Descriptive Statistics in the Field Study of Bread Demand in Sulaimani

Hanaw Ahmed Amin

College of Science, University of Sulaimani, hanaw.ammin@univusl.edu.iq , Sulaimani, Sulaimani, Iraq
*Corresponding author email: hanaw.ameen@univsul.edu.iq ; mobile: 07703150649

الإحصاء الوصفي في دراسة ميدانية لطلب على الخبز في السليمانية

هناو احمد امين

العلوم، جامعة السليمانية، hanaw.ammin@univsul.edu.iq، السليمانية، العراق

ABSTRACT

Background:

As it is clear, bread is the main food of the majority of the plates of food component, in a way that affects the desires and desires of food and even its business and movements. For the same reason, we did a survey among civilized the people of Sulaimani, whether they eat and buy more bread and are acceptable to them (wheat or barley).

Materials and Methods:

Focusing on the basics and framework of kind the facts of descriptive statistics from some details on the quantitative and qualitative data in the applications. Frequency distributions and Charts are plotted of the results with SPSS program.

Results:

According to frequencies, at most 54% of males were aged between 20 and over from extended families prefer 1 - 2 pieces of wheat bread can have in their meals daily.

Conclusion:

The suitable practice of descriptive statistics allows economic ministry managers to additional efficiently balance and control of economics strategies and plans. Note that these result gives further information about economic future policy.

Key words: Descriptive Statistics, Frequency Distribution, Sample Survey, SPSS.

1. INTRODUCTION

Quantitative research offers main numerical info to economical- market demand choice-producers that allow them to achieve responsibilities such as budget explanation, and distribution **of** economic control demand incomes. however, economical-demand statistics are important to quality improvement and production progress. Numerous markets quantity their presentation products consuming the grades of statistical analysis, as well as achieve excellence expansion plans to recover their proficiency. Economic-control demand measures are also useful to provide an amount to the market for the daily demand of bread. [1]

Within the economic-market demand framework, as in other parts, there are two major methodologies:
(1) organizing initial sample data using descriptive statistics and (2) making decision about the results by inferential statistics. [2]

According to William (2006), descriptive statistics are used to present quantitative descriptions in a practical formula. [4] In this article all descriptive statistics were used to represent measure of deviation or measure of central limit theorem tend to relief considerate the sense of analyzed the data of the demand for bread above means of tables, mutual argument and graphs to clarify a simple instant around the sample data set containing their procedures. [5]

In the studying of the demand for bread we included all types of variables such as nominal, ordinal, interval as well as methodologies of frequency distributions, measure of central, dispersion, and position before making inferential statistics with certain level of confidence results. [4]

1.1 Types of Statistical Analyses:

Fundamentally, there are two kinds of analyses contains in statistical results. One is descriptive, another is inferential. In descriptive statistics, only defines the results of the collected data. Inferential statistics consist of approaches to simplify data outcomes to the associated populations with definite near of confidence and a certain of impact of results. It is needed to reflect some groups of basics while showing such statistical calculations, the concepts of descriptive statistics definitions will be the main study of this article [6].

1.2 Descriptive Statistic Analysis:

Descriptive analysis includes explanation of data in expressions of occurrences, quantities, mean, median, quartiles, standard deviation, inter-quartiles range. Measurement of these data be contingent on form of variables whichever to be qualitative or quantitative. Qualitative variables are categorical, categorized and attributes such as gender, socio-economic position, hurt score, behavior sets. In the same time, quantitative variables are computable, continuous and numerical e.g. age, height, weight, hurt level. [7]

1.3 Types of Variables:

Previously studying any group of data, all types of variables should be classified. In statistics term there are variate score of amount for recording variables, which categorized in two main categories: [3]

- Categorical data

- Continuous data

1.3.1 Categorical Data Analysis:

Qualitative or distinct variables are also identified as categorical measures. Nominal variables are the simplest form, take in two groups that have no basic rank (e.g., factors of purchasing bread spreads like price, taste, availability and advertising). But dichotomous variables have two categories (e.g., male or female). Ordinal data have two or more categories that can be ranked or ordered, but there are no real values to the ranks [2]. Such as the decision in purchasing bread spreads scale with "daily," "weekly," "monthly," and "yearly").

When the data is qualitative in terms then picture of data should be complete in positions of occurrences and percentages. Usually, both frequencies and percentages will be common though accessing the results. [5]

1.3.2 Continuous Variable Analysis:

Generally descriptive statistics is applicate to describe the approach of a sample data. It is used to classify a numerical measure. [5] These measures are also identified as either interval or ratio. Interval variables can be weighted in array owning a scale assessment, without zero point like the variability between a temperature of 100 degrees and 90 degrees is the same variability as between 90 degrees and 80 degrees). Ratio variables same to interval variables but has a clear definition of 0.0 point such as age value. Both measurements can be measured by frequency, central tendency, variation, and position [7] (e.g. age of participants in our study).

1. 4 Measurement Type:

To understand the detailed meaning of analyzed data, the data can be described in many different ways. Measurements of descriptive statistics can be measured either by measures of central tendency or dispersion, [4] by using tables, graphs or common explanation a terms are in [1, Table1]:

Table-1: **Measurement type**

Measure	What does it include	How it is use
Measures of frequency	Frequency, ratio, rates, proportion, percentage	means how frequently a value occur
Measures of central	mean, median, mode	Indicate the scattering of values.
Measures of variation	Range, variance, standard deviation	Identifies the range of values
Measures of position	Percentile ranks, quartile ranks	Describes where values reduction in relative individually to other

1.4. 1 Measures of Frequency Distribution

Total occurrence is the time of a specific number arises in the survey. Even though, relative frequency is related to the number of times a specific value occurs relative to other values. The relative frequency may be stated in ratios, rates, proportion or percentage [1, Table2]. [7]

Table-2: Measure of frequency

Measures	Usage
Ratios	compares frequency values between two values with another for the same variable. This formula frequently used in medical, pharmacological and biological studies.
Rate	one value for a variable in relative to the total sample of values in a certain period.
Proportion	the portion of entire sample that has some value.
Percentage	Meaning multiplying any proportion by 100%.

The procedures of distributions are frequently stated in the system of frequency table, graphically presented by histograms shapes. Bar charts are used in the case of qualitative variables to mark the statistics further simply to explain.

1.4.2 Measures of Central Limit Theorem:

Central limit theorem measurements are considered as the greatest useful explanation of the features of any population. It is value defines the whole group of data as a unique result. [4] The three usual common scales are the arithmetic, median and mode [1, Table3]. [7]

Table-3: **the most common pattern of central tendency**

Measures	Usage
Mean	defined as the sum of values sub sample size (n), it is frequently used which described with normal interval and ratio data and the disadvantages of mean is affected strongly by outliers.[4]
Median	is the middle of a ranked in ascending or descending order. In the case an odd number of observations, the median is exact middle value in the dataset, but an even number of values, the median is the average of the two middle values. This arrangement makes median not affected by outliers [7]. Finally median also is applicate to with ordinal scales.[9]
Mode	Meaning the value that occurs furthermost in the sample data. A set of data possibly will have single mode, more than one mode, or no mode at all. It can be applicate to qualitative and quantitative measures. [8]

1.4.3 Measure of Dispersion:

Measures of deviation define the amount of similarity between variable value. Simply used with ordinal, interval, and ratio data that can be ranked [5]

Range is difference between largest and smallest value in data set. [4] Variance and standard deviation computes the variation out of a set of values from their arithmetic mean as shown in [1, Table1]. It is considered by taking average values of the squared differences of individually value and their mean.

A large amount of standard deviation means that the data facts are extent at different arrays of values, while a small standard deviation means that the data points are close to the mean [6].

Consequently, in case of no variability variance and standard deviation are zero. [7] Generally, 68% of the values classified one standard deviation on any side of mean, 95% of values classified in two standard deviations, and 99% of values classified three standard deviations of mean. [9]

The curve of the data division possibly affects measures of mean, median and mode. Once the curve balanced "bell-shaped", measures of central are centered [1, Figure1b].

While the curve positively skewed close the lower tail, mode stays the greatest public value, and median is the middle value, but mean pulled to the right tail of the distribution as in [1, Figure1c]. While the distribution negatively skewed from top to bottom, mean is drawing near left of the distribution in [1, Figure1a]. [1]
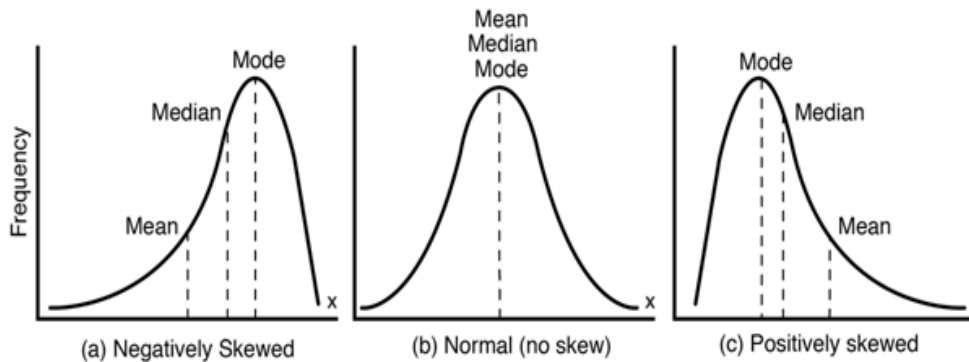
Figure (1): Three different shape of normal and skewed. (a) Negative skew

(b) Normal distribution. (c) Positive skew.

Outliers are extreme or unusual values, the mean is more sensitive to outliers than the median and mode. In presence of outliers, median is suitable practice for interval data. [7]

1.4.4 Measures of Position

A Measure of position defines the position of a value in relative to other values in a sample data set by three foremost techniques. Percentiles distribute values to 100 equal parts, deciles distribute values to ten equal parts, and quartiles distribute values to four equal quantities. Obviously, it can be known how they are related below: [2]

The minimal quartile, Q1 (25th percentile), is the location between the minimum 25% and maximum 75% of observations. The next quartile, Q2 (50th percentile), is the middle of the dataset. The higher quartile, Q3 (75th percentile), is the point between the lowest 75% and highest 25% of values. If the quartile located between two values, the average of those values indicates the quartile value. The interquartile range is the distance between the upper and lower quartiles and defines the median 50% of values when ranked from lowest to highest. Boxplots are commonly appropriate for explaining initial data graphically. [4] Boxplots are common and informal technique for detecting outliers [1, Figure2]. [7]
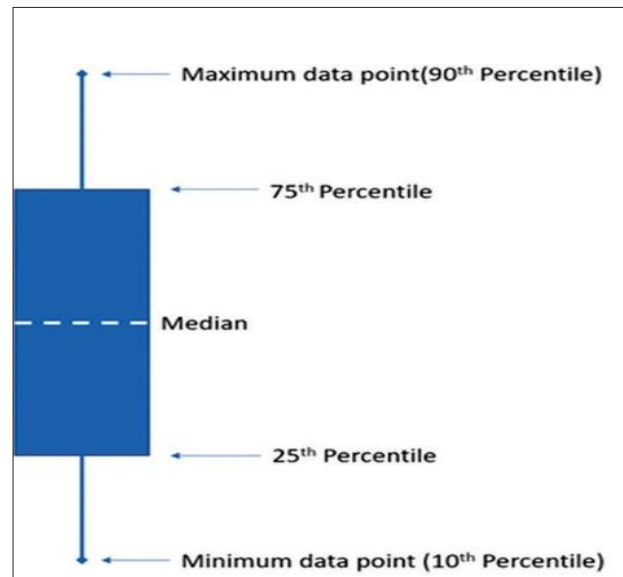
Figure (2): Box plot

1.5 Graphical Presentation of Data

Tabular or graphical performance of outcomes is an excessive skill that can practice to existent not easy and large sizes of results. The pie and bar graph are useful to show quantities and rates achieved for qualitative values. However, pie chart is not greatly well as it characterizes one variable only. Similarly, with large number of categories, the pie chart demonstration develops somewhat unclear. The colors of the part/block of the graph must be bright. Therefore, to distinct various sets of the variables, another design can be prepared whereas plotting them (Figure). Categorizing the data in a certain term of any frequencies or percentages. Both X (horizontal) and Y (vertical) axes must be classified. The name of the figure should be written under the graph. [6]

1.6 Tabular Presentation of Data

Many studies need to use tables as a means of illustration to help us for classifying a large quantity of data in an equivalent style, an attractive and easy to read. The data is set in rows and columns regarding to the features of data in a simple form, economies (save) space, enable evaluation, simplifying statistical analysis and reducing the probabilities of errors. Usually, tables cover explanation of fewer than three variables are reflected fit for data arrangement. Similarly, in the situation of having cell with zero occurrences, or taking a lot of classes with small quantity of numbers should not be available in studies. The most suitable distribution table is the one whose classes rang from five to fifteen. The title should be written above the table.

First researcher can analysis the dataset and stated what is main to be classified in figures or tables. Finally, tables and graphs can be stated in the main text of the article. [9]

1.7 Strategies for Determining Sample Size

There are several methodologies for defining the sample size. According to the Glenn D. Israel (1992) strategies, the strategy of a survey for small populations is used to the whole population as the sample. Even though, cost reflections is difficult for selecting whole populations, a survey is useful for small populations less than or equal to 200. A survey decreases sampling error and affords data on all the observations in the population, also approximately expenses such as questionnaire strategy and evolving the sampling organization are "stable," that is, they will be similar for samples of 50 or 200. To conclude, essentially the whole population would have to be experimented in small populations to get an essential equal of precision. [10]

2.APPLICATIONS

According to Demand for Bread (2023), the strategy "Census for Small Population" of Glenn D. Israel (1992) [10], the sample survey is used to determine a sample around of size (50) questionnaires in Sulaimani Kurdistan Iraqi region. As a good research practice to describe the best appropriate usage in practices the dynamic role of descriptive statistics was explained to the data analysis as well as given that the basis for comparing variables. It was optional with the support of a methodical approach to decrease the probabilities of giving misrepresentative results [2, Table 4].

Table-4: The sample questionnaire data

Data	Label	Variable Types
What is your gender?	Male Female	Dichotomous Nominal
What is your age range?	20 -25 25 – 30 30 – 35 35 +	Interval and continuous values
What is your family type?	Unclear Extended Single Parent	Nominal
How often do you Purchase Bread spreads?	Daily Weekly Monthly Yearly	Ordinal
What factors affect your decision in purchasing bread spreads?	Price Taste Availability Advertising	Nominal
What Bread spread would you prefer to buy?	Wheat Barley	Nominal
How many bread do you eat in a day?	1 – 2 piece 2 – 3 piece 3 – 4 piece More than 4	Ordinal

First type of descriptive statistics is measures of frequency gender group. Clearly the dichotomous nominal gender example variable in [2, Table5] has only two categories (male and female), a pie chart is plotted of the results with SPSS program. Also including frequency, percentage of each category, and seems to be that the total percentage of male is 54.3% of participants.

Table 5: Common statistical terms of gender

Gender	Frequency	Percent
Female	23	45.7%
Male	27	54.3%
Total	50	100.0

In fifty participants, the ratio of female to male is 27:23; conversely, the ratio of male to female is 23:27, presented by pie chart [2, Figure 3].
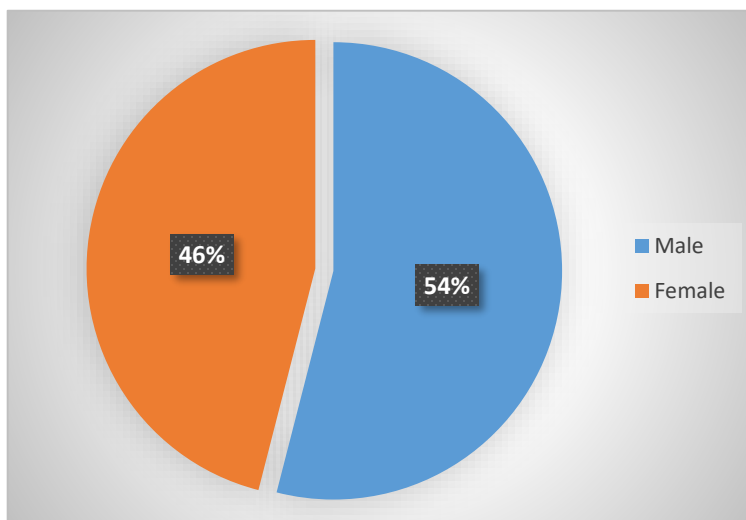

Figure (3): The ratio between male and female in pie format

The measures of frequency of continuous age values in [2, Table 6] presented by absolute frequency, rate, proportion and percentage and classified in to four class interval computing statistics from it with histogram chart.

Table-6: Measures of frequency of age values

Age	Frequency	Rate	Proportion	Percentage
20 –	23	23/50	0.46	$23/50 \times 100 = 46\%$
25 –	17	17/50	0.34	$17/50 \times 100 = 34\%$
30 –	8	8/50	0.16	$8/50 \times 100 = 16\%$
35 and over	2	2/50	0.04	$2/50 \times 100 = 4\%$
Total	50		1.00	100%

It is important to rely on different ages in the surveys, especially those surveys on public and civil issues, and in this regard, we have paid attention to different ages from teenagers to the elderly. Most of the things around them, especially food and medicine, change their rate and type according to the fogs, so most of the time they have to be relied on. In total of fifty participants there are 2 who show their age over 35, the rate age of over 35 years is 2/50 in the census. In proportion $2/50 = 0.04$ of participants are over 35 years. In the total of fifty the percentage of over 35 years is $2/50 \times 100 = 4\%$ of participants. But the higher percentage of age was 46% in the first class interval of age between 20-25.

Second type of descriptive statistics is mathematical computation regarding to age values as explained below in [2, Table 7]:

Table-7: Mathematical presentations of age values

Statistics		Computation	Common statistical terms and their definition
N		50	Number of all participant in the survey.
Mean		27.66	The mean age score.
Median		27	The middle score of a rank- order data.
Mode		28	The statistical value with the highest occurrence.
Std. Deviation		5.397	The root- mean square deviance from the center.
Variance		29.127	The square of the standard deviation.
Skewness		.725	Positively skewed to the right hand.
Range		22	The variance between max and min values.
Percentiles	25th	23.00	Position of lowest 25% and highest 75% of values.
	50th	27.00	Median is middle of the dataset
	75th	31.25	Position of lower 75% and upper 25% of values.
Interquartile rang (25%-75%)		31.25 – 23.00 = 8.25	Is the difference between the 75% and 25% quartiles.

In the above [2, Table7] the center distribution of age group is 27.66 years who participated in the study, the Medill of age value is 27.00 years according to median, but the most common value of age in the dataset is 28 years that because mode not accurately represent the center distribution.

Although the mode may be used with both quantitative and qualitative variables, such as in the case of gender the most frequently participated is 27 males.

The range of 22 years indicates the variability in the age. Although we have accurate range, most values are actually clustered around a clear middle which is equal to 28 years. Age curve

distribution is positively skewed to the right hand with degrees of 0.725, because existence of two outliers falls elderly range class within 35+, [2, Figure4] and boxplot [2, Figure5].
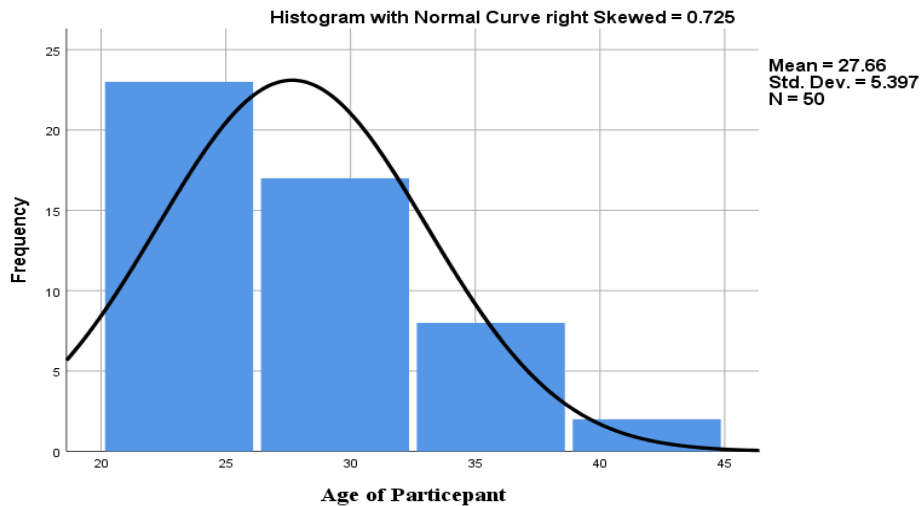
Figure (4): Age range between 20 to over 35 years

A high number of young people have responded that they are between the ages of 20 and 25, and SPSS has clearly shown through histogram.
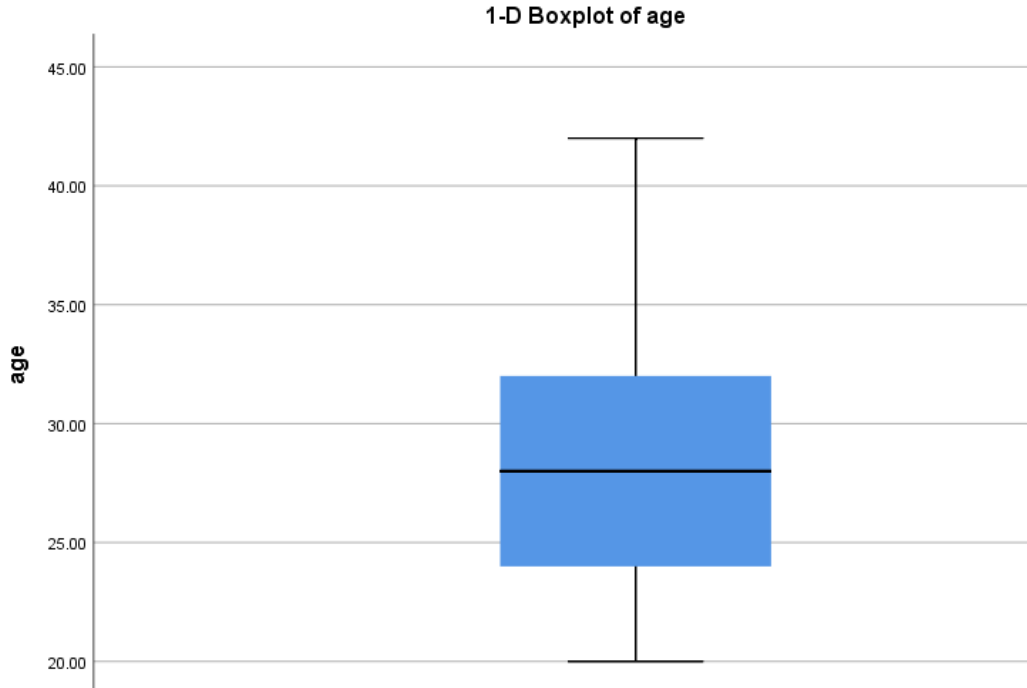


Figure (5): Boxplot of Age Scale

The following example is demonstrated the nominal data with three levels (unclear, extended, and single parent).

In our study we asked about family type which type of type they have and how they live to participate in the survey in [2, Table 8]:

Table-8: Family type

Family type	Frequency	Percent
Unclear	9	18%
Extended	30	60%
Single parent	11	22%
	50	100.0

In this way, the statistics showed that the high rate are extended and the low rate are unclear type in [2, Figure 6].
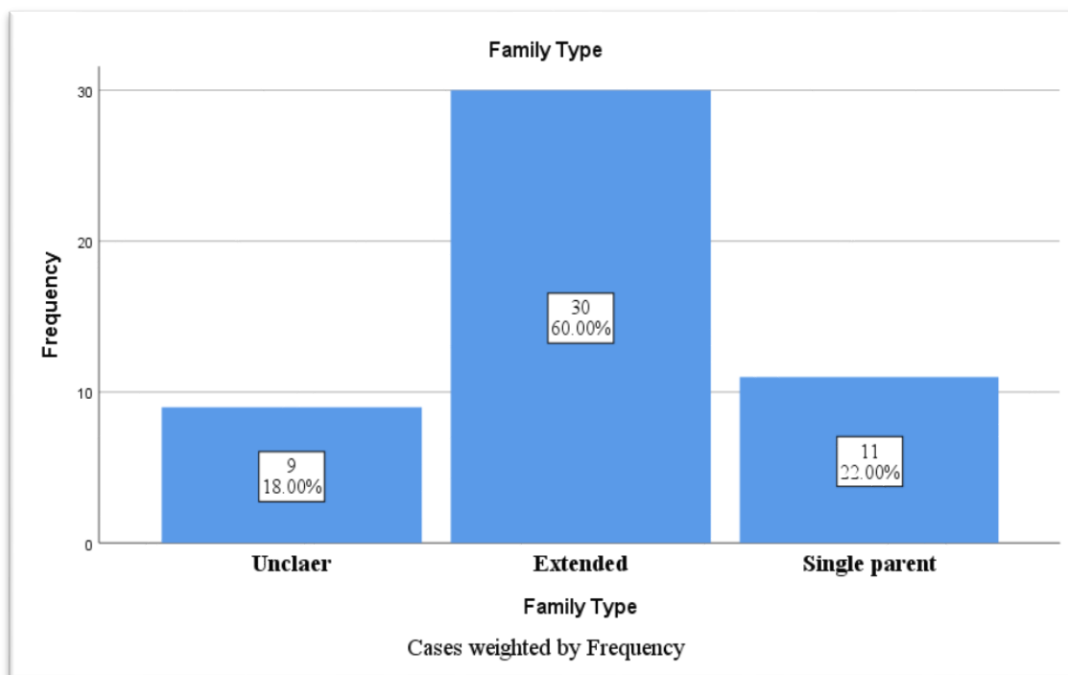
**Figure (6): Three different case of family type**

According to histogram chart shows that all the letters participated in the survey in a high rate and with little difference, the highest rate is extended and the low rate of single type.

According to the difference of meals, the rate of buying the type of food changes to the families and individuals the rates are high and low as in [2, Table 9] and [2, figure7].

Table-9: Bread purchase schedule

Purchasing Bread Spreads	Frequency	Percent
Daily	25	50.0
Weekly	13	26.0
Monthly	9	18.0
Yearly	3	6.0
Total	50	100.0

It seems that people care more about buying bread daily and being present at their tables, spss is a light in showing different rates in these subjects.
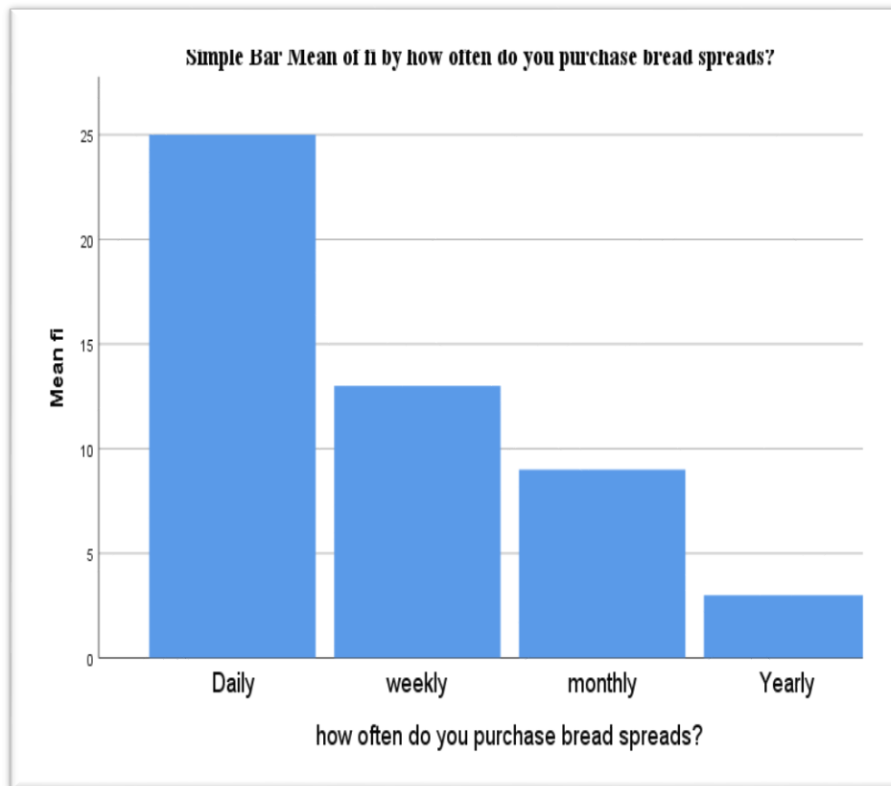

Figure (7): Spreads of purchasing bread

Everything daily has election conditions and the desires change according to the styles and are taken care of differently. The factors affected the participant decision in purchasing bread spreads is explained in [2, Table10]:

Table-10: Decision factor to purchasing bread

Decision Factor	Frequency	Percent
Price	22	44.0
Taste	12	24.0
Availability	10	20.0
Advertising	6	12.0
Total	50	100.0

We have an economic people think and care about the price of their food and evaluate them on this subject as they buy, other subjects such as taste, availability and advertising have decreased according to the statistics taken and graphically presented in [2, Figure 8].
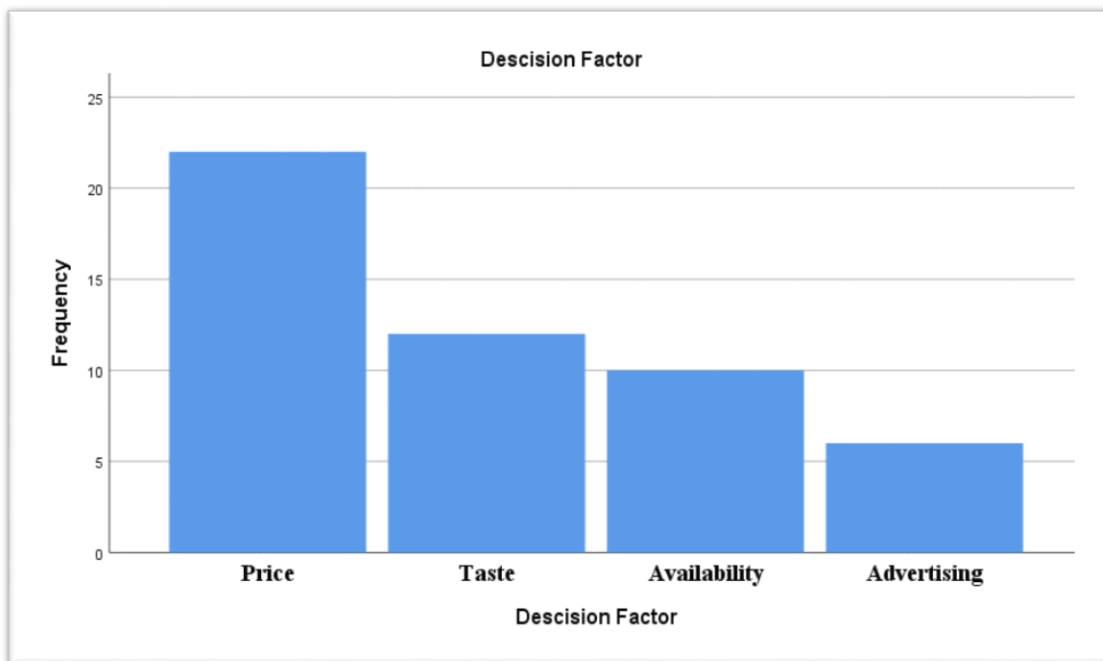

Figure (8): Decision factor

We see that by organizing things by spss on the program, the shape of the staircase is standing in a way that is high (price) and lower than the teste with availability and advertising. Things change every day until the foods are considered new flavors and styles and they change generally we have two kinds of bread that people care about.

Table-11: Preferring bread to buy

Preferring type of bread	Frequency	Percent
Wheat	40	44.0
Barley	10	24.0
Total	50	20.0

Although the odds of wheat are 4 times greater in the wheat group than barley group, it seems that people's desire to eat wheat bread greater than barley, [2, Table 11] and [2, Figure 9].
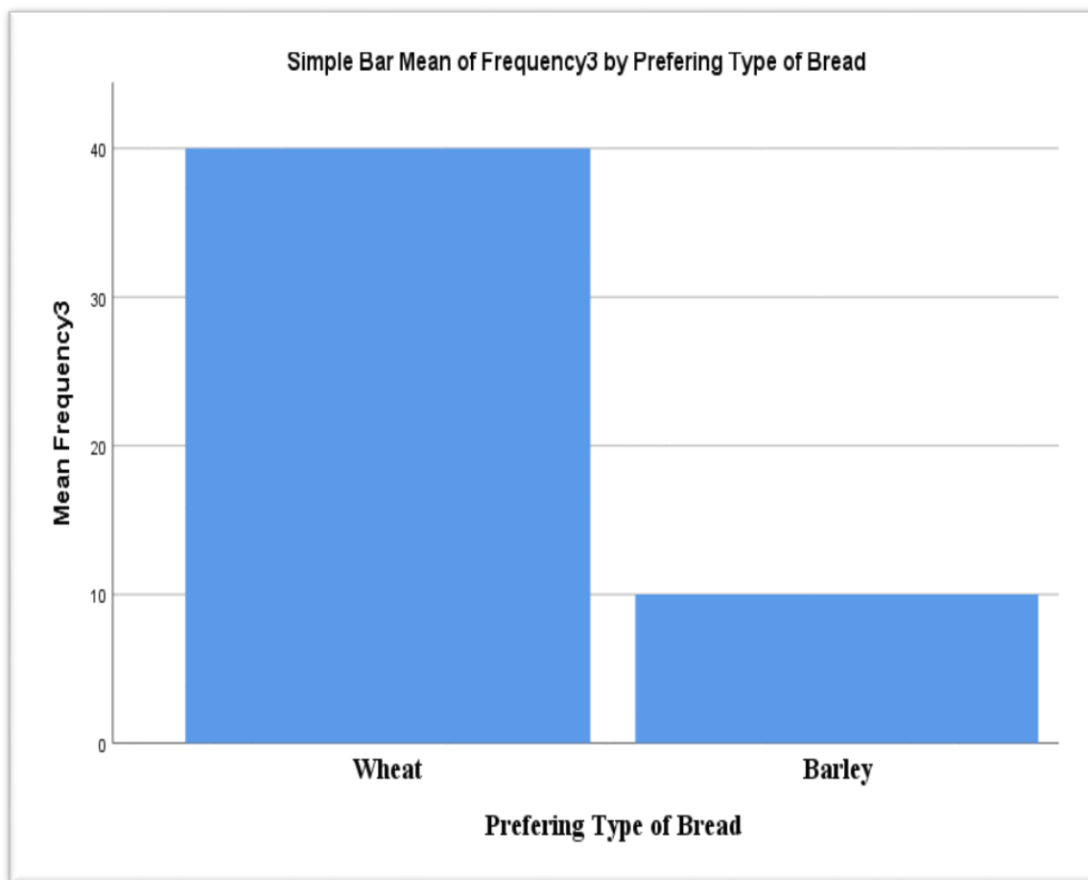
Figure (9): Preferring type

In the end, we tried to get the rate of people's interest in eating bread daily because it depends on people's desire for this type, [2, Table 12].
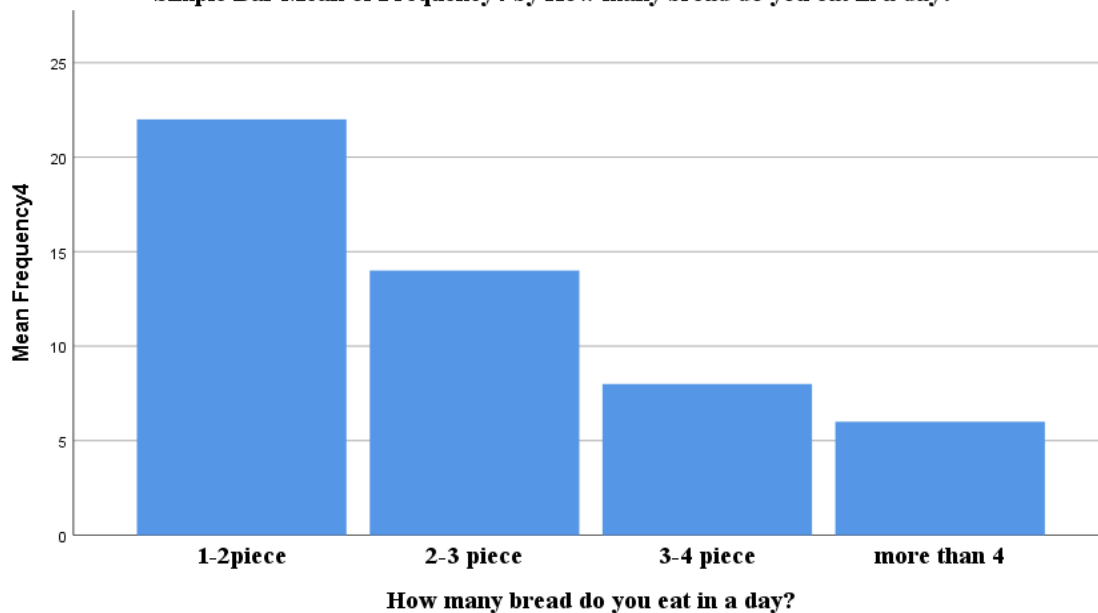
Table-12: Daily eating

Eating Pieces	Frequency	Percent
1 – 2	22	44
2 - 3	12	24
3 – 4	10	20
More than 4	6	12
Total	50	100.0

Graphically we see that between a piece to 2 piece most frequently eaten daily [2, Figure 10].

Figure (10): Pieces daily taken

Maybe taking a piece or two of bread is enough in a day of meals as the statistics taken give it to people how much they like to eat and think that a low percentage is enough in their meals.

Furthermore, the statistical terminology and methods used that comprise descriptive statistics are explained, including variable type, measure of frequency, central tendency (average), dispersion (spread), position and the concept of graphical distribution. Thus, descriptive statistics abbreviate data to clear classification, assist the economics-market research analysis to specific populations in a more manageable form.

Article

3. DISCUSSION

In table and pyramid of basic human needs, bread is at the first level, meaning that it's the main necessity of human beings, and in a way that now the importance of economics has become a competition in the type, taste, and form. Here we showed some attention to trading in Sulaimani city according to the type of people's desires.

According to the frequencies, at most 54% of male were aged between 20 and over from extended families prefer 1 - 2 pieces of wheat bread can have of their meals daily. Economic people think and care about the price of their food and evaluate them on this subject as they buy. Other subjects such as taste, availability and advertising have decreased according to the statistics taken and graphically presented and organized in the sample study.

 Then we should consider that taking the statistics of the surrounding things is important and human can organize more by reviewing the statistics of work, necessities, and life issues in terms of providing time, health, and economic issues, while we have programs like SPSS in this day it is much easier to reach these rates of subjects.

4. CONCLUSION

Descriptive statistics is used to review and existent data briefly and implicitly. It is usually used in numerous fields such as investigation, commercial, economic, community sciences, and healthcare. Descriptive statistics services investigators and predictors to define the central measures such (mean, median, mode), variation (range, variance, and standard deviation), and figure of the distribution of a dataset. It furthermore includes graphical demonstration of data to relief conception and considerate. The convenient use of descriptive statistics allows economic ministry managers and providers to additional efficiently weigh and control of economics policies and programs.

Article

Conflict of interests.

There is no conflict of interests

References

[1]   Z. Ali and S. Bhaskar, "Basic statistical tools in research and data analysis." *Indian Journal of Anaesthesia*, vol. 60, no. 9, p. 662, 2016, doi: 10.4103/0019-5049.190623.

[2]   G. Marshall and L. Jonker, "An introduction to descriptive statistics: A review and practical guide." *Radiography*, vol. 16, no. 4, 2010, doi: 10.1016/j.radi.2010.01.001.

[3]   A. Spriestersbach, B. Röhrig, J.-B. du Prel, A. Gerhold-Ay, and M. Blettner, "Descriptive Statistics." *Deutsches Ärzteblatt international*, 2009, doi: 10.3238/arztebl.2009.0578.

[4]   T. Edgar, "A Factorial Card Trick." *Math Horizons*, vol. 27, no. 1, pp. 23-25, 2019, doi: 10.1080/10724117.2019.1611058.

[5]   M. Huebner, W. Vach, and S. Le Cessie, "A systematic approach to initial data analysis is good research practice." *The Journal of Thoracic and Cardiovascular Surgery*, vol. 151, no. 1, pp. 25-27, 2016, doi: 10.1016/j.jtcvs.2015.09.085.

[6]   Hussain, Mehwish. "Descriptive statistics--presenting your results I." JPMA. The Journal of the Pakistan Medical Association 62.7 (2012): 741-743, doi: 10.1159/000098324. Epub 2006 Dec 29.

[7]   G. Fulk, "Descriptive Statistics, An Important First Step." *Journal of Neurologic Physical Therapy*, vol. 47, no. 2, pp. 63-63, 2023, doi: 10.1097/npt.0000000000000434.

[8]   P. Kaur, J. Stoltzfus, and V. Yellapu, "Descriptive statistics." *International Journal of Academic Medicine*, vol. 4, no. 1, p. 60, 2018, doi: 10.4103/ijam.ijam_7_18.

[9]   S. S. Sonnad, "Describing Data: Statistical and Graphical Methods." *Radiology*, vol. 225, no. 3, pp. 22-628, 2002, doi: 10.1148/radiol.2253012154.

[10]   G.D. Israel,"Determining sample size", Institute of Food and Agricultural Sciences, University of Floridapp.1-5 ,(1992): 2017.

Article

الخلاصة

المقدمة:

من الموضح في جدول وهرم الاحتياجات اساسية للإنسان، يكون الخبز في المستوى الأول. لهذا الغرض تم استخدام الاحصاء الوصفي لدراسة الطلب على الخبز المصنوع من نوعية (القمح او الشعير) في محافظة السليمانية مع دراسة رغبات و اهتماماتهم عن اي نوع اخرمن الطعام.

طرق العمل:

تم تركيز على أساسيات لفهم الإحصاء الوصفي وإلقاء بعض الضوء على البيانات الكمية والنوعية في التطبيقات. تم الحصول على النتائج و توزيعات التكرارية واشكال البيانية باستخدام برنامج SPSS.

الاستنتاجات:

وفقًا للتكرارات، كان هناك 54 ٪ من الذكور تتراوح أعمارهم بين 20 فما فوق من العائلات ممددة(الموسعة) يفضلون قطعة أو قطعتين من الخبز مصنوع من القمح في وجباتهم اليومية. انخفضت جوانب الأخرى مثل الذوق والتوافر والإعلان وفقًا للإحصاءات بيانياً في عينة الدراسة.

الكلمات المفتاحية:

الاحصاء الوصفي, التوزيع التكراري, المسح بالعينة, SPSS