

# Complaints Analytics

Water & Wastewater Use Case  
**Region of Peel**

September 2022



# What we will be talking about

- Who are we?
- What is Complaint Analytics?
- What steps did we take to build it?
- What's next?
- Questions and Answers



# Who are we?

The Regional Municipality of Peel (aka Peel Region or Region of Peel) is a regional municipality with a population of about 1.4 million in the Greater Toronto Area (GTA). The Region of Peel is comprised of the cities of Mississauga and Brampton, and the town of Caledon.

Felix da Silva is an Advisor, Analytics for the Corporate BI team at the Region of Peel. He was also the project manager for the Complaint Analytics work that is being discussed today.

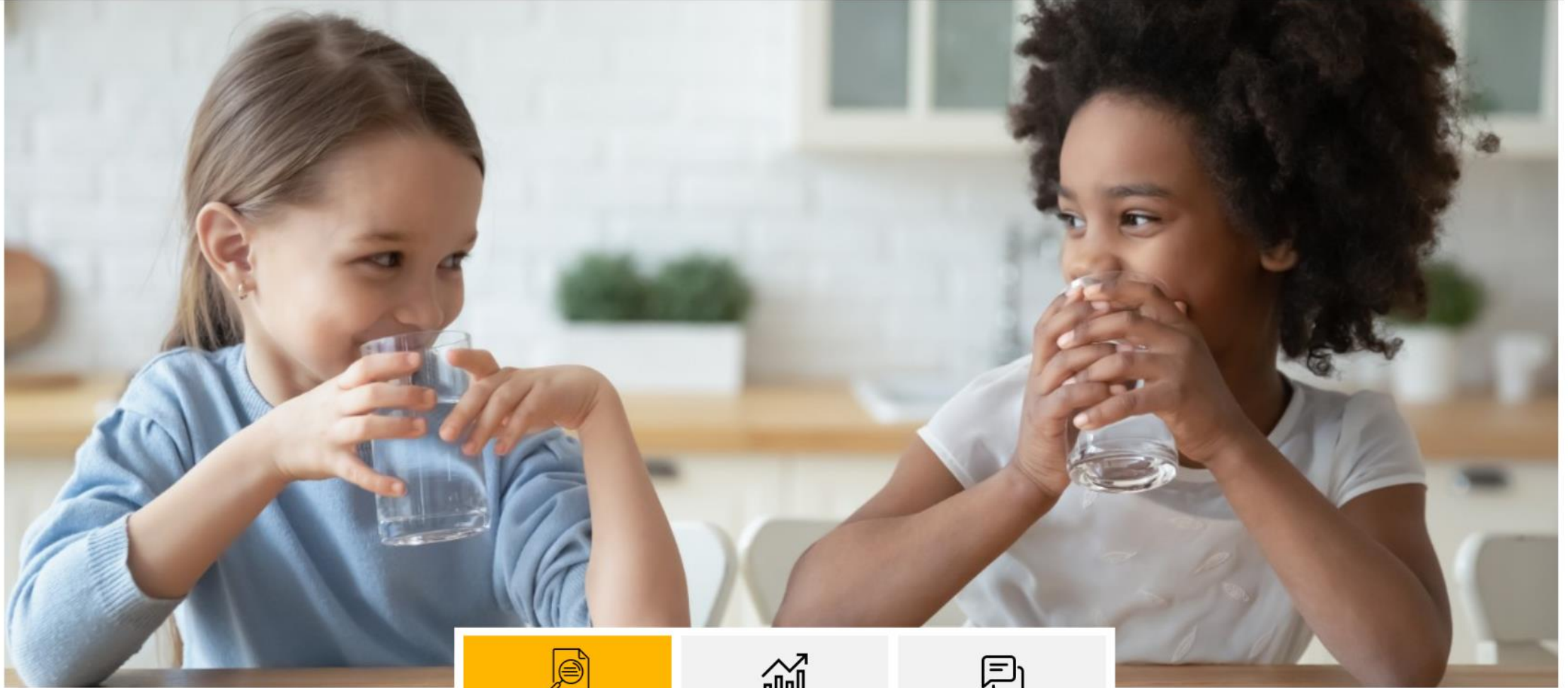
# What is Complaint Analytics?

A partnership between the corporate BI group and Water & Wastewater division to develop a repeatable and scalable EBI Use Case Playbook through the execution of a real-life scenario.

- Process Optimization from automation of processes for resident complaint management
- Improved Citizen Engagement by proactively making the relevant information available online or planning for a surge in complaints
- Repeatable process and lessons learned for the next use case through the development of the EBI Use Case Playbook
- Reusable and tested foundational components (Azure Cognitive Services, Data Bricks, etc.)

This proof of concept was completed in 2021 taking about 4 months (June to Oct).

# What does the end product look like?



Overview of Complaints



Complaint Intent & Trends



Complaint Process &  
Response

City	Ward	Date	
All	All	1/1/2021	7/29/2021
Division		Sentiment Score	
All		-1.00	1.00

Average sentiment across customer complaints. A higher score represents positive sentiment. Conversely, a lower score expresses negative sentiment.



Problem Type	# of Complaints	Avg Sentiment Score
Pressure	340	-0.20
Information	127	0.07
Discoloured	126	-0.09
Taste/Odour	124	-0.08
Residue	35	-0.06

## What are people concerned about?

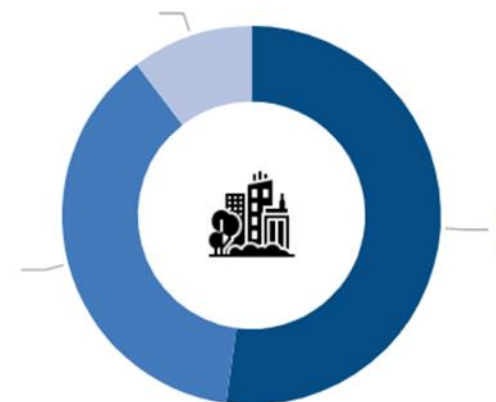
Understanding complaint intent & trends. When / where are they occurring?



**What are the overall Themes/Topics among the complaints?**

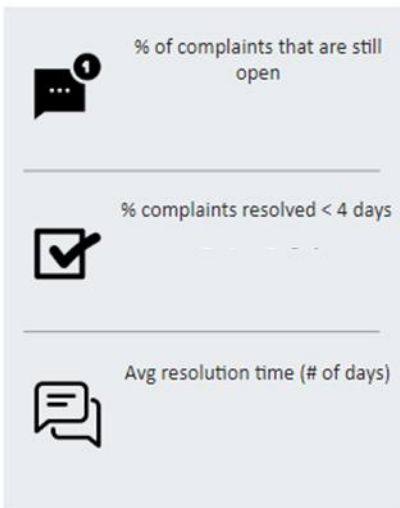


How many complaints are there by city?

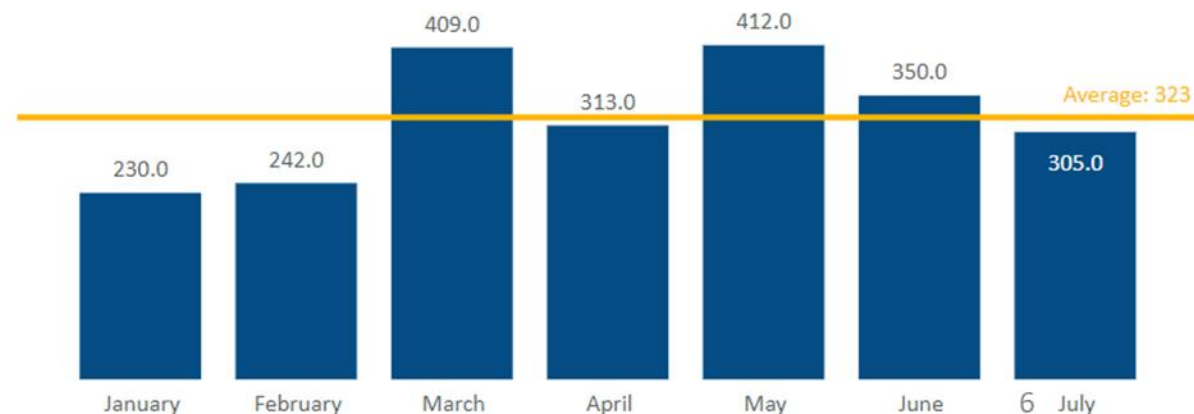


## What / How did we do?

### Exploring complaint process & resolution

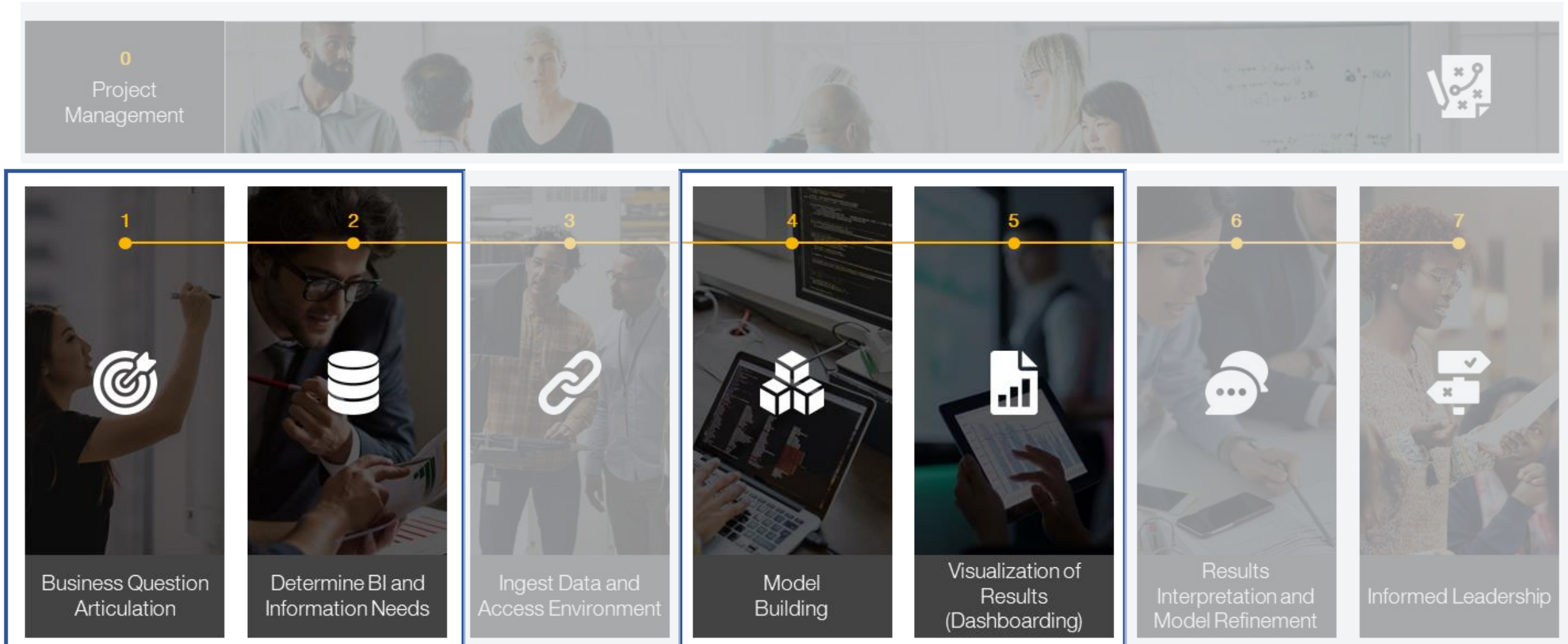


How many distinct service complaint numbers (ServNo) do we receive on average every month?



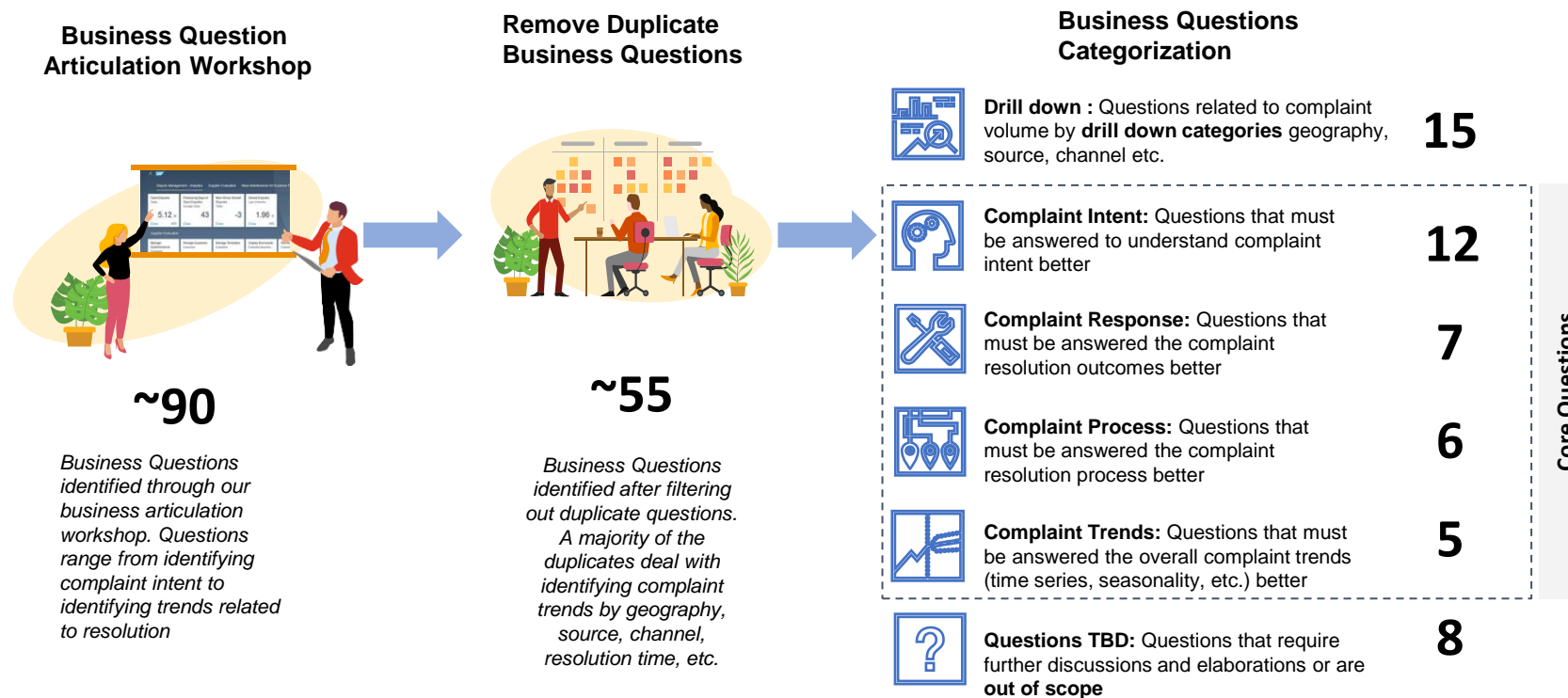


# 8 Steps for our Complaint Analytics Journey



# 1. Business Questions Articulation

We identified ~**90 questions** through our workshops. Barring duplicates, there were **30 core questions** that require prioritization



**Note:** Finalized set of questions to be determined based on data availability and condition



# 1. Business Questions Articulation

Based on our understanding of the problem, and the business stakeholders involved, we've identified 4 main personas of the dashboard and analysis



Allocate Resources  
Addressing Citizen Concerns

## Leadership

### Key User Stories to explore in this pilot:

- Understanding High-Level Patterns and Themes in order to:
  - Allocate resources more effectively
  - Provide clearer communication to constituents



Efficiency  
Processes  
Trends

## Supervisor / GM

### Key User Stories to explore in this pilot:

- Understanding trends, patterns, and themes in order to:
  - Allocate resources more effectively
  - Escalate and report to senior leadership
  - Make informed decision making, and getting feedback on past decisions



Root cause analysis  
Understanding why and intent

## Water and Wastewater Analytics Staff

### Key User Stories to explore in this pilot:

- Understanding root causes in order to:
  - Forward findings to the appropriate decision making teams for planning
  - Assess and evaluate effectiveness of past decisions, and provide recommendations
- Identifying data quality in order to:
  - Determine trustworthiness and reliability of insights generated



Citizen Experience  
Resolution effectiveness




## Service Experience

### Key User Stories to explore in this pilot:

- Understanding root causes and themes in order to:
  - Identify ways to prepare for different complaints (resourcing, and processes)
  - Identify whether there are potential areas of improvement as it pertains to service experience processes

## 2. Determine BI and Information Needs

A continuation of Step 1, this step is focused on the **technical feasibility of the analytics use case**, which will assist in finalizing the prioritized business question list, which should be prioritized based on a combination of technical feasibility, business needs, and user desirability

Master Datasets for the current phase																	
Dataset	Intended Audience	Datasets Covered															
		CUSTPROB	ADDR	CUSTCALL	HISTORY	PROBDEFN	ACTDEFN	COMPTYPE	VARCPROJ	TBL 3125	V_AIMSSRW QC	GIS Capital projects	GIS address points	CC&B	Access - Issues	Salesforce	Twitter
Hansen Master Construction		✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓			
Hansen Master Compliance / Operations		✓	✓	✓	✓	✓	✓	✓			✓		✓	✓			
MS Access												✓			✓		

### Backup datasets for additional insights

  
W&WW complaints (**700-800** complaints in last 7 years)

  
Customer sentiment and comments related to Peel's water and wastewater service excellence



Construction



Compliance



Operations



Billing



Table included in the master dataset



Tables that may not be included in the final master dataset (if data availability is low)

## 2. Determine BI and Information Needs

Leverage machine learning to clean up text based feedback, emails or comments to derive complaints and themes.



### Step 0: De-Identification & Custom Acronyms

Data should be de-identified, remove all personal information. This also helps remove unnecessary words not required for analysis such as Names.

Translate custom shorthand into long form such as WM = Water Main or COM for City of Mississauga



### Step 1: Normalization

Process of doing initial cleanup of the data, including:

- Lowercasing words
- Removing unknown characters
- Cleaning HTML tags (ex: \n>)



### Step 2: Tokenization

Tokenization breaks raw text into smaller chunks (typically individual words). This allows text to be fed to a model sequentially, and allows individual words to be analyzed for steps 3 and 4



### Step 3: POS (Part of Speech) Tagging

Part-of-Speech (POS) tagging is the process whereby words within a corpus are categorized and labelled according to their definition within a set context.

For example, the words in the sentence "I like apples" would be tagged as: pronoun, verb, noun.



### Step 4: Lemmatization

Lemmatization analyzes individual words, identified through tokenization and POS tagging, and identifies the word's root/base form.

Example:

- To walk → walk
- Walked → walk
- Walks → walk
- Walking → walk



### Step 5: Stop Word Removal

Stop words are words commonly contained in text which do not convey a significant amount of information.

Conjunctions, prepositions, and articles are typical examples of stop words which are removed to reduce the training set (data) size to focus on the more semantic parts of text.



### Step 6: Additional stop words

Also, additional stop words like Region of Peel and the list of staff names to be removed as not part of the analysis.

Algorithm-based

Rules-based

Rules-based

Algorithm-based

Algorithm-based

Rules Based

Dictionary Based

# 4&5 Model building and Visualization of Results

An iterative process switching from one to another on a weekly basis with the business.

## Sample Comment

We **understand the north lanes are closed** but there is no work happening.  
This situation **is creating a situation where congestion blocking access and people are making illegal traffic turns.**

Can I ask that you have someone review the impact of the closure to traffic flow. And **perhaps there can be a better layout.**

Let me know when you have reviewed

## Remediated Comment

**understand north lane close** work happen  
situation **create situation congestion block access people make illegal traffic**  
turn ask someone review  
impact closure traffic flow  
**perhaps good layout** let  
know review

## Cognitive Services Output

Sentiment = **Negative**

Positive Confidence = **0.17**

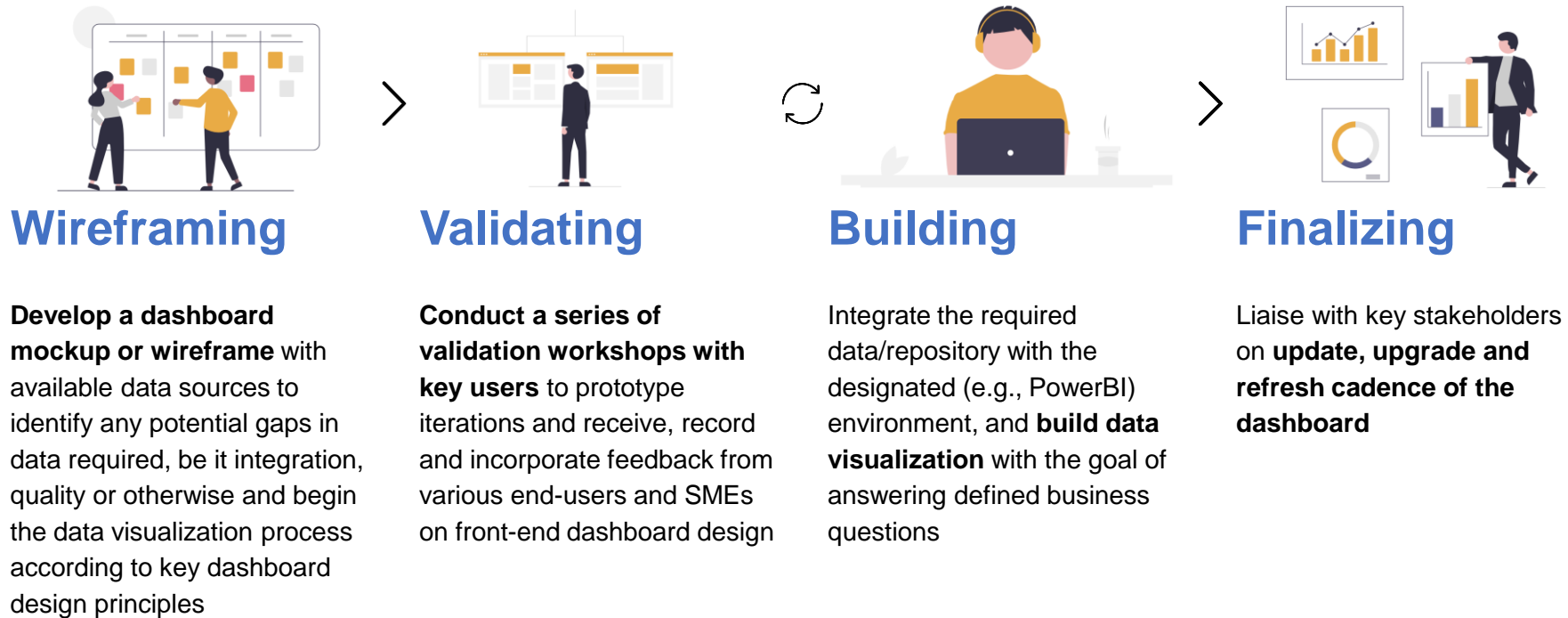
Negative Confidence = **0.78**

Neutral Confidence = **0.05**

Key Phrases = **north lane close work,illegal traffic,impact closure,traffic flow,good layout,situation congestion,access,people,so meone**

# 4&5 Model building and Visualization of Results

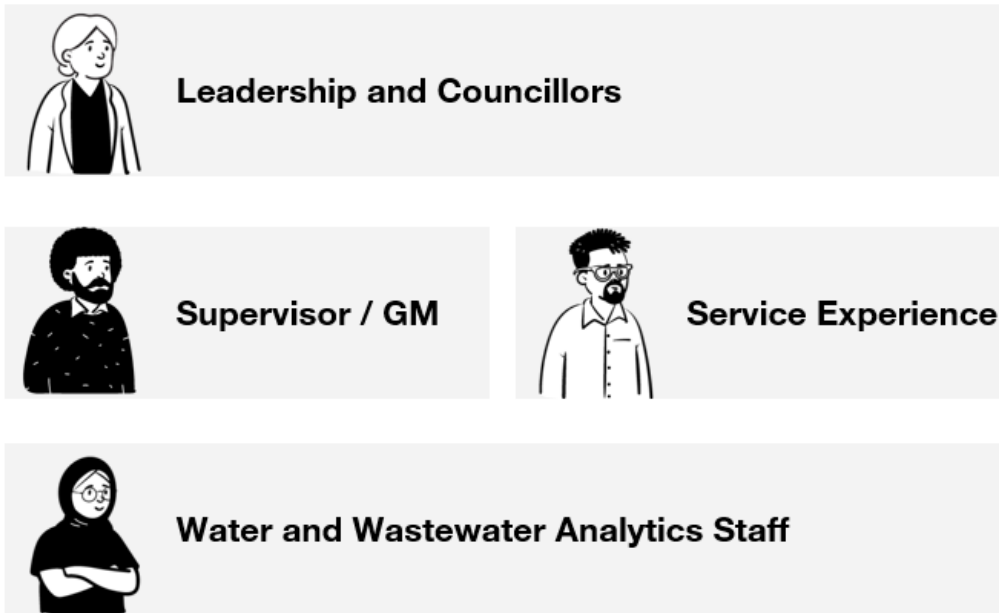
Taking all the prioritized questions, each persona and breaking it down to the number of visualizations and reports we intend to build





# 4&5 Model building and Visualization of Results

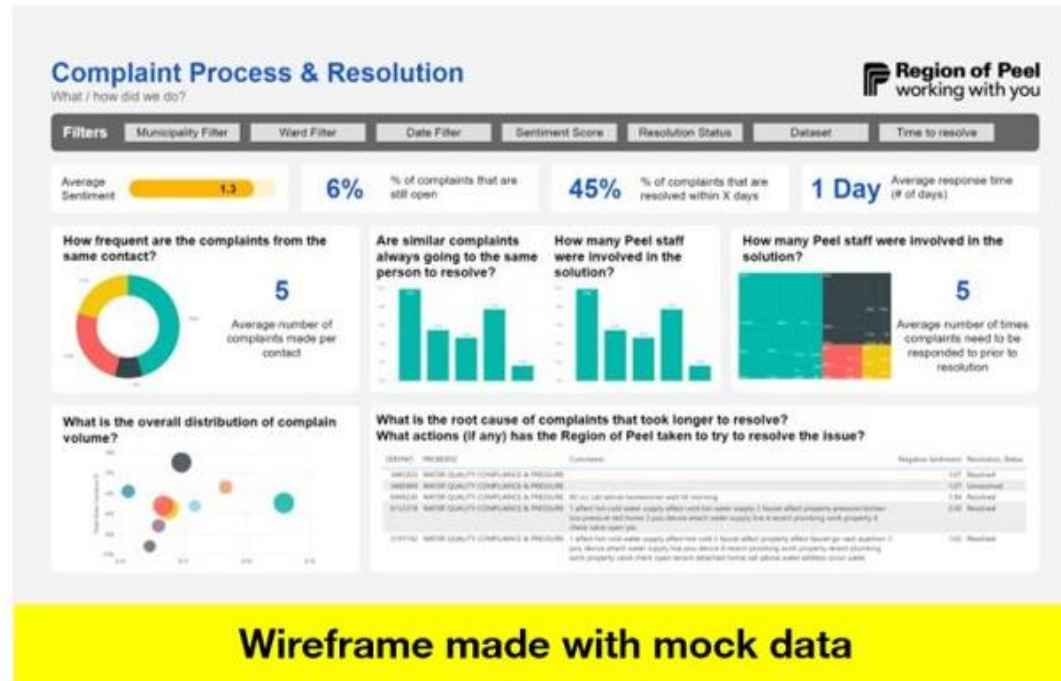
Each of the **9 pages (reports)** target different personas, but personas can look through other pages as well to fill in information gaps as needed



# 4&5 Model building and Visualization of Results



## Service Experience



## User Stories: What am I trying to do with the data?

### Improving Service Outcomes

I need to know complaint resolution effectiveness and outcomes so I can assess whether the resolution process is optimal, and to identify ways to better prepare for different complaints (resourcing, and processes)

### Tracking Satisfaction

I need to know if residents are satisfied with how we are communicating with them during the resolution process so that I can report our performance and progress to our senior leadership team

- Are complaints resolved consistently within a time frame?
- What percentage of complaints are still open and outstanding vs resolved?
- Are similar complaints always going to the same person to resolve? (TBD based on data)
- What was the time taken to resolve?
- What actions (if any) has RoP taken to try to resolve the issue?
- How many times did the complaint need to be responded to before it could be resolved?
- What is the root cause of complaints that took longer to resolve?
- How many Peel staff were involved in the solution? (TBD based on data)
- Who resolved the complaint, when? (RoP vs contractors)

- How frequent are the complaints from the same contact?
- Did complaint require resolution or was complaint just an expression of anger?
- Is it a repeat complaint? Are people complaining about the same issue?
- What is the overall distribution of complaint volume?
- What is the overall citizen sentiment?

**Business Questions:** What questions need to be answered for me to accomplish by goal?

# 4&5 Model building and Visualization of Results

Mapping back each prioritized question to a visualization to gather user feedback and user acceptance testing.

# = visual number reference



**Key Persona:** Leadership / Councillor

**Questions Answered:**

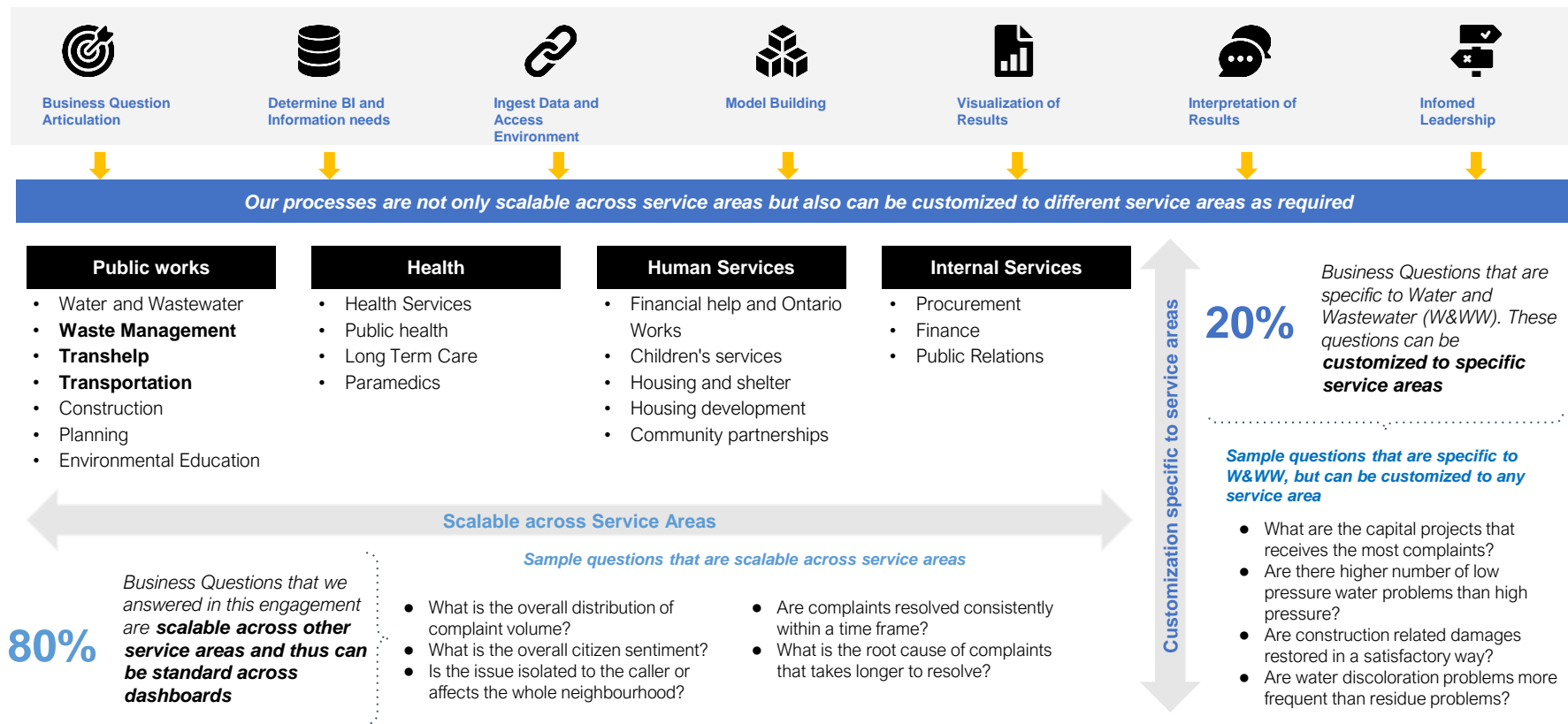
- 2** What is the overall citizen sentiment?
- 5** What are the overall Themes/ Topics among the complaints?
- 4** **6** Is the issue isolated to the caller or affects the neighborhood?
- 8** **9** What was the time taken to resolve?
- 10** Which months receives the most complaints?
- 8** Are complaints resolved consistently within a time frame?
- 7** What percentage of complaints are still open and outstanding vs resolved
- 3** Is the complaint about something in our control?

**Filters:** **1**

by Municipality, Date, Division (construction, compliance, operations), Sentiment score

# Scaling and centralizing complaints

80% of the business questions that we answered during the Complaint Analytics work with Water & Wastewater can be scaled across other service areas. The data source or system is irrelevant as long as it can be connected to or extracted. Centrization does not mean it has to be the same system.



# What's next?

- This proof of concept and was successful in that we created a repeatable process that could be applied to other areas
- Manually refresh the data for the reports and dashboards that was built as is
- From the business perspective, the focus is on the process, systems and data improvements as one of the recommendations that came out from this work
- Consistency in terminology should be across the board (resolved vs. responded) for centralization
- Started to address data related initiatives such as data catalogues, de-identification and dictionaries
- Established and tested technical foundational components (azure data lake, cognitive services and data bricks)
- Currently executing another use case using the EBI Use Case Playbook for Growth Management building upon the same foundation



# Questions and Answers

Thank you

Connect with me any time ([felix.daSilva@peelregion.ca](mailto:felix.daSilva@peelregion.ca))

