Water Physicochemical Analysis with Factor Analysis and Cluster Analysis of Ahmedabad city

Sweta Patel¹, Chetna Bhavsar²

{¹Research Scholar, ²Professor} Department of Statistics, School of Sciences, Gujarat University, Ahmedabad, Gujarat, INDIA.

Corresponding Addresses:

¹<u>sweta1611us@gmail.com</u>, ²<u>Chetna_bhavsar@yahoo.co.in</u>

Research Article

Abstract: Multivariate statistical techniques, cluster analysis (CA) and Factor analysis (FA) were applied to the data on water physicochemical parameters of Ahmedabad City of Gujarat in India. This study was carried out data of 2004, 2005, 2006, 2007 and 2010 years. This study evaluated and interpreted complex water quality data sets and apportioned of pollution sources to get better information about water quality and to design a monitoring network. Cluster Analysis and Factor Analysis was applied on the data and had some Clusters for physic chemical parameters and had some Factors for the same. Based on the study we can conclude that water quality assessment is a major aspect of human health. Government should keep track on that and gave pure water to public for drinking purpose. This study can help to improve our water quality analysis in future with the help of these clusters and factors.

Keywords: Cluster Analysis, Factor Analysis, Chlorinity, Salinity, Magnesium Hardness, Calcium Hardness.

1. Introduction

Without water, life cannot survive. Water and life are two sides of the same coin. Life initiates and grows in the lap of water. Water is very vital to all forms of life: from very small living creatures to very complex systems of animals and human being. The purity of water varies from place to place in nature. Rain water, if not contaminated by atmospheric pollutants, is highly pure while the sea water contains large amount of salt. Water for a variety of uses can be obtained from the sources like precipitation in the form of rain, snow and hail while surface water in the form of glaciers, streams, rivers and sea water. Besides these sources of water. there is also a natural rich source of water in the form of groundwater which is complementary to the surface water. Due to steady increase in the population urbanization, deforestation etc, the water resources have been adversely affected both qualitatively and qualitatively. Water pollution is one of the major problems in developing countries like India [1, 2, 3]. Improper policy is one of the most important factors that have caused severe environmental pollution and ecological degradation. Almost all developing countries are experiencing an increase of population, urbanization, depreciation etc. [3,4]. Pollution has become a major threat to existence of man on earth. Rapid industrialization, urbanization and human activities consequently cause water pollution which has brought a veritable water crisis [5-8]. Sirkantaswamy et al. [8] reported seasonal variation of drinking water

quality at Mysore, in Karnataka state. They found higher amount of chemical (total dissolved solids, alkalinity) and bacteriological parameters. They concluded that the drinking water quality varied from moderate contamination to larger extent of contamination. Kadam et al. [9] reported more than permissible limit of borewell drinking water in Ahmedpur area of Maharastra. Agnihotri and Singh [10] reported bacteriologically unfit quality of drinking water in Sagar city of Madhya pradesh. Papanna and Nagaraju [3] found 97% of the total water sample (60 bore wells) in Kollegal taluk of Karnataka within the desirable to permissible limit according to Bureau of Indian standards (BIS). Susiladevi et al. [11] studied ground water sample from 30 different sites like bore wells, tube wells, and hand pumps in and around Cuddalere town in Tamil Nadu state. They found that the water in some places were unfit for human consumption due to industrial waste disposal and sewage. In 2004, Suthar et al. [12] reported total hardness and calcium hardness within desirable limit while more than desirable limits of magnesium hardness, chlorinity and salinity were observed in some areas of Ahmedabad city. In 2005, Suthar et al. [13] reported higher amount of calcium hardness, magnesium hardness, chlorinity and alkaline nature of water in eastern part of Ahmedabad city. In 2006, Suthar et al. [14] found total hardness, magnesium hardness, calcium hardness, chlorinity and salinity either above the desirable limit or maximum allowable limits as per Gujarat Pollution Control Board (GPCB) standards in samples from Ahmedabad city. In 2007, Suthar et al. [15] observed higher values of calcium, magnesium, chlorinity and salinity above the desirable limits in Ahmedabad city. Suthar et al. [7] have recently found alterations in physico-chemical characteristics of drinking water collected from 17 areas of Ahmedabad city. In this scenario, to provide safe drinking water is a very big accountability for the governments.

2. Materials and Methods

Ahmedabad is the largest city in Gujarat state and sixth largest city (metro city) in India with a population of almost 5 million. It is located on the bank of Sabarmati River at an elevation of 55 meters (180 ft). It is located at 23.030 N and 72.580 E. It has a dry climate. Its

highest recorded temperature is 48° C and lowest is 15°C. The average rainfall is 932 mm. The present study is associated with water quality evaluated of Ahmedabad city of Gujarat state of years 2004, 2005, 2006, 2007 and 2010. The water analysis were collected and assessed by examining chemical characteristics by standard methods done by Prof. M. B. Suthar and his students of Department of Biology, K. K. Shah Jarodwala Maninagar Science College, Maninagar, Ahmedabad, India. In the year 2004, 2005, 2006, 2007 and 2010 the water samples were collected in the morning by students from their homes from the tap and labeled appropriately which later on brought to college laboratory. The drinking water quality was assessed by examining chemical characteristics. The parameters analyzed by standard methods were Total Hardness, Calcium Hardness, Magnesium Hardness, Chlorinity, Salinity and pH. They used total hardness tablets, calcium hardness tablets (both EDTA method) for total hardness and calcium hardness while Argentometric method for chlorinity respectively. Magnesium hardness and salinity were calculated from theses data. The pH was measured using Systronic pH meter 324 at 30°C. The data was compared with GPCB drinking water standards.

3. Statistical analysis

We have data of year 2004, 2005, 2006, 2007 and 2010. We run two Multivariate Techniques with SPSS 16.0 version software. Water quality data sets were subjected to univariate analysis: mean, maximum and minimum and multivariate analysis: Cluster analysis (CA) and Factor analysis (FA). These analyses required a preliminary step of the treatment of data which consisted of the normalization of the raw analytical data, so as to avoid misclassifications due to the different order of magnitude and range of variation of the analytical parameters (Aruga et al., 1995 Aruga, R., Gastaldi, D., Negro, G., Ostacoli, G., 1995. Pollution of a river basin and its evolution with time studied by multivariate statistical analysis. Analytica Chimica Acta 310, 15-25). Statistical computations were executed using the statistical software package, SPSS 16.0. The multivariate methods are summarized in the results and discussion.

3.1 Factor analysis (FA)

Factor analysis is a very powerful technique applied to reduce the dimensionality of a dataset consisting of a large number of interrelated variables, while retaining as much as possible the variability presented in dataset and with a minimum loss of information [J. F. Hair, Multivariate data analysis (3rd ed.). New York: Macmillan, (1992).]. This reduction is achieved by transforming the dataset into a new set of variables factors, which are orthogonal (non-correlated) and are arranged in decreasing order of importance. FA can also be used to generate hypotheses regarding causal mechanisms or to screen variables for subsequent analysis.

FA can be expressed as:

 $F_i = a_1 x_1 j + a_2 x_2 j + ... + a_m x_m$

- Where $F_i = factor$
- a = loading
- x = measured value of variable
- i = factor number
- i = sample number
- m = total number of variables
- There are three basic steps to factor analysis:

1. Computation of the correlation matrix for all variables.

2. Extraction of initial factors.

3. Rotation of the extracted factors to a terminal solution [Ho. Robert, Handbook of univariate and multivariate data analysis and interpretation with SPSS.

3.2 Cluster analysis (CA)

Cluster analysis is a major technique for classifying a mountain of information into manageable meaningful piles. It is a data reduction tool that creates subgroups that are more manageable than individual datum. In cluster analysis there is no prior knowledge about which elements belong to which clusters. The grouping or clusters are defined through an analysis of the data. Hierarchical CA, the most common approach, starts with each case in a separate cluster and joins the clusters together step by step until only one cluster remains [J. Lattin, D. Carroll and P. Green, Analyzing multivariate data. New York: Duxbury, (2003). J. McKenna, Environmental Modelling and Software, 18 (2003) 205.]. The Euclidean distance usually gives the similarity between two samples, and a distance can be represented by the difference between transformed values of the samples [M. Otto, Multivariate methods. In: R. Kellner, J. M. Mermet, M. Otto and H. M. Widmer, (Eds.), Analytical chemistry. Weinheim: Wiley-VCH. (1998).].

There are four basic cluster analysis steps:

1. Data collection and selection of the variables for analysis

- 2. Generation of a similarity matrix
- 3. Decision about number of clusters and interpretation
- 4. Validation of cluster solution

4. Results and Discussion

4.1 Data Analysis of Year 2004

Total 36 samples were collected and analyzed in the K. K. Shah Jarodwala Maninagar Science College, Ahmedabad laboratory. The sample source has no significant effect on these parameters as shown in Table 1. All the water samples were colourless, odourless and without any pleasant taste.

No.	Sample	No. of Samples	Total Hardness	Ca- Hardness	Mg- Hardness	Chlorinity	Salinity
1	Municipality	19	141.11 (28-304)	68.27 (8-212)	72.83 (20-156)	650.69 (127.8- 1491)	1175 (231- 2692)
2	Tubewell	17	148.31 (80-292)	69.88 (48-144)	78.57 (8-240)	533.59 (35.5- 1178.6)	964 (64- 2691)
	Drinking water standard (GPCB)	Desirable Limit- Maximum allowable limits	300-600	75-200	30-90	250-1000	450-1800

 Table 1: Parameters of water samples collected from different areas showing mean values of Municipality and Tube well samples of vear 2004

We have done factor analysis and Cluster Analysis on the above year 2004 data with SPSS 16.0 version and found some results below. This SPSS output lists the eigenvalues associated with each linear component (factor) before extraction, after extraction and after rotation. Before extraction, SPSS has identified 7 components within the data set (There should be as many eigenvectors as there are variables and so there will be as many factors as variables.) Factor 1 explains 47.261% of total variance. Before rotation, factor 1 accounted for considerably more variance than the remaining two (47.261% compared to 17.238% and 14.650%), however after extraction it accounts for only 44.664% of variance (compared to 18.616% and 15.869%) respectively in Table 2.

Figure 1 shows the Scree plot of the whole data set. From the Scree Plot we can directly visually say that the data have maximum four factors or components which consider the maximum amount of data and then it goes down.

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.308	47.261	47.261	3.308	47.261	47.261	3.126	44.664	44.664
2	1.207	17.238	64.499	1.207	17.238	64.499	1.303	18.616	63.280
3	1.025	14.650	79.149	1.025	14.650	79.149	1.111	15.869	79.149
4	.848	12.120	91.269						
5	.611	8.731	100.000						
6	9.392E-6	.000	100.000						
7	1.499E-6	2.141E-5	100.000						

Table 2: The table contains total Variance explained data which had Extraction Method was Principal Component of year 2004



Figure 1: This figure shows the Scree Plot of the data of year 2004

Table 4 shows the Component Matrix contains all the parameters divided into three different components. This output shows the component matrix before rotation. Where extraction method was Principal Component Analysis. By default SPSS displays all loadings; however, we requested that all loadings less than 0.5 be suppressed in the output and so there are blank spaces for many of the loadings. The first component (PCA 1) has uniform loadings from all the variables. So, it shows that PCA 1 includes Total Hardness, Magnesium Hardness, Chlorinity and Salinity. Those four parameters become First Principal Component. PCA 2 includes only one parameter that is Source (Tube well or Municipality). PCA 3 includes two parameters Calcium hardness and station (from where we collect the sample).

	Component		
	1	2	3
Total Hardness	0.857		
Calcium Hardness		-0.696	0.629
Magnesium Hardness	0.809		
Chlorinity	0.897		
Salinity	0.897		
Source		0.550	
Station			0.518

Table 3: Component Matrix before Extraction Method of year 2004.

Table 4 shows the Rotated Component Matrix that shows the factor loadings for each variable for the factors. We can see that the variable "Total Hardness (TH)" falls into factor 1 as the loading is the biggest in that row (0.749) compared to other factors. Here again simplify the output by suppressing loadings that are less than 0.5 for easier interpretation. So, here we get 3 factors from this output.

Factor 1: Total Hardness, Magnesium Hardness, Chlorinity, Salinity

Factor 2: Calcium Hardness

Factor 3: Source, Station

	Co	Component		
	1	2	3	
Total Hardness	0.749	0.520		
Calcium Hardness		0.985		

0.913

0.914

0.709

0.646

Magnesium Hardness 0.855

Chlorinity Salinity

> Source Station

Table 4: Rotated Component Matrix of year 2004

After the Factor Analysis we will run Cluster Analysis on the same data set and we got the output below. Table 6 shows the Agglomeration Schedule. Displays the objects or clusters combined at each stage (second and third column) and the distances at which this merger takes place. For example, in the first stage, objects 4 (Salinity) and 7 (Chlorinity) are merged at a distance of 0.000. From here onward, the resulting cluster is labelled as indicated by the first object involved in this merger, which is object 1 (Total Hardness). The last column on the very right tells in which stage of the algorithm this cluster will appear next. In this case, this happens in the third step, where it is merged with object 3 (Magnesium Hardness) at a distance of 10.604. The resulting cluster is still labelled 1 (Total Hardness) and so on.

	Table 5. Aggiomeration Schedule of data of year 2004									
Store	Cluster Combined		Coofficients	Stage Cluster	Naut Staga					
Stage	Cluster 1	Cluster 2	Coefficients	Cluster 1	Cluster 2	Next Stage				
1	4	7	0.000	0	0	3				
2	1	3	10.604	0	0	3				
3	1	4	29.010	2	1	4				
4	1	2	34.554	3	0	5				
5	1	5	43.915	4	0	6				
6	1	6	66.793	5	0	0				

So, from the Multivariate Techniques (Factor Analysis and Cluster Analysis we found three new factors which includes majority of variance data and we found a clusters for our easy interpretation.

4.2 Data Analysis of Year 2005

In the year 2005 the study contains total 30 water samples were collected and analyzed in the laboratory of Biology Department, K. K. Shah Jarodwala Maninagar Science College Ahmedabad. All the water samples were colourless, odorless and devoid of any unpleasant taste. Table 6 shows the parameters of water samples mean values and minimum and maximum values are shown in parenthesis. Compared to GPCB drinking water standard, the Total

Hardness in most of the samples were either within desirable limits or permissible limit. The Calcium Hardness was above the desirable limit in most of the samples.

Table 6: Parameters of water samples shows from Municipality and Tube well mean value; minimum and maximum values are shown in parenthesis in the year 2005. Units of Measurements: Total Hardness (as CaCO₃) mg/l; Calcium (as Ca) mg/l; Magnesium (as Mg) mg/l; Chlorides (as Cl) mg/l; Salinity g/l

		0) (- , , ,	., 0			
No.	Sample	No. of Samples	Total Hardness	Ca- Hardness	Mg- Hardness	Chlorinity	Salinity	pН
1	Municipality	20	188 (116-312)	107.2 (60-204)	80.8 (16-200)	732.7 (35.5-2414)	732.76 (35.5- 2414)	8.15 (7.7- 8.6)
2	Tube well	10	207.2 (100-380)	96.2 (60-164)	111 (40-240)	823.5 (213.6- 1207)	823.5 (213.6- 1207)	8.04 (7.8- 8.4)
	Drinking Water Standard (GPCB)		300-600	75-200	30-90	250-1000	450-1800	6.5-8.5

We run Factor Analysis on the data of year 2005. And we get the below results. This SPSS output lists the eigenvalues associated with each linear component (factor) before extraction, after extraction and after rotation. So here factor 1 explains 39.548% of total variance. SPSS extracts all factors with eigenvalues greater than 1, which leaves us with 3 factors. Before rotation, factor 1 accounted for considerably more variance than the remaining three (39.548% compared to 20.003% and 18.566%), however after extraction it accounts for only 38.459% of variance (compared to 20.100% and 19.557%) respectively in Table 9.

Table 7: The table contains total Variance explained data which had Extraction Method was Principal Component of the year 2005

Component	Initial Eigenvalues		Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings			
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.164	39.548	39.548	3.164	39.548	39.548	3.077	38.459	38.459
2	1.600	20.003	59.551	1.600	20.003	59.551	1.608	20.100	58.559
3	1.485	18.566	78.116	1.485	18.566	78.116	1.565	19.557	78.116
4	0.790	9.871	87.988						
5	0.625	7.817	95.804						
6	0.323	4.040	99.845						
7	0.012	0.155	100.000						
8	2.382E-17	2.977E-16	100.000						

Figure 2 shows the Scree plot of the whole data set. From the Scree Plot we can directly visually say that the data have maximum four factors or components which consider the maximum amount of data and then it goes down.



Figure 2: This figure shows the Scree Plot of the data of the year 2005

Table 8 shows the Component Matrix contains all the parameters divided into three different components. This output shows the component matrix before rotation where extraction method was Principal Component Analysis. The first component (PCA 1) has uniform loadings from all the variables. So, it shows that PCA 1 includes Total Hardness, Magnesium Hardness, Chlorinity and Salinity. Those four parameters become First Principal Component. PCA 2 includes Calcium Hardness and pH. PCA 3 includes two parameters Source (Tube well or Municipality) and station (from where collect the sample in Ahmedabad city).

	Co	Component		
	1	2	3	
Station			0.583	
Source			0.793	
Total Hardness	0.877			
Calcium Hardness	0.604	-0.664		
Magnesium Hardness	0.591		0.539	
Chlorinity	0.901			
Salinity	0.906			
pН		0.865		

Table 8: Component Matrix before Extraction Method of the year 2005

Table 9 shows the Rotated Component Matrix that shows the factor loadings for each variable for the factors. Here again simplify the output by suppressing loadings that are less than 0.5 for easier interpretation. So, here we get 3 factors from this output.

Factor 1: Total Hardness, Calcium Hardness, Chlorinity, Salinity

Factor 2: Station, pH

Factor 3: Magnesium Hardness, Source

Table 9:	Rotated	Component	Matrix	of yea	r 2005
----------	---------	-----------	--------	--------	--------

	Co	ompone	ent
	1	2	3
Station		-0.589	
Source			0.797
Total Hardness	0.829		
Calcium Hardness	0.728	-0.539	
Magnesium Hardness			0.711
Chlorinity	0.906		
Salinity	0.912		

After the Factor Analysis we will run Cluster Analysis on the same data set and we got the output below.

Table 10 shows the Agglomeration Schedule. Displays the objects or clusters combined at each stage (second and third column) and the distances at which this merger takes place. Here in the first stage, objects 6 (Chlorinity) and 7 (Salinity) are merged at a distance of 0.815 from here onward; the resulting cluster is labelled as 0 indicated by the first object involved in this merger. The last column on the very right tells you in which stage of the algorithm this cluster will appear next.

	Table 10. Aggiomeration Schedule of data								
Store	Cluster Combined		Coofficients	Stage Cluster	Next Stage				
Stage	Cluster 1	Cluster 2	Coefficients	Cluster 1	Cluster 2	INEXT Stage			
1	6	7	0.815	0	0	4			
2	3	5	15.488	0	0	3			
3	3	4	21.960	2	0	4			
4	3	6	23.188	3	1	5			
5	2	3	39.881	0	4	6			
6	1	2	42.472	0	5	7			
7	1	8	43.175	6	0	0			

Table 1	0: Agg	lomeration	Schedul	le of data

So, from the Multivariate Techniques (Factor Analysis and Cluster Analysis we found three new factors which includes majority of variance data and we found a clusters for our easy interpretation.

4.3 Data Analysis of Year 2006

In the year 2006 total 13 samples were collected and analyzed in the K. K. Shah Jarodwala Maninagar Science Collage laboratory. Table 11 shows are wise analysis of different physicochemical parameters. All water samples were odorless, colourless and devoid of any unpleasant taste. Compared with drinking water standards (WHO, ICMR and BIS), the Total Hardness is present more than desirable limits in seven samples and more than maximum permissible limits in six samples.

No.	Area	No. of Samples	Total Hardness	Ca- Hardness	Mg- Hardness	Chlorinity	Salinity	pН
1	Amrojwadi	2	541	128	413	317	572	8.0
1	Amraiwaui	5	(312-692)	(112-136)	(200-556)	(178-518)	(321-935)	(7.8-8.4)
2	Chadasar	1	780	128	652	315	569	8.1
2	Gilodasai	1	(780)	(128)	(652)	(315)	(569)	(8.1)
3	Comtinur	1	780	128	652	355	641	8.3
5	Gomupu	1	(780)	(128)	(652)	(355)	(641)	(8.3)
4	Iconnur	1	484	180	304	325	587	8.0
4	Isanpu	1	(484)	(180)	(304)	(325)	(587)	(8.0)
5	Moninggor	2	684	114	570	259	469	8.4
5	Mailinagai	2	(280-2088)	(100-128)	(180-960)	(164-355)	(296-641)	(8.3-8.6)
6	Dainur	1	484	270	214	553	999	8.3
0	Kaipui	1	(484)	(270)	(214)	(553)	(999)	(8.3)
7	Shah a alam	1	448	76	372	20	36	8.2
/	Silaii-a-alaiii	1	(448)	(76)	(372)	(20)	(36)	(8.2)
0	Thakkarbana nagar	2	478	134	344	132	239	7.9
0	ПаккаЮара падаг	2	(408-548)	(88-180)	(228-460)	(20-245)	(36-442)	(7.6-8.3)
0	Votvo	1	688	180	508	369	666	8.1
9	v atva	1	(688)	(180)	(508)	(369)	(666)	(8.1)
	Total	12	585	142	444	282	510	8.1
	Total	15	(280-1088)	(76-270)	(180-960)	(20-553)	(36-999)	(7.6-8.6)
	WHO	HDL	200	75	50	200		7.0-8.5
	WIIO	MPL	600	200	150	600		6.5-9.5
	ICMP	HDL	300			200		7.5-8.5
	ICMIK	MPL	600			1000		6.5-9.2
	DIC	HDL	200	75		250		7.0-8.3
	D13	MPL	600	200		1000		8.5-9.0

 Table 11: Parameters of water samples collected from different areas of Ahmedabad city in year 2006. (Area wise Mean values and minimum and maximum values are shown in parenthesis.)

Units of Measurements: Total Hardness (as CaCO₃) mg/l; Calcium (as Ca) mg/l; Magnesium (as Mg) mg/l; Chlorides (as Cl) mg/l; Salinity g/l; Abbreviation: HDL- Highest Desirable Limit, MPL – Maximum Permissible Limit. We run Factor Analysis on the data of year 2006. And we get the below results. In Table 12 factor 1 explains 39.772% of total variance. Before rotation, factor 1 accounted for considerably more variance than the remaining three (39.772% compared to 31.017% and 12.917%), however after extraction it accounts for only 32.324% of variance (compared to 30.285% and 21.097%) respectively.

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.182	39.772	39.772	3.164	39.548	39.548	3.077	38.459	38.459
2	2.481	31.017	70.789	1.600	20.003	59.551	1.608	20.100	58.559
3	1.033	12.917	83.706	1.485	18.566	78.116	1.565	19.557	78.116
4	0.666	8.331	92.037						
5	0.445	5.561	97.598						
6	0.192	2.402	100.000						
7	3.988E-7	4.985E-6	100.000						
8	-2.443E-16	-3.054E-15	100.000						

Table 12: The table contains total Variance explained data of year 2006 which had Extraction Method was Principal Component

Figure 3 shows the Scree plot of the whole data set. From the Scree Plot we can directly visually say that the data have maximum four factors or components which consider the maximum amount of data and then it goes down.



Figure 3: This figure shows the Scree Plot of the data in the year of 2006

Table 13 shows the Component Matrix contains all the parameters divided into three different components. The first component (PCA 1) has uniform loadings from all the variables. So, it shows that PCA 1 includes station (from where collect the sample in Ahmedabad city), Calcium Hardness, Chlorinity and Salinity. Those four parameters become First Principal Component. PCA 2 includes Total Hardness, Magnesium Hardness, and pH. PCA 3 includes only one parameters Source (Tube well or Municipality).

	Component			
	1	2	3	
Station	0.670			
Source	0.517		0.637	
Total Hardness		0.796		
Calcium Hardness	0.834			
Magnesium Hardness	-0.615	0.702		
Chlorinity	0.720	0.598		
Salinity	0.720	0.598		
pH		0.682		

Table 13: Component Matrix before Extraction Method of the year 2	2006
---	------

Table 14 shows the Rotated Component Matrix that shows the factor loadings for each variable for the factors. Here again simplify the output by suppressing loadings that are less than 0.5 for easier interpretation. So, here we get 3 factors from this output. We can see from both the table that after rotation PCA becomes Factor and that will change the parameters.

Factor 1: Calcium Hardness, Chlorinity, Salinity

Factor 2: Total Hardness, Magnesium Hardness, pH

Factor 3: Station, Source

Table 14: Rotated	Component Matrix	x of year 2006
-------------------	------------------	----------------

	Co	mpone	ent
	1	2	3
Station			0.669
Source			0.833
Total Hardness		0.980	
Calcium Hardness	0.681		0.623
Magnesium Hardness		0.962	
Chlorinity	0.985		
Salinity	0.985		

From the Factor analysis we can convert the huge data in to small factors which includes maximum variance of the data. In this year 2006 we can say that from the Factor Analysis that if we want to convert the whole data in to small factors we can club the above factors. After the Factor Analysis we will run Cluster Analysis on the same data set and we got the output below. Table 15 shows the Agglomeration Schedule. Displays the objects or clusters combined at each stage (second and third column) and the distances at which this merger takes place. For example, in the first stage, objects 6 (Chlorinity) and 7 (Salinity) are merged at a distance of 0.000 from here onward; the resulting cluster is labelled as indicated by the first object involved in this merger. The last column on the very right tells you in which stage of the algorithm this cluster will appear next.

Stars	Cluster Combined Cluster 1 Cluster 2		Casffisiants	Stage Cluster	Novt Stago	
Stage			Coefficients	Cluster 1	Cluster 2	Next Stage
1	6	7	0.000	0	0	3
2	3	5	0.609	0	0	6
3	4	6	7.395	0	1	4
4	1	4	10.672	0	3	5
5	1	2	12.652	4	0	6
6	1	3	22.822	5	2	0

Table 15: Agglomeration Schedule of data of year 2006

So, from the Multivariate Techniques Factor Analysis and Cluster Analysis we found three new factors which includes majority of variance data and we found a clusters for our easy interpretation.

4.4 Data Analysis of Year 2007

In the year 2007 total 36 samples were collected and analyzed in the K. K. Shah Jarodwala Maninagar Science Collage laboratory. Table 16 shows sample wise list of physicochemical parameters. The data suggest that most of the samples have Total Hardness, Chlorinity and Salinity within the highest desirable limit of GPCB. Most of the samples have high amount of Calcium and Magnesium Hardness above the highest desirable limit but less than maximum permissible limit of GPCB standards. The Water Quality Index (WQI) showed that almost all the samples were having the index value more than 100 suggesting that drinking water is unsafe as per GPCB standards adopted.

No.	Sample	No. of Samples	Total Hardness	Ca- Hardness	Mg- Hardness	Chlorinity	Salinity
1 Maniaira lite		20	170.61	106.03	64.70	154.56	279.0
1	Municipanty	50	(100-233)	(48-152)	(28-146)	(56-312)	(101-563)
2	Tube well	6	242.36	124.40	118.0	236.78	427.51
2	Tube well	0	(172-408)	(84-168)	(64-240)	(56-540)	(101-974)
	Total	26	182.56	109.56	6.640	168.28	303.77
	rotar	30	(100-408)	(48-168)	(28-240)	(56-540)	(101-974)

Table 16: Sample source wise list of physicochemical parameters studied in the year 2007

We run Factor Analysis on the data of year 2007. This SPSS output lists the eigenvalues associated with each linear component (factor) before extraction, after extraction and after rotation. Before extraction, SPSS has identified 8 components within the data set in Table 17. Before rotation, factor 1 accounted for considerably more variance than the remaining two (56.872% compared to 15.697% and 12.946%), however after extraction it accounts for only 36.427% of variance (compared to 33.123% and 15.965%) respectively.

 Table 17: The table contains total Variance explained data which had Extraction Method was Principal Component of year 2007

Component	Initial Eigenvalues		Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings			
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4.550	56.872	56.872	4.550	56.872	56.872	2.914	36.427	36.427
2	1.256	15.697	72.568	1.256	15.697	72.568	2.650	33.123	69.550
3	1.036	12.946	85.514	1.036	12.946	85.514	1.277	15.965	85.514
4	0.622	7.777	93.291						
5	0.526	6.575	99.866						
6	0.011	0.134	100.000						
7	6.242E-7	7.802E-6	100.000						
8	4.513E-10	5.641E-9	100.000						

Figure 4 shows the Scree plot of the whole data set. From the Scree Plot we can directly visually say that the data have maximum four factors or components which consider the maximum amount of data and then it goes down.



Figure 4: This figure shows the Scree Plot of the data of year 2007

Table 18 shows the Component Matrix contains all the parameters divided into three different components. The first component (PCA 1) has uniform loadings from all the variables. So, it shows that PCA 1 includes Source (Tube well or Municipality), Calcium Hardness, Chlorinity and Salinity, Total Hardness, Magnesium Hardness and WQI. PCA 2 has no parameters before extraction method. PCA 3 includes only one parameters station (from where collect the sample in Ahmedabad city).

	Component			
	1	2	3	
Station			0.782	
Source	0.602			
Total Hardness	0.941			
Calcium Hardness	0.574	-0.543		
Magnesium Hardness	0.789	0.523		
Chlorinity	0.819			
Salinity	0.820			
Water Quality Index	0.927			

Table	18:	Component	Matrix	before	Extraction	Method
I GOIC	10.	component	111111111	001010	Linuction	11100100

Table 19 shows the Rotated Component Matrix that shows the factor loadings for each variable for the factors. Here again simplify the output by suppressing loadings that are less than 0.5 for easier interpretation. So, here we get 3 factors from this output. We can see from both the table that after rotation PCA becomes Factor and that will change the parameters. Here in second Factor we have Calcium Hardness, Chlorinity and Salinity as we have no parameters in the second PCA before rotation.

Factor 1: Source, Total Hardness, Magnesium Hardness, WQI

Factor 2: Calcium Hardness, Chlorinity, Salinity

Factor 3: Station

Table 19: Rotated Component Matrix of year 2007

	Component			
	1	2	3	
Station			0.868	
Source	0.665			
Total Hardness	0.685	0.548		
Calcium Hardness		0.718	0.571	
Magnesium Hardness	0.961			
Chlorinity		0.902		
Salinity		0.902		

From the Factor analysis we can convert the huge data in to small factors which includes maximum variance of the data. In this year 2007 we can say that from the Factor Analysis that if we want to convert the whole data in to small factors we can club the above factors. After the Factor Analysis we will run Cluster Analysis on the same data set and we got the output below. Table 20 shows the Agglomeration Schedule. Displays the objects or clusters combined at each stage (second and third column) and the distances at which this merger takes place. For example, in the first stage, objects 6 (Chlorinity) and 7 (Salinity) are merged at a distance of 0.000. From here onward, the resulting cluster is labelled as indicated by another object involved in this merger. The last column on the very right tells you in which stage of the algorithm this cluster will appear next.

Table 20. Aggiomeration Schedule of data of year 2007						
Store	Cluster C	Combined	Coofficients	Stage Cluster First Appears		
Stage	Cluster 1	Cluster 2	Coefficients	Cluster 1	Cluster 2	Next Stage
1	6	7	0.000	0	0	5
2	5	8	2.906	0	0	3
3	3	5	6.530	0	2	4
4	3	4	24.075	3	0	5
5	3	6	24.630	4	1	6
6	2	3	32.630	0	5	7

 Table 20: Agglomeration Schedule of data of year 2007

So, from the Multivariate Techniques Factor Analysis and Cluster Analysis we found three new factors which includes majority of variance data and we found a clusters for our easy interpretation.

4.5 Data Analysis of Year 2010

In the year of 2010 the study focuses on drinking water in some areas of Ahmedabad city. Table 21 shows a comparison of tube well water and municipal supplied water indicate that municipal supplied water is much better than tube well water.

minimum values in parenthesis								
	No. of	Total	Calcium	Magnesium	Chlorinity	Solinity	Electrical	
	Samples	Hardness	Hardness	Hardness	Chiofinity	Saminy	Conductivity	
Municipal	40	210.25	133.35	76.9	290.5	524.31	0.92	
		(108-600)	(60-320)	(20-280)	(144-1494)	(144-1494)	(0.25-2.3)	
Tuba wall	16	235.5	132.50	103.0	376.5	679.441	1.77	
Tube well	10	(188-392)	(64-272)	(20-264)	(100-760)	(180-1371)	(0.24-6.6)	
Total	56	217.4	133.11(60-	84.36	315.07	568.63	1.16	
		(88-600)	320)	(20-280)	(80-828)	(144-1494)	(0.4-6.6)	

Table 21: Sample source wise list of physicochemical studied parameters in the year 2010 shows mean values and maximum and

In the present study, samples from tube well have significantly higher amount of chlorinity, EC. They suggest possibilities of ground water pollution. It might be due to sewage or industrial sources as areas on eastern part of Ahmedabad city have industrial blocks. They might contribute to ground water pollution. Therefore, proper disposal of industrial waste with periodical monitoring of ground water is recommended.

We run Factor Analysis on the data of year 2010. This SPSS output lists the eigenvalues associated with each linear component (factor) before extraction, after extraction and after rotation. Before extraction, SPSS has identified 8 components within the data set in Table 22. So in Table 34 factor 1 explains 48.230% of total variance. It should be clear that the first few factors explain relatively large amount of variance (especially factor 1) whereas subsequent factors explain only small amount of variance. SPSS then extracts all factors with eigenvalues greater than 1, which leaves us with 2 factors. Before rotation, factor 1 accounted for considerably more variance than the remaining one (48.230% compared to 20.350%), however after extraction it accounts for only 36.448% of variance compared to 31.132%.

 Table 22: The table contains total Variance explained data which had Extraction Method was Principal Component of year 2010

Component]	Initial Eigenval	ues	Extrac	ction Sums of Sq	uared Loadings	Rotati	ion Sums of Squ	ared Loadings
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.376	48.230	48.230	3.376	48.230	48.230	2.551	36.448	36.448
2	1.424	20.350	68.580	1.424	20.350	68.580	2.249	32.132	68.580
3	.966	13.801	82.381						
4	.766	10.942	93.323						
5	.467	6.677	100.000						
6	2.329E-7	3.327E-6	100.000						
7	-4.331E-16	-6.188E-15	100.000						

Figure 5 shows the Scree plot of the whole data set. From the Scree Plot we can directly visually say that the data have maximum four factors or components which consider the maximum amount of data and then it goes down.



Figure 5: This figure shows the Scree Plot of the data of year 2010

Table 23 shows the Component Matrix contains all the parameters divided into three different components. This output shows the component matrix before rotation where extraction method was Principal Component Analysis. The first component (PCA 1) includes Total Hardness, Magnesium Hardness, Chlorinity, Salinity and Electrical Conductivity. PCA 2 has no parameters before extraction method. Calcium Hardness have value less than 0.5 so it is showing blank space in the below table.

Sweta Patel, Chetna Bhavsar

	Component		
	1	2	
Total Hardness	0.617	-0.539	
Calcium Hardness			
Magnesium Hardness	0.874		
Chlorinity	0.874		
Salinity	0.766	0.592	
Electrical Conductivity	0.766	0.592	

1 abit 23. Component Maura before Extraction Method in the year 201	Table 23:	Component	Matrix before	Extraction	Method ir	the '	vear	2010
--	-----------	-----------	---------------	------------	-----------	-------	------	------

Table 24 shows the Rotated Component Matrix that shows the factor loadings for each variable for the factors. Here again simplify the output by suppressing loadings that are less than 0.5 for easier interpretation. So, here we get 2 factors from this output. We can see from both the table that after rotation PCA becomes Factor and that will change the parameters. Here in second Factor we have Electrical Conductivity and Salinity as we have no parameters in the second PCA before rotation.

Factor 1: Total Hardness, Calcium Hardness, Magnesium Hardness, Chlorinity Factor 2: Salinity, Electrical Conductivity

	Component		
	1	2	
Total Hardness	0.819		
Calcium Hardness	0.608		
Magnesium Hardness	0.846		
Chlorinity	0.846		
Salinity		0.948	
Electrical Conductivity		.948	

|--|

From the Factor analysis we can convert the huge data in to small factors which includes maximum variance of the data. In this year 2010 we can say that from the Factor Analysis that if we want to convert the whole data in to small factors we can club the above factors.

After the Factor Analysis we will run Cluster Analysis on the same data set and we got the output below. Table 25 shows the Agglomeration Schedule for the year 2010. Displays the objects or clusters combined at each stage (second and third column) and the distances at which this merger takes place. For example, in the first stage, objects 6 (Salinity) and 7 (Electrical Conductivity) are merged at a distance of 0.000. From here onward, the resulting cluster is labelled as indicated by another object involved in this merger. The last column on the very right tells you in which stage of the algorithm this cluster will appear next.

			88			
Store	Cluster C	Combined	Coofficients	Stage Cluster	First Appears	Novt Store
Stage	Cluster 1	Cluster 2	Coefficients	Cluster 1	Cluster 2	Next Stage
1	6	7	0.000	0	0	4
2	4	5	0.000	0	0	3
3	2	4	50.221	0	2	4
4	2	6	60.512	3	1	5
5	2	3	62.444	4	0	6
6	1	2	82.663	0	5	0

Table 25: A	Agglomeration	Schedule of	data of 2010
-------------	---------------	-------------	--------------

So, from the Multivariate Techniques Factor Analysis and Cluster Analysis we found two new factors which includes majority of variance data and we found a clusters for our easy interpretation.

5. Conclusion

Water is the most common and important resource on the earth. The hydrologic cycle is entirely adequate to meet human needs for fresh water, because it processes several times as much water as we required today. However the availability of water varies from place to place and time to time. As a result, there is a persistent scarcity of water in many parts of the world. Exponential growth in populations creates an everincreasing demand for additional water for irrigation, industry and municipal use. This five year study represents an attempt to evaluate the status of ground water of Ahmedabad city used for drinking purpose. Ground water is a precious natural resource. From the foregoing discussion, it is inferred that concentration of most parameters are generally within the highest permissible limit. This research work is attempted to assess the drinking water quality. Ground water is a precious natural resource. From the foregoing discussion, it is inferred that concentration of most parameters are generally within the highest permissible limit. The present study reveals that water is not safe for drinking in industrial area, only it is useful for domestic purpose. So, people should be made aware of the water quality importance on sanitation and economical water treatment methods like filtration and boiling would prove beneficial to avoid water born diseases and other water related disease. We can conclude that inadequate balance of all the physicochemical parameters leads to severe diseases like Osteoporosis, Nephrolithiasis (Kidney stones), Colorectal Cancer, Hypertension, Stroke, Coronary artery disease, Insulin resistance obesity Type II diabetes mellitus, metabolic syndrome etc. The remedial measure must matter most immediately to safeguard and conserve the precious water resources from pollution for future generation. This is a prime solution to pollution and future imminent water wars [16-22]. Based on the result of analysis, it is suggested that further investigations of water may be carried out in future. Public should be made aware of drinking water quality and careful management of precious natural resources. Government and non-government agencies should setup immediate and long term quality monitoring programs. Proper water treatment is necessary. There is need for continuing monitoring for the water quality especially for drinking and other domestic use. Government of India can maintain the limits of physicochemical parameters before supplying to citizens for the prevention of its ill effects on human.

6. Acknowledgements

Authors like to thank Dr. R.R. Shah (Principal), Management and departmental staff (Biology) of K. K. Shah Jarodwala Maninagar Science College, Ahmedabad for facility and encouragement. The sincere and voluntary help of S. Y. B. Sc (CZ) for collection and laboratory analysis of water samples is highly appreciated. The authors are especially thankful to Prof. M. B. Suthar for helping throughout the research.

References

- 1. Sharma, R.B. and Sharma, R.C.: Biochem. And Cell Arch., 10(2): 267-273 (2010).
- 2. Suthar, M. B. and Suthar, T. M.: Biosci. Guardian, 1 (1): 1-23 (2010).
- Papanna, C. and Nagaraju, D.: Asian J. Environ. Sci., 5(1): 11-13 (2010).
- Singh, R., Raghuvanshi, S. P. and Chandra, A.: Ind. J. Environ. Ecoplan., 17 (1-2): 39-44 (2010).
- Singh, D. and Joshi, B.D.: Indian J. Environ. Ecoplan., 17 (1-2): 89-92 (2010).

- Shrivastava, B. K. and Kumar, A.: Asian J. Exp. Chem., 4 (1-22): 90-91 (2009).
- Suthar, M.B., Ravat, N.M., Mesariya, A.R. and Kanjariya, K.V.: Nat. Environ. Poll. Tech., 9(2): 399-408 (2010).
- Srikantaswamy, S., Shakunthalata, Bai, Gholami, S. and Mahadev, J.: Asian J. Environ. Sci., 3 (2):104-110 (2008).
- Kadam, R. M., Reddy, N.J.M. and Nagpurne, V.S.: Asian J. Environ. Sci., 3(2): 189-190 (2008).
- Agnihotri, V. K. and Singh, P.K.: Asian J. Environ. Sci., 5(1): 8-10 (2010).
- Sushiladevi, M., Pugazhendy , K., Jayachandran, K. and Jayanthi, C.: Indian J. Environ. Ecoplan., 15 (3): 511-517 (2008).
- 12. Suthar, M.B., Mesariya, A.R. and Ravat, N.M.: Elect. J. Environ. Sci., 1: 23-27 (2008).
- Suthar, M.B., Mesariya, A.R., Ravat, N.M. and, Kalola, S.K.: Int. J. Bioscience Reporter, 6(1): 55-59 (2008).
- 14. Suthar, M.B., Mesariya, A.R. and Ravat, N.M.: Ind. J. Environ. and Ecoplan., 15(1): 171-176 (2008).
- 15. Suthar, M.B., Mesariya, A.R., Kanjariya, K.V. and Ravat, N.M.: Bulletin Environ. Sci., 26(1): 51-56 (2008).
- Khandwala, R.V. and Suthar, M.B.: Int. J. Biosci. Reporter, 5(1): 1-6 (2007).,
- Sushiladevi, M., Pugazhendy , K., Jayachandran, K. and Jayanthi, C.: Indian J. Environ. Ecoplan., 15 (3): 511-517 (2008).,
- 18. Navin , K.C., Twarakavi, and Kalurachchi, J.J.: J. Environ. Managt., 81: 405-419 (2006).
- Prabhavathi, M., Reddy, S.K. and Indirani, R.: Ind. J. Environ. Ecoplan., 17 (1-2): 343-350 (2010).
- Raghuvansi, S., Mishra, A.K. and Arya, M.: The Asian J. Exper. Chem., 5(1): 7-11 (2010).,
- W.A.I.: Drinking Water and Sanitation. Coverage, Financing & Emerging Concerns. Water Aid. India, New Delhi (2005).
- Shah, A.N., Ghariya, A. S., Puranik, A. D. and Suthar, M.B.: Elect. J. Environ. Sci., 1: 49-56 (2008).