



gLOCAL
EVALUATION WEEK
2021

Data science, a solution for today's Monitoring and Evaluation challenges

By Janvier Mwitirehe, PhD Candidate in Data Mining

Kigali, June 1, 2021

GEI
Global
Evaluation
Initiative

Contents

- Introduction
- Data Science
- M&E Challenges,
- Addressing challenges
- Case study

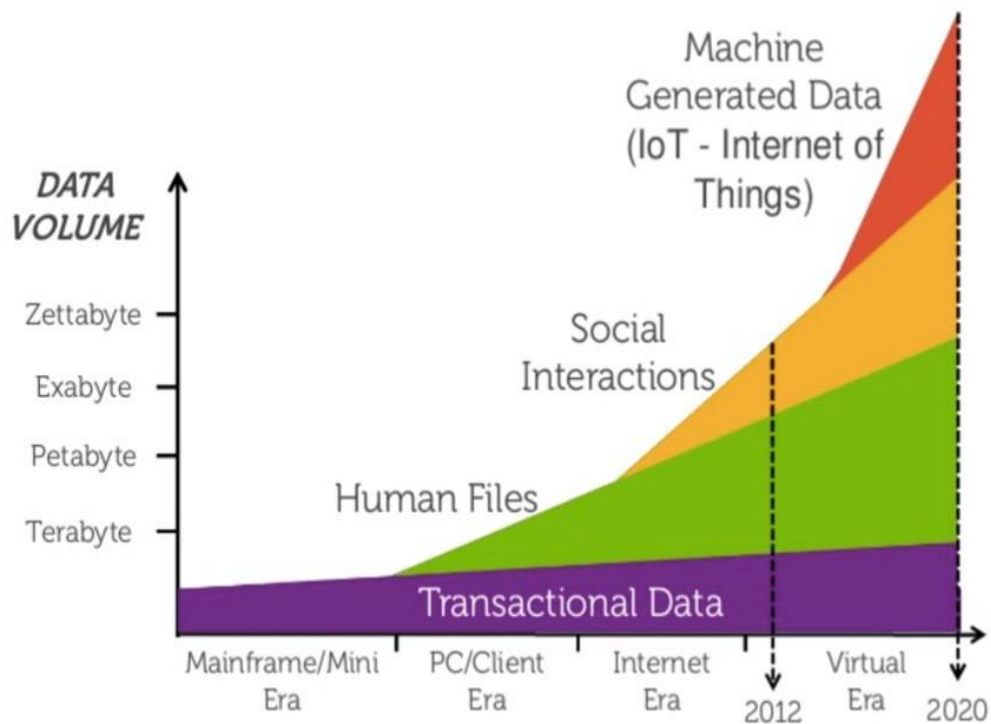
Introduction

Login in <https://www.mentimeter.com>, use a given number and chose the correct answer for the following statements.

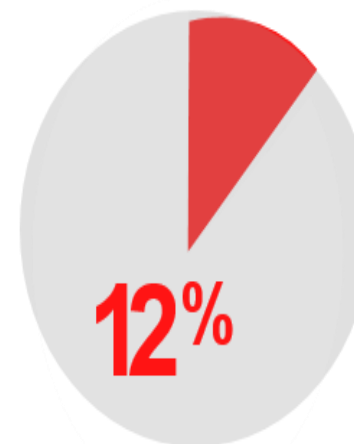
Compared to the situation before COVID-19,

1. Data accessibility has worsened
2. Data size has decreased
3. The quality of data is to be more questioned
4. M&E activities can still be performed reliably
5. SMART indicators' reliance may not provide enough insights
6. How do you feel when talking about COVID-19?

The Explosion of Data



The Useful Data GAP



Executives who feel they understand the impact data will have on their organizations

Produce Data

Use Data

Data Science

- <https://www.youtube.com/watch?v=f9AqD83qHGg>

- Data science :
 - is a multi-disciplinary field for data mining and big data
 - to extract knowledge and insights from structured and unstructured data [Wikipedia]
- Big data: 5Vs



Source : <https://www.edureka.co/blog/big-data-characteristic>

M&E Challenges

- Inaccurate results despite the increasing amounts of program related data in different forms
- Inability to make reliable predictions to inform future planning and new program designs
- Delayed implementation of corrective actions recommended by the evaluation undertakings
- Inability of some users to understand and use monitoring and evaluation results





- Rigorous planning/designs: M&E system, M&E plan, timeframes
- Performance management; Results Based Management
- Robust evaluation methods
- Program leadership and oversight: adaptive management, accountability

- Comprehensive Data
 - Inclusive
 - Enough/adequate
- Storage infrastructures:
 - space
 - processing capacities of computers
- Real-time data : COVID-19 worsened the situation
- Data analytics:
 - evaluating complex and behavior change programs
 - impact evaluation using predictive modelling.



shutterstock.com • 485510518

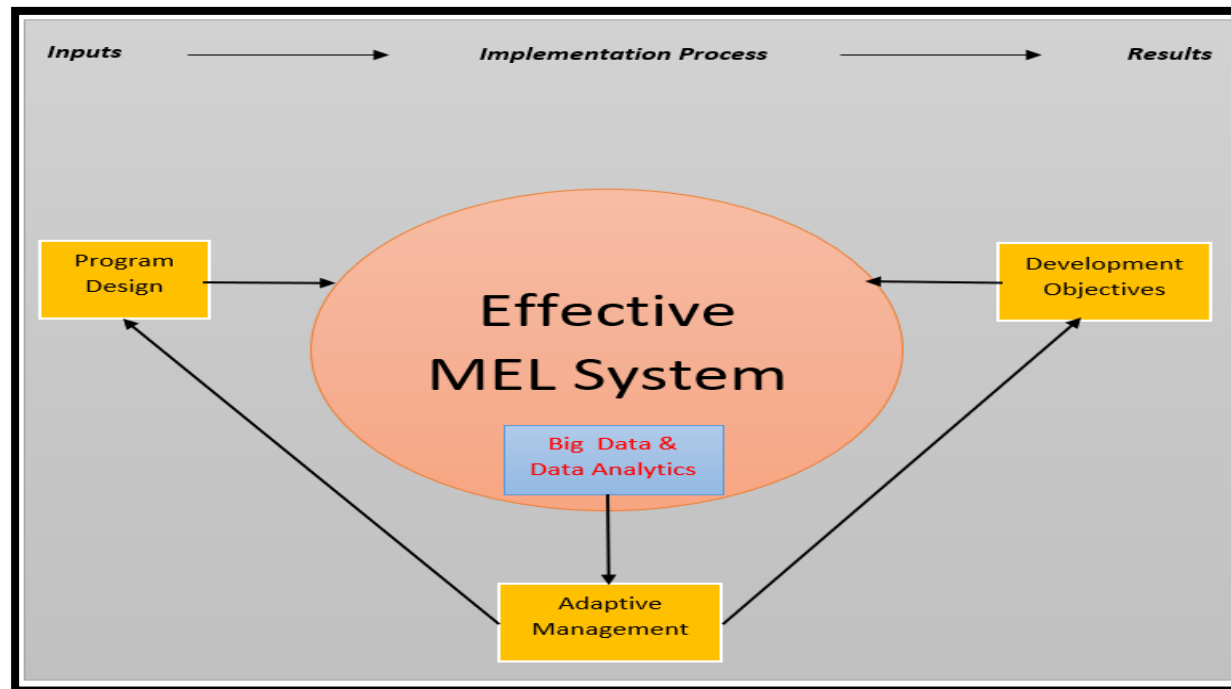
Addressing challenges

- Integrating big data in M&E system
- Using Machine learning to extract reliable data insights
- Performance dashboards as reporting tool

Through a PhD Research:

“Leveraging Big Data to Build an Effective Monitoring, evaluation and Learning (MEL) System for a Water Utility Company”.





Conceptual Framework

MEL systems are effective if they can address the challenges and gaps above; and to do that, they should **integrate and leverage big data** to create **exploratory, predictive, detection, and evaluation models** that can be **visualized in the forms of dashboards** to help stakeholders make informed and real-time actions.

SMART indicators should be:

- **specific** (precise and unambiguous),
- **measurable** (in numerous ways),
- **achievable** (of a reasonable cost),
- **relevant** (serve to assess performance)
- **time bound** (with term dates for measurement),

CREAM indicators should be:

- **clear** (indicators should be precise),
- **relevant** (appropriate to the subject and evaluation),
- **economic** (can be obtained at a reasonable cost),
- **adequate** (the ability to provide enough information on performance),
- **monitorable** (easily monitored, and amenable to independent validation)

(Dennis, 2014).

Case study

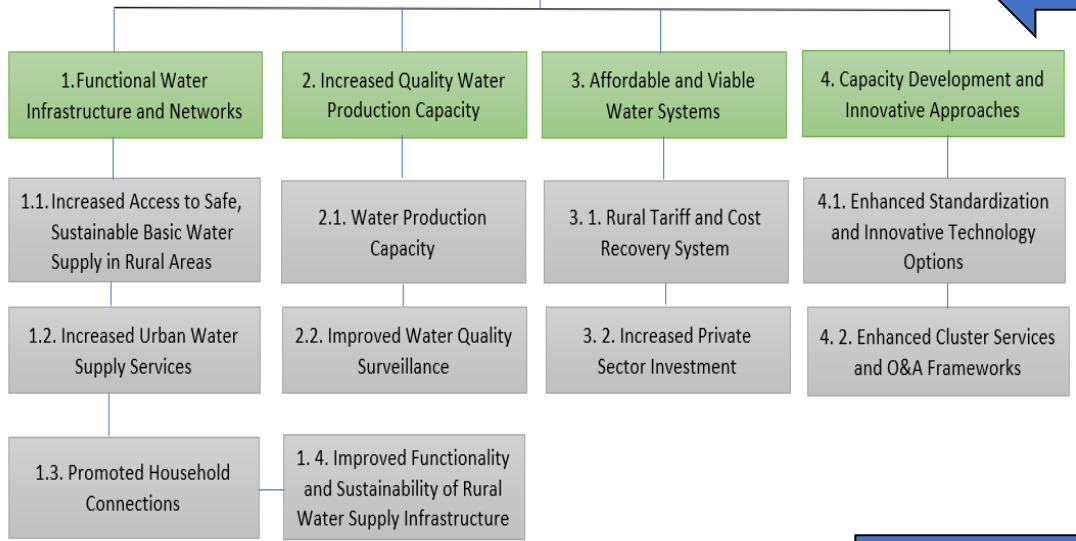
The research aims at strengthening the capacities of the current MEL system to better support the water utility to streamline its processes and evidence-based decision-making for an efficient and effective achievement of targets.



1. To develop a **methodological approach** to integrate and manage massive data from multiple sources and with multiple formats as well as working with real-time data;
2. To create an **interactive dashboard** in Microsoft Power Business Intelligence (BI) that can map interventions, perform calculations, assess the performance, identify trends, and provide real-time information on key performance indicators (KPIs)
3. To set up a **distributed and online reporting framework** that provides most accurate and reliable evidence to support program adaptive management actions, thereby satisfying information needs of stakeholders.
4. To evaluate the **use, efficiency and effectiveness** of the developed reporting framework in supporting adaptive management actions.

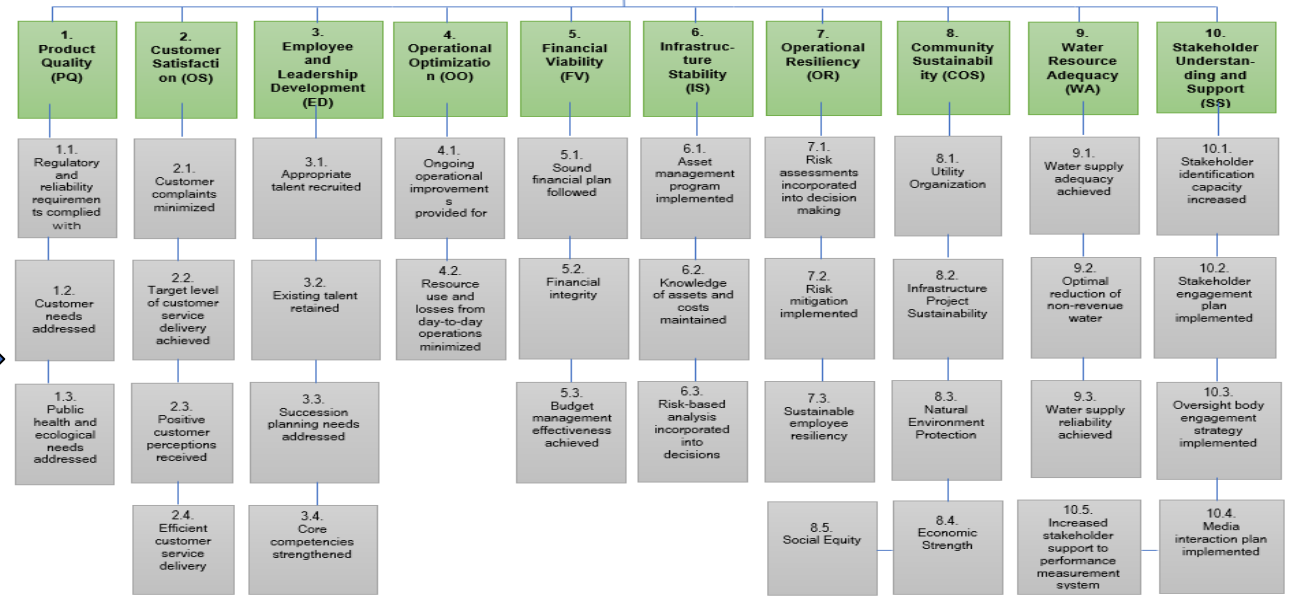


Sustainable, Equitable, Reliable and Affordable Access to Safe Drinking Water for All Rwandans as a Contribution to Improving Public Health and Socio-Economic Development



✓ 4 IRs and 10 Sub-IRs
 ✓ 12 sub-IRs related to Water access (IR 1)

Sustainable, Equitable, Reliable and Affordable Access to Safe Drinking Water for All Rwandans as Contribution to Improving Public Health and Socio-Economic Development



✓ 10 IRs and 35 Sub-IRs
 ✓ 104 KPIs

Results Framework built using info from Water and Sanitation Strategic plan 2018-2024

Revised Results Framework built with info from Water Research Foundation (2017)

| | text | favorited | favoriteCount | replyToSN | created | truncated | replyToSID | id | replyToUID | statusSource | screenName | retweetCount | isRetweet |
|----|---|-----------|---------------|------------------|---------------------|-----------|---------------------|----------------------|---------------------|--|--------------|--------------|-----------|
| 1 | @georgesmherrmann Dear George, There is no technical pr... | FALSE | 0 | georgesmherrmann | 2021-05-28 17:53:02 | TRUE | 1398196352182591488 | 1398336525247647745 | 60255263 | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 0 | FALSE |
| 2 | @Bwengepro @ruasham @GicumbiDistrict @RwandaNorth ... | FALSE | 4 | Bwengepro | 2021-05-28 06:33:49 | TRUE | 1398161693201645569 | 1398165597112971264 | 960478870414569473 | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 1 | FALSE |
| 3 | @Fina33299137 @RwandaInfra @Radiorwanda_RBA @IGIH... | FALSE | 2 | Fina33299137 | 2021-05-26 14:48:02 | TRUE | 1397514258397765638 | 1397565196378546182 | 1397126483769008137 | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 0 | FALSE |
| 4 | Kanombe: Nyarugunga, Nonko na Kabuga uyu mumsi muraz... | FALSE | 3 | NA | 2021-05-26 14:01:00 | TRUE | NA | 1397553356386754561 | NA | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 1 | FALSE |
| 5 | Rubavu: Bitewe n'imitingito yabaye muri Rubavu, hari imiyo... | FALSE | 6 | NA | 2021-05-26 08:01:31 | TRUE | NA | 1397462892681863171 | NA | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 2 | FALSE |
| 6 | Gicumbi: Uruganda rutunganya amazi rwa Nyamabuye @GI... | FALSE | 5 | NA | 2021-05-25 12:25:37 | TRUE | NA | 1397166967493210112 | NA | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 0 | FALSE |
| 7 | Itangazo rigenewe abafatabuzi ba wasac. @GicumbiDistrict ... | FALSE | 7 | NA | 2021-05-23 11:21:39 | FALSE | NA | 1396426091993124864 | NA | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 3 | FALSE |
| 8 | Ruhango: Uyu mumsi Wasac yajeje Amazi meza ku baturage ... | FALSE | 10 | NA | 2021-05-21 16:16:09 | TRUE | NA | 1395775428648919048 | NA | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 3 | FALSE |
| 9 | Kanombe: imirimo yo guhuza umuyoboro w'amazi irakomej... | FALSE | 23 | NA | 2021-05-21 13:02:24 | TRUE | NA | 1395726672415047680 | NA | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 4 | FALSE |
| 10 | @UwanyiligiraB Umuyoboro twawojeje. Ubu amazi arasa n... | FALSE | 0 | UwanyiligiraB | 2021-05-20 12:44:25 | FALSE | 1395260220910280710 | 1395359758434160650 | 1349631371707379712 | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 0 | FALSE |
| 11 | @kforeman1920 You will be supplied this evening. | FALSE | 0 | kforeman1920 | 2021-05-19 16:55:42 | FALSE | 1395030868998098950 | 1395060606005350411 | 3058462720 | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 0 | FALSE |
| 12 | @EzekielDukuze Murakoze, Mwaduha contact number yany... | FALSE | 3 | EzekielDukuze | 2021-05-19 07:31:44 | FALSE | 1394916387991302145 | 1394918678945312775 | 1201396721118527488 | Twitter... | wasac_rwanda | 0 | FALSE |
| 13 | Nyagatare: Aho ni Gihengeri aho bita kurutwe, kubera imvu... | FALSE | 7 | NA | 2021-05-18 13:31:35 | TRUE | NA | 1394646854160224259 | NA | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 3 | FALSE |
| 14 | Kimironko: amazi twayohereje Kimironko umuyoboro w'ama... | FALSE | 5 | NA | 2021-05-16 15:47:07 | TRUE | NA | 1393956182713319429 | NA | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 2 | FALSE |
| 15 | @NdahiroPapy @gatjmv @Rwandapolic @UrugwiroVillage... | FALSE | 3 | NdahiroPapy | 2021-05-16 05:45:35 | TRUE | 1393601604369190926 | 1393804803206168576 | 3253185387 | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 0 | FALSE |
| 16 | @Anizeyimana @Rkangirashyamba @KARANGWASewase ... | FALSE | 4 | Anizeyimana | 2021-05-13 15:44:51 | TRUE | 1392804348636368896 | 1392868451685961730 | 252042772 | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 0 | FALSE |
| 17 | Happy Eid-al-Fitr to all Muslims. May Allah bless you always... | FALSE | 7 | NA | 2021-05-13 15:17:44 | FALSE | NA | 1392861626177671170 | NA | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 0 | FALSE |
| 18 | @Anizeyimana @RwandaInfra @claverGate @RURA_RWA... | FALSE | 3 | Anizeyimana | 2021-05-13 12:28:15 | FALSE | 1392733517356875779 | 1392818975231430659 | 252042772 | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 0 | FALSE |
| 19 | @BuyoyaMugambi Murakoze Patrick! | FALSE | 0 | BuyoyaMugambi | 2021-05-13 12:23:40 | FALSE | 1392800614762045441 | 13928178226635693154 | 1288490448609320960 | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 0 | FALSE |
| 20 | @BenitaTunga Murakoze, twohereje abakozi babishinzwe ... | FALSE | 0 | BenitaTunga | 2021-05-13 06:12:59 | FALSE | 1392722482747084803 | 1392724534793486336 | 1356351754561085449 | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 0 | FALSE |
| 21 | @zeckley @RwandaWater @RURA_RWANDA Bwana Kabera... | FALSE | 1 | zeckley | 2021-05-12 10:08:51 | TRUE | 1392208582314233871 | 1392421504177119237 | 744550814694412289 | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 0 | FALSE |
| 22 | @PieterWerve @MTNRwanda Dear Customer, our system ... | FALSE | 1 | PieterWerve | 2021-05-12 09:49:08 | TRUE | 1392393829639245827 | 1392416542793609220 | 803200481438875648 | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 0 | FALSE |
| 23 | @TV1Rwanda @NgororeroDistr @RwandaGov Iki kibazo tur... | FALSE | 2 | TV1Rwanda | 2021-05-11 14:17:45 | TRUE | 1392059255973416960 | 1392121753376264201 | 781487013635690497 | Twitter... | wasac_rwanda | 0 | FALSE |
| 24 | @semukanyam @GasaboDistrict kubera imvura yaguye yan... | FALSE | 1 | semukanyam | 2021-05-09 15:19:28 | TRUE | 1391352975210201089 | 1391412512172695553 | 283658930 | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 0 | FALSE |
| 25 | @semukanyam @GasaboDistrict Mwirire Abakozi bacu bar... | FALSE | 1 | semukanyam | 2021-05-09 15:14:57 | TRUE | 1391352975210201089 | 1391411374094434305 | 283658930 | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 0 | FALSE |
| 26 | @semukanyam @GasaboDistrict Mwaduha contact zanyu k... | FALSE | 1 | semukanyam | 2021-05-09 12:20:24 | FALSE | 1391352975210201089 | 1391367447593246724 | 283658930 | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 0 | FALSE |
| 27 | Kanombe: Kubera imirimo yo kubaka umuhanda, umuyobor... | FALSE | 13 | NA | 2021-05-09 09:06:35 | TRUE | NA | 1391318670220177408 | NA | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 2 | FALSE |
| 28 | @Manny42351276 Mwaduha contact zanyu kuri DM tukar... | FALSE | 0 | Manny42351276 | 2021-05-05 16:45:45 | FALSE | 1389914842237186050 | 1389984674869616648 | 1264493461325086721 | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 0 | FALSE |
| 29 | Nyagatare: Twagize ikibazo ku Umuyoboro munini wa uvan... | FALSE | 4 | NA | 2021-05-02 12:56:11 | TRUE | NA | 1388839736354160647 | NA | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 2 | FALSE |
| 30 | @KagaboRwa @UrugwiroVillage @RwandaInfra @Pudence... | FALSE | 1 | KagaboRwa | 2021-04-30 09:47:43 | TRUE | 1388030700901781506 | 1388067533752766466 | 1241659889035993089 | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 1 | FALSE |
| 31 | @KaginaUL @claverGate @RwandaInfra @CityofKigali @P... | FALSE | 0 | KaginaUL | 2021-04-27 10:13:55 | TRUE | 1386979123948556289 | 1386986960791773187 | 1285190725 | Twitter... | wasac_rwanda | 0 | FALSE |
| 32 | @GratienNsanzi1 Murakoze kutumenyeshya, mwaduha co... | FALSE | 0 | GratienNsanzi1 | 2021-04-27 06:05:58 | FALSE | 1386910950377369604 | 1386924565671161858 | 1347856138788876291 | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 0 | FALSE |
| 33 | Imirimo yo gusana umuyoboro w'amazi wari wangiritse mu ... | FALSE | 10 | NA | 2021-04-23 07:32:17 | TRUE | NA | 1385496735208857600 | NA | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 2 | FALSE |
| 34 | @NdengeraF Mwatwohereza nimero yanyu kuri Direct Mess... | FALSE | 0 | NdengeraF | 2021-04-23 07:25:44 | FALSE | 1385245963980247042 | 1385495804389523456 | 1171681647995097088 | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 0 | FALSE |
| 35 | @kagire Mwaduha address yanyu tugakurikirana iki kibazo... | FALSE | 4 | kagire | 2021-04-21 09:12:19 | FALSE | 1384385096975781890 | 1384797132725555201 | 45608325 | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 0 | FALSE |
| 36 | @MillerTuyizere @RwandaHealth @RwandaLocalGov @Ang... | FALSE | 0 | MillerTuyizere | 2021-04-21 07:48:20 | TRUE | 1384749258654588928 | 1384775996885377030 | 1084834014622556160 | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 0 | FALSE |
| 37 | @GMuhamoud @CitvofKioali @EnvironmentRw @Rwandaao... | FAI SF | 0 | GMuhamoud | 2021-04-21 07:37:33 | TRUE | 1384736692884713475 | 1384773282784817155 | 4754164401 | <a href="http://twitter.com/download/android" rel="nofoll... | wasac_rwanda | 0 | FAI SF |

| Sub-IR | Water Quality Indicators | Target | Data analytics | Methodology for performance measurement | Data used measure the performance | Tools used to measure the performance |
|--|--|---|-------------------------|---|---|--|
| <i>Regulatory and reliability requirements complied with</i> | Degree of drinking water quality compliance | Consistent DWQI=100% for more than five consecutive years. | Exploratory, predictive | i. Calculate the weighted arithmetic Drinking Water Quality Index (DWQI). In this method, water quality rating scale, relative weight, and overall WQI were calculated as below: <ul style="list-style-type: none"> ✓ $qi = (Ci/Si) * 100$; where qi, Ci, and Si indicated quality rating scale, concentration of i parameter, and standard value of i parameter, respectively. ✓ Relative weight was calculated by $Wi = 1/Si$; where the standard value of the i parameter is inversely proportional to the relative weight. ✓ Finally, overall WQI was calculated according to the following expression: $WQI = \frac{\sum qiwi}{\sum wi}$ ii. Monthly and annual average DWQIs including yearly predictions per each water treatment plant will visualized in the Power Bi dashboard. | Monthly water production per treatment, drinking water standards, daily chemicals used, such as pH, manganese, iron, salinity levels, arsenic concentration, etc. | Data were initially stored in Hadoop data lake, then replicated in Azure Spark HDInsight to be connected to Microsoft Power BI where computations and visualizations were done |
| | Urban access to improved sources of drinking water | 100% Households with access to clean drinking water within 200m | Exploratory | i. Visualize targets vs actual water connections ii. Visualize information related to HH access on clean drinking water available from WASAC and partner surveys and researches | Target water connections by water treatment plant Water connections per water treatment plant National survey reports (EICV, DHS, RURA, etc.) | Same as above |
| | Rural access to improved sources of drinking water | 100% Households with access to clean drinking water within 50m | Exploratory | i. Visualize targets vs actual water connections ii. Visualize information related to HH access on clean drinking water available from WASAC and partner surveys and researches | Target water connections by water treatment plant Water connections per water treatment plant National survey reports (EICV, DHS, RURA, etc.) | Same as above |

| Sub-IR | Water Quality Indicators | Target | Data analytics | Methodology for performance measurement | Data used measure the performance | Tools used to measure the performance |
|---------------------------------|--|--|---------------------------------------|---|---|--|
| <i>Customer needs addressed</i> | Achievement of target drinking water flow and pressure | 100% monthly drinking water flow and pressure between 0.6 and 1.5 bars | Exploratory, predictive and detective | <ul style="list-style-type: none"> i. Monthly and annual average water flow/pressure including yearly predictions per each water treatment plant will be visualized in the Power Bi dashboard. ii. Using machine learning anomaly detection model to identify and predict drinking water flow and pressure falling outside the regulated boundary of 0.6 and 1.5 bars. By so doing, One-Class Support Vector Machine based and PCA-based anomaly detectors will be compared to distinguish between the two classes (within and outside the boundary). | Drinking water flow and pressure data- | <ul style="list-style-type: none"> 1. Data accessed in Power BI 2. Data accessed in Azure Machine Learning studio from Hadoop data lake |
| | Reliability of water supply | Average number of days the water was not supplied per a week is 3 or more days | Exploratory | <ul style="list-style-type: none"> i. Visualizing daily water supply trend to identify instances of a failure to supply water for at least three consecutive days. ii. This exercise also diagnosed the consistence of the water quantity supplied by each treatment plan | Water supply data, in rural (Nyagatare) versus urban (Kimisagara) | <ul style="list-style-type: none"> 1. Data accessed in Power BI 2. Data accessed in Azure Machine Learning studio with results published in Power BI |
| | A system of addressing customer complaints | The utility maintains a reliable and timely responsive customer service center | Exploratory | <ul style="list-style-type: none"> i. Collect 1620 sampled tweets which are tagged “@wasac_rwanda” and #wasac_rwanda. That is, what people say to WASAC and about it respectively. ii. Using R, the analysis was performed to determine when individuals tweet the most, looked at what time day users tweet most frequently, and determined what words were used most commonly when individuals were speaking negatively about WASAC iii. Application with Twitter was set through http://developer.twitter.com/en/apps. An account was also created. iv. | Twitter data | Data stored in R but Results reflected in Power BI dashboard. |
| | Well-functioning water supply interruption programs | 100% of annual water supply interruptions that respected at least 48-hour prior notice | Exploratory | Idem | Twitter data | Data stored in R but Results reflected in Power BI dashboard. |
| | Well-functioning water supply rationing programs | 100% of annual rationing programs that respected the requirements of at least 72-hour | Exploratory | Idem | Twitter data | Data stored in R but Results reflected in Power BI dashboard. |

Reporting framework

Naturally, many people are not naturally good at story telling. Therefore, visualization is needed:

- ✓ **Visual reports:** traditional 80% text and 20% graphics is more likely to be reversed;
- ✓ **Executive summaries:** most of the viewers don't read more than one- or two-page reports
- ✓ **Online reporting:** viewers would respond better to an online report like website-based reports

Washington Evaluators

2014

Washington Evaluators (WE) is a local affiliate of the American Evaluation Association (AEA). WE was founded over 30 years ago as a professional society devoted to fostering state-of-the-art knowledge and information sharing.

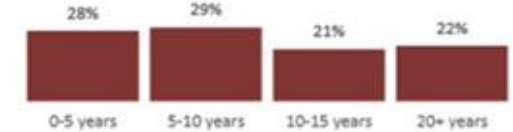
Who We Are

225+ members from a variety of evaluation backgrounds

We work throughout the DC metro region



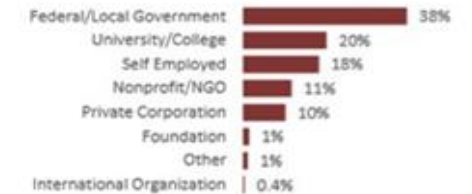
We encompass a range of experience levels



We're primarily working professionals and 3 in 4 of us are female

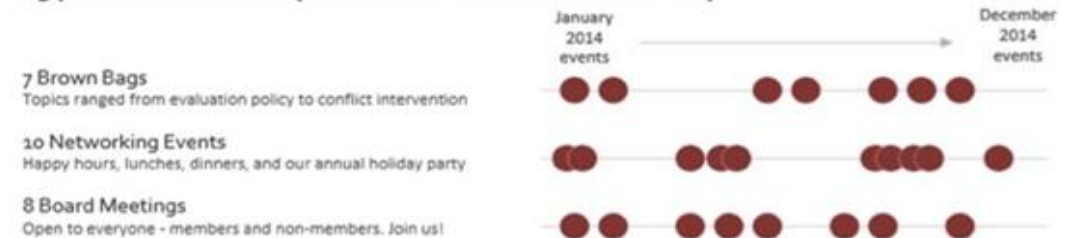


38% of us work for Federal or local government



What We Do

25 professional development and social activities in 2014



Join Us

Membership
\$25 professional
\$15 student

Contact
washingtonevaluators.org
washeval@gmail.com
@washeval



Microsoft Power BI Dashboard

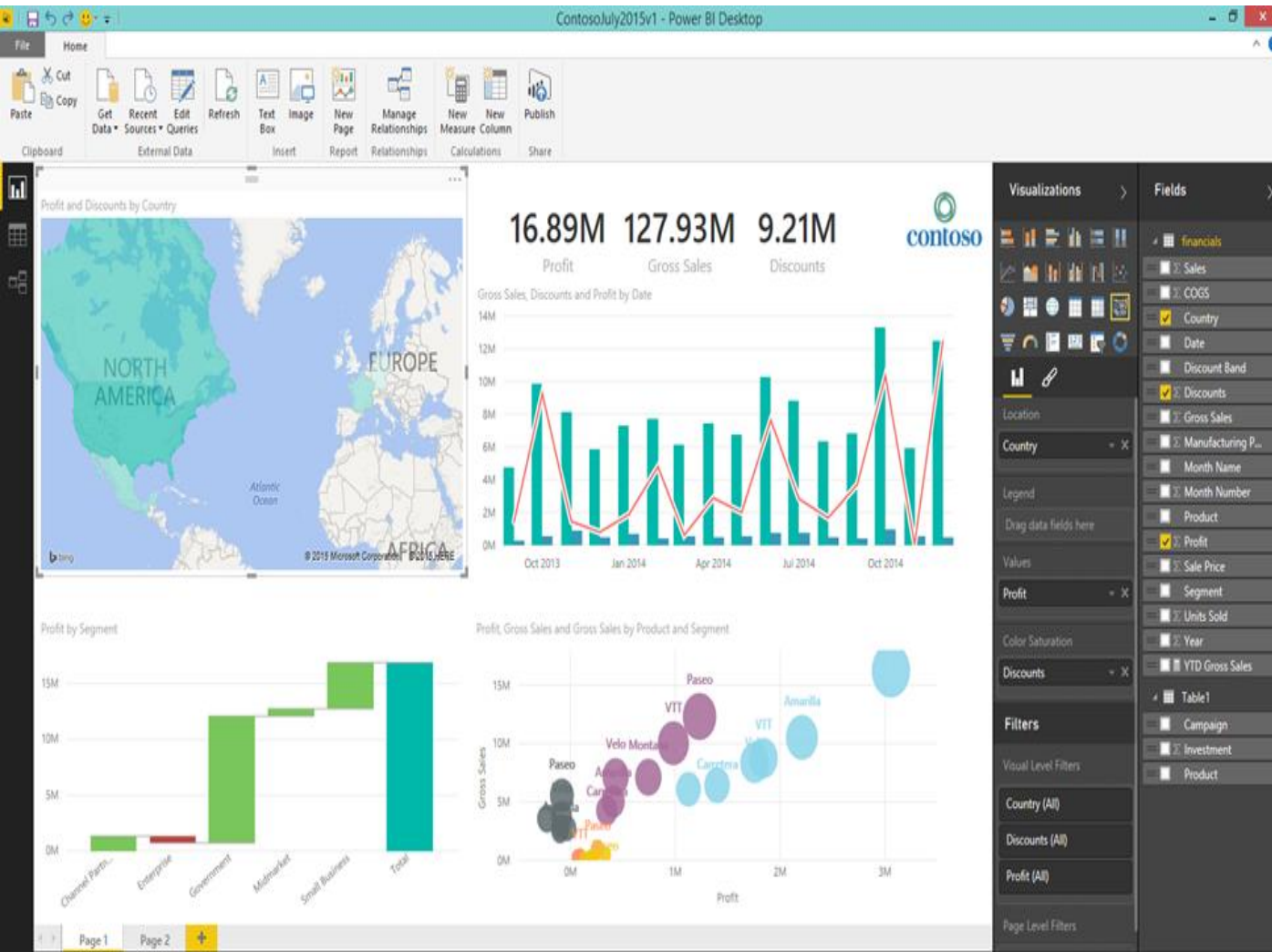


Tableau Dashboard



Country

Cities

Years

Trees by Cities

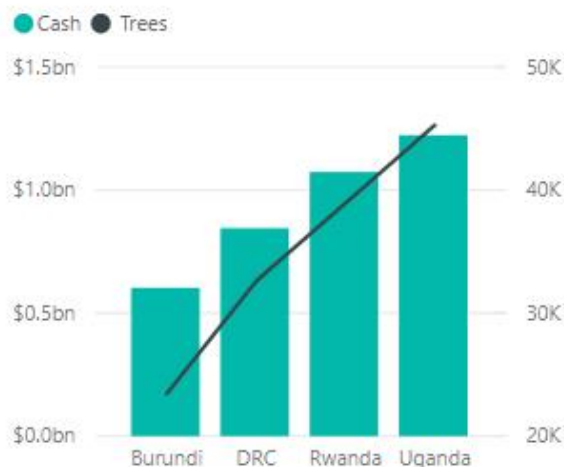
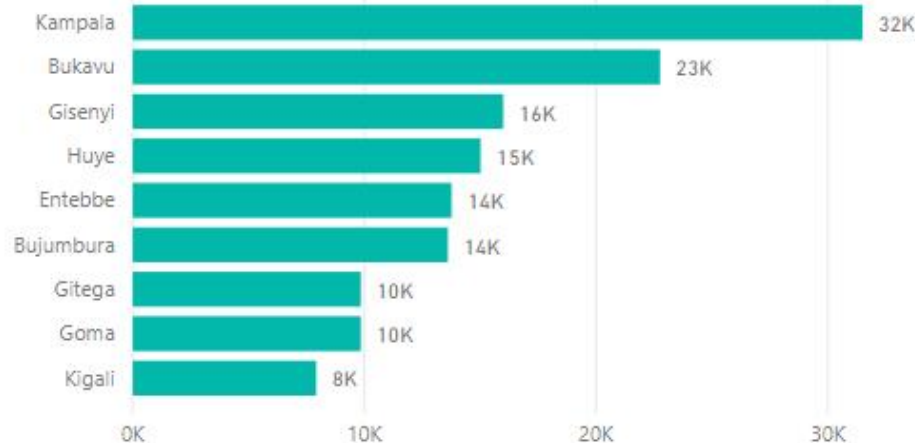
Cash and Trees by Country

Multiple selecti...

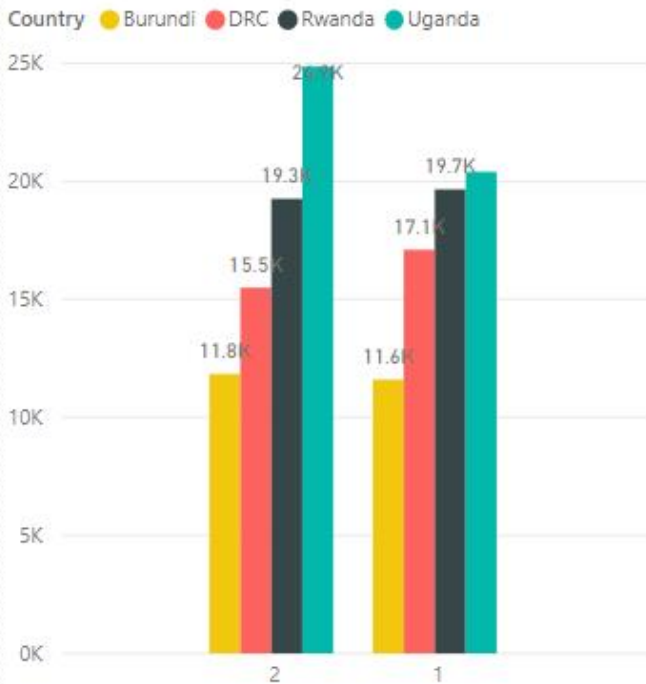
- Bujumbura
- Bukavu
- Entebbe
- Gisenyi
- Gitega
- Goma
- Huye
- Kampala
- Kigali

- 1
- 2

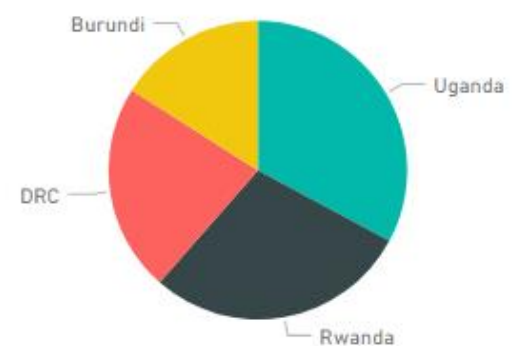
\$3.74bn
Cash



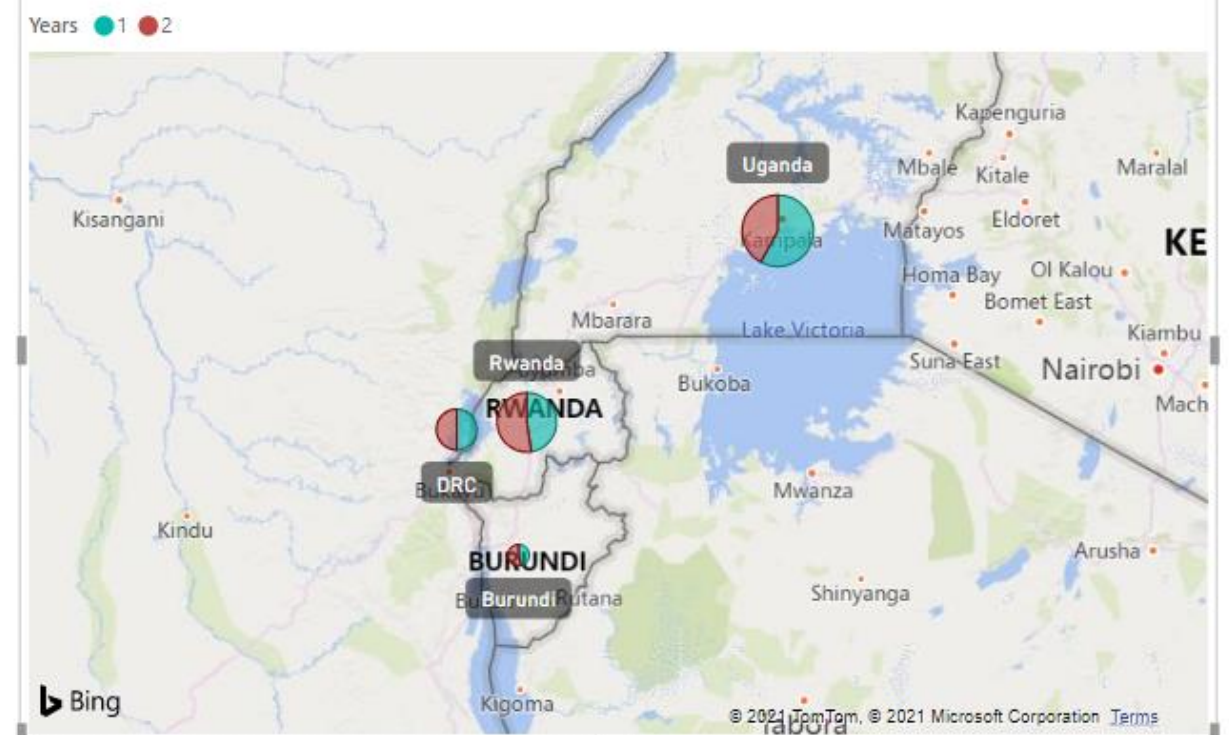
Trees by Years and Country



Quantity by Country



Average of Longitude, Average of Latitude, Cash, Trees and Quantity by Country and Years





Conclusion

- ❖ *“Data science is the cocktail of the day, and the party is just getting started”*
(Dima Alazzi, 2019)
- ❖ Without embracing data science, current Monitoring, evaluation and learning (MEL) systems may not work effectively.
- ❖ MEL is ideally understood as **dialogue** on development between all stakeholders.
- ❖ In that respect, an **effective MEL system** :
 - ✓ identifies all stakeholders,
 - ✓ ensures their information needs are covered
 - ✓ ensures adequate data is used for adaptive management
 - ✓ Clarifies what needs to be sustainable and how to design it

Thanks

