3.1. Introduction

Water plays a major role in meeting the day-to-day needs worldwide. Based on gender, age, season and other factors the demand and consumption pattern varies.

Due to the non-controlled cities growth and scarcity of water resources, the water systems could not supply the required daily amount of water needs. Moreover, the region is under construction and that amplifies water demand.

This chapter presents the collected variables and their sources, besides to the techniques employed for collecting the data set. Before the survey conducted, individual discussions and field visits were held to understand the situation of water supply and use in the region.

Estimation of water use determinants requires a reliable measure of water consumption and information on the consumers and their houses. The study was conducted using a questionnaire that contains the purpose of the survey, followed by questions on water usage practices and behaviors. The survey questionnaire is built basing on previous studies and covers all relevant parameters. For better understanding of question by users, it is redacted in two languages and aims only single houses.

The present chapter is divided into statistical description of variables and data preparation. Any statistical analysis starts with standard data preparation techniques. Basic descriptive statistics are produced to note any missing/ abnormal values.

3.2. Description of the variables (Raw data)

Both quantitative and qualitative data are collected and in order to account factors that could explain variations in indoor water use, the selected sample considers many variables. Water consumption is "**the dependent"** variable, while socio-economic parameters, indoor habits of residents, physical characteristics of buildings units and climatic factors are "**the independent"** variables in this study (Figure 3.1). The difference between *dependent* and *independent* is that:

- **Dependent variable:** is variable that may depend on other factors. For example, household water use may change depending on weather factors.
- **Independent variable:** is a variable that does not depend on other factors. For example, weather factors like temperature does not change depending on household water use.

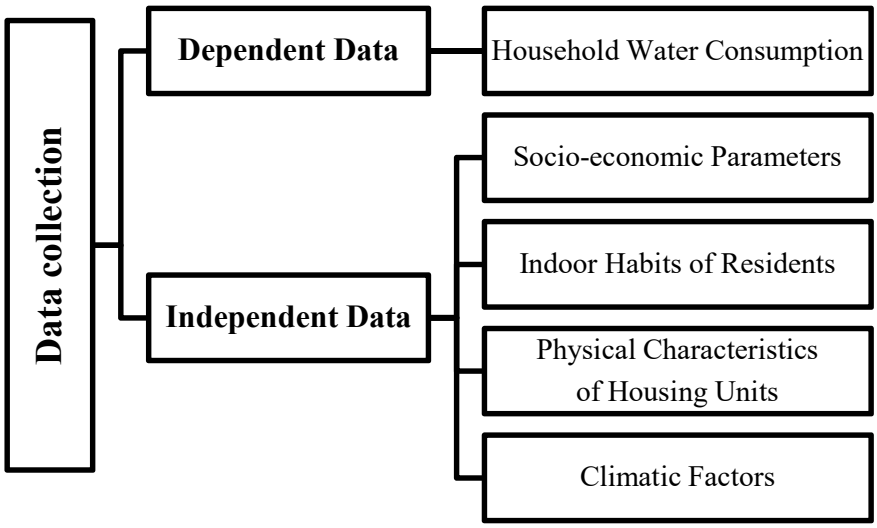

Fig 3.1 Type of collected data

The collected data were chosen based on previous studies as shown in the literature review chapter and they could be categorized into:

3.2.1. Water Consumption (WCP)

Household water consumption is obtained from **"Algérienne des eaux (ADE)"**, authority of water distribution and management, of Souk-Ahras. According to Sadoulet and De Janvry, 1995 the term "Household" varies widely across cultures. Their study defines household as "*the group of people living together and sharing the same kitchen or, in the case of piped water households, using the same piped water"*.

WCP data represents trimester values from 2012 to 2017 for 363 single houses selected from more than 500 houses in Sedrata city. The sample size of data is minimized because of the lack of socio-economic variables and physical characteristics of buildings information. WCP is measured by cubic meters.

3.2.2. Socio-economic Parameters (SEP), Indoor Habits of Residents (INH) and Physical Characteristics of Buildings Units (PHC)

To gather socio-economic parameters, indoor habits of people and physical characteristics of houses, the public inquiry represents the best tool.

People are selected randomly to get a representative sample on Sedrata from different types, where a questionnaire is distributed and filled during the period between March and April 2018 (to people living in the study area) based on a list created from preliminary literature review. The questionnaire paper contains questions on all relative parameters (Annex 01).

3.2.2.1. Socio-economic Parameters (SEP)

SEP should be considered in detail in any WCP investigation. In fact, this link between SEP and WCP is logical and confirmed by many authors (Chapter 01).

a. **Family Composition and Gender:** Family composition refers to total number of individuals (HOUS) beside the number of females (FEM) and males (MAL) living in the house. It is expected that a larger household size will increase the tendency to have a higher WCP because household need more water if there is more people. Thus, this variable is expected to have a positive sign to indicate that females are likely to have a higher WCP. This hypothesis is because women usually deal with domestic activities.
b. **Age of Family Members:** This means age of residents in each house. It is divided into four classes of age: under 8 years old (AG1), between 9 to 15 years old (AG2), between 15 to 35 years old (AG3) and older than 35 years old (AG4).
c. **Education Level of Residents:** Educational status of the residents is considered by classifying the level of education into four categories: primary school (PRS), medium school (MDS), high school (HGS) and university level (UNIV). It is expected that higher level of education would decrease WCP thanks to understanding water importance.
d. **Household Income (INC):** This variable is a combination of all incomes from all household members. It is measured in monetary unit, Algerian Dinar, DA/Month.
e. **Car Possession and Usage:** Car possession refers to the existence of cars, their numbers (CARN) and frequency of washing cars per month (WCAR).

3.2.2.2.Indoor Habits (INH)

This INH represents the individual's personnel and hygienic practices inside the housing units and includes frequency of using toilets per day (UTLT), frequency of washing clothes per week (WCL), frequency of washing dishes per day (WDISH) and frequency of showering for females and males per week (FSHW) and (MSHW), respectively.

3.2.2.3.Physical Characteristics of Building units (PHC)

Besides of residents related characteristics, the PHC has also a possible impact on WCP and describes the *building* and *the garden*.

a. **Building:** It refers to type of building, total area (TAR), building area (size) of the houses (BAR) and number of rooms (ROMN) in every habitation unit. Surface area is measured by squared meters.
b. **Garden Possession and Usage:** Garden possession signifies the existence of garden and their size (GAR), number of times watering the garden (GWAT). It is measured by squared meters.

3.2.3. Climatic Factors (CLF)

Climate is one of the most influential factors on water demand because it dictates consumption. Accordingly, the effects of season and weather must be accounted for any evaluation of water savings. Meteorological data may be directly factored WCP. Most commonly, mean temperature and mean precipitation data are used. Climatic data of Sedrata are obtained from both meteorological station of Souk Ahras and online databases.

a. **Mean Precipitation (PRE)** Precipitation could affect outdoor water use by reducing garden watering for example, and it is measured by millimeter (mm).
b. **Mean Temperature (TEM)** Temperature like precipitation could affect both indoor and outdoor household water use. This variable is measured by (°C).

Other information was collected from official documents, published reports, journals, books and related literature. The previous variables will explain the variation in domestic WCP and table 3.1 summarizes the collected data.

Table 3.1 Variables Description

Type		Variable	Acronym
Dependent		Household Water Consumption (m^3)	WCP
Independent	**Socio-economic indicators (SEP)**	**Family composition and gender**	
		Household size	HOUS
		Number of Female	FEM
		Number of Male	MAL
		Age of family member	
		Under 8 years old	AG1
		Between 9 to 15 years old	AG2
		Between 15 to 35 years old	AG3
		Older than 35 years old	AG4
		Education level of residents	
		Primary School	PRS
		Medium School	MDS
		High School	HGS
		University	UNIV
		Household Income (DA)	INC
		Car possession	
		Number of Cars	CARN
		Number of washing Car (month)	WCAR
	Indoor Habits (INH) Frequency	Washing Clothes (week)	WCL
		Washing Dish (day)	WDISH
		Using Toilets (day)	UTLT
		Female takes shower (week)	FSHW
		Male takes shower (week)	MSHW
	Physical Characteristics of Building (PHC)	**Building**	
		Total Area (m2)	TAR
		Building Area (m2)	BAR
		Number of Rooms	ROMN
		Garden possession	
		Garden Area	GAR
		Number of Garden Watering	GWAT
	Climatic Factors (CLF)	Mean Precipitation (mm)	PRE
		Mean Temperature (°C)	TEM

3.3. Preliminary Data Analysis

In quantitative research, when primary data is collected from surveys, a preliminary data analysis is a critical step required before the actual data analysis such as regression and parametric or non-parametric statistics.

After collecting data, a pre-processing of data (preliminary analysis) was prepared to eliminate outliers, incomplete dataset and to make sure that the subsequent analyses are all valid (Xue et al., 2017). Dataset was pre-processed in the following steps (procedures):

a. Dataset selection

The first step is the selection of the required data. In this work, both of water consumption values that obtained from the authority service and data collected from the questionnaire paper were first amassed together in one Excel sheet. The total number of selected houses was only **363 single houses,** because some residents of the study area were not available during the study visits.

b. Dataset examination

The main purpose of the second step is to check missing values and outliers, where it can happen by chance in any data distribution. Missing values are common in a large dataset, while outliers may indicate measurement errors dataset. In this research, values equal to zero in water consumption are considered as missing values, because it means no residents and no water consumed. Furthermore, housing units with low water consumption (WCP< 2 m^3) are also excluded from further analysis because such values are not realistic and may affect consequently the results. In addition, after evaluating the questionnaire papers the incomplete and incoherent answers are removed from the study.

c. Dataset summarizing

Accordingly, after the two previous steps, **162** houses were chosen and excluded from the first selected 363 houses. The final number of houses used for the study is **201** houses contained a dataset of **4824** valid water consumption values. Also, water consumption values are combined with household information. More details about household parameters are presented in next parts of this chapter.

3.4. Dataset Descriptive Statistics

All descriptive statistics were calculated by "**the Statistical Package of Social Sciences (SPSS 19)"** software and **"STATISTICA 8"** software. More details about SPSS and STATISTICA will be presented in the next chapters.

3.4.1. Water Consumption (WCP)

Description of statistical characteristics of annual water consumption are presented in table 3.2. Results from table 3.2 shows that mean value of water consumption varies between 115,55 m^3 and 139,03 m^3. For min and max values, water consumption varies from 11 m^3 to 36 m^3 and from 301 m^3 to 784 m^3 respectively.

Table 3.2 Descriptive statistics of annual water consumption (2012_2017)

Statistical Parameters	WCP 2012	WCP 2013	WCP 2014	WCP 2015	WCP 2016	WCP 2017
Mean	115,55	130,73	122,57	139,03	138,70	117,10
Median	112,00	125,00	117,00	127,00	126,00	107,00
Mode	144	144	144	120	80[a]	75
Std. Deviation	55,387	63,238	54,745	80,998	72,878	55,981
Variance	3067,715	3999,037	2996,995	6560,606	5311,146	3133,839
Skewness	1,208	2,333	0,648	3,234	1,390	0,983
Std. Error of Skewness	0,172	0,172	0,172	0,172	0,172	0,172
Kurtosis	2,679	12,673	0,245	20,233	2,719	0,838
Std. Error of Kurtosis	0,341	0,341	0,341	0,341	0,341	0,341
Minimum	30	36	11	36	11	30
Maximum	384	578	313	784	429	301
Sum	23226	26277	24637	27945	27880	23536

The analysis of the average per capita water consumption shows that the highest consumption is during the year 2015.

Also, it is evident that there is perceivable variation in WCP availability among the six years.

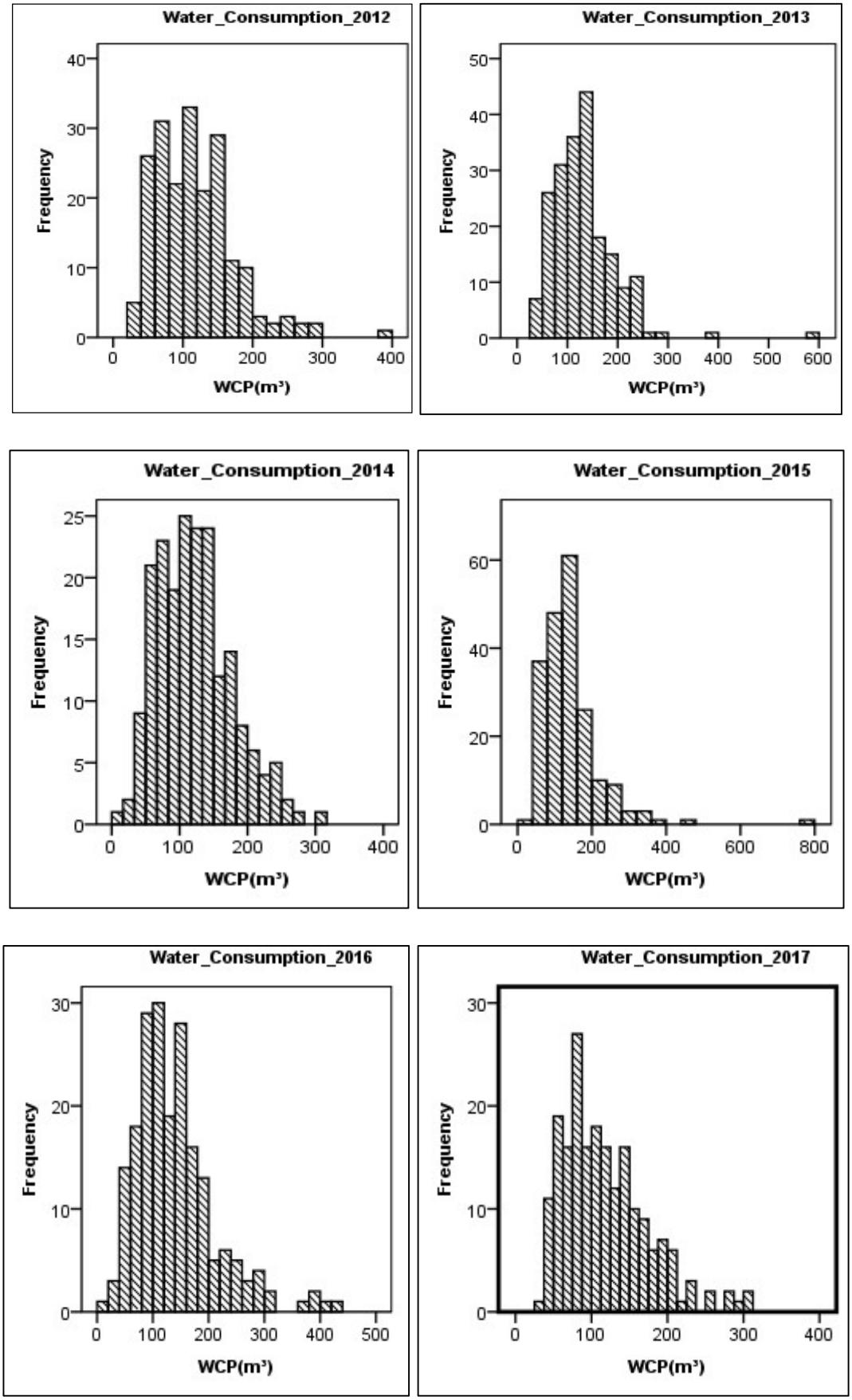

Fig 3.2 Water consumption distribution during the period 2012-2017

In addition to annual consumption, the mean trimester WCP and total WCP from 2012 to 2017 is presented in Table 3.3. Mean WCP values were obtained by calculating the mean value from all trimester WCP of every year, i.e. first value (Mean WCP 1st trimester) was obtained by: WC of 1st trimester in 2012 + WC of 1st trimester in 2013+ WC of 1st trimester in 2014+ WC of 1st trimester in 2015+ WC of 1st trimester in 2016+ WC of 1st trimester in 2017.

Table 3.3 Mean trimester and total water consumption (2012_2017)

Statistical Parameters	**Mean WCP 1st_Trimester**	**Mean WCP 2nd_Trimester**	**Mean WCP 3rd_Trimester**	**Mean WCP 4th_Trimester**
Trimester	**31 Jan to 31 Mar**	**31 Mar to 31 Jun**	**31 Jun to 31 Sep**	**31 Sep to 31 Dec**
Mean	171,78	167,51	204,50	219,89
Median	163,20	160,00	201,00	199,00
Mode	216	216	216	210
Std. Deviation	74,527	74,987	88,120	111,418
Variance	5554,222	5622,985	7765,150	12414,062
Skewness	0,897	1,016	1,331	1,953
Std. Error of Skewness	0,172	0,172	0,172	0,172
Kurtosis	1,012	2,008	3,369	7,910
Std. Error of Kurtosis	0,341	0,341	0,341	0,341
Minimum	38	42	67	41
Maximum	432	523	605	913
Sum	34528	33671	41105	44198

Examination of table 3.3 and figure 3.3 reveals that WCP vary through seasons, where WCP tend to increase with hot seasons.

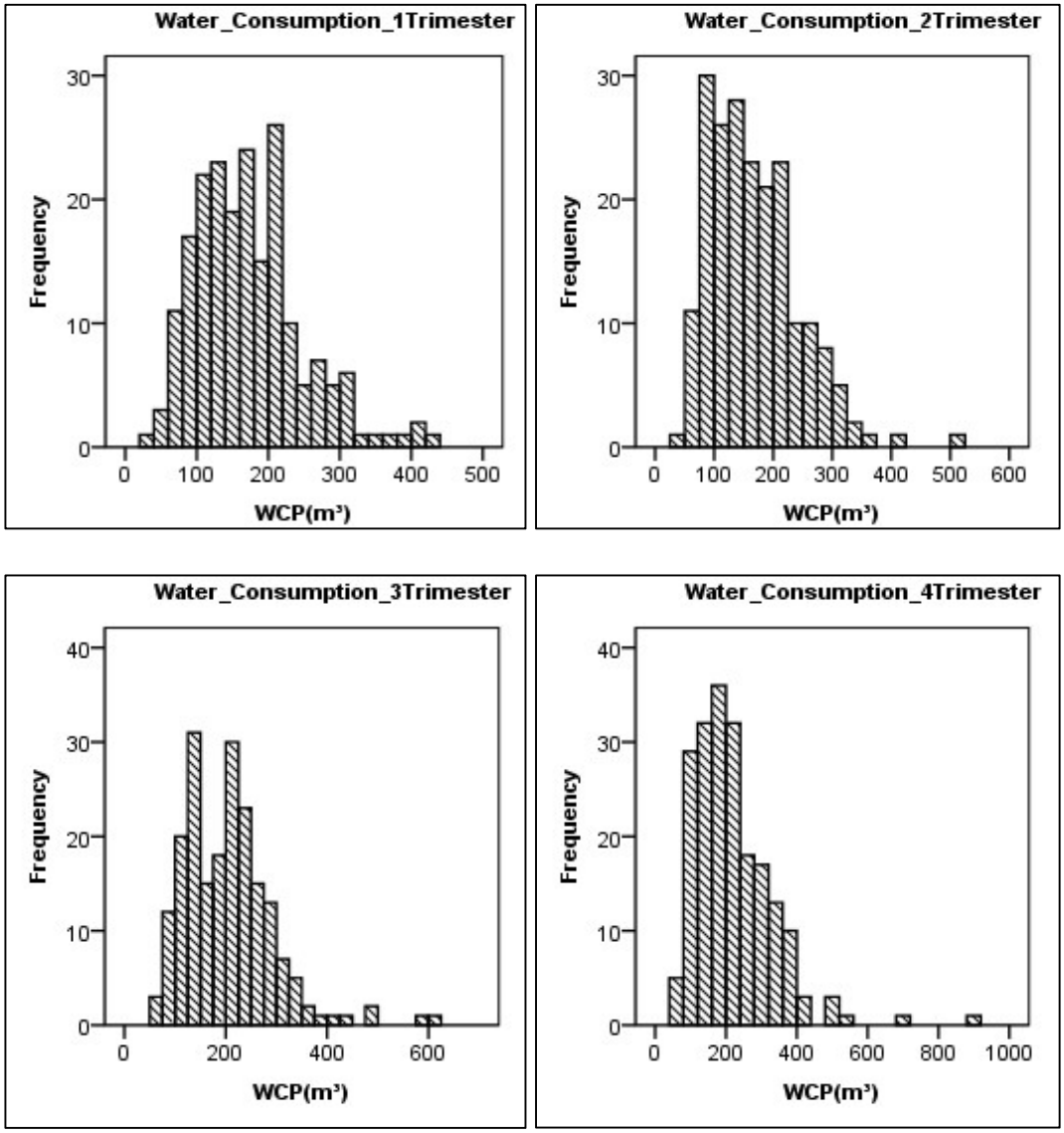

Fig 3.3 Mean trimester water consumption (2012_2017)

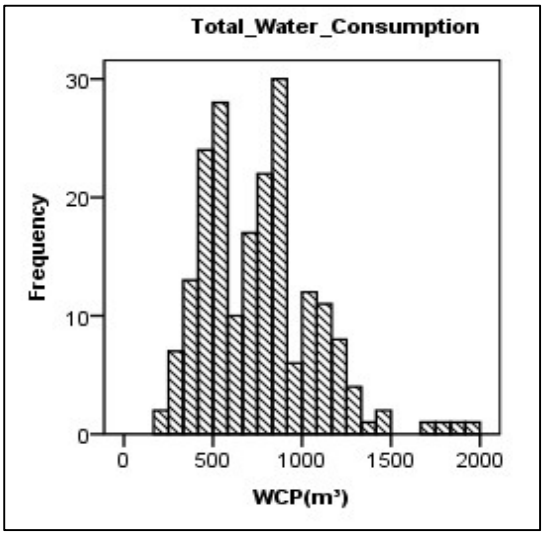

Fig 3.4 Mean total of water consumption (2012_2017)

3.4.2. Socio-economic Parameters (SEP)

Table 3.4 summary the descriptive statistics of the socio-economic parameters

Table 3.4 Descriptive statistics of socio-economic parameters

Statistical Parameters	FEM	MAL	AG1	AG2	AG3	AG4	PRS	MDS	HGS	UNIV	INC	CARN	WCAR
Mean	3,00	2,00	2,00	1 ,00	2,00	1,00	1,00	1,00	1,00	1,00	53905	2,00	2,00
Median	3,00	2,00	1,00	1,00	2,00	1,00	1,00	1,00	1,00	1,00	40000	1,00	1,00
Mode	3,00	2,00	1	0	1	1	0	1	1	1	40000	1	1
Std. Deviation	1,26	0,89	0,78	0,67	1,12	0,71	0,58	0,67	0,77	0,73	19976	0,69	1,19
Variance	1,58	0,79	0,60	0,45	1,25	0,50	0,34	0,45	0,60	0,53	3,99E8	0,47	1,41
Skewness	0,24	0,38	0,05	0,51	0,46	-0,32	0,73	0,22	0,81	0,09	1,09	-0,33	0,85
Std. Error of Skewness	0,17	0,17	0,17	0,17	0,17	0,17	0,17	0,17	0,17	0,17	0,17	0,17	0,17
Kurtosis	-0,51	-0,16	-0,47	-0,73	-1,21	-0,98	-0,4	-0,78	-0,86	-1,11	-0,13	-0,72	-0,05
Std. Error of Kurtosis	0,34	0,34	0,34	0,34	0,34	0,34	0,34	0,34	0,34	0,34	0,34	0,34	0,34
Minimum	1	0	0	0	1	0	0	0	1	0	$35^{E}3$	0	0
Maximum	6	5	**3**	**2**	4	2	2	2	3	2	$11^{E}4$	3	4
Sum	619	438	257	133	425	242	98	165	323	190	$10835^{E}3$	257	252

3.4.2.1.Household Size (HOUS) and Gender

In term of household size and composition, there is evidence to suggest that Household size is an important indicator for water consumption. People with larger families are expected to use more water.

The data shows that consumption patterns vary also according to gender distribution. Results from table 3.5 and figure 3.5 indicate that the sample consists of **619** (**58.56%**) females and **438** (**41.44%**) males.

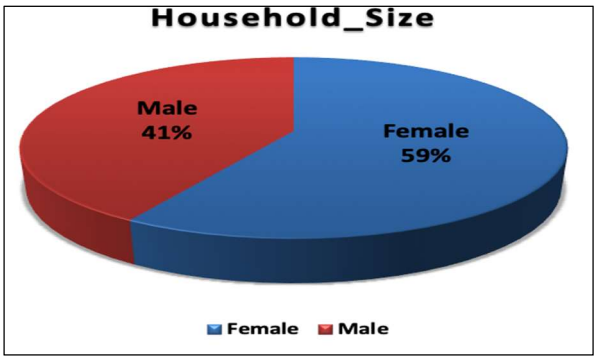

Fig 3.5 Gender in the study area

Table 3.5 Household size and gender

Gender	Frequency	Percentage
Female	619	58,56 %
Male	438	41,44 %
Total	**1057**	**100 %**

The majority of houses **(33.83%)** has six members, as shown in table 3.6 and figure 3.6. The minimum and the maximum household members ranges between 2 to 8, respectively. Generally, houses with family members more than seven persons having more than one floor and maybe are not single families (collective houses).

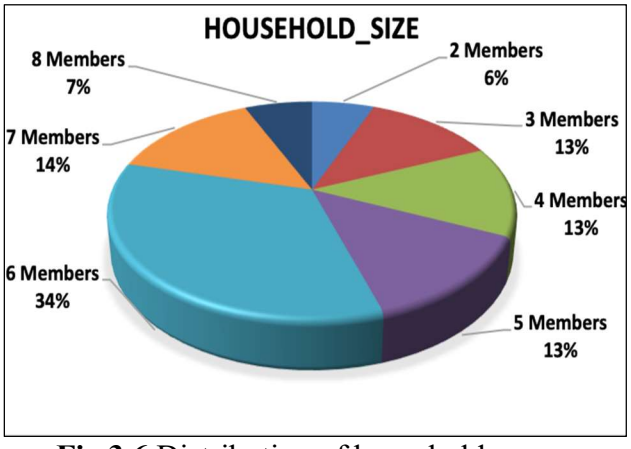

Fig 3.6 Distribution of households

Table 3.6 Distribution of residents (family size)

Family Size	Frequency	Percentage
2 Members	12	5.97 %
3 Members	25	12.44 %
4 Members	27	13.43 %
5 Members	27	13.43 %
6 Members	68	33.83 %
7 Members	29	14.43 %
8 Members	13	6.47 %
Total	**201**	**100**

Table 3.4 shows that minimum number of females (FEM) per house is 1, while the maximum number of females is 6. The distribution buildings according to females can be categorized into six main groups (figure 3.7):

- 20 building units with one female (9,95%)
- 52 building units with 2 females (25,87%)
- 53 building units with 3 females (26,37%)
- 51 building units with 4 females (25,37%)
- 18 building units with 5 females (8,96%)
- 7 building units with 6 females (3.48%)

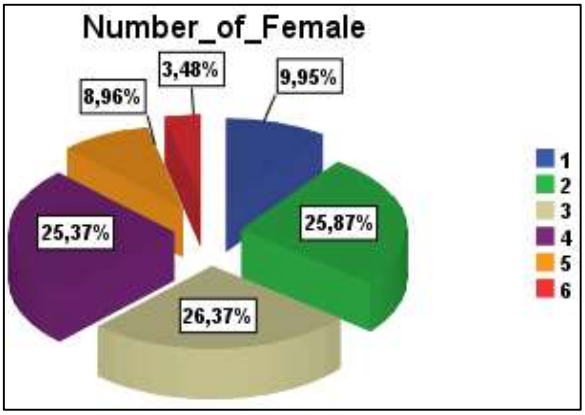

Fig 3.7 Number of female

Similarly to females, number of males (MAL) by house could affect water use in houses too. Examination of table 3.4 shows that minimum and maximum number of males in a house are 0 and 5, respectively.

The mean is equal to 2 males. Observation of Figure 3.8 gives the following classification of houses according to number of males:

- 1 house has no male (0,50 %)
- 44 building units have 1 male (21,89 %)
- 90 building units have 2 males (44,78 %)
- 51 building units have 3 males (25,37 %)
- 14 building units have 4 males (6,97 %)
- 1 building unit has 5 males (0,50 %)

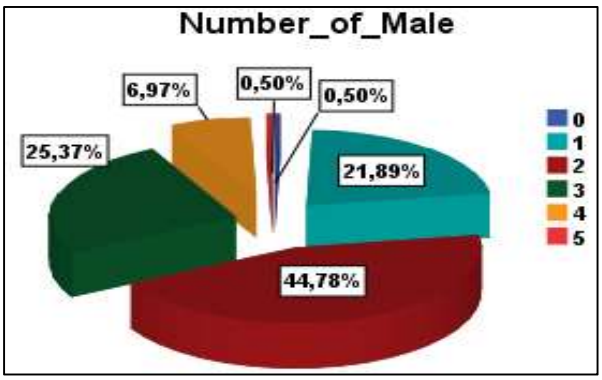

Fig 3.8 Number of males

3.4.2.2.Age of Family Members

Four main categories of age are obtained and their percentages are varied. Children category represents **24%** of the total sample (under 8 years old), with total number of **257 child**, for age category 9-15 years represents **13%** with total number equal to **133** and by this it represents the smallest age group. In addition, **40%** belong to residents having age between 15-35 years old (**425** persons) and it is the largest of the four groups. Finally, **23%** of the samples (with total number equal to **242**) belong to persons with 35 years and more. (Table 3.4 and Figure 3.9).

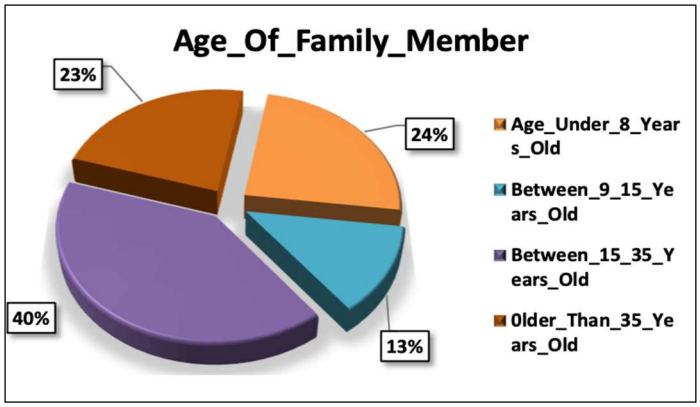

Fig 3.9 Repartition of residents age categories

a. Residents with Age under 8 years old (AG1)

Table 3.4 shows that the minimum, maximum and mean of AG1 that are 0; 3 and 2 respectively. Histogram in figure 3.10 demonstrates the distribution of children. The majority (92) of houses have a single kid in this age range.

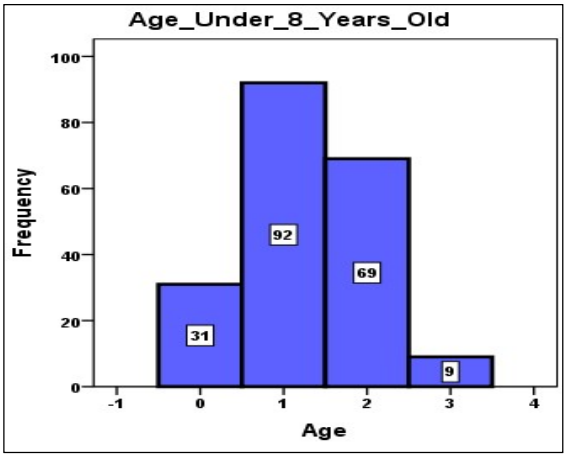

Fig 3.10 Distribution of residents with age under 8 years old

b. Residents with Age between 9-15 years old (AG2)

Table 3.4 shows the minimum, maximum and mean of AG2 that are 0; 2 and 1, respectively. Histogram in figure 3.11 illustrates the distribution of second category of residents. Results shows that most of people have one or none persons in this range of age.

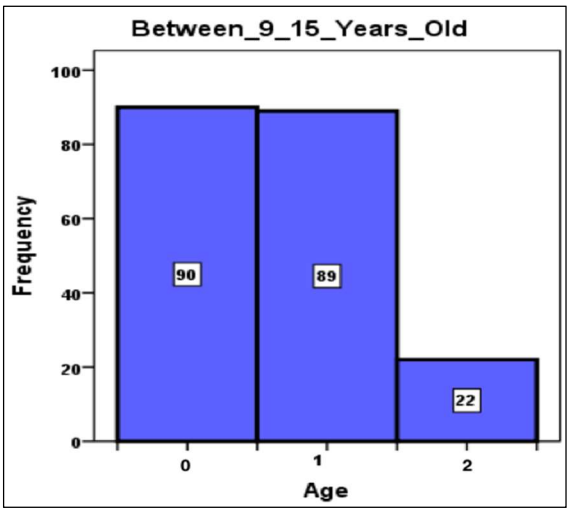

Fig 3.11 Distribution of residents with age between 9-15 years old

c. Residents with Age between 15-35 years old (AG3)

The minimum, maximum and mean of category 3 of age are 1; 4 and 2, respectively (table 3.4). Histogram in figure 3.12 below shows the distribution of third category of residents. Most of houses (83) have one person in this category.

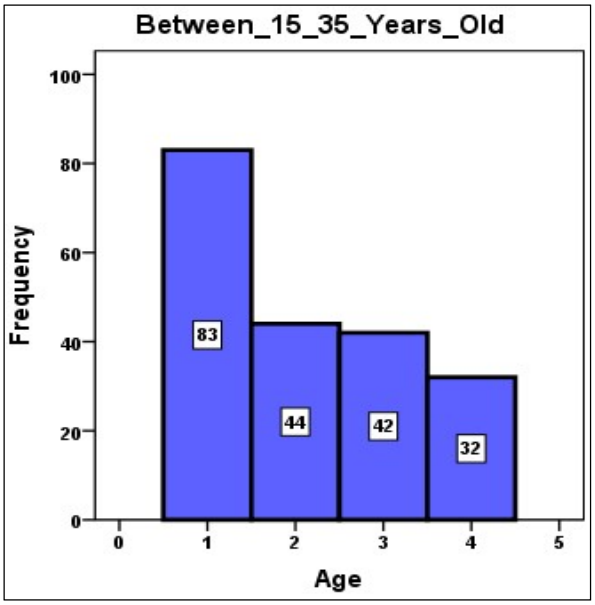

Fig 3.12 Distribution of residents with age between 15-35 years old

d. Residents with Age older than 35 years old (AG4)

Table 3.4 shows the minimum, maximum and mean of AG4 that are 0; 2 and 1, respectively. Histogram in figure 3.13 below shows the distribution of forth category of residents. The main reason behind this result is that the parents fall in this range.

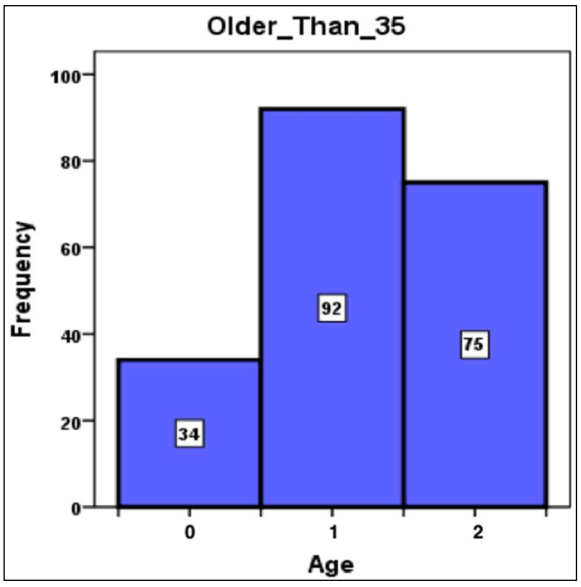

Fig 3.13 Distribution of residents with age more than 35 years old

3.4.2.3.Education Level of Residents

It should be notice that in Algeria , the study program passes by four main grades : primary school in 5 years, medium in four years , secondary (or high school) in three years and finally university program where the lowest title is obtained in three years (bachelor degree) and vary according to field of study and the degree attended. This study program is free and a typical person would pass all the cycle before acceptance in university in 12 years.

Education level affects all daily decisions and tasks, because educated or the literate persons are more aware about the consequences of shortage of water. The distribution of the sample based on educational status of the household is summarized in table 3.4 above and pie chart in figure 3.14 below.

Table 3.4 shows that people in the study zone has at least primary level and quarter of population have higher education (university). 87 % of residents have completed or in medium degree. And more than half have at least a high school level (66%). For university level, some residents are already graduated, others are still studying in the university and the rest of them are employed.

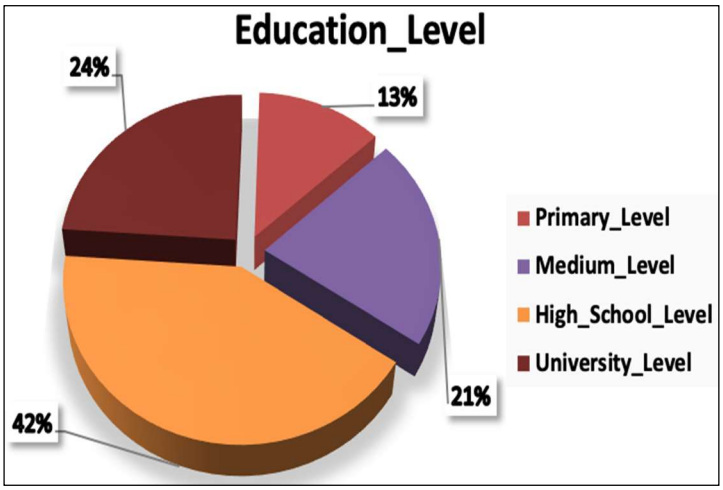

Fig 3.14 distribution of residents based on their educational level

a. Primary School (PRS)

Pie chart below demonstrates the repartition of habitation according to their primary level attendants (Figure 3.15).

More precisely 55, 7 % of houses do not have any pupils in primary school, 39,80% have one and only 4,5% have two in primary school.

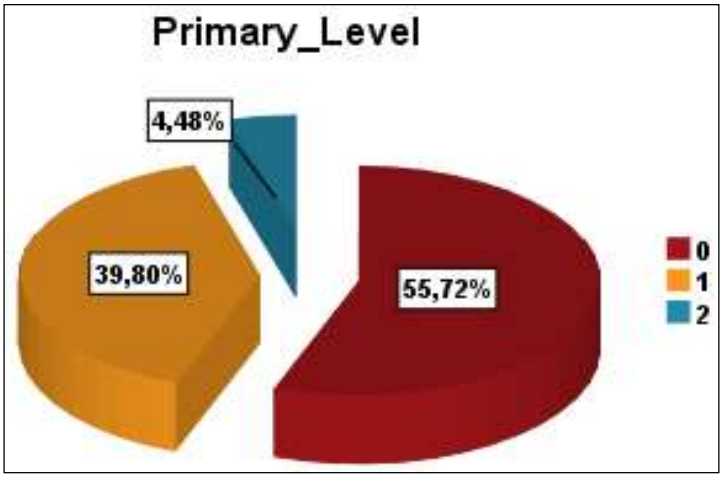

Fig 3.15 Repartition of number of habitations according to primary level