Developing Deterioration Prediction Model for the Potable Water Pipes Renewal Plan

Case of Jubail Industrial City, KSA

Ali Madan Al-Ali

GIS Supervisor

Marafiq-Saur Operation & Maintenance Company (MASA)

Overview

- Background
- Scope and Objective of the Study
- Preparation and Methodology
- Identify Criticality of Residential Zones
- Identifying Influence Threat Factors
  - Asset Data
  - Hydraulic Data
- Statistical Analysis
  - Research Question
  - Logistic Regression and ANOVA Analysis
- Probability of Failure Prediction
- Prioritization of Critical Pipes
- Discussion and Conclusion

Background

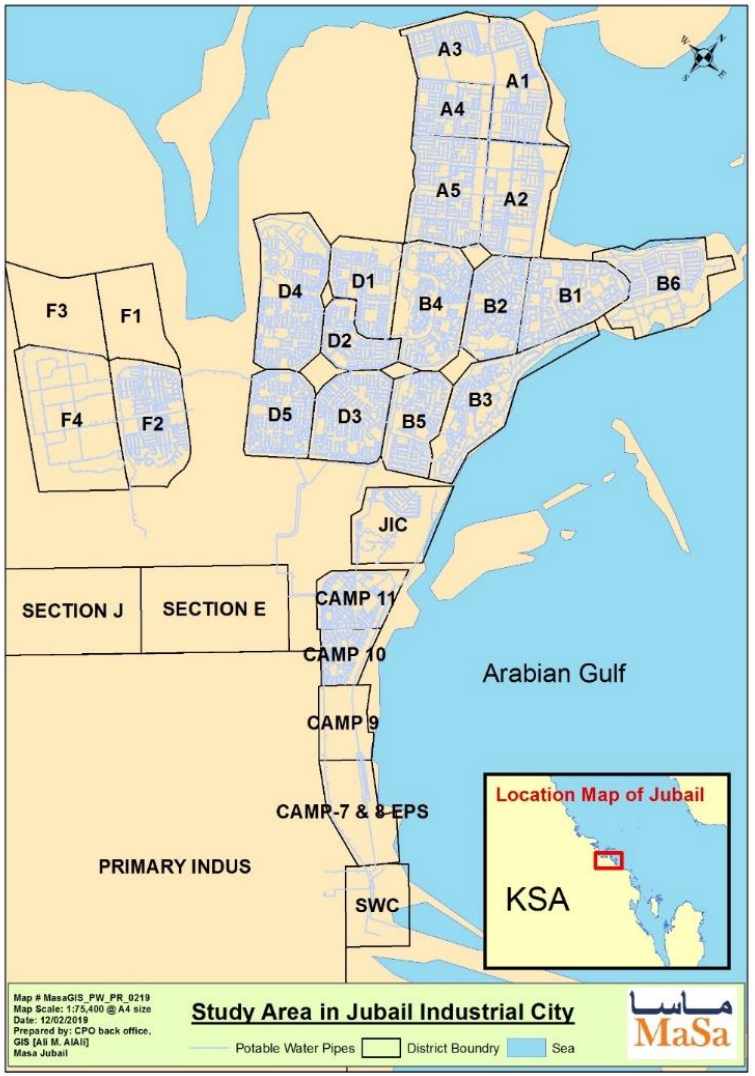
- **Marafiq** established in terms of Royal Decree M/29 of 18 October 2000 (22 Rajab 1421 Hijra) as a joint-stock company. Marafiq is owned by its four major shareholders:

شركة مرافق الكهرباء والمياه بالجبيل وينبع
Power and Water Utility Company for Jubail and Yanbu

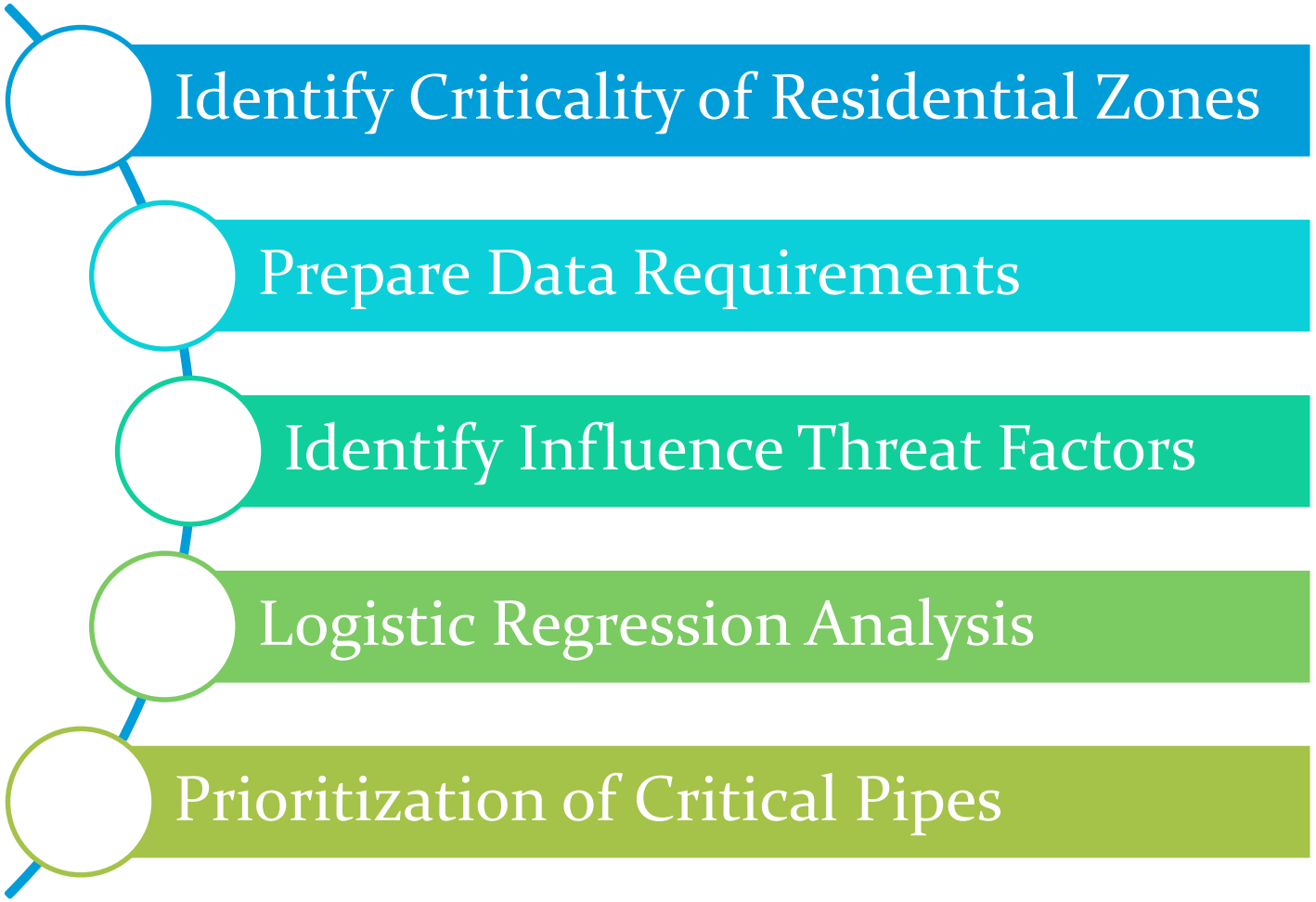
- **Marafiq and Saur O&M JV Company (MaSa)** is the first private water organization in Saudi Arabia based in Jubail and Yanbu founded in 2012. It serves among the largest Petro-Chemical industrial cities in the world. MaSa's role is to operate and maintain Marafiq's utilities in Jubail and Yanbu, these include:
  - Potable Water Facilities
  - Sea Water Cooling Facilities
  - Reclaimed Water Facilities

Scope and Objective of the Study

- **Community area** in Jubail industrial city including: Deffi, Fanateer, East Corridor, Jalmudah and Southern part of Mutrafiah.
- Quantity of PW Pipes Segments = 29,658
- Network Length = 928.25 km
- Pipes Built Years = 1980 to 2017
- Annual Renew Target = 1.7% (AWWA;2017)
- Life Cycle = 59 years
- Total Count of PW Pipe Breaks (Jan. 2012 to April. 2018) = **1,053**
- **Study Aim:** Identifying the most critical pipe segments (8.5% of the total network) that needs to be replaced during the next 5 years. Total length of 78.9km.



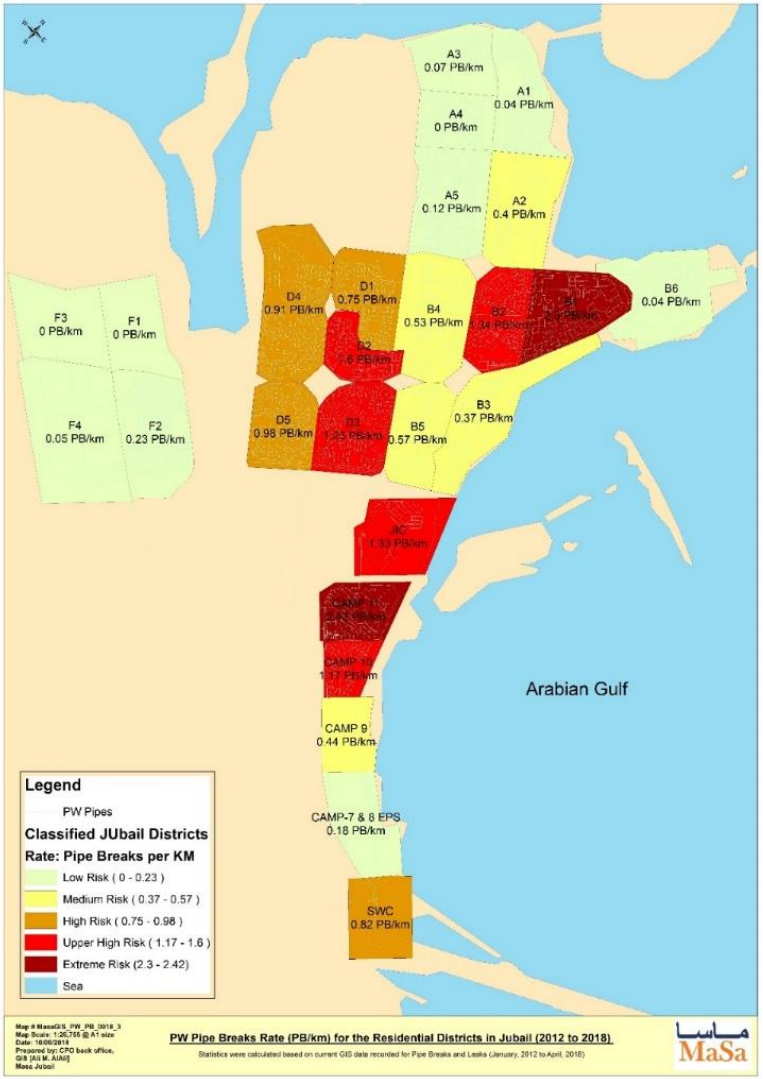
Process of Calculating Pipes Risk



Identify Criticality of Residential Zones

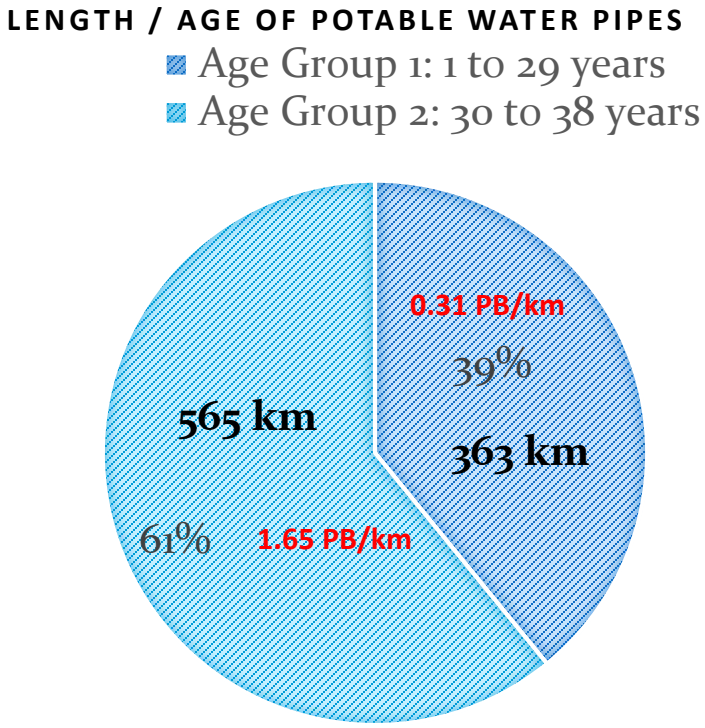
Rate: No. of Pipe Breaks / KM of Pipes for each District (2012 to 2018)

AREA NAME	AREA CODE	PB Qty	PW Pipes Length (KM)	Rate: PB/km
Al-Rayhan	F1	0	0.00	0
Al-Yasameen	F3	0	0.00	0
AL-ULA	A4	0	33.97	0
DAREEN	B6	2	41.31	0.04
RIYADH SECTOR	A1	2	41.86	0.04
Al-Rabea	F4	1	19.68	0.05
TAIF	A3	2	26.27	0.07
YANBU	A5	8	62.26	0.12
CAMP-7 & 8 EPS	C7 & 8EPS	1	5.57	0.18
Al-Khozamah	F2	13	56.10	0.23
SUDAYER	B3	17	44.99	0.37
TAIBAH	A2	16	39.48	0.4
CAMP 9	C9	8	18.15	0.44
NAJD	B4	33	61.49	0.53
AL AHSA	B5	24	41.94	0.57
AL ANDALUS	D1	29	38.56	0.75
SWC	SWC	4	4.87	0.82
AL FAROUK	D4	66	71.78	0.91
AL FERDOS	D5	48	48.61	0.98
CAMP 10	C10	25	21.36	1.17
AL FAIHA	D3	67	54.20	1.23
JIC	JIC	29	21.69	1.33
MAKKAH	B2	64	47.54	1.34
AL KODS	D2	50	31.07	1.6
AL HIJAZ	B1	132	57.18	2.3
HUAWALAT	C11	91	37.56	2.42



Influence Threat Asset Factors

- **Age of Pipes (Years)**



- **Pipes Material**

Table 3: Calculations of length and pipe breaks categorized by material type

Material Type	Length (km)	Total Pipe Breaks	Rate (PB/km)
AC	59.66	28	0.47
DI	41.21	15	0.36
FRP	0.09	0	0.00
GRP	52.01	25	0.48
PVC	129.24	171	1.32
RCP	20.17	1	0.05
SCP	20.14	9	0.45
uPVC	605.74	598	0.99
Grand Total	**928.3**	**847**	***Avg = 0.91***

Influence Threat Asset Factors

- **Diameter Size**

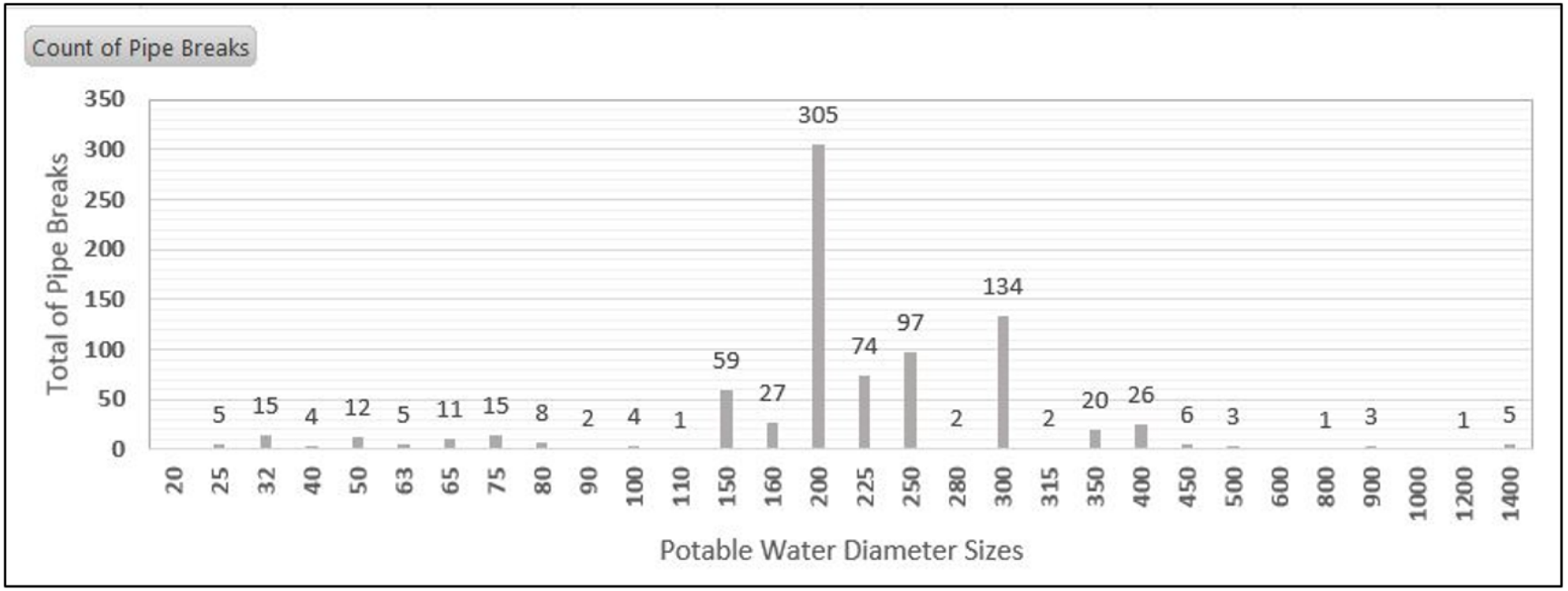


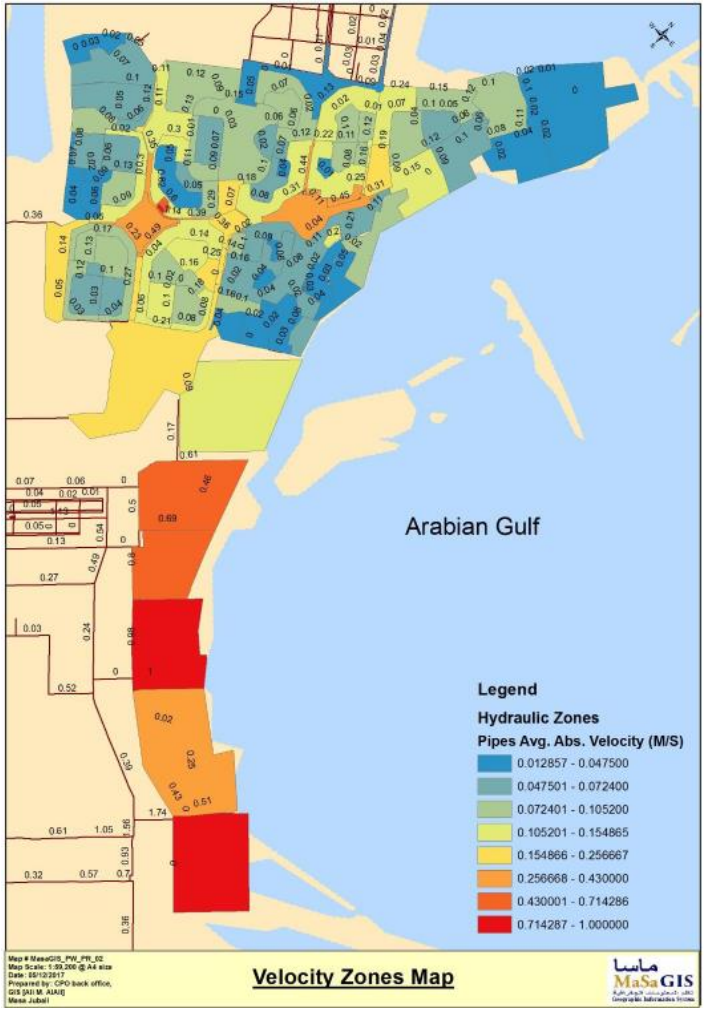
Figure 2: Total potable water network pipe breaks according to the diameter sizes of the pipes

Influence Threat Hydraulic Factors

- **Velocity**
  - Lack of complete hydraulic model
  - Velocity for the main lines only
  - 20% of the total network length (187.2 km).
  - 202 pipe break events (23.8%) occurred on these main lines

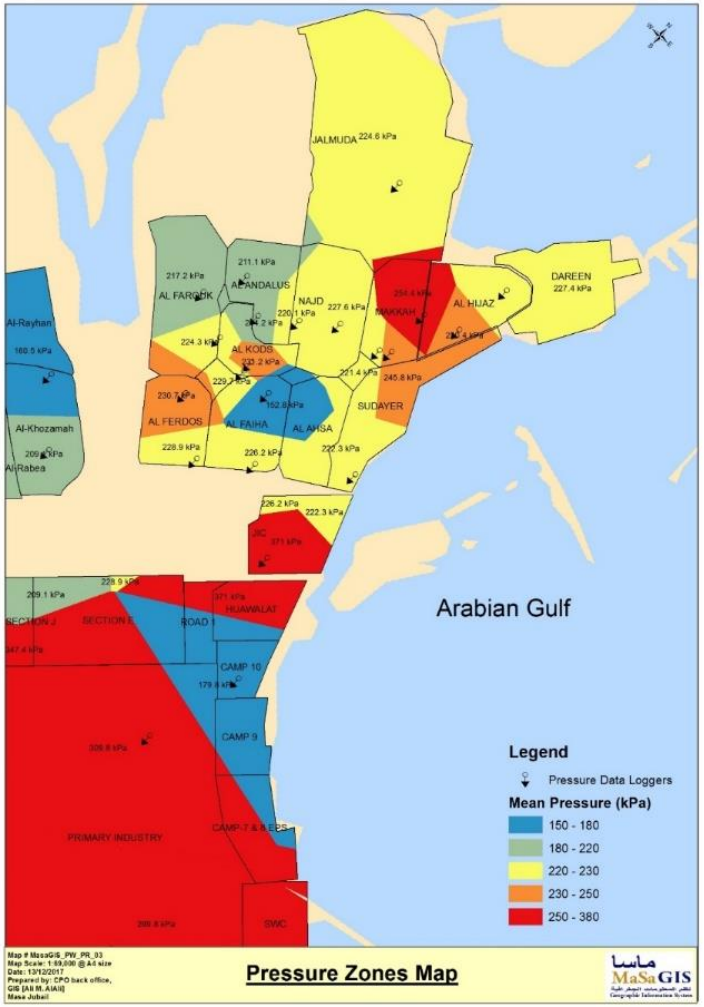
Table 4: Length and pipe breaks of main lines categorized by velocity level

Velocity Group	Length (km)	Total Pipe Breaks	Rate (PB/km)
Low Velocity (< 0.10 m/s)	116	135	1.16
High Velocity (=> 0.10 m/s)	71.8	67	0.93
Grand Total	**187.8**	**202**	***Avg = 1.07***



Influence Threat Hydraulic Factors

- **Pressure**



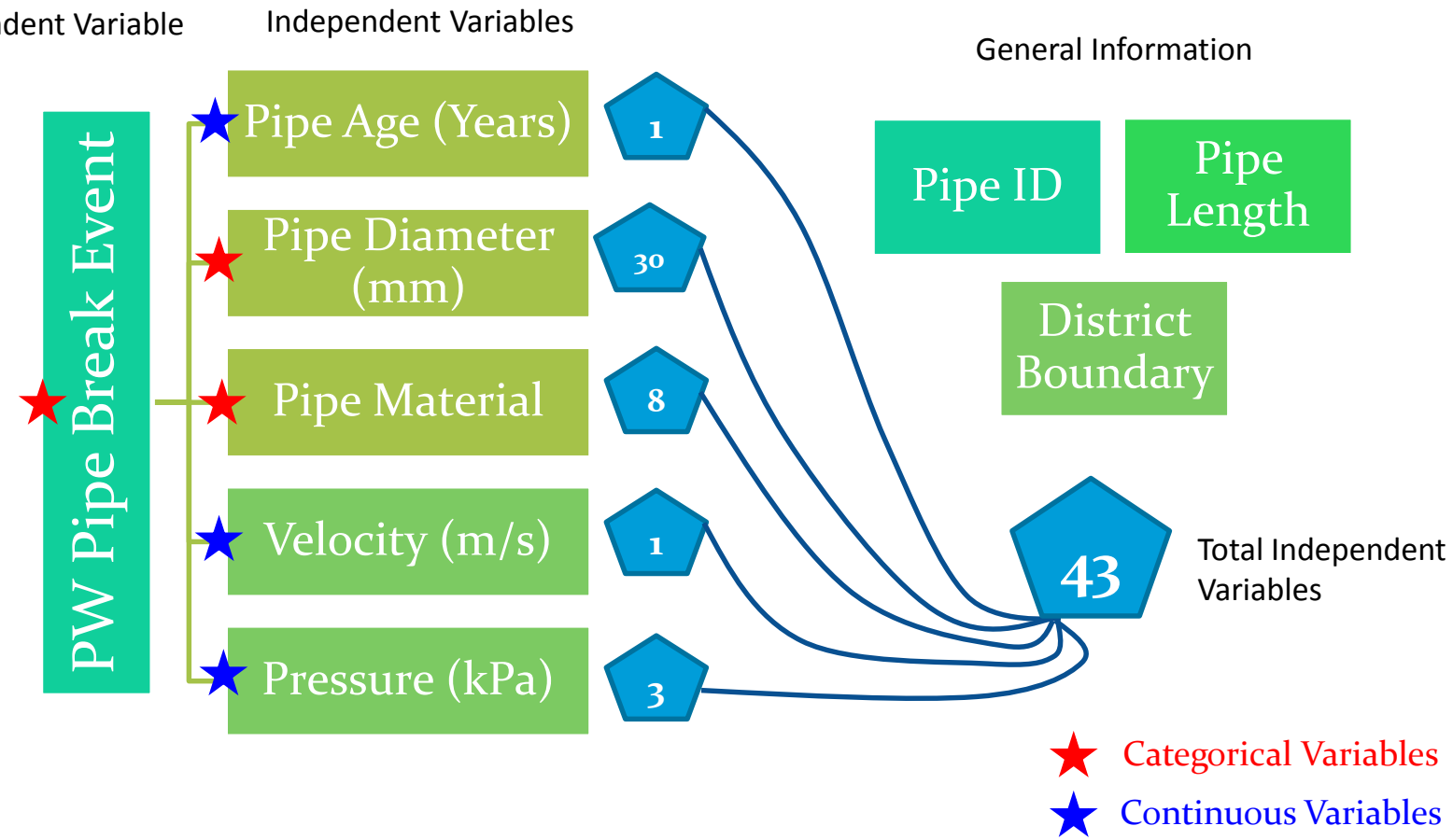
- Field data loggers = 23
- Processed logs = 6,543,563
- Thiessen polygons

Table 5: Network and Pump stations Data loggers used to measure Pressure (kPa) parameter

Seq.	Logger Location	Logger Type	P_Max	P_Min	P_Mean	Date_From	Date_To	No. of Logs
1	T-230, Near Petrokemya, FH-90	KPI Logger	393.90	188.06	299.84	01/10/2017	31/10/2017	89,280
2	T-154, WWPS-7, Near SAFCO	KPI Logger	400.36	182.07	309.84	01/10/2017	30/10/2017	89,280
3	Ferdaus, T-Ahzab, FH-32	KPI Logger	280.73	125.54	230.79	01/10/2017	31/10/2017	89,280
4	T-Dammam/Dammam 17, FH-16	KPI Logger	262.06	117.75	227.46	01/10/2017	31/10/2017	89,280
5	Makkah, T-Sarat/T-Zamzam, FH 1-1	KPI Logger	295.40	153.78	254.41	01/10/2017	31/10/2017	89,280
6	Sudayer, Hawiyah 2, FH-15	KPI Logger	264.75	128.54	222.37	01/10/2017	31/10/2017	89,280
7	RC Building, backside visiter building	Network Logger	483.05	-0.97	371.08	12/07/2017	09/11/2017	345,826
8	T-Dammam, near Dammam 26, FH-1/3	Network Logger	290.96	67.02	230.41	17/07/2017	22/11/2017	368,714
9	T-Andulus/Andulus 9, FH-33	Network Logger	273.03	-1.38	211.18	17/07/2017	09/10/2017	242,039
10	T-Ferdaus/Ferdaus 20, FH-5	Network Logger	301.16	-1.38	228.99	13/07/2017	21/11/2017	378,028
11	T-Faiha/T-Khamees, near Faiha 27, FH-201	Network Logger	290.68	-1.65	226.28	13/07/2017	27/11/2017	394,700
12	Camp 11, T-Huwaylat/T-Dairie, St. 46, FH-01	Network Logger	251.80	61.09	179.81	12/07/2017	09/10/2017	255,979
13	T-Najd/Najd 16, FH 1-8	Network Logger	349.15	75.43	227.60	17/07/2017	22/11/2017	368,662
14	T-Faiha/Faiha 7, FH-62	Network Logger	273.31	-8.27	152.87	13/07/2017	27/11/2017	394,642
15	Kods 8/T-Khalil, in front of fire station, FH-30	Network Logger	292.75	91.84	214.29	17/07/2017	21/11/2017	366,384
16	T-Andulus/Andulus 23, FH-43	Network Logger	294.27	-3.72	220.20	17/07/2017	09/10/2017	241,924
17	Farooq, T-Karamah/T-Batra, FH-327	Network Logger	279.24	-0.55	217.25	13/07/2017	21/11/2017	377,771
18	T-Farooq/T-Sedieg, FH-37	Network Logger	280.62	72.95	224.38	13/07/2017	21/11/2017	377,773
19	Fanateer PS, Discharge line A	PS Logger	441.61	48.61	221.47	18/07/2017	27/11/2017	380,257
20	Fanateer PS, Discharge line B	PS Logger	448.50	52.06	245.81	18/07/2017	27/11/2017	380,204
21	Deffi PS, Discharge line A	PS Logger	315.44	26.54	229.78	18/07/2017	27/11/2017	380,240
22	Deffi PS, Discharge line B	PS Logger	317.16	-2.07	231.25	18/07/2017	27/11/2017	380,191
23	Jalmudah, T-6, after EXTRA mall, FH	PS Logger / RTU	339.56	32.75	225.05	17/07/2017	03/09/2018	374,549

Data Sources Parameters

Requirements for the Logistic Regression Statistical Analysis

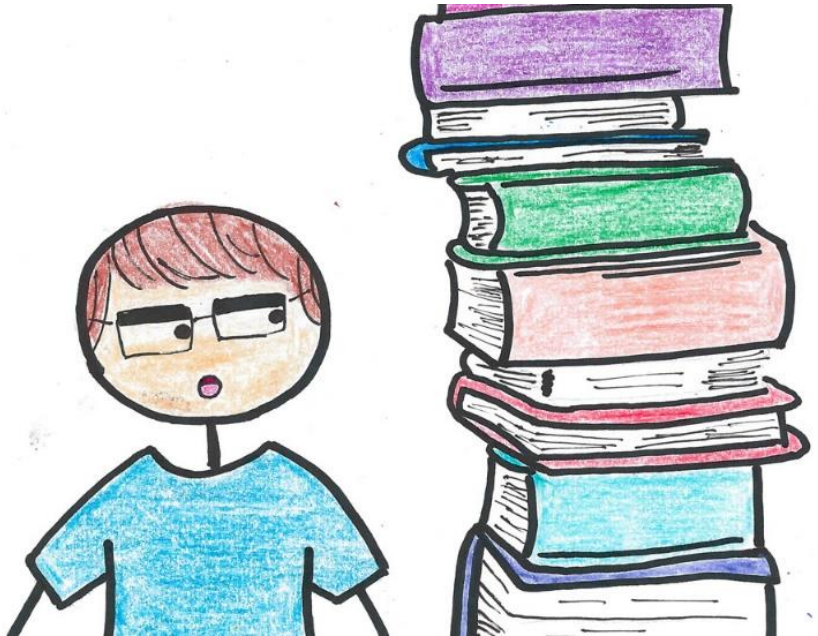


Research Question

- Failure predictions to reduce the predicted costs.
- Analysis of existing asset and failure data.
- Statistical logistic regression analysis.
- Get the prediction equation based on the explanatory variables.

- **The research question is:**

What is the impact of age, diameter, material, velocity and pressure on the probability of pipe breaks?



Statistical Analysis

- Using ANOVA and Logistic Regression to identify strength and significance of the predictors

Overall Likelihood index of Failure

= f (age, diameter, material, velocity, pressure)

- 43 Independent Variables
- Coefficients of 2 predictors could not be defined by the model due to singularities.

- **First Logistic Regression Analysis:**
  - Only 7 independent variables were statistically significant
  - **Negative Coefficients:** Small diameter pipes
  - **Positive Coefficients:** Age (years) and Pressure (mean)

- **Analysis of Variance (ANOVA):**
  - Compare each coefficient against full model sequentially to explore the possibility of an 'interaction effect' among levels of independent variables on the dependent variable.
  - ANOVA showed that the probability of seeing a difference in Resid. Dev "Pr(>Chi)" indicated possible improvement in the model fit upon adding some variables is greater than what is expected by chance alone.
  - Additional 15 significant independent variables (Diameters, Materials and Maximum Pressure).

Final Logistic Regression Analysis Result

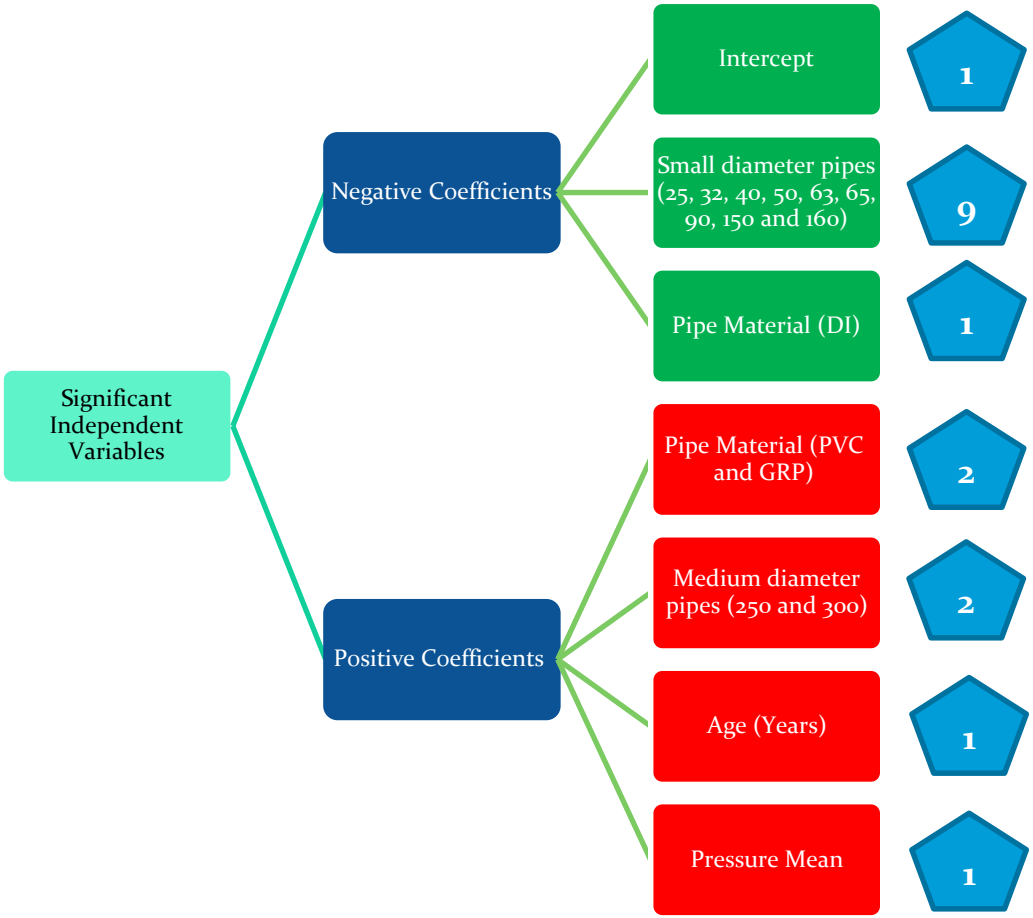
- Direct logistic regression was performed again to assess the impact of significant factors after performing ANOVA on the initial logistic regression model as these additional factors showed possible improvement in the model fit on the likelihood that pipe break will occur.

- The low p-value out of the final logistic regression model indicated that the model fit improved and the significant predictors increased from 7 to 16 independent variables which are statistically significant suggesting a strong association between them with the probability of pipe break event.

Table 6: Model result of fitting logistic regression analysis in R

```
Call:
glm(formula = PB_Count.f ~ Renewal.data$Age_Years + DIA_25.f +
    DIA_32.f + DIA_40.f + DIA_50.f + DIA_63.f + DIA_65.f + DIA_90.f +
    DIA_110.f + DIA_150.f + DIA_160.f + DIA_225.f + DIA_250.f +
    DIA_300.f + DIA_400.f + DIA_450.f + M_AC.f + M_PVC.f + M_DI.f +
    M_GRP.f + Renewal.data$P_Max + Renewal.data$P_Mean, family = binomial(link = "logit"),
    data = model1data)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.2615  -0.2056  -0.1398  -0.1026   3.8441

Coefficients:
                          Estimate Std. Error z value Pr(>|z|)
(Intercept)            -5.8679067  0.2118612 -27.697  < 2e-16 ***
Renewal.data$Age_Years  0.0894120  0.0045638  19.592  < 2e-16 ***
DIA_25.fDIA = 25       -2.9881975  0.4525545  -6.603 4.03e-11 ***
DIA_32.fDIA = 32       -2.4958960  0.2680098  -9.313  < 2e-16 ***
DIA_40.fDIA = 40       -3.2727430  0.4533895  -7.218 5.26e-13 ***
DIA_50.fDIA = 50       -2.5612037  0.2960059  -8.653  < 2e-16 ***
DIA_63.fDIA = 63       -2.1198955  0.4535722  -4.674 2.96e-06 ***
DIA_65.fDIA = 65       -1.6706617  0.2864155  -5.833 5.44e-09 ***
DIA_90.fDIA = 90       -1.8579021  0.7148429  -2.599 0.009349 **
DIA_110.fDIA = 110     -2.4019933  1.0069747  -2.385 0.017063 *
DIA_150.fDIA = 150     -2.0850377  0.1400334 -14.890  < 2e-16 ***
DIA_160.fDIA = 160     -0.7561222  0.2006584  -3.768 0.000164 ***
DIA_225.fDIA = 225     -0.1349843  0.1239883  -1.089 0.276292
DIA_250.fDIA = 250      0.3030320  0.1144329   2.648 0.008094 **
DIA_300.fDIA = 300      0.7815801  0.0978849   7.985 1.41e-15 ***
DIA_400.fDIA = 400     -0.2992226  0.2292412  -1.305 0.191800
DIA_450.fDIA = 450      0.9825621  0.4434521   2.216 0.026711 *
M_AC.fAC Material       0.4279170  0.2267588   1.887 0.059147 .
M_PVC.fPVC Material     0.4424478  0.0851588   5.196 2.04e-07 ***
M_DI.fDI Material      -1.1501682  0.2630056  -4.373 1.22e-05 ***
M_GRP.fGRP Material     1.0838486  0.2364036   4.585 4.55e-06 ***
Renewal.data$P_Max     -0.0011517  0.0008658  -1.330 0.183472
Renewal.data$P_Mean     0.0048819  0.0011928   4.093 4.26e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 9113.0  on 29863  degrees of freedom
Residual deviance: 7238.7  on 29841  degrees of freedom
AIC: 7284.7

Number of Fisher Scoring iterations: 9
```

Final Logistic Regression Analysis Result



Less likely to have pipe breaks

More vulnerable to pipe breaks

Note: AC material variable along with the other diameters, maximum pressure and velocity variables showed high p-values in the logistic regression model fitting results which indicate that all remaining variables are not statistically significant.

Calculate Probability of Failure Prediction

The equation of the final prediction model (Variable Pipe_Breaks) is:

Pred (Pipe_Breaks = 1) = exp(z) / [1 + exp(z)]

Where;

$z = b_0 + b_1x_1 + b_2x_2 + \ldots\ldots + b_nx_n$

b_0 = the intercept constant

b_n = the regression coefficient of the n variables

Then;

z = -5.8679067 + 0.0894120 X Age_Years + -2.9881975 X DIA_25 + -2.4958960 X DIA_32 + -3.2727430 X DIA_40 + -2.5612037 X DIA_50 + -2.1198955 X DIA_63 + -1.6706617 X DIA_65 + -1.8579021 X DIA_90 + -2.0850377 X DIA_150 + -0.7561222 X DIA_160 + 0.3030320 X DIA_250 + 0.7815801 X DIA_300 + 0.4424478 X M_PVC + -1.1501682 X M_DI + 1.0838486 X M_GRP + 0.0048819 X P_Mean

The final prediction model was tested on N = 837 pipes with previous real failure history where the mean of 0.047756 was used as decision boundary.

The results showed that **74.3% of the pipe breaks were predicted correctly**.

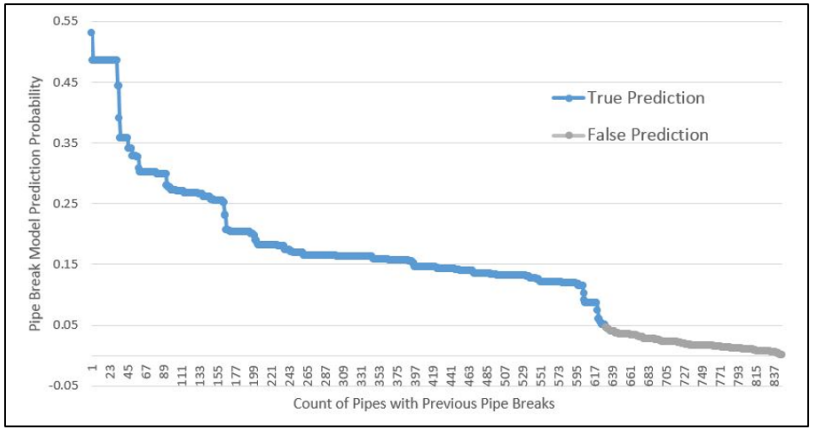
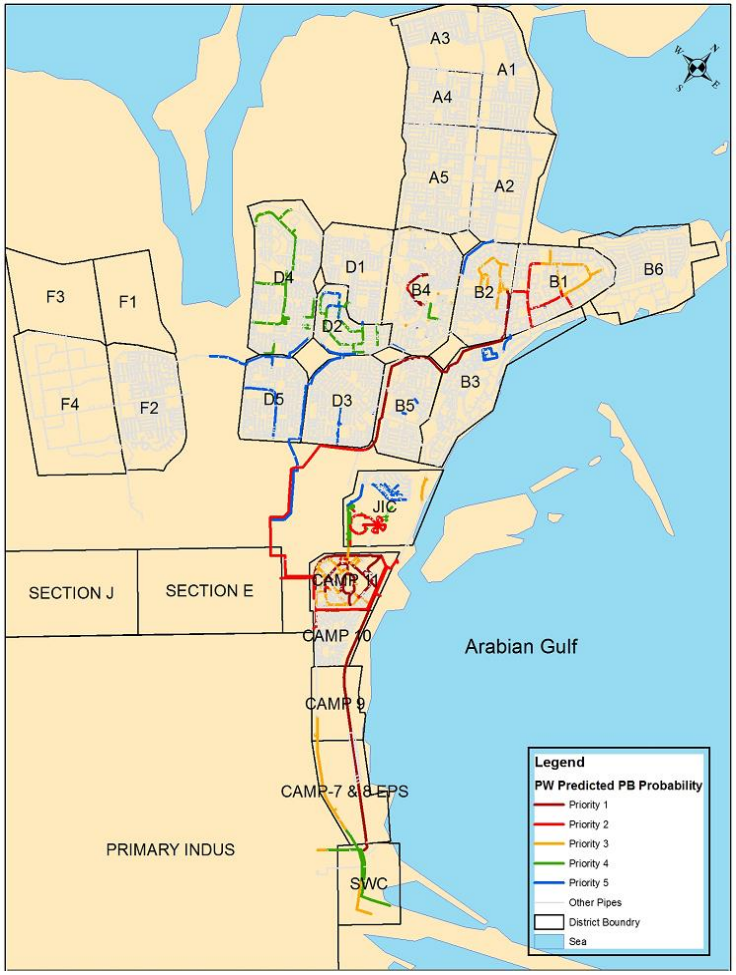


Figure 3: Graph of the model predicted probability (0 to 1) as result of logistic regression prediction equation tested on real sample.

Prioritization of Critical Pipes



Run the developed model equation on the 29,658 records of pipe segments to get the Pipe Break Probability.

Table 7: Priority levels for the annual critical pipes renewal plan

Priority Levels	Length (km)	Quantity of Pipes	Predicted PB Probability Range
Priority 1	17.28	110	0.358 to 0.530
Priority 2	15.69	150	0.302 to 0.358
Priority 3	14.65	305	0.267 to 0.302
Priority 4	15.68	259	0.251 to 0.267
Priority 5	15.46	206	0.204 to 0.251
Total	**78.76**	**1030**	**0 to 1 pipe break probability**

Probability prediction can be recalculated using simulation after changing/updating any parameters such as Age, Pressure, etc.

Conclusion

- Out of 43 independent variables, 16 predictors showed to have impact.
- Age, some diameter classes (250mm and 300mm), some material types (PVC and GRP) and the pressure mean showed positive correlation.
- Some variables showed tendency to decrease pipe breaks such as smaller diameter sizes and pipes made from DI material.
- The result of final model of this paper gave more detailed answer to the initial analysis of critical areas and provided higher resolution plan.
- The use of GIS tool as a master repository for all key analysis information was very useful and efficient.
- The study can be advanced in the future by improving some of the current parameters (such as velocity) and adding more explanatory variables.
- Additional parameters could include water temperature, ground water, improper bedding, low stiffness, corrosion issues, operating condition, roots from trees, leakage and water loss, history of water quality complaints and bad joining.

Any questions?

Acknowledgment

Co-Authors:

- Jean Laurent, TSD Manager (Masa)
- Jean Philippe Dulot, CPO Manager (Masa)

Other Supporters:

- Mahmoud Saleh Al-Theeb VP,O & M-J (Marafiq)
- Hamad Saad Al-Suhaim, EGM (Masa)
- Jean Noel Viot, GM Performance & Reliability (Masa)
- Saeed K. Al-Abdullah, Manager Corp. Comm. (Marafiq)
- GIS Team in CPO department (Masa)
- Potable Water Team (Masa)

ماسا MaSa مرافق MARAFIQ