

DATA CLEANSING

It's commonly said that,

"Users of data spend 80% of their time cleaning and manipulating data and only 20% of their time actually analyzing it."

What is Data Cleaning?

Data cleansing or data cleaning is the process of identifying and removing (or correcting) inaccurate records from a dataset, table, or database and refers to recognising unfinished, unreliable, inaccurate or non-relevant parts of the data and then restoring, remodelling, or removing the dirty or crude data. Data cleaning may be performed as batch processing through scripting or interactively with data wrangling tools.

After cleaning, a dataset should be uniform with other related datasets in the operation. The discrepancies identified or eliminated may have been basically caused by user entry mistakes, by corruption in storage or transmission, or by various data dictionary descriptions of similar items in various stores

Overview of our approach

Spatial Data: External



- •SG diagrams
- •Deeds register data
- •Cadastral for the municipality
- •Bureau Data

Spatial Data: Internal



- •Town planning schemes
- •Land use management data
- Rezoning
- Valuation roll analysis

Revenue System



- •Recon of:
- valuation roll to deeds register database
- Properties to valuation roll with billing engine
- •Electricity and water meters

Exceptions

- •Land parcels in SG diagram but not in deeds d/base and/or
- •in the deeds d/base and not in valuation roll and/or
- •Billing engine but not in valuation roll
- •Properties billed for 1+ services
- Data quality assessment



- Land Audits linked Rezoning, Deeds Register, Surveyor
 General's General Plan and Building Plans
- Integration of valuation roll
- ✓ Property Rates Act Compliance

- ✓ Reconciliation of valuation roll with revenue systems records
- ✓ Data integrity analysis and quality
- ✓ Each geo-coded customer record is added with a SG Code from the corresponding Cadastral data.

Outcomes

Ntiyiso Consulting will undertake a 3-phased approach in Data Cleansing



a. Data quality check

- Identify missing information
- Testing general accuracy and data validation
- · Check for duplicates and reasons thereof

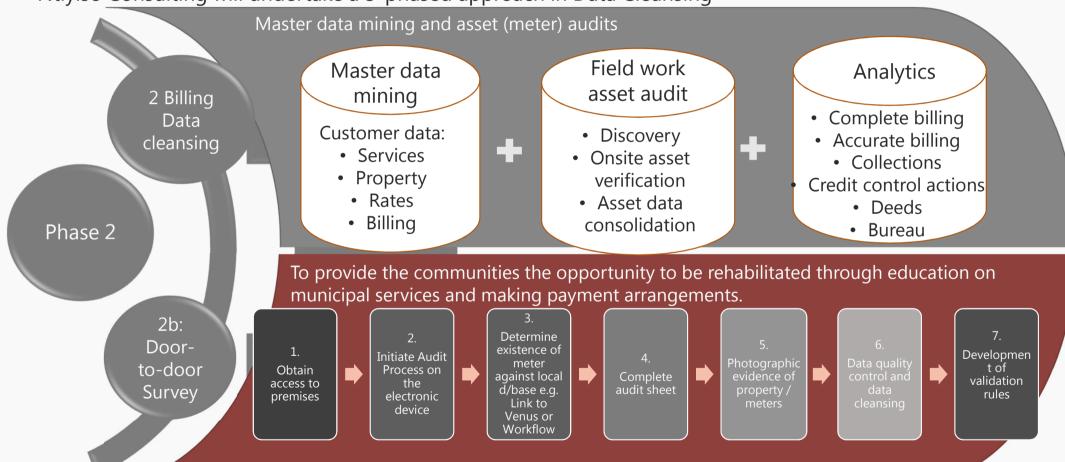
b. Data enrichment

- Ntiyiso has a partnership Consumer Profile Bureau (CPB) that uses technologically advanced systems containing large sources of consumer data that assist with Verification of ID, Verifying of addresses, Tracing of consumers, Consumer spending patterns, Credit analytics, Contact information, etc.
- SG Diagrams Spatial Alignment
- Deeds Data

c. Analyse enriched data

- Compare given customer information vs bureau data
- Compare title deed owner type relating to the property against debtor type
- Consolidate customer profiles
- Compare title deed owner relating to the property against debtor
- Identify Indigent customers using the municipality's indigent policy
- Identify debt to be written off using the municipality's write-off policy
- Compare billed amounts against tariffs/user charges for accuracy

• Ntiyiso Consulting will undertake a 3-phased approach in Data Cleansing



Ntiyiso Consulting will undertake a 3-phased approach in Data Cleansing

Phase 3:
Monitoring
&
Verification

a. Bureau

Ntiyiso has a partnership with bureaus and other similar service providers that uses technologically advanced systems containing large sources of consumer information that would assist with:

- Verification of ID
- Verifying of addresses
- Tracing of consumers
- Consumer spending patterns
- Credit analytics
- Contact information, etc.

b. Exceptions Flagging and Reporting

- c. Corrected accounts will be monitored and the following will be checked
 - Historical Consumption and payments
 - Kilowatt
 - Kilo Litres
 - Payment history

PRACTICAL EXPERIENCE OF DATA CLEANSING PROJECT

The ideal of this presentation is not to share only theoretical knowledge of a Data Cleansing project but rather to share with members the achievements and benefits of a Data Cleansing Project.

Ntiyiso Consulting has been appointed to conduct a Data Cleansing Project in **Knysna Local Municipality**. The project is still underway but the following slides will show some of the benefits already achieved and plan to achieve.



PRACTICAL EXPERIENCE OF DATA CLEANSING PROJECT

The project contains the following phases:

Phase 1: Customer Data Cleansing

Phase 2: (a) Billing Data Cleansing

Phase 2: (b) Door-to-Door Survey

Phase 3: Annual review of billing

information



PROGRESS ON THE PROJECT SO FAR

1. GOVERNANCE MATTERS

- Meet the officials of the Municipality
- Introduce the team responsible for the project, which includes
 - Project Manager
 - Specialist on GIS mapping
 - Data Analysts
 - Financial Specialist
 - Fieldwork Specialists
- The legal compliance of the tender were discussed Draft Project Initiation
 Document (PID) and Service Level Agreement (SLA)
- The duration and project plan were shared with officials. It includes payment schedules as well
- Any logistical matters were addressed © 2017 Ntiyiso Consulting. All Rights Reserved.

PROGRESS ON THE PROJECT SO FAR

2. OBTAIN ALL RELEVANT INFORMATION

- Details of financial system
- Details of GIS information
- Details of water and electricity meters

We collect as much as possible information from the Municipality

- Copy of Master Data file
- Copy of Valuation Roll
- Debtor Book with age analysis
- Meter reading information
- Any other information that could assist the project

PROGRESS ON THE PROJECT SO FAR

3. FIELD SURVEY AND CHANGE MANAGEMENT

- A critical aspect of the project was to employ local labour for the door-to-door survey process.
- Employed Fieldworkers in each of the 11 Wards.
- The Public Participation department/section also assisted with the identification
- Marketing Department assisted with the Change management process to advertise the various platforms of the media (including local news papers, municipal website and social media platforms (i.e. Facebook)
- In addition to the advertisement of the project the consultants attended all the ward committee meetings during February and March 2019 to inform the ward committees of the project.

4. CORRECTION OF PROPERTY / GIS INFORMATION

The property details of Knysna was compared with the information from the Surveyor-General's office, and the following was discovered:

- a) The street names of 32 645 properties/accounts were verified.
- b) A total of 18 845 properties/accounts do have the correct street addresses. (i.e. 58% of listed properties/accounts have the correct street names)
- c) A total of 13 727 street addresses was wrong in the Knysna billing system and has to be corrected (i.e. 42% of the listed properties/accounts have to be corrected)
- d) A total of 73 street names still need to be investigated (0.22%).

There was a total of 1 439 wrong spelling of street addresses. After corrections it reduce the number of street names to 473.

This is some examples of street names that need to be corrected:

	•	
Account No_	Correct Street Name	Street Name in billing system
	ALBATROSS STREET	ALBATROS
	HART ROAD	HART RAOD
	BARACUDA STREET	BARRACUDA STREET
	CALLANDER STREET	KALENDER STR
	DAFFODIL AVENUE	DAFODIL AVENUE
	BURCHELL STREET	BURCHEL LANE
	BONGOLETHU STREET	BONGOLETHA STREET
	DD NGXALE STREET	DD NXGALE STREET
	HOOPOE AVENUE	HOEPOE
	KALLOSSIE STREET	KALOSIE
	AGAPANTHUS STREET	AGAPANTUS STR
	BOUGAINVILLEA STREET	BOUGENVILLA
	BARRACUDA STREET	BARACUDA STREET
	KRWETSHE STREET	RWETSHA
	MANKONKWANE STREET	MANKONKWANA STREET
	ADRIATIC CLOSE	ADRIATIC, COLA BEACH,
	ADRIATIC CLOSE	ADIATIC CLOSE
	AGAPANTHUS AVENUE	AGAPHANTUS AVENUE
	CAPTAIN DUTHIE AVENUE	CAPT W.A. DUTHIE AVENUE
	BERN CLOSE	BENN CLOSE
	LENTEGEUR STREET	LENTEGEURS STREET
JI .	AZALEA STREET	AZALEAS STREET
	BLUEBELL STREET	BLUEBEEL STREET

Validation of Identity Numbers:

- A total of 24 601 ID numbers were submitted to Consumer Profile Bureau to verify the correctness. The purpose was to get the following type of information i.e.
- Valid ID's
- · People who are deceased
- Invalid ID's
- Not traceable ID's.
- From the results that were returned it was discovered that 2 577 of ID numbers in the billing system belongs to residents who are <u>already deceased</u>.
- It implies that outstanding balances on these accounts belonging to deceased residents could not be collectable and needs to be investigated with possible writeoff of debt of outstanding accounts.

© 2017 Ntiyiso Consulting. All Rights Reserved.

14

VALIDATION OF NAMES OF ACCOUNT HOLDERS:

It appears that there are **5 054** names that differ with the names on the Knysna billing records. However, these name differences do not implicate that it is wrong but could be the conformation of the names that differs. Typical examples of some differences are reflected in the following table:

Acc Number	ID number	Name as per billing file	Name from CPB
		VALU-VALU E	VALUVALU EMILY
		MPITIMPITI N	MPITI MPITI NOTEMBILE
		MABI E	MABIE ELIZABETH
		McMASTER MD	MC MASTER MARY
		WOLFE GB	WOLFF GILLIAN
		MANEVIL A	MANEVILLE ARJANA
		TERBLANCE SK	TERBLANCHE SOPHIA
		MACNICOL MJ	MAC NICOL MARGET
		T.H.F. McCulloch	MC CULLOCH TERENCE
		SLABBERT EMA	BRUWER-SLABBERT ELSJE

VALIDATION OF CELL PHONE NUMBERS:

 The consultant also obtained information of residents' cell phone numbers and the following results were obtained:

Cell Phone number matching	Number
Matching cell phone numbers with Knysna file	4 146
Different Cell phone number with Knysna file	3 959
Cell phone numbers not appearing on Knysna file	13 894
	21 999

• Officials from the Revenue Department is in a process to upload all the cell phone numbers that does not appear on the billing information, which will greatly assist in contractibility of residents.

DOOR-TO-DOOR SURVEY

- In order to conduct the door-to-door survey 22 Fieldworkers have been appointed. There will be 2 Fieldworkers per ward deployed.
- They were issued with the following items to equip them during the survey:
 - One Rugged data logger device with charger (as per example)
 - Protective vests to identify them and make them visible.
 - An ID card with photo, name and project details (as per example)
 - Clip board and writing pad to make notes
 - Pepper Spray (for protection)
 - Water bottle
 - Letter signed by the Municipal Manager regarding the project
 - Maps per ward to conduct the survey.

FIELDWORKERS TEAM



DOOR-TO-DOOR SURVEY

 The following pictures are examples of the highly technologic devices issued to each fieldworker, as well as the identity cards that they have to wear all the time





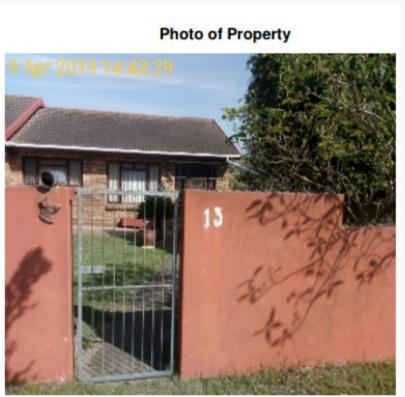
DOOR-TO-DOOR SURVEY

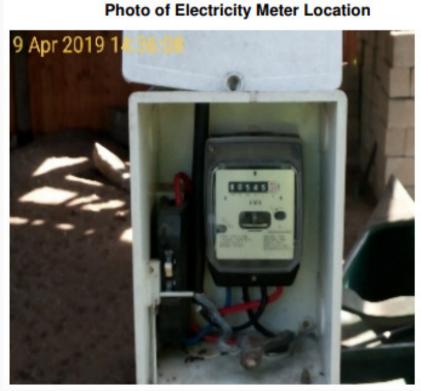
- The door-to-door survey results are uploaded daily to the Data Management Centre in order to compile the results from the survey.
- This information includes a large number of information obtained from the survey and details will be submitted to Knysna in a job card format (which incudes photos) of the house visited, a photo of the water and electricity meter on the premise.
- Data Analyst is comparing survey information with Billing information.
 These results are shared with the client.
- The information will also be supplied in an excel format to Knysna officials.

EXAMPLE OF JOB CARD CREATED FROM THE SURVEY

Audit Ref No: 4584879		Comments	GPS Coordinates Electricity Meter	Electricity Moley Status	Electricity Meter Location	Electricity Meter Reading	Electricity Meter No	Electricity Meter Type			Comments	Owner Postal Address	Owner Email	Owner Telephone No2	Owner Telephone No	Owner ID Type	Owner ID No	Owner First Name	Owner Surname	Question		Comments	_	Occupant Postal Address	_	_	Occupant Telephone No	Occupant ID Type	_	Occupant First Name	Occupant Surname	is the occupant available to take questions?	_	Outestion		Comments	Number of Bins	Water/Electricity Account No	Property Usage	Question	Cance Mila	Audit Date 09 Apr 2019		II-	Instruction Perform Prof	Date Issued 2019-04-08	Ward		ч.	Fare wer tult Red 20190610	CONSULT		フーイン協	
Fage 113 Inspected by: Device: Nilyiso208 on 4/9/2019 2:15:12 PM	WATER METER INFO			Normal	Inside Property	21	01080171232	Prepaid	ELECTRICITY METERI INFO	EL DONOLOGIC ANTIGO DAGO						National ID				response	Charles area	1	You Wanter? You				0726677634	National ID	7911271102080	VUJENA	oteph	aks quessions? Yes	_	Baccorrea	OCCUPANT INFO			1080171232		Response	1000 Part 200	 9 Time On Site Time Off Site	AUDIT FINDINGS	The state of the s		Issued To Davice: Niyiso206 Device Search Code 1180	GPS Long: 23.10008	C03900050001731300000		PROPERTY & SERVICES AUDIT REPORT	10	MAI ON MONOR ALL	KNYSNA MINICIDALITY	* KNYCNA MIINICIDALITY

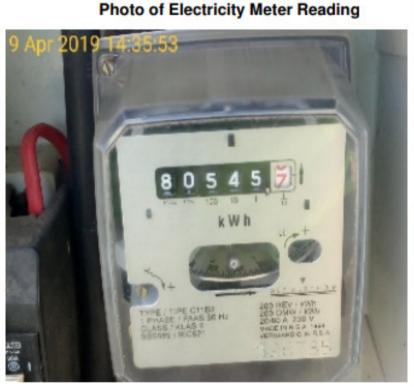
EXAMPLES OF SOME PHOTOS TAKEN DURING THE SURVEY FROM THE SURVEY





EXAMPLES OF SOME PHOTOS TAKEN DURING THE SURVEY FROM THE SURVEY





EXAMPLES OF SOME PHOTOS TAKEN DURING THE SURVEY FROM THE SURVEY





CHALLENGES / PROBLEMS

- Residents do not want to give co-operation in the door-to-door survey.
- Fieldworkers gave feedback that several residents do not want to give them access to their properties (which could merely be from a safety point of view)
- Fieldworkers also experienced some political sensitivity during their visits in some areas, which could be attributed to the upcoming elections in May 2019.
- Fieldworkers also experienced reluctance from some residents to give their personal details such as ID numbers, cell phone numbers or email addresses.
- Alternative methods are considered to deploy entire Fieldworker team into one area (not per ward) to instill motivation and support to each other

CONCLUSION

- Data Cleaning is a critical requirement for any municipality to ensure Revenue sustainability
- Must have correct information of residents and businesses
- Such project help tremendously in analysing your debt book identifying inactive debtors, deceased debtors, even indigents if the scope requires it, etc.
- It helps the municipality to create confidence to their residents if their information is correct.
- It helps enormously with debt collection process you have less disputes, less wasted expenditure in chasing wrong debtors (or finding the responsible debtor)
- Municipalities must review their Credit Control policies especially with regarding the deceased properties.

ABOUT NTIYISO CONSULTING



OVERVIEW

- Management consulting firm
- 100% black owned and managed
- Over 14 years in existence
- Over 100 consultants



LOCATIONS

- Centurion
- Bedfordview
- Durban
- Malamulele Limpopo
- Plans to establish a regional office in Western Cape



SERVICES

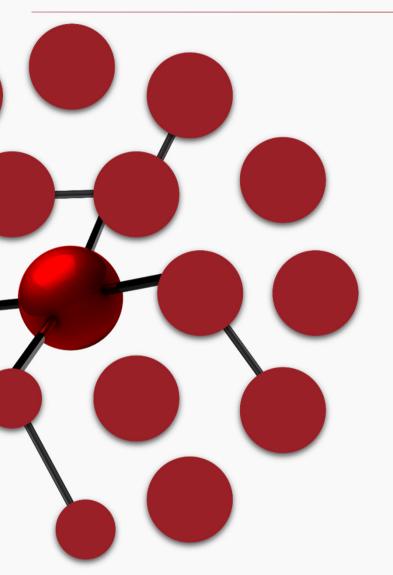
- Revenue enhancement and management
 - Debt Collection
 - Revenue Enhancement
 - Financial Management
 - Operational Support
 - Data Cleansing
- Governance and strategy
- Fundraising and development of catalytic projects
- Customer education and stakeholder engagement



CLIENTS

- City of Tshwane
- City of Ekurhuleni
- City of Joburg
- Emfuleni LM
- Ngqushwa LM
- Dr Kenneth Kaunda LM
- Greater Giyani LM
- Collins Chabane LM
- Knysna LM
- Soon Steve Tshwete LM
- Soon Bitou LM
- Soon Breedevallei LM

CONTACT US



ADDRESS

Unit 3, Stanford Office Park, 12 Bauhinia Street, Highveld Technopark Centurion

PHONE & WEB

T: +27 12 940 5435 F: +27 12 940 5436

www.ntiyisoconsulting.co.za

CONTACT PERSON/S

Tondy Nkuna – Managing Director Tondy.Nkuna@ntiyisoconsulting.co.za

Hannes Dednam – Associate Partner Hannes.Dednam@ntiyisoconsulting.co.za

