Data and the Leakage Team Forum

Data and the Leakage Team Forum

- Chaired by Pauline Walsh, CEO, Affinity Water and Chair, Leakage PIC
- The application of data science in the delivery of Yorkshire Water's leakage ambitions
  - Chris Smith, Data Science Manager, Yorkshire Water
  - Carol Smith, Data Science Manager, Yorkshire Water
- Data and leakage at Dwr Cymru/Welsh Water
  - Chris Rees, Leakage Strategy Manager, Dwr Cymru
  - Dr Willow Smallbone, Data and Analytics Manager, Dwr Cymru

The application of Data Science in delivering Yorkshire Water's Leakage ambitions

Carol White & Christopher Smith

Yorkshire Water – Data Science

YorkshireWater

Data Science at Yorkshire Water

Carol White

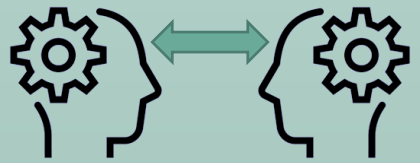
Yorkshire Water – Data Science

YorkshireWater

Data Science at Yorkshire Water

Upskilling & BAU

Collaborative working approach to upskill colleagues and enable knowledge transfer



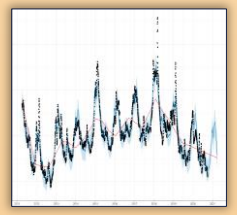
Partnership working

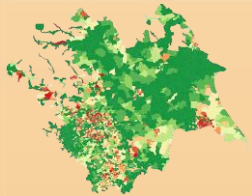
Mentoring

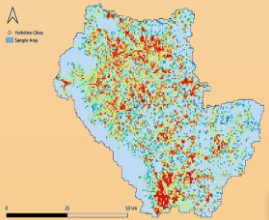
Knowledge transfer (2-way)

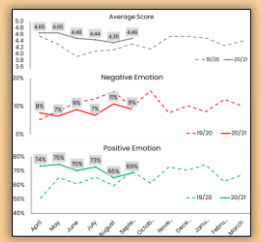
Tactical

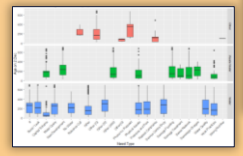
Stand-alone projects, dashboards, forecasting (reactive to immediate business needs)





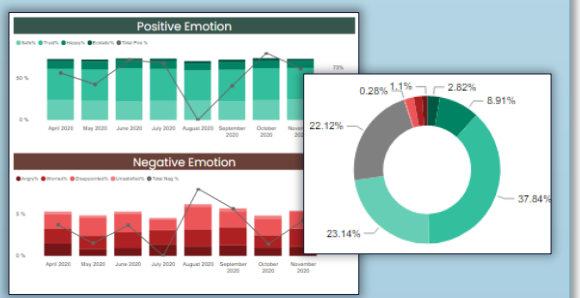






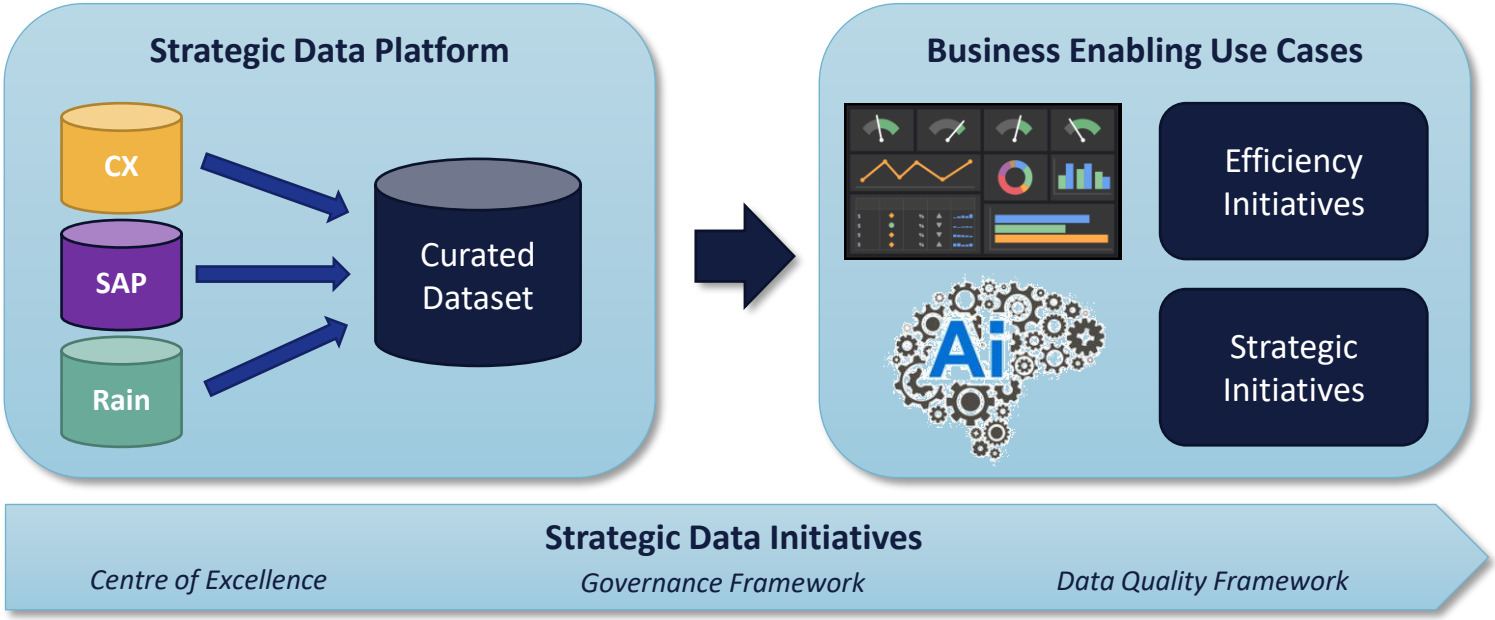
Strategic

Fully productionised models (automated & deployed)



Data Transformation

Data Centric Organisation



YorkshireWater

Approach

Key guiding principles

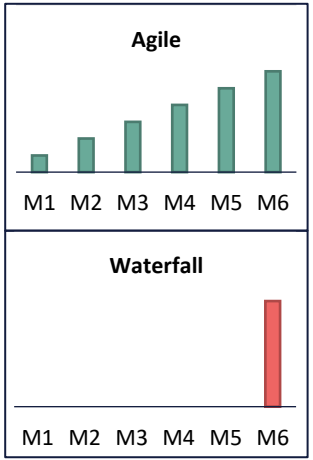
- Actionable insight
- Enable strategic change
- Sustainable
- Accessible to all
- Customer-focused

Agile delivery

- Minimise risk
- Frequent re-prioritisation
- Deliver MVPs
- Adaptive

Key Considerations

- Technical enablers
- Multiple delivery streams
- Business benefit & prioritisation
- Data ownership & governance



CASE STUDY

Customer Emotion Model

- Custom-built Natural Language Processing model to capture emotion from customer feedback
- Fully automated user-friendly dashboards
- Partnered with business change team to realise benefits
- Enables colleague self-service and proactive intervention
- Potential to apply to new data in data mart

Potential Applications of Data Science in the Leakage Arena

Christopher Smith

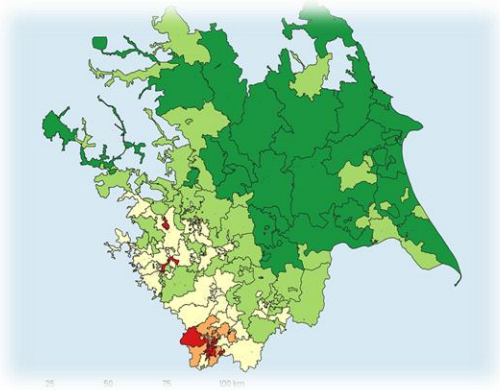
Yorkshire Water – Data Science

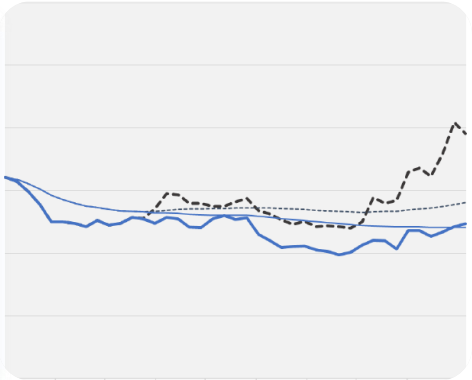
YorkshireWater

Hindsight: Benefit Assessment

What

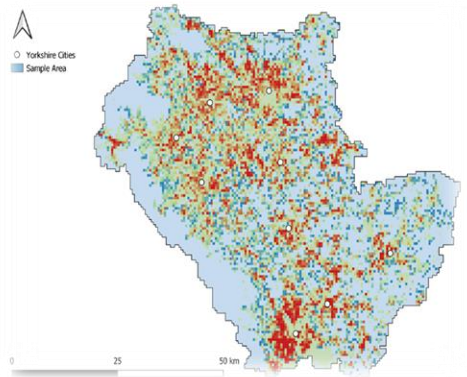
- Quantifying the benefit of existing & new approaches to leakage management
  - ALC
  - Enhanced Telemetry
  - Pressure Optimisation

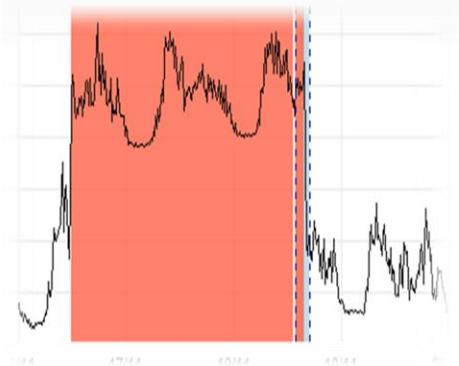




Why

- Evaluating the effectiveness of leakage activity
- Assessing the cost-benefit of a proposed solutions
- Establishing the potential for wider deployment – full benefit potential
- Identifying the opportunities for further improvement





Key Considerations

- Transient nature of leakage
- Trial versus BAU
- Attributes that determine benefit

YorkshireWater

Foresight: Forecasting

What

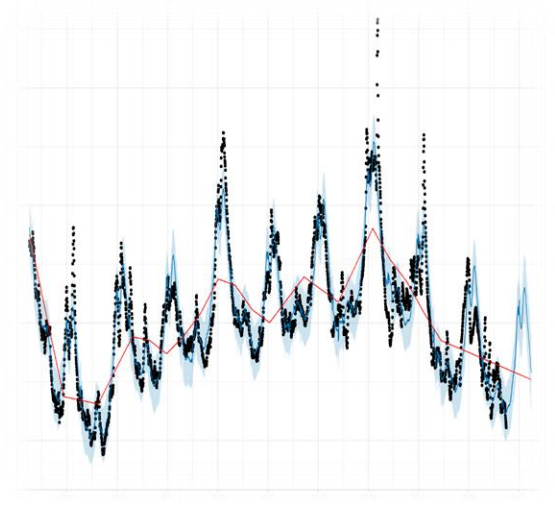
- Statistical modelling to predict leakage performance in-year and beyond

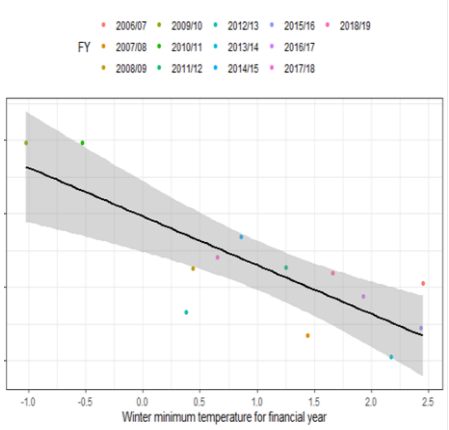
Why

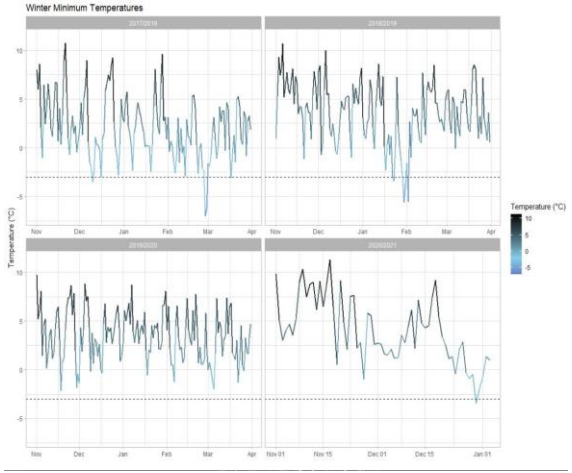
- Business planning
- Determining operational targets
- Monitoring risk

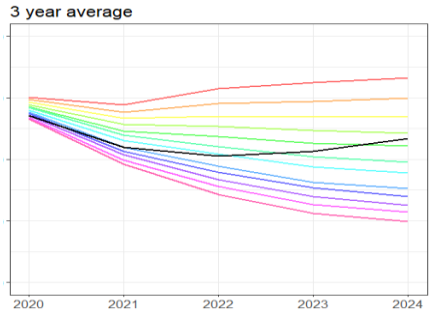
Key Considerations

- Understanding relationship between external factors and leakage
- Understanding benefit of activity (AMP forecast)
- Embracing & reflecting uncertainty









YorkshireWater

Insight: Analytical-driven Interventions

What

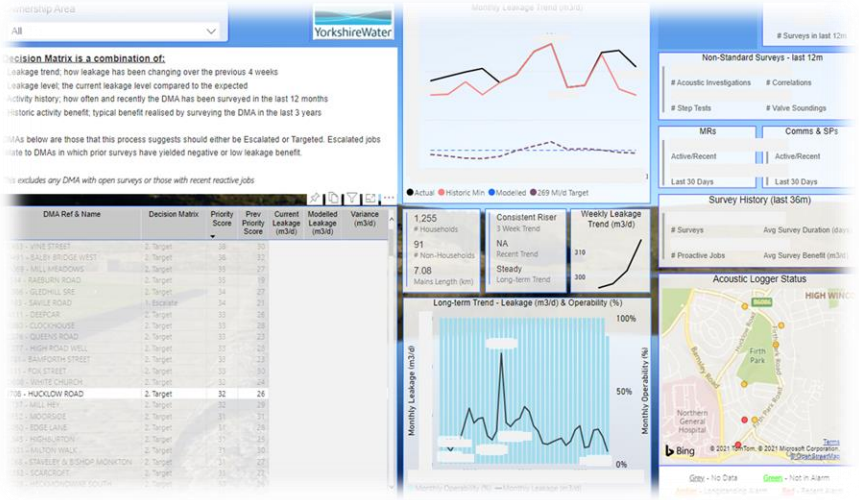
- Analytical solutions to enable the prioritization & evaluation of leakage activity

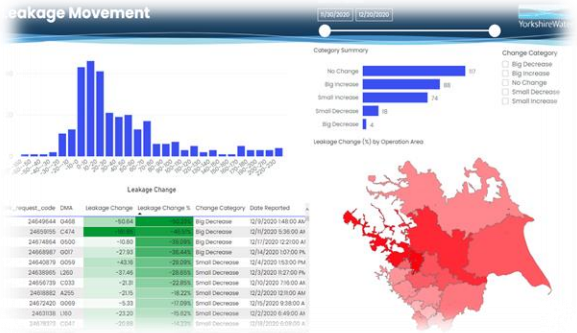
Why

- Assisting in decision-making process
- Identifying problem DMAs for escalation
- Providing performance KPIs

Key Considerations

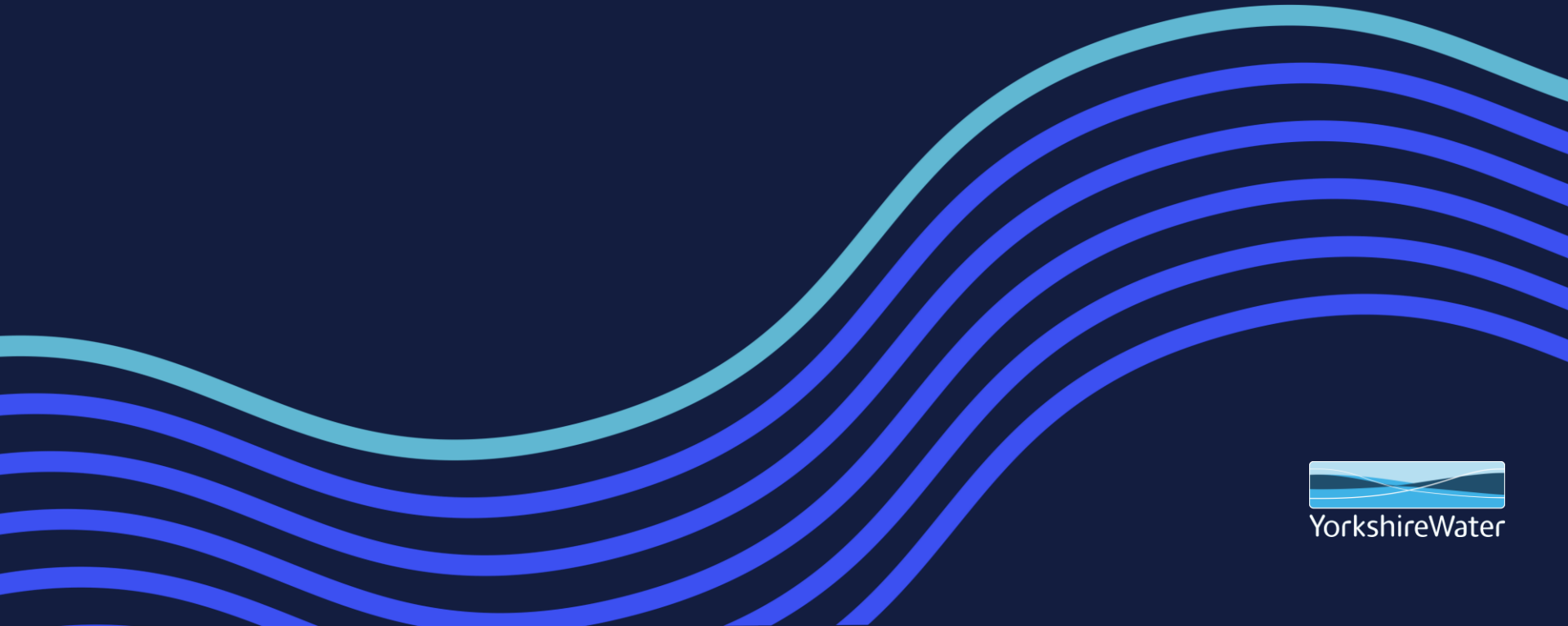
- Deliver little & often
- Incorporate attributes of benefits
- Integrated view of the network
- Data availability & veracity
- Data velocity



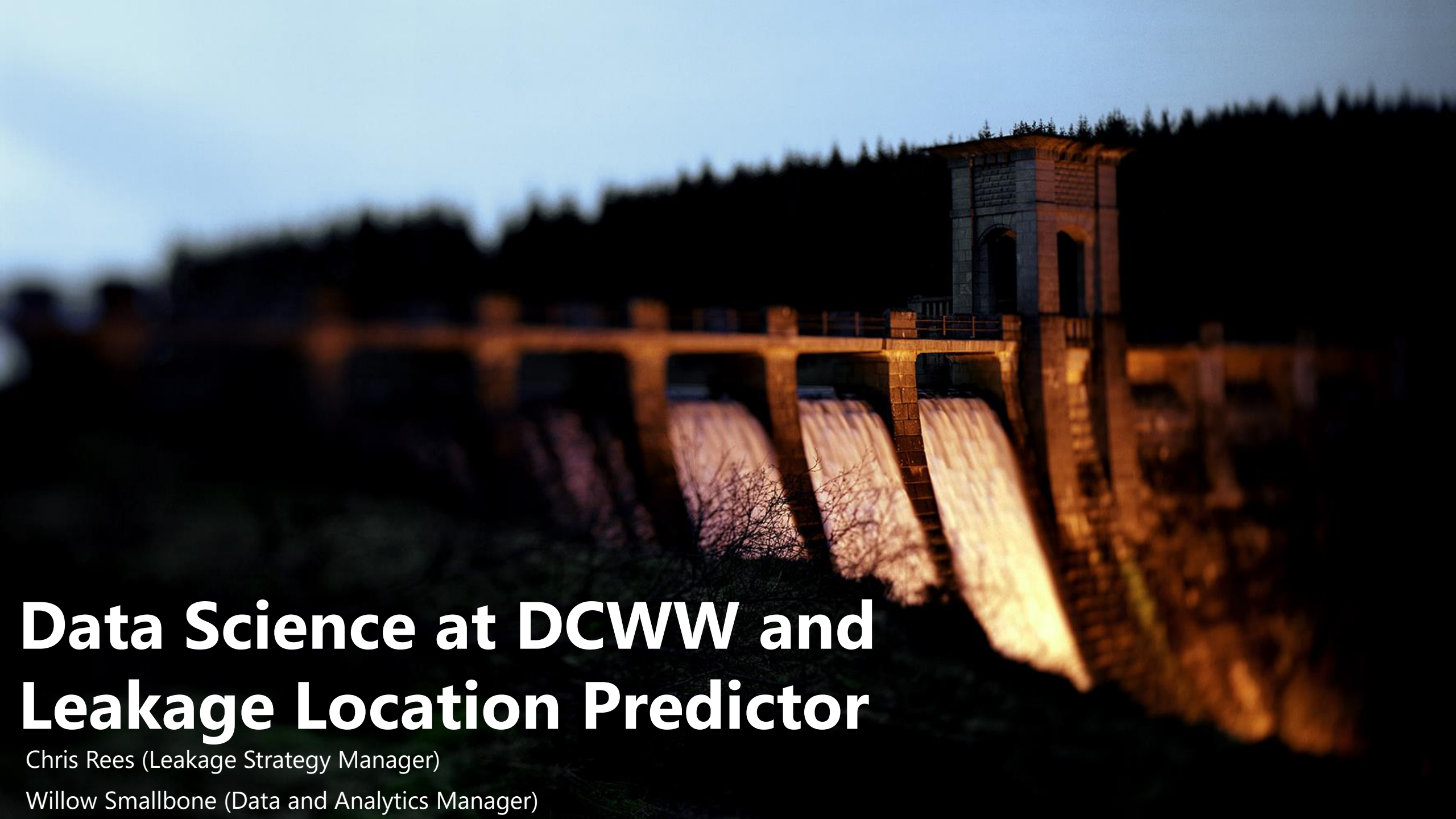


YorkshireWater

Thank you



YorkshireWater



Data Science at DCWW and Leakage Location Predictor

Chris Rees (Leakage Strategy Manager)

Willow Smallbone (Data and Analytics Manager)

Data Team

- Director of Operational Services
  - Head of Data
    - Data Governance
    - Data Engineering
    - Data Science
    - Business Intelligence
    - Data Migration

Dŵr Cymru
Welsh Water

Data Team

The Journey

Our Journey with Data and Analytics

Some Highlights…

2014
Chief Statistician Appointed
1

2015
Statistics and OR Team Formed
3

2016
First Data Engineer Appointed
4

2016
First Data Science Graduate Programme
6

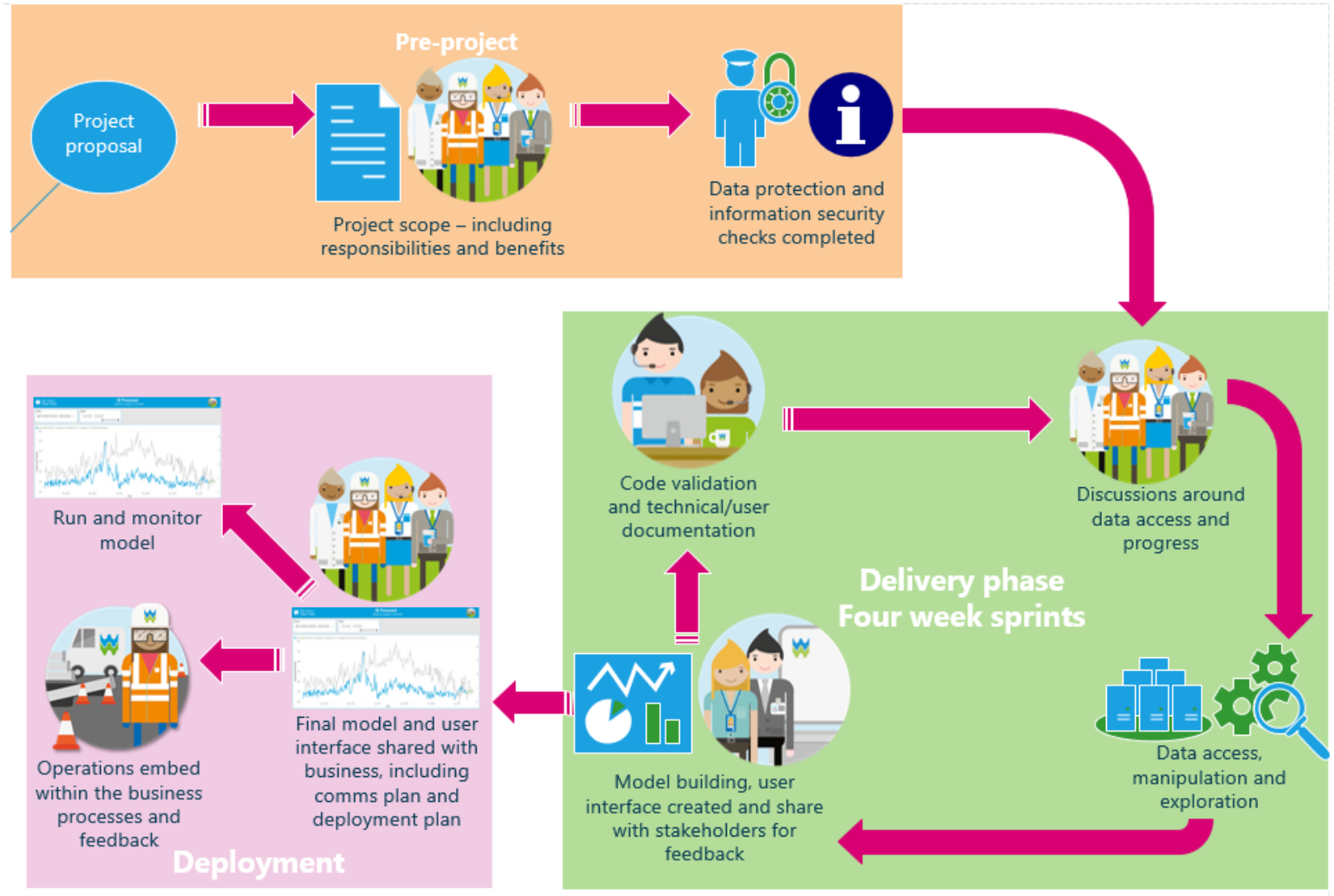
2017
Data Governance Team Formed
13

2018
Data Team Formed
16

2019
Business Intelligence CoE Formed
22

Project Lifecycle

Data Science



Dŵr Cymru
Welsh Water

Project Nemo

Data Science Idea

Project Nemo

An Overview

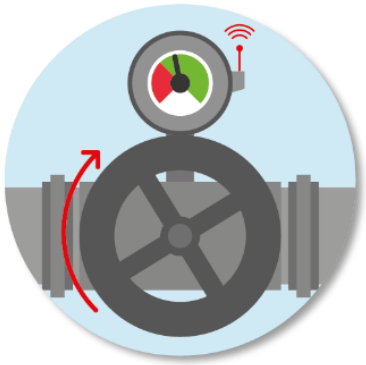
Current challenge

When a pipe leaks/bursts, the exact location of the leak is not known. It can take some time to identify the exact location of the leak and is quite a manual task, delaying any resolving action to be taken.

Identify locations where a leak is most likely to be and provide an optimised route for locations to be inspected within a DMA

- Predict the locations of a leak
- Provide a minimal route in which to inspect the most likely leakage locations
- Create an application to be used by inspectors that outputs the locations to be inspected and provides the minimal route which should be taken





Project Nemo

Benefits

Reduced leakage

Improved productivity

Reduced environmental impact

Safer working environment

More engaged workforce

Reduced CML impact

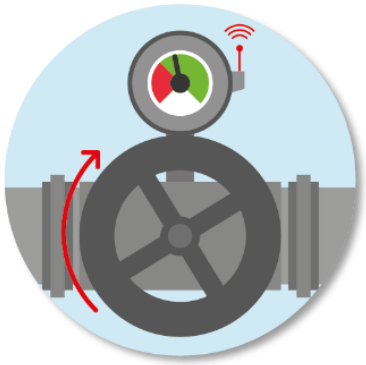
Reduced cost by optimised route

Data driven investment

Project Nemo

Delivery approach

Phase 1	Proof of concept MSc project with Cardiff University	Produce a model to identify leak location
Phase 2	DMA Clustering	Group DMAs based on similar characteristics
Phase 3	Leakage Location Predictions	Produce a model for each group of DMAs to predict leak location
Phase 4	Route Optimisation	Provide the best route to visit the predicted leak locations
Phase 5	Application	Develop an application for operations to use

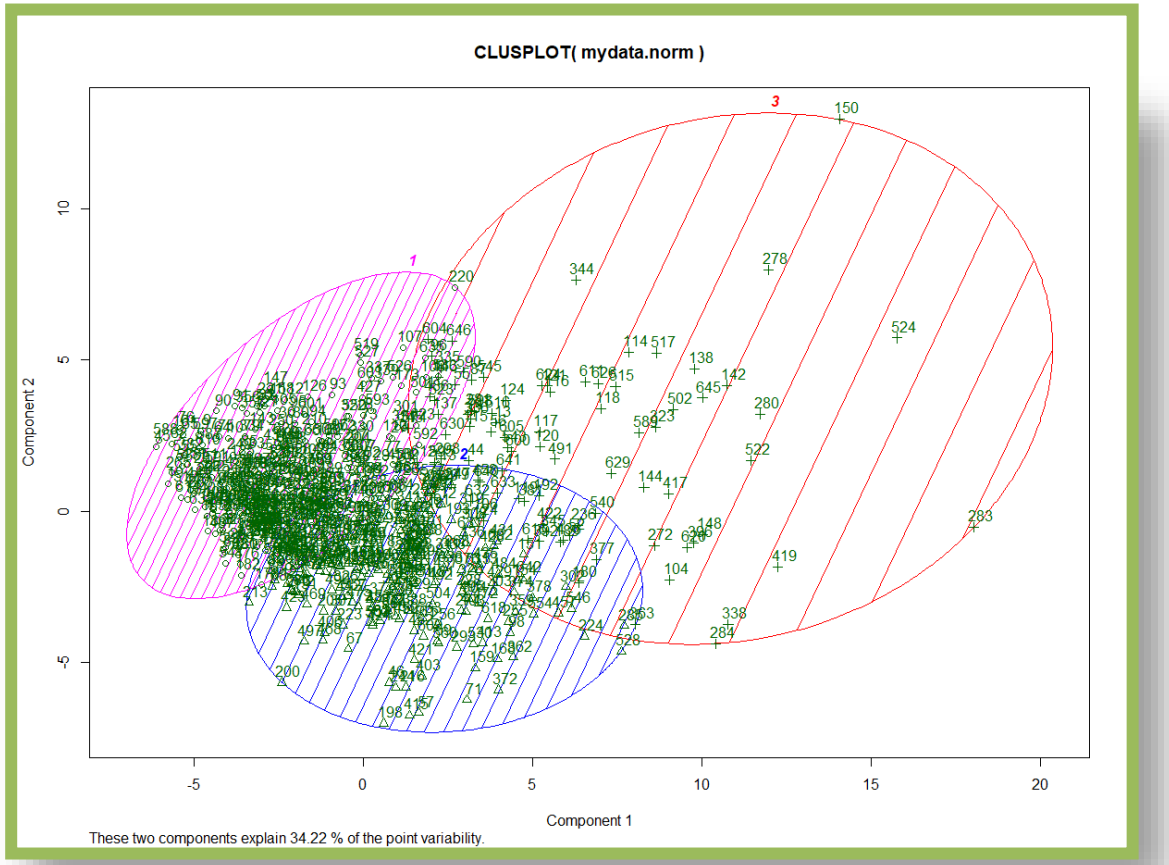


Project Nemo

DMA Clustering

Group DMAs based on similar characteristics

- Clustering DMAs based on the similar characteristics relating to leakage
- Identified similar DMAs into three clusters
- Descriptive statistic analysis is performed to gain further insight into the leakage data outputs

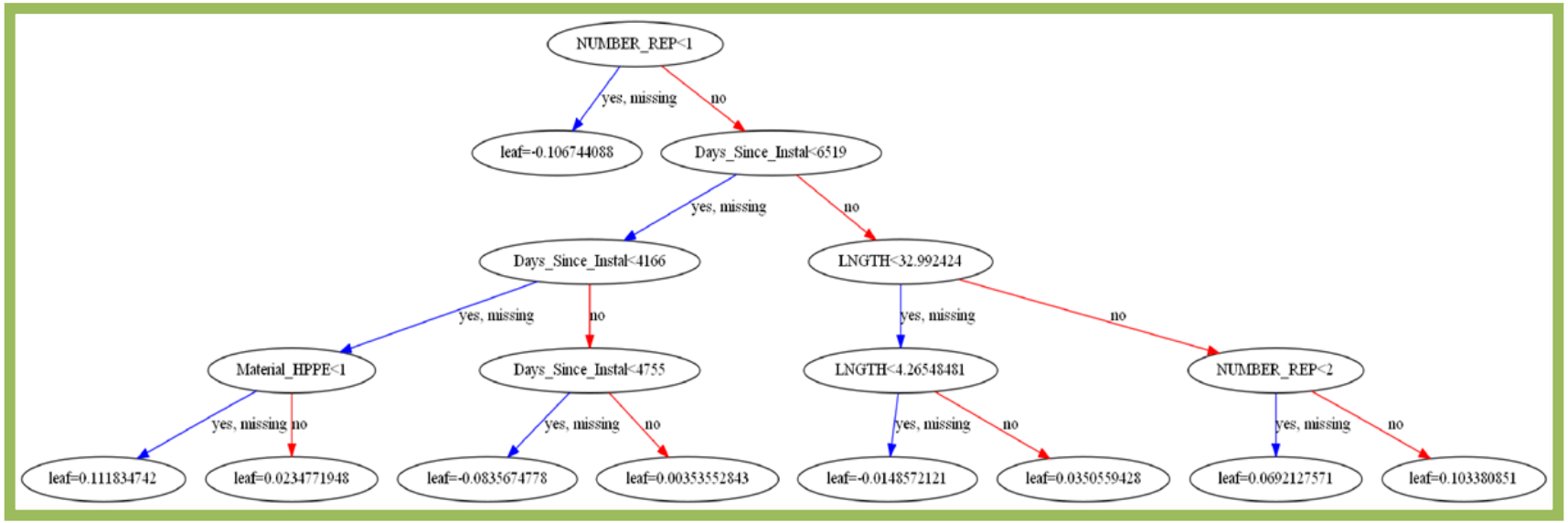


Project Nemo

Predictive Model

Predict the pipes that are most likely to be the cause of a given leak for each group of DMAs

- Created a model for each group of DMAs by using multiple data sources in order to predict the pipes that are most likely to be the cause of a given leak
- Tested several different statistical algorithms and determined the methodology that provided the most accurate results

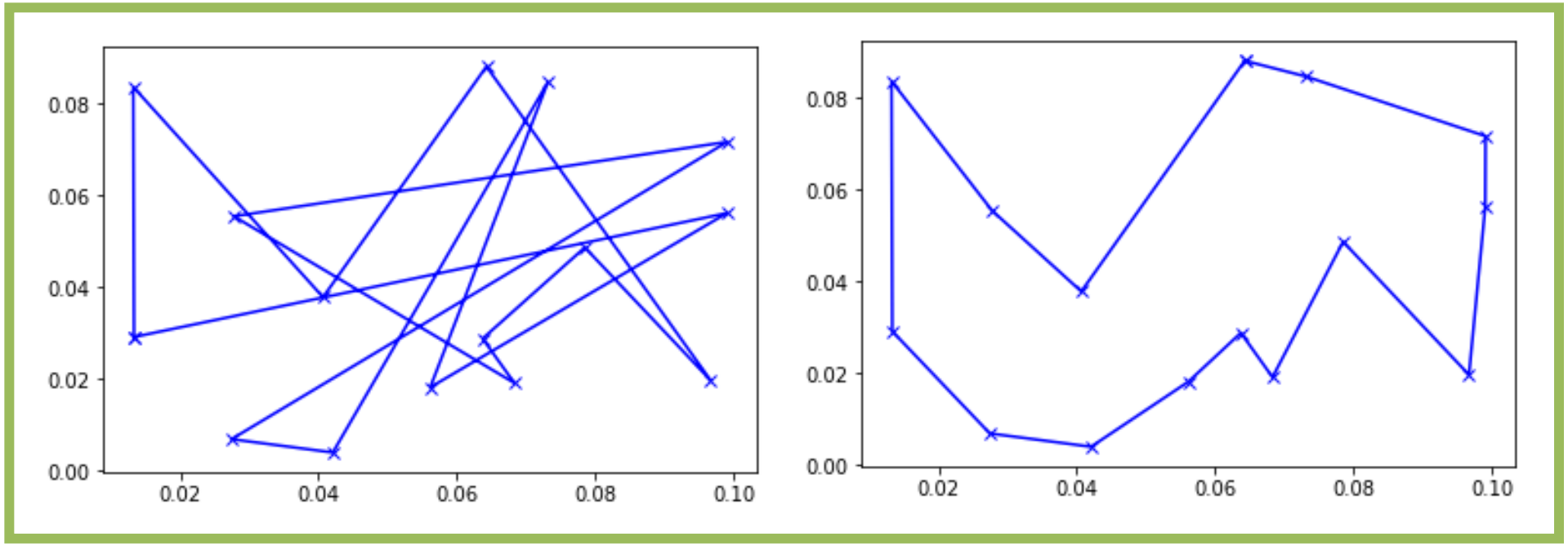


Project Nemo

Route optimisation

Create the optimal route for visiting potential leak locations

- The problem that is being solved is a variation of the Travelling Salesperson Problem
- A genetic algorithm has been identified as the best algorithm for the problem
- Weighted distances are used to take the risk of each pipe having a leak into consideration

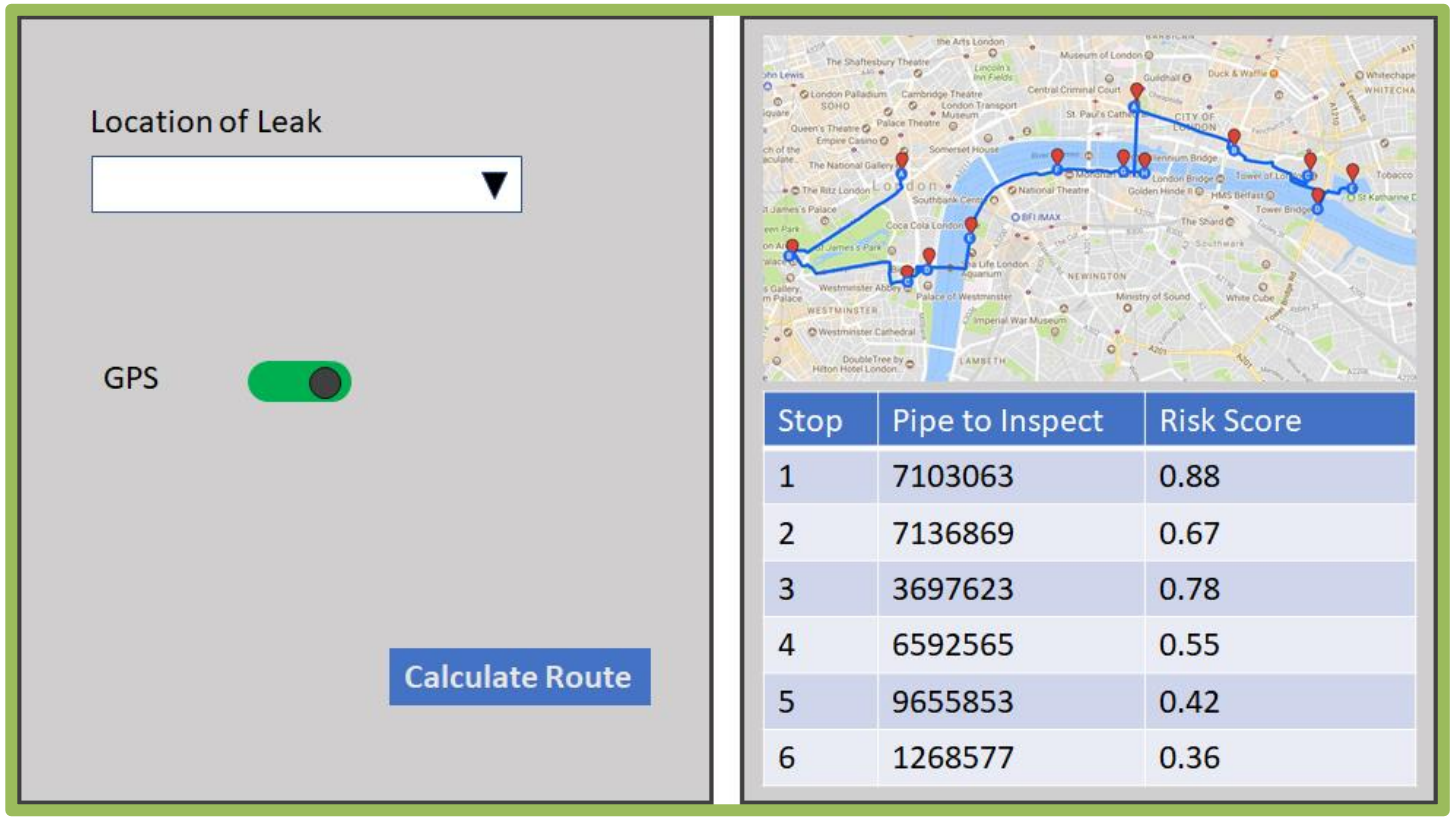


Project Nemo

Application mock-up

Develop a user interface

- Develop a user interface for operations to use to enable them to follow the optimal route to check all potential leak locations



Stop	Pipe to Inspect	Risk Score
1	7103063	0.88
2	7136869	0.67
3	3697623	0.78
4	6592565	0.55
5	9655853	0.42
6	1268577	0.36

Dŵr Cymru
Welsh Water

Any Questions?