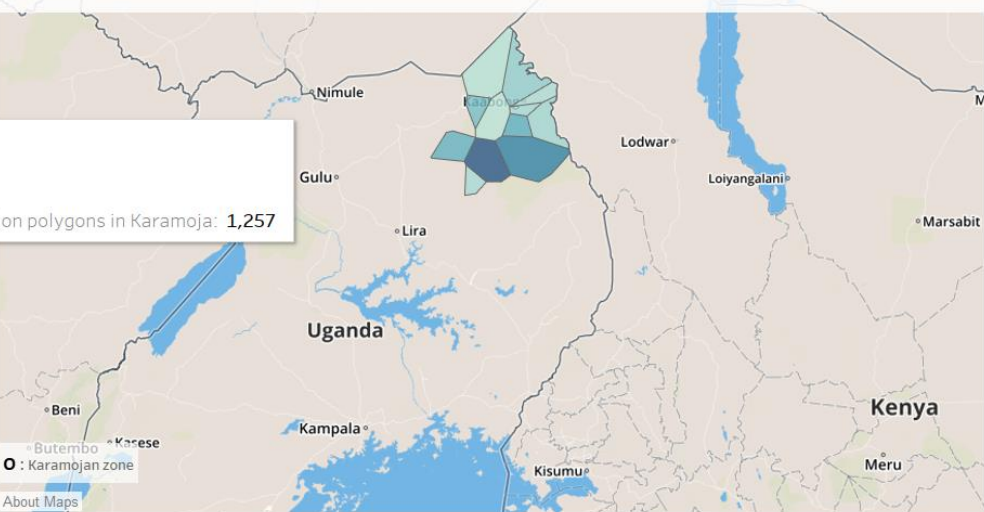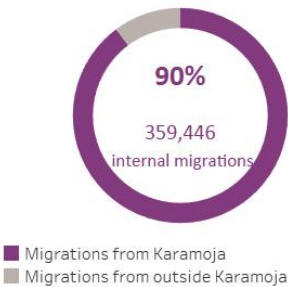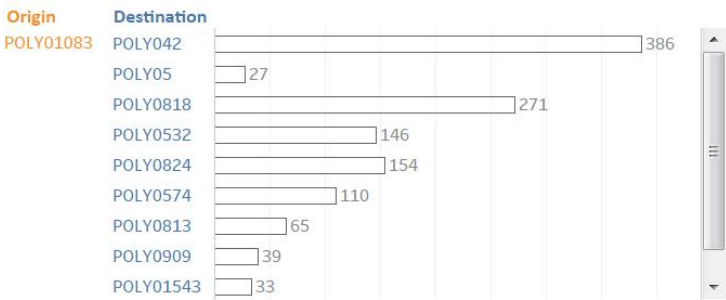
Origin zones from where people migrate to Karamoja

Origin MTN polygon ID: POLY01083
Population: 67,510
Subscribers: 370
Migrations from POLY01083 towards destination polygons in Karamoja: 1,257

O : Karamojan zone
About Maps

Destination zones where the people (coming from the selected oirgin zones) settle

O : Karamojan zone
About Maps

Migrations flows from the Origin(s) towards the impacted Destination(s)

Origin	Destination	
POLY01083	POLY042	386
	POLY05	27
	POLY0818	271
	POLY0532	146
	POLY0824	154
	POLY0574	110
	POLY0813	65
	POLY0909	39
	POLY01543	33

Origin of the total migrations towards Karamoja

90%

359,446
internal migrations

■ Migrations from Karamoja
■ Migrations from outside Karamoja

Migration flows towards the destination zone(s) selected over the time

486	783	663	639	720
201705	201706	201707	201708	201709

Dalberg
Data Insights

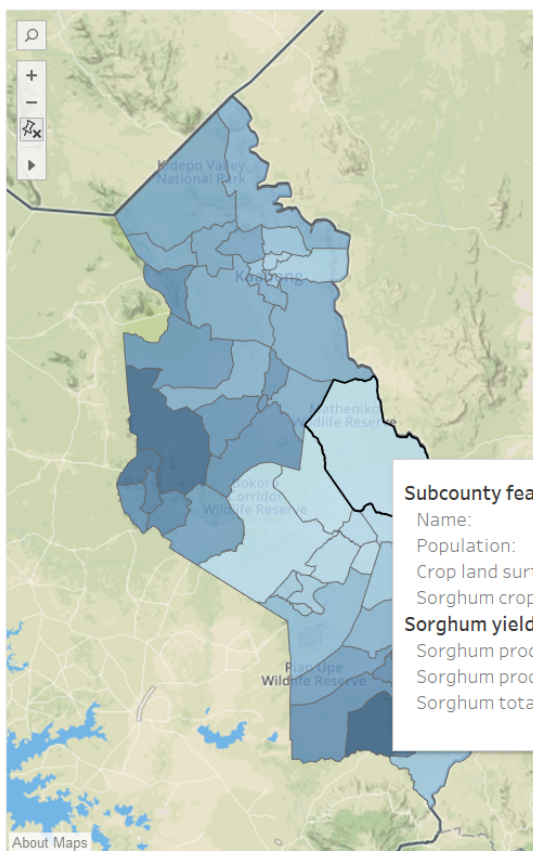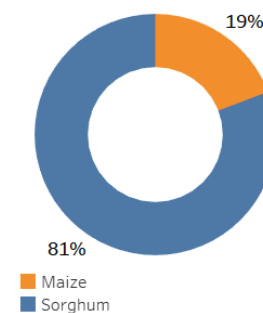**Production of** ◉ Sorghum ○ Maize **measured in** ○ Kg (Total Production) ◉ Kg/Ha (Production per surface unit) ○ Kg (Production per capita)
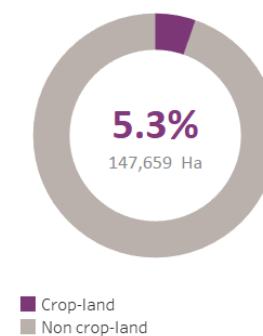
Methodology 💡

Subcounty	District	Population	Total Area (Ha)	Crop Area (Ha)	
NAMALU	NAKAPIRIPIRIT	25,905	55,307	5,794	560.3
ALEREK	ABIM	12,573	125,576	1,266	520.4
ABIM	ABIM	14,377	18,861	999	501.3
NAKAPIRIPIRIT ..	NAKAPIRIPIRIT	1,418	467	29	482.4
ABIM TOWN CO..	ABIM	11,489	11,594	788	439.8
MORUITA	NAKAPIRIPIRIT	16,588	83,929	1,161	430.6
MORULEM	ABIM	16,834	25,036	925	407.5
KAKOMONGOLE	NAKAPIRIPIRIT	15,960	31,271	3,265	399.0
LOBALANGIT	KAABONG	22,907	30,257	1,890	386.5
LOREGAE	NAKAPIRIPIRIT	26,644	92,868	5,760	384.3
		541	19,168	698	382.9
		704	88,096	10,398	373.8
		389	2,397	1,657	369.3
		771	59,758	13,588	367.9
		244	106,718	7,024	354.2
		301	64,240	5,080	334.3
		470	76,961	793	329.8
KATHILE	KAABONG	56,870	22,846	6,891	312.3
KATIKEKILE	MOROTO	11,772	27,745	145	311.0
KAPEDO	KAABONG	33,951	30,456	2,365	301.2

Subcounty features
Name: RUPA
Population: 41,493
Crop land surface: 2,217 (Ha)
Sorghum crop surface: 1,989 (Ha)
Sorghum yield
Sorghum production per surface unit: 114.3 (Kg/Ha)
Sorghum production per capita: 5.5 (Kg)
Sorghum total production: 227,299 (Kg)

About Maps

Maize vs. Sorghum

19%
81%

◼ Maize
◼ Sorghum

Percentage of Crop land

5.3%
147,659 Ha

◼ Crop-land
◼ Non crop-land

Agenda

- Dalberg Data Insights – Our Big Data approach

- Social Impact & Migration

- **Challenges – Data Biases, Partnerships and Sustainability**

Structure and manage a complex economically sustainable ecosystem

Data partners

MTN   airtel

Financial partners

KINGDOM OF BELGIUM
Foreign Affairs,
Foreign Trade and
Development Cooperation

UN CDF

Dalberg Data Insights

Implementation partners

UN CDF

GLOBAL PULSE

KCCA
KAMPALA CAPITAL CITY AUTHORITY
For a better City

Hunger Fighters

Create leadership on data

Build products to address social needs

- **Identify key topics and needs**, incl. tests from emerging countries
- **Idenitfy champions** per topic across countries, public sectors and industries
- **Secure gradual funding process** involving a PE approach, e.g. seed, rounds of fundings, IPO
- **Build robust supply chain (incl. data supply) and commercial model**

Identify needs from private corporations

- **Idenitfy champions** per topic across industrial sectors
- **Build alliances** with data providers (e.g. Nielsen) and secure ecosystem
- **Address key challenges from the regulatory and competitve environments,** e.g. limitations of direct data commercialization from telecom operators

Secure innovation cycle

- **Create new approaches and technologies** to handle larger volumes of data, e.g. AI, deep learning
- **Mix data from an increasing number of different sources**
- **Expand the coverage of the value chain**, e.g. create bots and automate decision-process, including elements of the value chain at the end-users
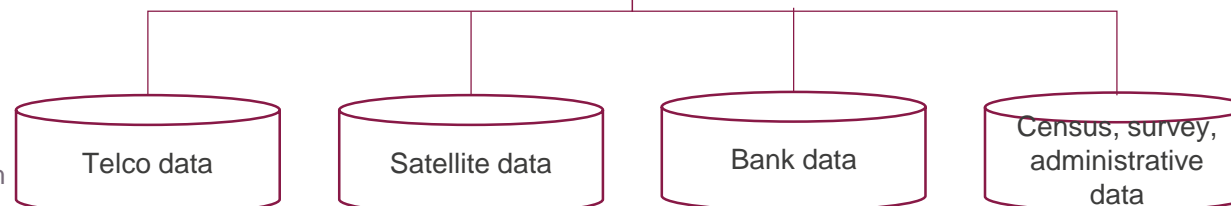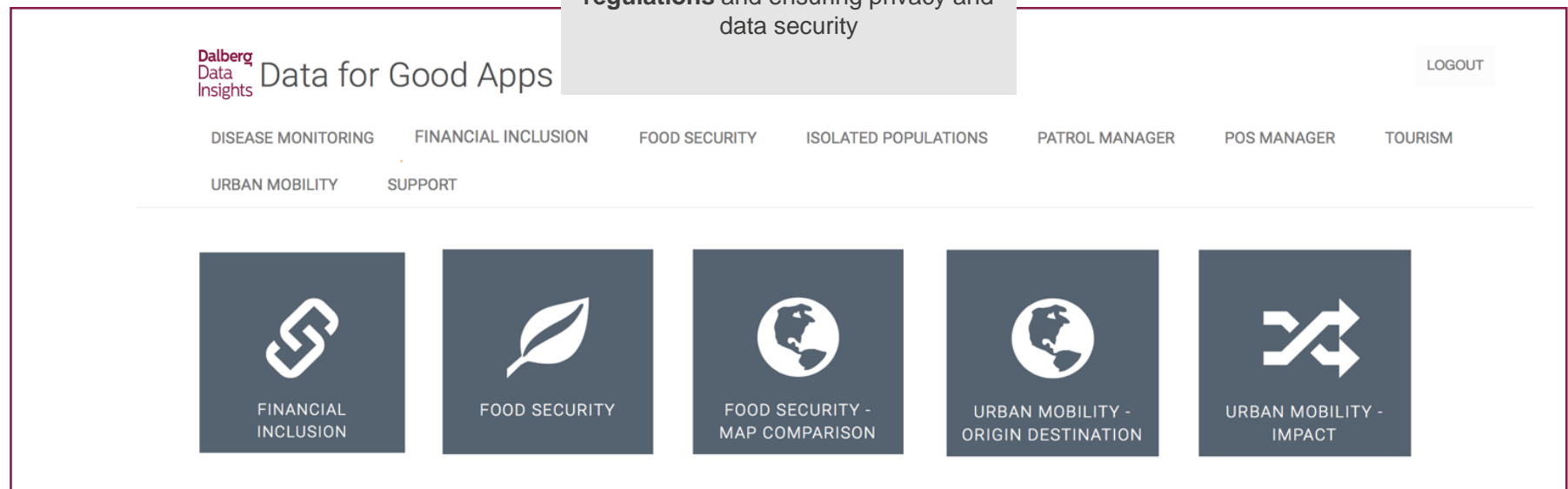
Dalberg

Testing a concrete example of sustainability – Data-as-a-Service

Platform **to integrate multiple data source**s and **allow development of analytics tools and aggregates** that can be shared in a secure manner with 3rd parties

Platform to communicate value to external stakeholders and raise awareness about the value of new data sources to support data driven decisions

Sustainable Platform generating revenue streams and covering cost from hosting and maintenance, data access etc

Platform **complying with EU / Local regulations** and ensuring privacy and data security

Dalberg Data Insights — Data for Good Apps — LOGOUT

DISEASE MONITORING   FINANCIAL INCLUSION   FOOD SECURITY   ISOLATED POPULATIONS   PATROL MANAGER   POS MANAGER   TOURISM

URBAN MOBILITY   SUPPORT

FINANCIAL INCLUSION

FOOD SECURITY

FOOD SECURITY - MAP COMPARISON

URBAN MOBILITY - ORIGIN DESTINATION

URBAN MOBILITY - IMPACT

Telco data

Satellite data

Bank data

Census, survey, administrative data

Dalberg

Main office DDI

5, Place du Champ de Mars
1050 Brussels
Belgium

Create a sustainable digital ecosystem and start productization

	From	To
Topics / Sectors	**Pilot use cases** using aggregated public open data and some private data sources to show value and opportunities / ecosystem	**Platforms of mosty open algorithms with a network of technical partners** accessing and integrating multiple public and private data sources to address **scalable topics**
Regulation	**Research** environment	Supportive **set of laws and regulations**
Data providers	**Research** partner	**Market for data**, where data providers see data as a commodity, including economic / financial flows
End-users	**Co-developing** third parties	**Ecosystem of end-users** of operational tools, involving specific processes, e.g. resilience officer for smart cities

Dalberg