Team ID	PNT2022TMID11378
Project Name	Efficient Water Quality Analysis and Prediction using Machine Learning

## **ANALYSING THE DATA:**

	CODE	LOCATIONS	STATE	Temp	D.O. (mg/l)	PH	(µmhos/cm)	B.O.D. (mg/l)	NITRATENAN N+ NITRITENANN (mg/l)	COLIFORM (MPN/100ml)	(MPN/100ml)Mean	
0	1393	DAMANGANGA AT D/S OF MADHUBAN, DAMAN	DAMAN & DIU	30.6	6.7	7.5	203	NAN	0.1	11	27	201
1	1399	ZUARI AT D/S OF PT. WHERE KUMBARJRIA CANAL JOI	GOA	29.8	5.7	7.2	189	2	0.2	4953	8391	201
2	1475	ZUARI AT PANCHAWADI	GOA	29.5	6.3	6.9	179	1.7	0.1	3243	5330	201
3	3181	RIVER ZUARI AT BORIM BRIDGE	GOA	29.7	5.8	6.9	64	3.8	0.5	5382	8443	201
4	3182	RIVER ZUARI AT MARCAIM JETTY	GOA	29.5	5.8	7.3	83	1.9	0.4	3428	5500	201
ata.d	lescribe(	()										
	y	ear										
count	1991.0000	000										
mean	2010.0381	172										
std	3.0573	333										
min	2003.0000	000										
25%	2008.0000	000										
50%	2011.0000	000										
75%	2013.0000	000										
max	2014.0000	000										
ata.i	info()											
angel ata d	Index: 19	991 entries, 0 to :	1990 :	Non-Nu	ll Cour	nt D	type					
		/1)										
		/TTV (umbos/sm)										
		(4) -1-2(44)		1991 n	on-null	l i	nt64					
	std min 25% max class: angel ata c class: angel ata c c class angel ata c c class angel ata c c class angel ata c c c class angel ata c c c c c c c c c c c c c c c c c c	2 1475 3 3181 4 3182 ata.describe(	1 1399 WHERE KUMBARJRIA CANAL JOI  2 1475 ZUARI AT PANCHAWADI  3 3181 RIVER ZUARI AT BORIM BRIDGE  4 3182 RIVER ZUARI AT MARCAIM JETTY  ata.describe()  year  count 1991.000000  mean 2010.038172  std 3.057333  min 2003.000000  25% 2008.000000  50% 2011.000000  75% 2013.000000  max 2014.000000  ata.info()  class 'pandas.core.frame.DataFlangeIndex: 1991 entries, 0 to: atac columns (total 12 columns)  # Column   0 STATION CODE  1 LOCATIONS 2 STATE 3 Temp 4 D.O. (mg/1) 5 PH 6 CONDUCTIVITY (µmhos/cm) 6 CONDUCTIVITY (µmhos/cm) 7 B.O.D. (mg/1) 8 NITRATENAN N + NITRITENANN 9 FECAL COLIFORM (MPN/100m1) 10 TOTAL COLIFORM (MPN/100m1)	1 1399 WHERE KUMBARJRIA GOA CANAL JOI 2 1475 ZUARIAT PANCHAWADI 3 3181 RIVER ZUARIAT BORIM BRIDGE 4 3182 RIVER ZUARIAT MARCAIM JETTY GOA 3 3182 RIVER ZUARIAT GOA 3 3182 RIVER ZUARIAT MARCAIM JETTY GOA 3 3182 RIVER ZUARIAT GOA 3 3182 RIVER ZUARIAT MARCAIM JETTY GOA 3 3182 RIVER ZUARIAT MARCAIM JETTY GOA 3 3182 RIVER ZUARIAT MARCAIM JETTY GOA 3 3182 RIVER ZUARIAT GOA 4 3182 RIVER ZUARIAT GOA 5 2013.000000 5 25% 2013.000000 5 2011.00000 5 2011.00000 5 2011.00000 5 2011.00000 5 2011.00000 5 2011.00000 5 2011.00000 5 2011.00000 5 2011.00000 5 2011.00000 5 2011.00000 5 2011.00000 5 2011.00000 5 2011.00000 5 2011.0	1 1399 WHERE KUMBARJRIA GOA 29.8 CANAL JOI  2 1475 ZUARIAT PANCHAWADI GOA 29.5  3 3181 RIVER ZUARIAT BORIM BRIDGE GOA 29.7  4 3182 RIVER ZUARIAT GOA 29.5  ata.describe()  year  count 1991.000000  mean 2010.038172  std 3.057333  min 2003.000000  25% 2008.000000  50% 2011.000000  max 2014.000000  max 2014.000000  ata.info()  class 'pandas.core.frame.DataFrame' > angeIndex: 1991 entries, 0 to 1990  ata columns (total 12 columns):  # Column Non-Nui  ata.columns (total 12 columns):  # Column Station (total 12 columns):  #	1 1399 WHERE KUMBARJRIA GOA 29.8 5.7 CANAL JOL  2 1475 ZUARIAT PANCHAWADI GOA 29.5 6.3  3 3181 RIVER ZUARIAT GOA 29.7 5.8  4 3182 RIVER ZUARIAT GOA 29.5 5.8  4 3182 RIVER ZUARIAT GOA 29.5 5.8  ata.describe()  year  count 1991.000000  mean 2010.038172  std 3.057333  min 2003.000000  25% 2008.000000  50% 2011.000000  75% 2013.000000  max 2014.000000  ata.info()  class 'pandas.core.frame.DataFrame' > angeIndex: 1991 entries, 0 to 1990  ata columns (total 12 columns):  # Column Non-Null CourageIndex: 1991 entries, 0 to 1990  ata columns (total 12 columns):  # Column Non-Null courageIndex: 1991 non-null 1991 non-	1 1399 WHERE KUMBARJRIA CANAL JOI  2 1475 ZUARIAT PANCHAWADI GOA 29.5 6.3 6.9  3 3181 RIVER ZUARIAT BORM BRIDGE GOA 29.7 5.8 6.9  4 3182 RIVER ZUARIAT GOA 29.5 5.8 7.3  ata.describe()  year  count 1991.000000  mean 2010.038172  std 3.057333  min 2003.000000  25% 2008.000000  50% 2011.000000  75% 2013.000000  max 2014.000000  ata.info()  class 'pandas.core.frame.DataFrame' > angeIndex: 1991 entries, 0 to 1990  ata columns (total 12 columns):  # Column Non-Null Count Described State Columns (total 12 columns):  # Column Non-Null Count Described State Columns (total 12 columns):  # Column Non-Null Count Described State Columns (total 12 columns):  # Column Non-Null count Described State Columns (total 12 columns):  # Column Non-Null count Described State (total 12 columns):  # Column Non-Null count Described State (total 12 columns):  # Column Non-Null count Described State (total 12 columns):  # Column Non-Null count Described State (total 12 columns):  # Column Non-Null count Described State (total 12 columns):  # Column Non-Null count Described State (total 12 columns):  # Column Non-Null count Described State (total 12 columns):  # Column Non-Null count Described State (total 12 columns):  # Column Non-Null count Described State (total 12 columns):  # Column Non-Null count Described State (total 12 columns):  # Column Non-Null count Described State (total 12 columns):  # Column Non-Null count Described State (total 12 columns):  # Column Non-Null count Described State (total 12 columns):  # Column Non-Null Count Described State (total 12 columns):  # Column Non-Null Count Described State (total 12 columns):  # Column Non-Null Count Described State (total 12 columns):  # Column Non-Null Count Described State (total 12 columns):  # Column Non-Null Count Described State (total 12 columns):  # Column Non-Null Count Described State (total 12 columns):  # Column Non-Null Count Described State (total 12 columns):  # Column Non-Null Count Described State (total 12 columns):  # Column Non-Null Count Described State	1 1399 WHERE KUMBARJRIA GOA 29.8 5.7 7.2 189  2 1475 ZUARIAT PANCHAWADI GOA 29.5 6.3 6.9 179  3 3181 RIVER ZUARIAT GOA 29.7 5.8 6.9 64  4 3182 RIVER ZUARIAT GOA 29.5 5.8 7.3 83  ata.describe()  year  count 1991.000000  mean 2010.038172  std 3.057333  min 2003.000000  25% 2008.000000  50% 2011.000000  max 2014.000000  ata.info()  class 'pandas.core.frame.DataFrame' > angeIndex: 1991 entries, 0 to 1990 ata columns (total 12 columns):  # Column Non-Null Count Dtype  ata columns (total 12 columns):  # Column Non-null object	1 1399 WHERE KUMBARJRIA GOA 29.8 5.7 7.2 189 2 2 1475 ZUARIAT PANCHAWADI GOA 29.5 6.3 6.9 179 1.7 3 3181 RIVER ZUARIAT GOA 29.7 5.8 6.9 64 3.8 4 3182 RIVER ZUARIAT GOA 29.5 5.8 7.3 83 1.9  ata.describe()  year count 1991.000000 mean 2010.038172 std 3.057333 min 2003.000000 25% 2008.000000 50% 2011.000000 max 2014.000000 max 2014.000000  ata.info()  class 'pandas.core.frame.DataFrame' > angeIndex: 1991 entries, 0 to 1990 lata columns (total 12 columns): # Column	1 1399 WHERE KUMBARJRIA GOA 29.8 5.7 7.2 189 2 0.2 CANAL JOI  2 1475 ZUARIAT PANCHAWADI GOA 29.5 6.3 6.9 179 1.7 0.1  3 3181 RIVER ZUARIAT BORM BRIDGE GOA 29.7 5.8 6.9 64 3.8 0.5  4 3182 RIVER ZUARIAT MARCAIM JETTY GOA 29.5 5.8 7.3 83 1.9 0.4  ata.describe()  year count 1991.000000 mean 2010.038172 std 3.057333 min 2003.000000 25% 2011.000000  ata.info()  class 'pandas.core.frame.DataFrame'> angeIndex: 1991 entries, 0 to 1990 leate a columns (total 12 columns):  # Column (total 12 columns):  # Fooloo(total 12 columns):  # Fooloo(total 12 columns):  # Fooloo(total 12 columns):  # Fooloo(total 12 columns):  # Panon-null object  1 p991 non-null object  1 p991 non-null object  2 STATE 1991 non-null object  3 Temp 1991 non-null object  3 Temp 1991 non-null object  3 Temp 1991 non-null object  4 D.O. (mg/1) 1991 non-null object  5 PH 1991 non-null object  1 p991 non-null object  1 p991 non-null object  2 STATE 1991 non-null object  3 Temp 1991 non-null object  4 D.O. (mg/1) 1991 non-null object  5 PH 1991 non-null object  6 CONDUCTIVITY (µmhos/cm) 1991 non-null object  9 FECAL COLIFORM (MPN/100ml) Mean 1991 non-null object  10 TOTAL COLIFORM (MPN/100ml) Mean 1991 non-null object  10 TOTAL COLIFORM (MPN/100ml) Mean 1991 non-null object  10 TOTAL COLIFORM (MPN/100ml) Mean 1991 non-null object	1 1399 WHERK KUMBARJRIA GOA 29.8 5.7 7.2 189 2 0.2 4953   2 1475	1 1399 WHEREKUNBARJRIA CANALOL