PART 3

New Experiments Three Factors On The Box plot, Box notch & Outlier Values

Mahmud M. Hussin
Statistics Dept. College of Admin. & Econ.
Baghdad University

الخلاصة

يعتبر الرسم الصندوقي (box plot) من أفضل وأهم طرق الرسم على الاطلاق لعرض البيانات وكذلك لاجراء المقارنة بين عدة مجاميع وحسب الغرض المطلوب من البيانات ممكن المقارنة لجميع خواص العينة أو لوحدة من الصفات والرسم يحدد خمس قيم أساسية من البيانات وهي أكبر قيمة وأصغر قيمة والربيعات الثلاث أضافة الى حدود الثقة حول الوسيط (Notch) كما تلاحظه من الشكل رقم (1A) في نهاية البحث .تجربتين حديثة لكل تجربة ثلاث متغيرات تم تطبيقهما في هذه الدراسة لاجراء بعض التعديلات على الرسم الصندوقي (box plot) .لأيجاد تأثير بعض التحويرات على قرارت الاشخاص بالنسبة الى قرار طول الرسم الصندوقي ; في التجربة الاولى تجربة حدود الثقة للوسيط (The first experiment box notch experiment) تم تغيير العرض للصندوق وكذلك حدود الثقة وكذلك الطول للصندوق (box width, box notch, box length) في التجربة الثانية تجربة القيم الشاذة (And the second experiment outlier values experiment) تم تغيير العرض للصندوق وكذلك الطول للصندوق وكذلك موقع اقيم الشاذة (box width, box length, outlier values) المحددة في الرسم الصندوقي. كل تجربة في ملف يحتوي على (64) شكل كل شخص يساهم بتجربة واحدة ويقارن أطوال الاشكال (box plots) في أحد الملفات مع طول الصندوق الثابت (standard plot) في كلا التجربتين كان طول الصندوق الثابت (standard plot) هو مستوى الطول الذي يقع في منتصف المدى ولم يستخدم في كلا التجربتين وكذلك طلب من الاشخاص أولآ أن يحدد في كل مقارنة من ال (64) مقارنة أيآ من أطول الصندوقين أقصر لتحديد الاتجاه هل هو الثابت أم الذي في الملف ويقدر نسبة مئوية طول الصغير من الكبير وبشكل سريع أول تصورأو أدرك يلاحظه وليس قياس دقيق. هذه التجارب تم تصميمها كأمتداد الى التجارب التي تحتوي على متغيرين في المجموعة (group B experiments, the ratio experiments) والتي تم تطبيقها سابقآ من قبل الباحث Hussin, M.M. (1989, 2006, 2007)

عندما أقترح Hussin, M. M. (1989) في الدراسات المستقبلية في النقاط (1, 2, 3) (ممكن أن نأخذ عدد أكثر من المتغيرات بل أضافة الى المتغيرات الحالية وفي تجارب جديدة ولأكثر من متغيرين مثلآ (box notch, outlier values) وكذلك نأخذ مستويات أوسع لأطوال المتغيرات كذلك أن تغير الطول الى الرسم الصندوقي الثابت (standard box length) سوف يؤثر على أدراك الشخص من طول الصندوق (box length) عند المقارنة بين هذه الازواج من الرسوم الصندوقية (box plots) . النتائج لهذه التجارب تؤكد على أن هذه التحويرات أو التعديلات تؤثر على أدراك الأشخاص لأطوال هذه الرسوم الصندوقية (box plots) . نعتقد أن هذا التأثير في قرارات الأشخاص لهذه التحويرات على الطول الى الرسم الصندوقي (box plots) ممكن أن تكون كنتائج للتفاعلات (interactions) بين متغيرات الرسم

الصندوقي والتي هي (box length, box width, box notch, and outlier values) وهذه التفاعلات تكون ما يسمى بالوهم المرئي (visual illusion) وهذا يؤثر على قدرة الأشخاص لأتخاذ قرارت دقيقة بالنسبة الى أطول الرسم الصندوقي (box lengths) . وكذلك النتائج تؤكد على أن الاقتراح بأضافة عدد من المتغيرات وكذلك عدد مستويات الاطوال الى المتغيرات بل أضافة الى تغيير الطول الى الرسم الصندوقي الثابت (standard box plot length) في الفقرات (1, 2, 3) للدراسات للمستقبل كانت صحيحة وبينت نتائج هذه التجربتين أن زيادة عدد المتغيرات الى ثلاثة جعل الصعوبة بشكل أكبر وحتى تختلف في نوع هذه الصوبة كما نوهنا سابقآ عند مناقشة النتائج لهذه التجارب هذه الاضافات الى الرسم الصندوقي أضافة معلومات أضافية الى الشكل ولكن الكلفة كانت الصعوبة. هذه التجارب أجريت على طلاب قسم الاحصاء المرحلة الثالثة والرابعة/ كلية الادارة والاقتصاد/ جامعة بغداد .

Summary

New two experiments of the three factors, in this study were constructed to investigate the effects, of the fixed variations to the box plot on subjects' judgments of the box lengths. These two experiments were constructed as an extension to the group B experiments, the ratio experiments the experiments with two variables carried out previously by Hussin, M.M. (1989, 2006, 2007). The first experiment box notch experiment, and the second experiment outlier values experiment. Subjects were asked to judge what percentage the shorter represented of the longer length in pairs of box lengths and give an estimate of percentage, one being a standard plot and the other being of a different box lengths and also varying with respect to other elements such as, box width & box notch or box width & outlier values. When Hussin, M.M. (1989) suggested in the future research points (1, 2, 3), can take account of wider range of the variables levels and the changing length of the standard box plot effects on the subjects' perception of the box length and further investigations could be made into variations such as box notch with box length; box notch with box length; whisker length, and also outlier values with these variations. However, both experiments were used the stander box length as the middle box length levels were not used in the experiments. The results of these two experiments suggest that these variations effected the subjects' perception of box length, as a results of the interactions between box plot features by creating visual illusions as which effect the subjects ability to accurately judge box length, both experiments were run in statistics department, Baghdad University.

1 -Introduction.

Exploratory data analysis can help to show the unexpected features of the data or can allow us to make simple or detailed comparisons between distributions of the data sets. Graphical methods are an essential part of the exploratory data analysis. These techniques can give us a clear idea about the patterns of data set distributions. Graphical methods are used not only to summarize data, but also as diagnostic aids in analysis, and to decoding of quantitative information from the graphs. These tools represent a great part in exploratory data analysis in statistics, and have a long history of use in preparing pictures of data and presentation.

The box plot Tukey (1977) is one of the important tools of the graphical methods. This tool can give the viewers a fast idea of fixed features of the distribution, the shape and the spread of the data. This tool can be applied to make simultaneous comparisons between the distributions of several sets of data. The idea of the box plot is simple (see example no.2), it is a graphical display which uses five values obtained from the data set, the upper and lower (hinges), the median, and the upper and Lower adjacent values. This paper is concerned to investigate whether fixed variations to a basic box plot affect subjects' judgments of the box length (midspread). Features studied in first experiment making the box width proportional to sample size and box notch equal confidence interval around the median and in the second experiment box width with outlier values. Subjects were asked to make comparisons between two box lengths and to give what percentage the shorter was of the longer. These experiments were carried out at Baghdad University. There were two statistical methods used to analyze the data of these two experiments, the analysis of variance techniques and the median polish techniques. The results of these experiments suggest that these two variations are affected subjects judgments of the box length.

2- Previous work on box plot

Recent experimental work on graphical perception was carried out by Cleveland et al, 1982, 1984, 1986, 1987 and Simkin & Hastie, 1987 for detail see Hussin, M.M.(1989,2006).And also new recent Six studies have been carried out in Box plot. First study by Mc Culloch in (1981), run an experiment to study the effect of three variables, box length, box width, and viewing time on the subjects judgments of the box length. Mc Culloch concluded that the subjects reaction time of the box length affected by the two variables, box width, and viewing time. The interaction between two variables box length and box width affected subjects' judgments of box length. Box

width added more information to the box plot but made the interpretation of the box plot more difficult.

Second study in (1982) by Knight, were run four experiments in box plot, to examine the effects of varying four features of a box plot on subjects' judgments of box length. These variables investigated in four separate experiments were box width, box notch, whisker length, and outlier values. Knight found that box width, box notch, and whisker length affected the subjects' judgments of the box length, but the outlier values did not affect the subject judgments of box length. The outlier value, the observation their position is beyond the whisker length.

The third study by Hussin , M.M. in (1989), investigated the effects of vary three features of a box plot on subjects judgments of box length, in two different groups of experiments and carried out at Keele and Baghdad Universities. In the group A experiments (comparative experiments), three experiments to study the effects of three variations to the box plot Tukey (1977) on the subjects judgments of the box length, these variation box width, whisker length, and box width with box notch. Subjects were asked to make comparisons between pair of box plots; one of the pair is the standard plot and the other from the booklet. Subjects were asked to respond if the box length of the box plot of the booklet is shorter or longer from the standard plot, and to give a rating of how confident in their judgments by giving a score of 50% to 100%. The subjects were asked to give their answers as a first impression. Hussin, M.M. (1989) found for the box width experiment at Keele, that there is a significant interaction between the two variables, box width / length. The two variables box width and box length affected subjects judgments of box length, but the box length more than box width. And for the box width experiment at Baghdad found that the results are similar to the Keele experiment results.

Hussin, M.M. found for the whisker length experiment at Keele, that there is a significant interaction between two variables box length and whisker length. These two variables are important and affected subjects judgments of box length, but box length more affected than whisker length. And for the Baghdad experiment found that there is no significant interaction between two variables, whisker length and box length. Which is different from Keele experiment, also these two variables are important.

For the third experiment of three variables box notch experiment he found that for Keele experiment, there is a significant interaction between these three variables of this experiment box width, box notch and box length, and only one interaction was significant of the two way interaction, box length and box width interaction. These three variables are important

and affected subjects' judgments of box length. The subjects have more difficulty with judgments in this experiment than in all the other two factor experiments, make more error judgments, and have little confidence in their answers. He found for the same experiment at Baghdad that similar results for three way interaction was significant, but different for two-way interaction the box notch and box length was significant. And these three variables are important for the experiment.

In group B experiments (ratio experiments), which contains four experiments, two of them for length judgments and the other two for area judgments. Two of them were carried out at Keele and all of them at Baghdad. These experiments seek to examine, which features of box plot affected subjects judgments of box length (midspread). Subjects were asked to give percentages for how much shorter , smaller, represented of the longer or larger the length or area of box plot from the pair of box plots, one of the box plot being standard and the other with one or more of the features changed, these compared side by side on A4 sheet of paper. Their effects on judgments were estimated by the error size. Absolute value of the error = [judged percentage - true percentage].

The box length experiment at Keele was build to study the effects of the box length and whisker length variables on the subjects' Judgments. Hussin, M.M., found that there is interaction between these two variables. The box length variable is important and affected subjects judgments of box length, and more important than whisker length. The subjects tend to increase the midmeans of the absolute error with increase in box length, and similar results for the same experiment at Baghdad.

For the box plot three variables, box length, box width, whisker length, Baghdad length experiment. He found that, there are some interactions between these three variables, these interactions affected subjects' judgments of box length. The subjects in this experiment with change in three variables faced more difficulty than in any other experiments using two variables that would means these variables add more difficult to interpret the box plot, and affected the subjects' ability to make accurate judgments.

And for the box plot two factors area experiment at Keele, box length, whisker length, he found that, there is some interaction between these two variables, also the box length very important, but the whisker length less important for this experiment and this result for area experiment very reasonable the whisker length variable is not relevant, the subjects tend to overestimate with small areas and underestimate with large areas.

And for the same experiment at Baghdad. He found that, there is not significant interaction between these two variables, whisker length and box length. The box length variable is very important and dominated the experiment, by affected subjects of box length; the subjects tend to increase the errors with increases on the box area and also the same with keele experiment, this result for area experiment very reasonable and important the whisker length variable is not relevant.

And for the box area, three variables experiment, at Baghdad, Hussin, M.M. concluded that, there is significant interaction between these three variables, and also between any two of them, box length/width, and whisker length ; The two variables box length/width are important, and affected subjects judgments of box length, but whisker length was not important on its own in this experiment and also the same with previous area experiments this result for area experiment very reasonable and important the whisker length variable is not relevant.. Subjects faced difficulty with this experiment more than with other experiments and made large errors. Area judgments are more difficult than length judgments, these results agree with Cleveland & McGill (1984), and the power law results, and also Weber's law might help to explain the results.

The fourth study by Sim, C.H. ;Gan, F.F. ;Chang, T.C. in 2005, they focus on the detection of possible outliers based on the box plot procedures .The outliers in a set of data are defined to be a subset of observations that appear to be inconsistent with the remaining observations. They indicate that the commonly constructed box plot is in general inappropriate for detecting outliers in the normal and especially the exponential samples .And they suggest that the graphical box plot be constructed based on the knowledge of the underlying distribution of the dataset and by controlling the risk of labeling regular observations as outliers.

The fifth study by Hussin, M.M. (2006), run two experiments to investigate the effects of vary two features of a box plot on subjects' judgments of box length, these variation box width, whisker length and carried out at Baghdad University. Subjects were asked to make comparisons between pair of box plots; one of the pair is the standard plot and the other from the booklet. When in this study the standard plot (box length) is the smallest one of the box length levels in these two experiments, which is different from all these previous studies in the box plots, and also he found the results of these two experiments different from all these previous studies in the box plots the interactions between the two variables and box length very highly significant and the whisker length very important by it self and with the interaction and these two variables very important with the box length variable which this result

different from the others but agree with the suggestions points. when Hussin, M.M. (1989) suggested in the future research points (1, 2),can take account of a wider range of variables levels of the experiments, and the changing length of the standard box plot effects on the subjects perception of the box length.

Sixth study two experiments by Hussin, M.M. (2007), to examine the effects of varying two variables of a box plot on subjects' judgments of box length. These variables were box width, whisker length and carried out at Baghdad University. In this study the standard plot (box length) is the middle one of the box length levels in these two experiments, which is different from all the previous box plot experiments. when Hussin, M.M. (1989) suggested in the future research points (1, 2),can take account of a wider range of variables levels of the experiments, and the changing length of the standard box plot effects on the subjects perception of the box length. Subjects were asked to make comparisons between pair of box plots; one of the pair is the standard plot and the other from the booklet. He concluded that the results of these two experiments similar to the results of Hussin, M.M. (2006), and different from all the previous studies in the box plots the interactions were highly significant between the two variables box width and whisker length with box length, and the whisker length very important by itself as a box width and these two variables very important with the box length variable which this result different from the others and agree with the suggestions of the further research points.

3 -The errors in box plot judgments and The problem suggested for this study.

It might be possible to explain the biases in subjects' judgments of the box plots found in previous studies of box plots, Hussin, M.M. (1989, 2006, 2007). One possible explanation of the errors in the subjects judgments of the box plots are the interactions between the box plot variables with the box length variable. When the subjects made comparisons between two box plots one of them standard plot and the other with some change in the box plot variables such as box width or whisker length or box notch with the box length. The subjects might underestimate the box length when boxes are wider as changed the width or have longer whisker lengths and vice versa, this results might be similar to Baldwin's (1895) figures, when he found that the line lengths closer to the large square look shorter than that line lengths closer to the small square .

Another possible explanation of the errors by using Cleveland & Mc Gill's (1984) theory. It might be that the subjects make area judgments instead of length judgments in the case of varying box width or box notch in combination with box length. And that make the judgments more difficult, because the area judgments more difficult than length judgments, power law (Stevens, 1975), and Cleveland & Mc Gill (1984).

Moreover, in the case of varying box notch with box length, the subjects might face difficulty because subjects need to make two length judgments, one for box length and the other for box notch, in addition to the interaction between these two variables. Also as Lovie (Lovie 1985) discussed the nature of the box plot is not a simple graph by which to make quick judgments.

Another law also might help to explain the problem in judging box plot, in 1834, Weber proposed what we call now Weber's Law (Stevens, 1975) and we can give simple idea of this law is that when we need to make comparison between lengths of two things we need first to determine the difference between them by fixed percentage, and not on the overall sizes of the two lengths. Also Stevens (1975) proposed power law might help to explain the errors in the subjects' judgments of the box length of the box plot , and this law used to determine the accuracy in the judgments of different physical aspect objects, such as area, volume, Length , or... etc. The law state that the accuracy of these aspect judgments can be ordered as follow, length, area, and volume. for more detail see Hussin, M.M.(1989,2006).

The problem suggested in this study was to examine of the effects of certain variables in the box plot on subjects' judgments of box length. To find in the first experiment which one or two of the box plot variables, box width, and box notch effect on the subjects' judgments of the box length as a relevant factor, and in the second experiment which one or two of the box plot variables, box width, and outlier values effect on the subjects' judgments of the box length. Subjects were asked to make comparisons between pair of the box lengths of box plots placed side by side; we tried to make accurate judgments and to avoid any effects of the orientation on the subjects' judgments. When in this study the standard plot (box length) is the middle one of the box length levels in these two experiments similar to the Hussin, M.M. (2007) experiments and this level is not used in the experiments, but the box length levels of these experiments different from Hussin, M.M. (2007), and different from all the previous experiments in the box plots and a stander plot and also as a three variables experiments in group B ratio experiments the are the first time used. When Hussin, M.M. (1989) suggested in the future research points (1, 2, 3), can take account of wider range of the variables levels and

the changing length of the standard box plot effects on the subjects' perception of the box length and further investigations could be made into variations such as box notch with box length; box notch with box width; whisker length, and also outlier values with these variations. The same thing might happen in this study inaccuracy or biases in the subjects' judgments might occur with box length, box width, box notch and outlier values , as results of the visual illusion effect created by these interactions between the box length, and these three variables, box width and box notch and outlier values.

4- Method of the experiments.

This section will discuss the method used for these two experiments. The subjects were asked to make comparisons between box lengths of pair of vertical box plots placed side by side. Each box plot was in the center of an A4 sheet of paper, this applied to both the standard plot and the comparison plot. To give percentages for how much shorter, longer the length of the box plot(midspread) was than this in the booklet of the standard box plot. Also discuss the results of the experiments by using the analysis of variance with trend analysis.

4.1- Design.

These two experiments were constructed to examine the effect of box width and box notch in the first experiment, and box width and outlier values in the second experiment, on the subjects' judgments of the box length. The first experiment box notch experiment, contains the sixty four box plots, which were generated from levels combinations of the three factors, box length with four levels, and box width with four levels and box notch with four levels. Each box plot was on an A4 sheet of paper and also a standard plot. And the second experiment outlier values experiment, contains the sixty four box plots, which were generated from levels combinations of the three factors , box length with four levels, box width with four levels, and outlier values with four levels.

The three variable levels for the first experiment determined to fit with the size of an 4A sheet of paper as you find in the table below and the box length levels for second experiment only were determined by Cleveland & McGill (1984)؛ who used the formula :$L_j = 10 \times 10^{(j-1/12)}$, (j=1 —-n), then we suggested 1 unit = 3 mm

These values are equally spaced on a log scale and range from 10 to... N units, chosen values in order started by 26 units represented 4 box length levels after we divided the length by 2, we chose the standard box length for the two experiment in the middle range of the box length levels but not used in the experiments. which is different from the standard plot of the other experiments of box plots. These box length

levels selected to fit the box plot on an A4 sheet paper, and present as large a range of plots as possible, with the other levels of variables, box width, and box notch and outlier values and the position of outlier as a percentage % of the box lengths, see variables levels in table no.1A.

Table No.1A. levels of variables of these two experiments.

First experiment

Length	Width	Notch
L1 = 84	D1 = 50	N1 = 34
L2 = 87	D2 = 60	N2 = 37
L3 = 93	D3 = 80	N3 = 43
L4 = 96	D4 = 90	N4 = 46
Second experiment	**Width**	**U % Outlier Values L%**
L1 = 39	D1= 25	O1= 10 20 30 10
L2 = 47	D2= 40	O2= 10 10 20 30
L3 = 57	D3= 55	O3= 15 30 15
L4 = 70	D4= 70	O4= 15 15 30

Standard plots variables levels for experiments

First exper. L = 90 D = 70 N = 40 Second exper. L=50 D=50 W= 70.

4.2 -Materials.

There were two booklets, there are sixty box plots in each of the booklet. The first

sheet in the booklet contained two examples of practice plot so that the subjects understood the experiment. Subjects were also given an instruction sheet, an answer sheet , and a standard box plot, the booklets were given to subjects in the lecture room , and each subject was given a booklet of one of the experiments. The instruction sheet asked subjects to compare the box plots from the booklet with the standard box plot. The subject was asked to give a percentage of how shorter or longer the length of the box plot was than that in the booklet of the standard box plot, and the standard box plot length was in the middle range levels without this being mentioned to the subjects. The subjects were also asked to write (T) or (B) respectively on the answer sheet if they thought that the length of the box plot on the booklet was longer than the standard box plot or vice versa. This provided a check on the direction of their judgments. The instructions asked subjects to make quick visual judgments rather than measurements. Examples of standard box plot, instruction sheet, answer sheet are not included, because the problem of the space.

4.3- Subjects: Subjects taking part in these two experiments were undergraduate third and fourth years from statistics department, Baghdad University, they were not familiar with the box plot, but had some knowledge of data analysis. There were (57) subjects taking part in

box notch experiment and (51) subjects were taking part in outlier values experiment; the subjects who had not understanding the instructions had their answers excluded from analysis.

5 -*The statistical technique used to analyze the data .*

The analysis of variance technique was used to analyze the data of these two experiments, these experiments were designed as repeated measures, and for such data the analysis of variance technique appears to be appropriate, for more detail see Hussin, M.M. (1989, 2006). The assumptions of the design can be summarized as :

$$X_{ij} \sim N(\mu_i, \sigma^2).$$

There were three models can be used for the analysis of variance technique fixed effects model, random effects model, and mixed effects.

The design of these two experiments were repeated measures design, and the model for this design is the special case from mixed mode1. In this design subjects are observed at all combinations of the independent variables, and the model for the first experiment the box notch experiment is ;

$$Y_{ijmk} = U .... + L_i + D_j + N_m + LD_{ij} + LN_{im} + DM_{jm} + LDN_{ijm} + E_{ijmk} . (1)$$

'K' th observation (subjects).

'I' th level of box length factor (q) levels.

'j' th level of box width factor (r) levels.

'm' th level of box notch factor (p) levels.

In this model (1), the box length , box width and box notch are the fixed effects factors, and the subjects are a random effects factor. For this design as subjects are observed at all observations of the variables, it is expected that the observations on the same subjects will tend to be correlated, or be dependent. For this reason, this design needs more assumptions of homogeneity of the variance- covariance matrix.

1- The variances are : $\sigma^2_{x1} = \sigma^2_{x2} = \sigma^2_{x3} = \ldots = \sigma^2_{xn}$.

2- The covariances are: $\sigma_{x1\,x2} = \sigma_{x1\,x3} = \sigma_{x2\,x3} = \ldots = \sigma_{xn-1\,xn}$.

If this assumption is not met, it is impossible to use the usual F test, without some modifications. For this reason the conservative test provides approximation, but some times this test is negatively biased, (Winer, 1962, P. 306), for more detail see Hussin, M.M. (1989, 2006).

6- The first experiment box notch experiment results.

Now will discuss the results of the box notch experiment, and the fact that we choice the univariate analysis of variance, as the problems with the assumptions of normality as we found that some of data sets were light- tailed and some were skewed, or double peaked at upper and lower extremes, and also the violation of the variance- covariance matrix, we found the adjusted univariate analysis of variance with trend analysis is more power full and suitable than the multivariate analysis. This was recommended by Winer (1962, p. 306) by Rogan et al (1979, p. 269- 286), and by Huynh 1970, Huynh, 1978, Huynh et al, 1979, and also by Charles S. Davis (2002). The trend analysis also was found to study more specific aspects of the differences in patterns or shapes for the simple main effects of the variables in the analysis, and the polynomial contrasts is the best way to do this job.

There are three variables in this experiment; box length, box width and the box notch with four levels for each of them, the model is the equation no.1. Now let start analyze the results of the analysis of variance in Table no.1, and started with the interaction effects of these three variables (LDN_{ijm}). It was found that the (F) value of this interaction effects (LDN_{ijm}) was equal to (2.866) and the tail probability for the usual (F) test was equal to (0.000). This means that the interaction effects have a high level of significance. But to use the usual (F) test for this design is highly restrictive because a design having correlated observations will affect the results in a positive bias in the usual (F) test. That is, the variance- covariance matrix should confirm the assumption of homogeneity of this test. Checking this assumption for the interaction by using the sphericity tail probability for the (LDN_{ijm}) revealed that the assumption of sphericity was not met, thus the conservative test provided an approximate test with the number of degrees of freedom for the (F) value reduced by (E) Epsilon. But even with this test whatever the reduce in the degrees of freedom as a large degree of heterogeneity in the variance - covariance matrix , still this test interaction effects have very high level of significance with this conservation test because it has highly level of significance. Also it is very clear to recognize the interactions between the two variables with the box length from the two plots of the box notch experiment plot no.1, and plot no.2, and also can be seen from these two plots, that the averages of the absolute values of the errors increases with the middle length levels as nearer from the standard box length and decreases with shorter and longer levels as far from the standard length, these results opposite to the results of two variables box width experiment Hussin, M.M. (2007) as similar to the standard box plot length and it might be these results because of difficulties from the change

of the three variables affected the subjects' judgments .

And to examine the differences in trends of the variables' effects for the interaction, by using the polynomial contrasts, we found from Table no.1 six polynomial components of this interaction significant. That means there are significant differences between all the trends of the interaction of these three variables, box length, box width and box notch as you can see from the Table no.1 and from the two plots the plot no.1, and plot no.2. These three variables all of them are responsible for this interaction , and three of them are important for the subject' judgments or three of them have affected subjects' judgments and all of them were important for the experiment .One possible explanation of these results were that as the box width and box notch changed , subjects judged area instead of length for area judgments and area is more difficult than length judgments,
as the power law (Steven , 1975) suggested and Cleveland et al (1984) found from their results . And also it might be as a result of the visual illusion effects on the subjects judgments created by the interaction between these three variables box width, box notch and box length as Cleveland et al (1987) accepted in their replies to the comments on their results. This result of three variables interaction box width, box notch, and box length the ratio experiment agree to the same interaction of the three variables of Keele and Baghdad box notch experiments, but in group A comparative experiments.

Now let examine the results of the two way interactions effects of the two variables and started with the box length and box width interaction (LD_{ij}). It was found that the (F) value of this interaction effects (LD_{ij}) is equal to 2.992 and the tail probability for the usual (F) test is equal to (0.002). This means that the interaction effects have a very high level of significance. But checking the assumption for the interaction by using the sphericity tail probability for the (LD_{ij}) revealed that the assumption of sphericity was not met. But even with this test, still interaction effects have very high level of significance with conservation test equal (0.004). And also it is very clear to recognize the interaction between these two variables from plot of box length and box width in box notch experiment plot no.1. And to examine the differences in trends for this interaction, by using the polynomial contrasts , we found from Table no.1 one polynomial component of this interaction significant but with very high level of significance. That means there are significant differences between Cubic trend and quadratic trend, of this interaction of these two variables, box length and box width. These variables two of them are responsible for this interaction , and two of them have affected subjects' judgments.

This result of the interaction agrees with McCulloch's (1981), and Hussin, M.M. (1989) , group A experiment, box width experiment, box notch three variables at Keele and Baghdad box width results and disagree with box notch three factors Baghdad and Knight's (1982) results, and also agree with Hussin, M.M. (2006, 2007) . And disagree with Hussin, M.M. (1989), three factors experiment results, group B experiment at Baghdad. But this experiment result and Hussin, M.M. (2006, 2007) the interaction highly significant more than all the other experiments results and one explanation for these results are the standard box plot lengths in these experiments represent the middle level of the length levels and shorter in (2006) than all the other box plots length levels of the experiment, which is different from all other experiments, and that agree with the suggestions of the further research points (1 ,2, 3) to build these two experiments.

And now let examine the results of the interaction of the variables box length and box notch (Ln_{im}). It was found that the (F) value of this interaction effects (Ln_{im}) is equal to 4.309 and the tail probability for the usual (F) test is equal to (0.000). This means that this interaction effects have a very high level of significance. But checking the assumption for this interaction by using the sphericity tail probability for the (Ln_{im})revealed that the assumption of sphericity was not met. But even with this test, still interaction effects have very high level of significance with conservation test equal (0.000). Also it is very clear to see the interaction between these two variables from plot of box length and box notch in box notch experiment plot no.2. And we found from the table no.1 four polynomial components of this interaction significant. That means there are significant differences between all components of this interaction of these two variables, box length and box notch. These variables two of them are responsible for this interaction , and two of them have affected subjects' judgments. This results of this interaction agree with Baghdad, but disagree with keele, the box notch three factors experiment (box length, box width ,and box notch). Group A comparative experiment, and agree with Knight's (1982) results .

In the same time we found the (F) value of the interaction effects between box width and box notch(Dn_{jm}) is equal to 3.113 and the tail probability for the usual (F) test is equal to (0.001). This means that the interaction effects have a very high level of significance. But we found the assumption of sphericity was not met. And still interaction effects have very high level of significance with conservation test equal (0.002). This mean there was interaction between these two variables. And we can see from the Table no.1 four or five polynomial components of this interaction significant. That means there are significant differences

between all components of this interaction of these two variables, box width and box notch. These variables two of them are responsible for this interaction, and two of them have affected subjects' judgments. This results of this interaction disagree with Baghdad, and also disagree with Keele, the box notch three factors experiment(box length, box width, and box notch). Group A comparative experiment.

To examine the results of the main effect in Table no.1. It was found that the F-value of the length variable equal to (19.88), and the tail probability for this variable is equal to (0.000), and that means very high level of significance. The sphericity assumption is not met and the Mauchly's W test equal (0.566)and their significance level (0.000) . We do need to use the conservative test for this length variable, and still main effects have very high level of significance with conservation test equal (0.000) . Now let us examine trend analysis for the length variable main effect , we found two components linear and the quadratic were significant with very high level of significance(0.000) ,and that can be seen from plot no.1 box notch experiment. However, the F-value still has a very high level of significance (0.000), and the length variable is very important for the experiment.

The F-value of the width variable equal to (3.427), and the tail probability for this variable is equal to (0.018), and that means high level of significance, and there is homogeneity of variance - covariance matrix. The sphericity assumption is met and the Mauchly's W test equal (0.868)and their significance level (0.170). The conservative test for this width variable is not important. Now the trend analysis for the width variable main effect, we found only the quadratic is significant with high level of significance ,and that can be seen from plot no.1 box notch experiment. This results disagree with the results of the Keele and Baghdad the box notch three factors experiment (box length, box width , and box notch). Group A comparative experiment. In this experiment box notch, width variable is important by it self and with the others variables by the interactions and had affected subjects' judgments of box length.

Finally let us examine the box notch main effect, it was found that the F-value was equal to 0.686 in Table no.1, and the tail probability for the F-value was equal to 0.562. Now we don't need to use the conservative test and the trend analysis because the main effect of this variable is not significant. And that would mean this box notch variable was not important by itself but with others interactions as we explained before .This results of the box notch variable agree with Hussin, M.M. (1989) result, of box notch experiments group A, at Keele , and Baghdad

7- The second experiment outlier values experiment results.

This section for outlier values results of the analysis of variance with the trend analysis, the model for the experiment equation no. (2), three variables, box length with four levels and box width with four levels and outlier values with four levels.

$$Y_{ijnk} = U.... + L_i + D_j + O_n + LD_{ij} + LO_{in} + DO_{jn} + LDO_{ijn} + E_{ijnk} .....(2)$$

' k ' th observation (subjects)

' I ' th level of box length factor (q) levels .

' j ' th level of box width factor (r) levels .

' n ' th level of outlier values factor (p) levels .

This model repeated measures design, the box length, box width, and outlier values are the fixed effects factors, and the subjects are a random effects factor. Now let us consider Table no.2 , and begin with the interaction effect of the three variables box length, box width, and outlier values(LDO_{ijn}). The F value for this test was equal to 2.731, and the tail probability for the F– value was equal to 0.000 . This means that this interaction is significant with very high level of significance, and the sphericity tail probability for the interaction was equal to 0.000. The sphericity assumption was not met, and the conservation test was used, and still the tail probability for the F - value of this test very high level of significance was equal to 0.001, with the Greenhouse & Geisser. Therefore, this means there was a very high level of significance for the interaction effects. To examine the trend analysis of this interaction. We found five trend components were significant in table no.2, with a very high levels of significance and also can see that very clear from the plot no.3 and plot no.4 of outlier values experiment, for the box width / box length, and box length /outlier values.

That means this interaction (LDO_{ijn}) was arise from the differences between all components of trends for these three variables. These variables all of them important for the experiment and effected on the subjects' judgments. In this experiment the subjects faced problems in their judgments, the reason might be that the joining of these three variables creates an interaction or perceptual problem as Lovie (1985) argued, and the visual illusion might then affect the subjects' judgments as Cleveland et al (1987) accepted.

And now let examine the two way interactions effects and start with the interaction of box length/box width interaction (LD_{ij}). It was found that the (F) value of this interaction effects(LD_{ij}) was equal to

4.222 and the tail probability for the usual (F) test was equal to (0.000). That means this interaction effects have a very high level of significance. But Checking the assumption for the interaction (LD_{ij}) by using the sphericity test the tail probability for this test revealed that the assumption of sphericity was not met, thus the conservative test provided an approximate test. But even with this test, still the interaction effects have a very high level of significance (0.000). Also it is very clear to recognize the interaction between these two variables from outlier values experiment plot no.4 the plot of box width/box length, and that would means this interaction very important on the subjects' judgments. And the second interaction the box length /outlier values (LO_{in}), we found the F-value of this interaction equal to (3.066)and the tail probability equal to (0.001), and that means with high level of significance, in the same time the sphericity test is not met ,but even with the conservative test still the interaction effects have a very high level of significance (0.005), and we can show that from plot no.3. And that would means this interaction very important on the subjects' judgments as the first one. And also the third interaction box width/outlier values(DO_{jn}) have a very high level of significance(0.000) and the F-value equal (4.135), and even after using the conservative test still the interaction effects have a very high level of significance (0.000). And also this interaction very important on the subjects' judgments as the first one and the second and all three variables important on the experiment by their interactions between them.

And to consider the results of the main effects of these three variables in Table no.2. It was found that the F-value of the box length is equal to 265.65 and value larger from all other values of the box length in all pervious experiments in box plot , and the tail probability of this test is equal to 0.000. The sphericity tail probability of the F-value is equal to 0.000, the assumption of the sphericity is not met .The conservation test should be used , and still the F-value has a very high level of significance higher than all the other in previous work in this area of box plot. One possible explanation of this result is that the box length of the standard box plot is the middle length levels of the other box lengths of the booklet, and also can be seen very clear from the plot no.3 and plot no.4, that the average of the absolute values of the errors increases with the middle length levels and decreases with shorter and longer levels. These results opposite to the results of the experiment of two variables M. M. Hussin(2007) and we can recognize that this variable very important by itself more than other variables

and also three variables experiments more difficult for the subjects than the two variables experiments . Now let us examine the trend test of this effect. It was found that three trend components are significant with very high levels of significance, this result can be seen from the two plot no.3 and plot no .4 and also from Table no. 2, and these components are responsible for the high significance of box length main effect. This variable very important for the experiment, and Wber's Law might help to explain the results.

To examine the other two main effects of box width and outlier values, it was found that the F-value of the box width was equal to 2.174 in Table no.2, and the tail probability for the F-value was equal to 0.093, the width variable was not significant. Now we don't need to use the conservative test because the variable main effect was not significant. And also the same things for the main effects of the outlier values the F-value equal to 0.529 and the significance level was not significant it was equal to (0.665). These two variables box width and outlier values were not important by itself for this experiment, but important in their interactions together with box length as we mention before, and their effect were not in obvious pattern as length variable as we seen from two plot no.3 and plot no.4. And these results were agree with Knight's (1982) results for the outlier values main effect two variables experiment box length/outlier values and there were no other experiment for outlier vales, but was disagree with Knight's (1982) results for the box width main effect two variables box length/box width and also disagree with M. M. Hussin (1989, 2006, 2007) two variables experiments ,but agree with M. M. Hussin (1989) box notch experiment three variables(box length, width, notch) group A experiments at Keele and Baghdad . One possible explanation of these results it might be that this experiment three variables and it was more difficult than the others experiments with two variables.

8- Conclusion

As we mentioned before, these two new experiments three variables investigate the three feature variations, box width, box notch, outlier values and the combinations between them ,together with the box length. We suggested in the future research points (1, 2, 3), Hussin, M.M. (1989), that the variations on a basic box plot are important and affected subjects' judgments. And changing the standard box plot length influence a subjects' perception of box length, and also changing the length levels of the box plots and other features from the experiments of Hussin, M.M. (1989) group B. We found that from the results of these two experiments, the conclusions of these two experiments can be summarized as follows:

1- These variations on a basic box plot very important to add more information to the box plot, but the coast more difficulties arises. Subjects in these two experiments with three variables, faced more difficulties than two variables experiments and some times the results opposite to them as we mentioned to them before and clear from plots (1, 2, 3, 4) if we compare with Hussin, M.M. (1989, 2006, 2007).

2- The three variables interactions of the box plot features are very important and impaired subjects' judgments of box length and make the subjects' judgments very difficult . And some time the results unexpected as we found the interactions between box width and outlier values and also the interaction between box length and outlier values, that might be as a visual illusion effects created by these interactions, or as Lovie (1985) put it perceptual problems with judging box plot.

3-The standard box plot length very important and influence subjects' judgments as wefound from this experiment results and the same results of Hussin, M.M.(2006,2007), and the results agrees with Hussin, M.M.(1989) suggestion in future research points(1, 2, 3).

4- The outlier values variable was a very important variable and influences a subjects' judgments of the box length, even this variable irrelevant variable in the experiment. But it is not by itself but by the interactions with other two variables in the experiment, even these interactions unexpected.

5- The box notch variable was a very important variable and influences a subjects' judgments of the box length. But it is not by itself but by the interactions with other two variables in the experiment, and in the same times these interactions expected

6- The box plot length (midspread) was the most important variable to affect subjects' judgments in the box plot, because the box length variable the relevant variable in the experiment so this result very reasonable , these results agree with Knight's (1982) results, and with Hussin, M. M.'s (1989,2006,2007) results for all box plot experiments.

7- The box width was a very important variable in the experiment, but it might be lease than to the box length and that's very fair and reasonable, and even some times with three variables experiments we found the width variable was not significant, but the interactions between these three variables were very important and make the subjects' judgments more difficult, as we mentioned before it might be create visual illusion.

BIBLIOGRAPHY

1- Baldwin, J.M. (1895). The effect of size contrast upon judgments of position in the retinal field, Psychological Review, 2, p.244-59. Cited in Robinson, 1972.

2- Chambers, J.M., Cleveland, W.S., Kleiner, B. and Tukey, D.A.(1983). Graphical methods for data analysis, Duxbury (soft cover), Boston and Wadsworth (hard cover), Belmont, California.

3- Charles S. Davis (2002).Statistical Methods for the Analysis of Repeated Measurements. Springer-Verlag New York, Inc.

4 - Cleveland, W.S., Harris, C.S. and McGill, R. (1982). Judgments of Circle Size on Statistical Maps, Journal of the American Statistical Association 77 p.541-547

5 -Cleveland, W.S. & McGill, R. (1984). Graphical perception: Theory, Experim- entation, and application to the development of graphical methods. Journal of the American Statistical Association, 79, p.531-554.

6 -Cleveland, W.S. & McGill, R. (1986). An experiment in graphical perception. International Journal of Machine Studies, 25, p.491-500.

7 - Cleveland, W.S. & McGill, R. (1987). Graphical perception: The visual decoding of quantitative information on graphical displays of data,
(with discussion). Journal of the Royal Statistical Society A, 150, part3, p. 192-229.

8 - Greenhouse, S.W. and Geisser, S. (1959). On methods in the analysis of profile data. Psychometrika, 24, p.95-112.

9 - Hussin, M. M. (1989). Some studies of a graphical method in statistical data analysis; subjective judgments in the interpretation of Box plots- Unpublished Ph.D. Thesis, Keele university , U.K. .

10- Hussin, M. M. (2005), Statistical Data Analysis Using SPSS, Amman, Dar Alhamed .

11- Hussin, M. M. (2006),Part one, An Experiments on the Box plot, Journal of Administration and Economy, Volume 12, number 42, P1-26.

12 - Hussin, M. M. (2007),Part two, An Experiments on the Box plot, Journal of Administration and Economy, Volume 13, number 45, P 315-332.

13 - Huynh, H. (1978). Some approximate test for repeated measurements designs. Psychometrika, Vol.43, No2, p. 161-175.

14 - Huynh, H. and Fedlt, L. (1970). Conditions under which mean square ration in repeated measurements designs have exact F-distributions. Journal of the American Statistical Association, Vol. 65, No.332, p. 1582-1589.

15 - Huynh, H. and Mandeville, G.K. (1979). Validity conditions in repeated measures designs. Psychological Bulletin, Vol.86, No.5, p.964 -973.

16- Knight, L.C. (1982). "Do additions to the box plot confuse the viewer "؟ Unpublished B Sc dissertation - University of Keele.

17 - Lovie, P.(1985). Subjective judgment in data analysis. Research Report, Department of Mathematics, University of Keele. Also presented in International Seminar on Cognitive processes in Mathematics, Keele, 1985.

18 - Lovie, A.D. & Lovie, P. (1987). Discussion of the paper by Drs Cleveland & Mc Gill. J.R.Statist. Soc. A., p.210-229.

19 - Loynes, R.M. (1987), Discussion of the paper by Drs Cleveland &McGill, J.R.Statist. Soc. A., p.210-229 .

20 - Mc Culloch, J. (1981). The variable- width box plot from different perspective. Unpublished B.A. dissertation, University of Liverpool.

21 - McGill, R., Tukey, J.W. and Larsen, W.A. (1978). "Variations of box plots", The American Statistician, 32, p. 12-16.

22 - Rogan, J.C., Keselman, H.J. and Mendoza, J.L. (1979).Analysis of repeated measurements. British Journal of Mathematical and statistical Psychology, 32, p.269-286

23 - Simkin, D. and Hastie, R. (1987).An information processing analysis of graph perception. Journal of the American Statistical Association, Vol.82, No.398, p.454-465.

24- Sim, C.H. ;Gan, F.F ;Chang, T.C.(2005). Outlier Labeling with Box plot Procedures. Journal of the American Statistical Association ,Volume 100, Number 470, June ,PP. 642-652.

25 - Stevens, S.S. (1975). Psychophysics. Wiley, New York.

26- Titchener, E.B. (1906). Experimental psychology. New York, Macmillan. Cited in Murch, G.M. (1973). Visual and Auditory perception, New York, The Bobbs-Mervill Company, Inc.

27 - Tukey, J.W. (1977). Exploratory data analysis. Reading mass: addison- Wesley.

28 - Winer, B.J.(1962). Statistical Principles in Experimental Design. New York: Mc Graw-Hill.

Box notch Experiment No.1 Table no.1

Mauchly's Test of Sphericity

Measure: errors

Within Subjects Effec	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon Huynh-Feldt
length	.566	31.114	5	.000	.819
width	.868	7.758	5	.170	.978
notch	.675	21.497	5	.001	.835
length * width	.310	61.832	44	.040	.950
length * notch	.217	80.754	44	.001	.904
width * notch	.435	43.950	44	.477	.982
length * width * notch	.000	558.458	377	.000	.791

Tests of Within-Subjects Effects

Measure: errors

Source		F	Sig.
length	Sphericity Assumed	19.88	.00
	Greenhouse-Geisser	19.88	.00
	Huynh-Feldt	19.88	.00
width	Sphericity Assumed	3.427	.02
	Greenhouse-Geisser	3.427	.02
	Huynh-Feldt	3.427	.02
notch	Sphericity Assumed	.686	.56
	Greenhouse-Geisser	.686	.53
	Huynh-Feldt	.686	.54
length * width	Sphericity Assumed	2.992	.00
	Greenhouse-Geisser	2.992	.00
	Huynh-Feldt	2.992	.00
length * notch	Sphericity Assumed	4.309	.00
	Greenhouse-Geisser	4.309	.00
	Huynh-Feldt	4.309	.00
width * notch	Sphericity Assumed	3.113	.00
	Greenhouse-Geisser	3.113	.00
	Huynh-Feldt	3.113	.00
length * width * notch	Sphericity Assumed	2.866	.00
	Greenhouse-Geisser	2.866	.00
	Huynh-Feldt	2.866	.00

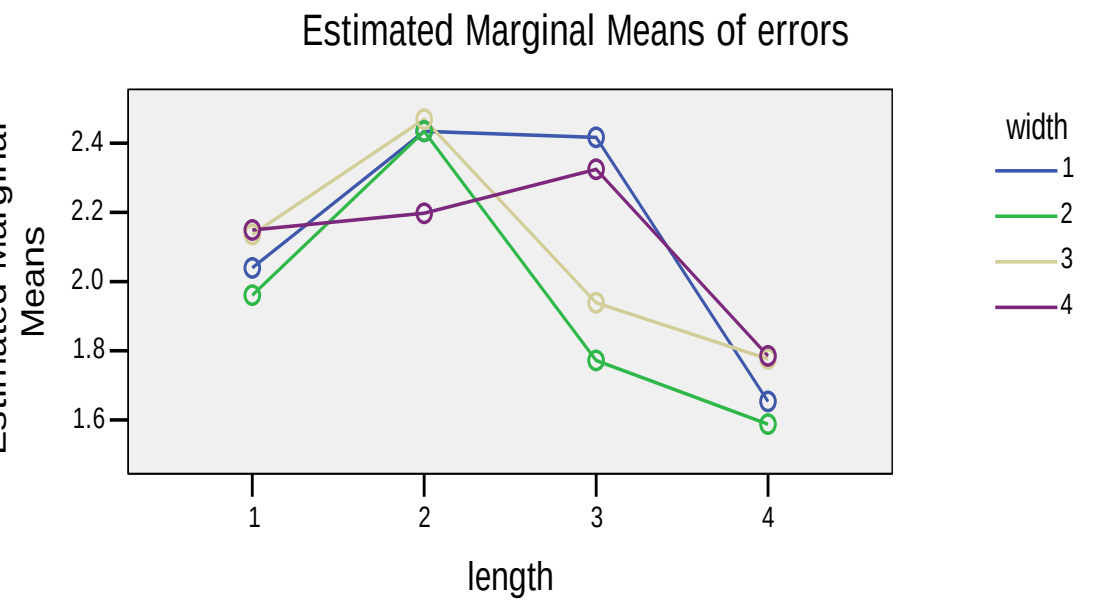
Tests of Within-Subjects Contrasts

Measure: errors

Source	length	width	notch	F	Sig.
length	Linear			44.218	.000
	Quadratic			22.590	.000
width		Quadratic		5.154	.027
		Cubic		3.494	.067
length * width	Quadratic	Linear		3.276	.076
	Cubic	Quadratic		19.874	.000
length * notch	Linear		Linear	5.955	.018
			Quadratic	23.965	.000
			Cubic	4.420	.040
	Cubic		Cubic	5.610	.021
width * notch		Linear	Cubic	4.281	.043
		Quadratic	Quadratic	4.681	.035
			Cubic	6.626	.013
		Cubic	Linear	3.724	.059
			Cubic	5.638	.021
length * width * notch	Quadratic	Linear	Linear	12.925	.001
		Cubic	Linear	10.416	.002
	Cubic	Quadratic	Quadratic	12.891	.001
		Cubic	Linear	.008	.929
			Quadratic	3.736	.058
			Cubic	8.237	.006

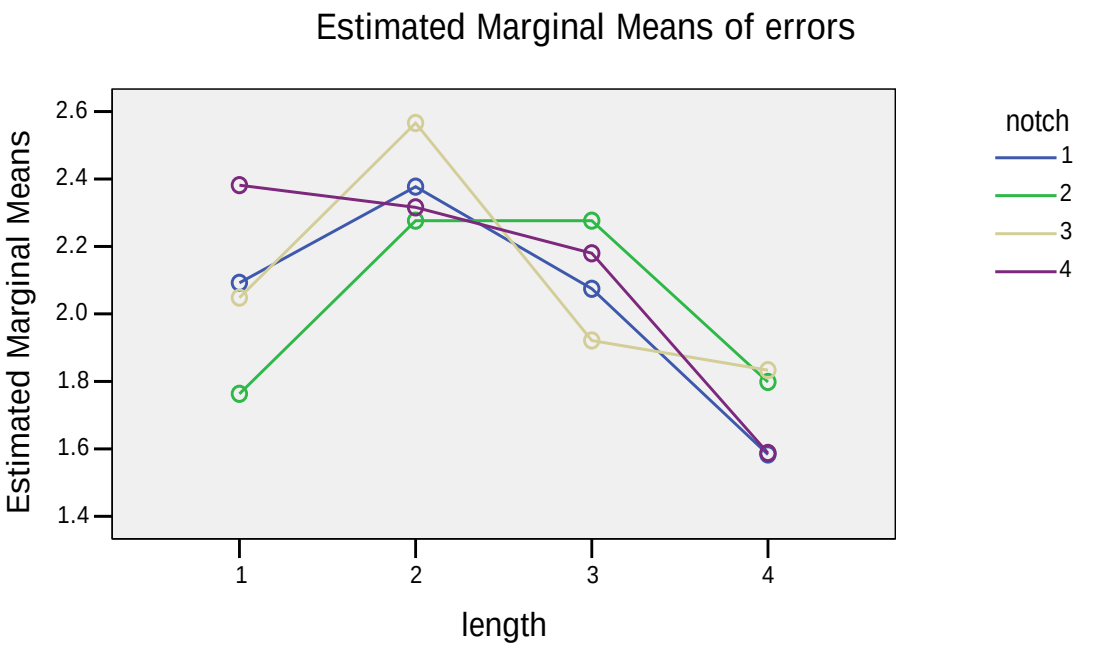
Plot

no.1Length&Width



Plot no.2

Length&Notch



Outlier Values Experiment no.2 Table no.2

Mauchly's Test of Sphericity

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon Huynh-Feldt
length	.375	47.757	5	.000	.621
width	.814	10.007	5	.075	.947
outliers	.921	4.022	5	.546	1.000
length * width	.229	69.011	44	.010	.855
length * outliers	.171	82.762	44	.000	.858
width * outliers	.246	65.711	44	.019	.913
length * width * outliers	.000	556.454	377	.000	.755

Tests of Within-Subjects Effects

Measure: MEASURE_1

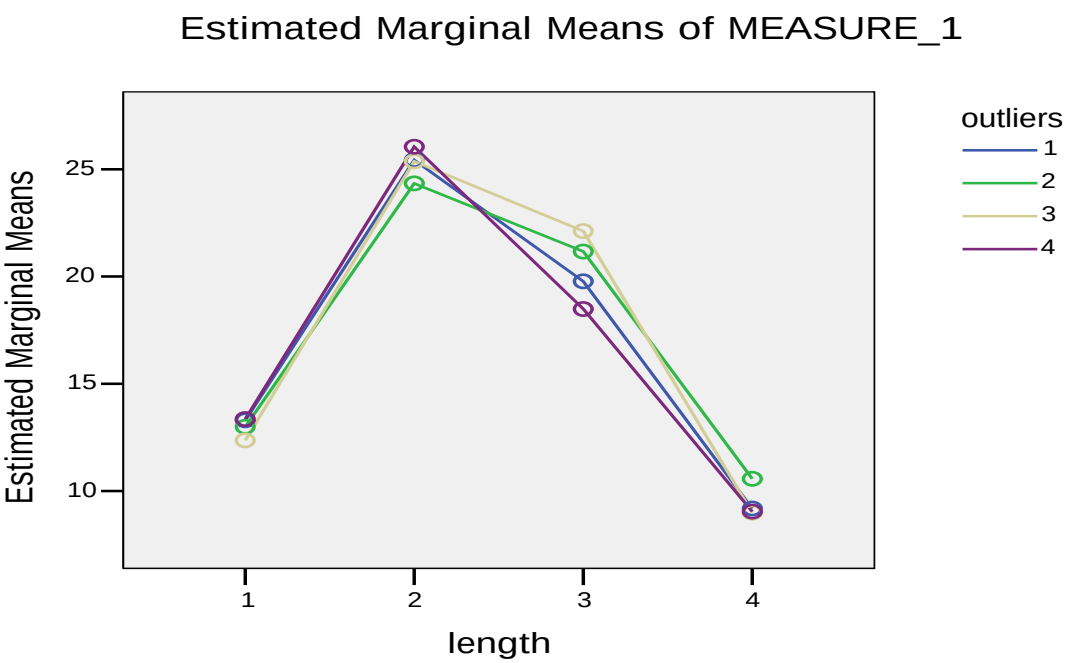
Source		F	Sig.
length	Sphericity Assumed	265.6	.000
	Greenhouse-Geisser	265.6	.000
	Huynh-Feldt	265.6	.000
	Lower-bound	265.6	.000
width	Sphericity Assumed	2.174	.093
	Greenhouse-Geisser	2.174	.101
	Huynh-Feldt	2.174	.097
outliers	Sphericity Assumed	.529	.663
	Greenhouse-Geisser	.529	.654
length * width	Sphericity Assumed	4.221	.000
	Greenhouse-Geisser	4.221	.000
	Huynh-Feldt	4.221	.000
	Lower-bound	4.221	.045
length * outliers	Sphericity Assumed	3.066	.001
	Greenhouse-Geisser	3.066	.005
	Huynh-Feldt	3.066	.003
width * outliers	Sphericity Assumed	4.135	.000
	Greenhouse-Geisser	4.135	.000
	Huynh-Feldt	4.135	.000
length * width * outliers	Sphericity Assumed	2.731	.000
	Greenhouse-Geisser	2.731	.001
	Huynh-Feldt	2.731	.000

Tests of Within-Subjects Contrasts

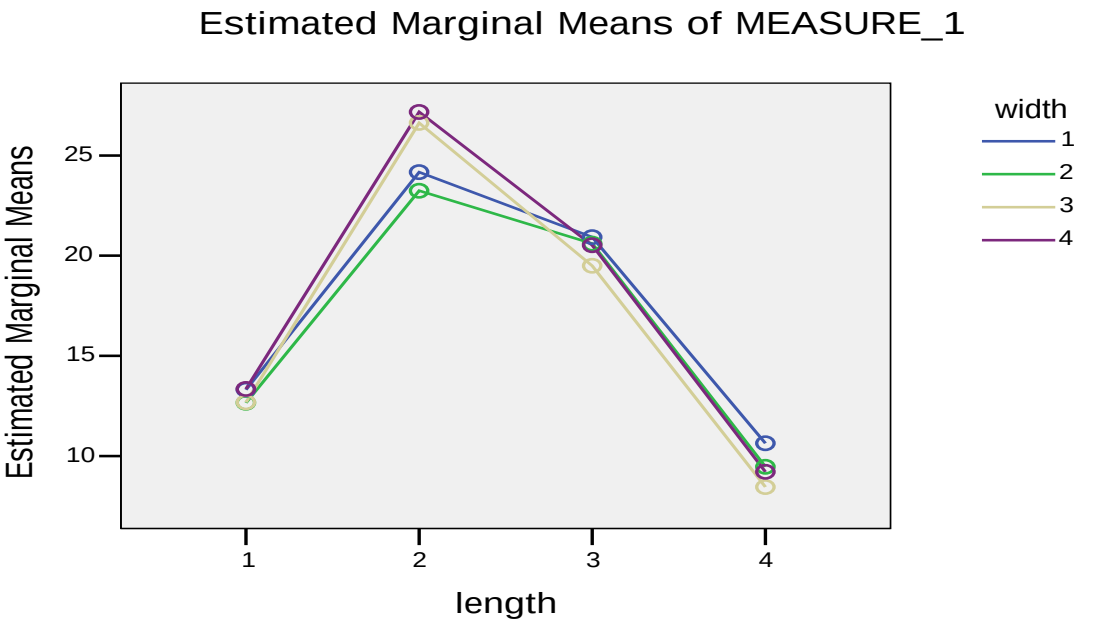
Measure: MEASURE_1

Source	length	width	outliers	F	Sig.
length	Linear			108.13	.000
	Quadratic			383.04	.000
	Cubic			55.490	.000
width		Quadratic		5.615	.022
length * width	Linear	Linear		7.777	.007
	Quadratic	Linear		7.700	.008
	Cubic	Linear		7.696	.008
		Cubic		7.255	.010
length * outliers	Linear		Quadratic	11.073	.002
	Quadratic		Cubic	6.176	.016
	Cubic		Quadratic	9.009	.004
width * outliers		Linear	Linear	4.783	.033
			Cubic	9.832	.003
		Quadratic	Linear	8.862	.004
			Quadratic	11.943	.001
length * width * outliers	Linear	Linear	Cubic	10.880	.002
	Quadratic	Linear	Quadratic	16.979	.000
		Quadratic	Cubic	9.949	.003
		Cubic	Quadratic	9.213	.004
	Cubic	Cubic	Cubic	4.913	.031

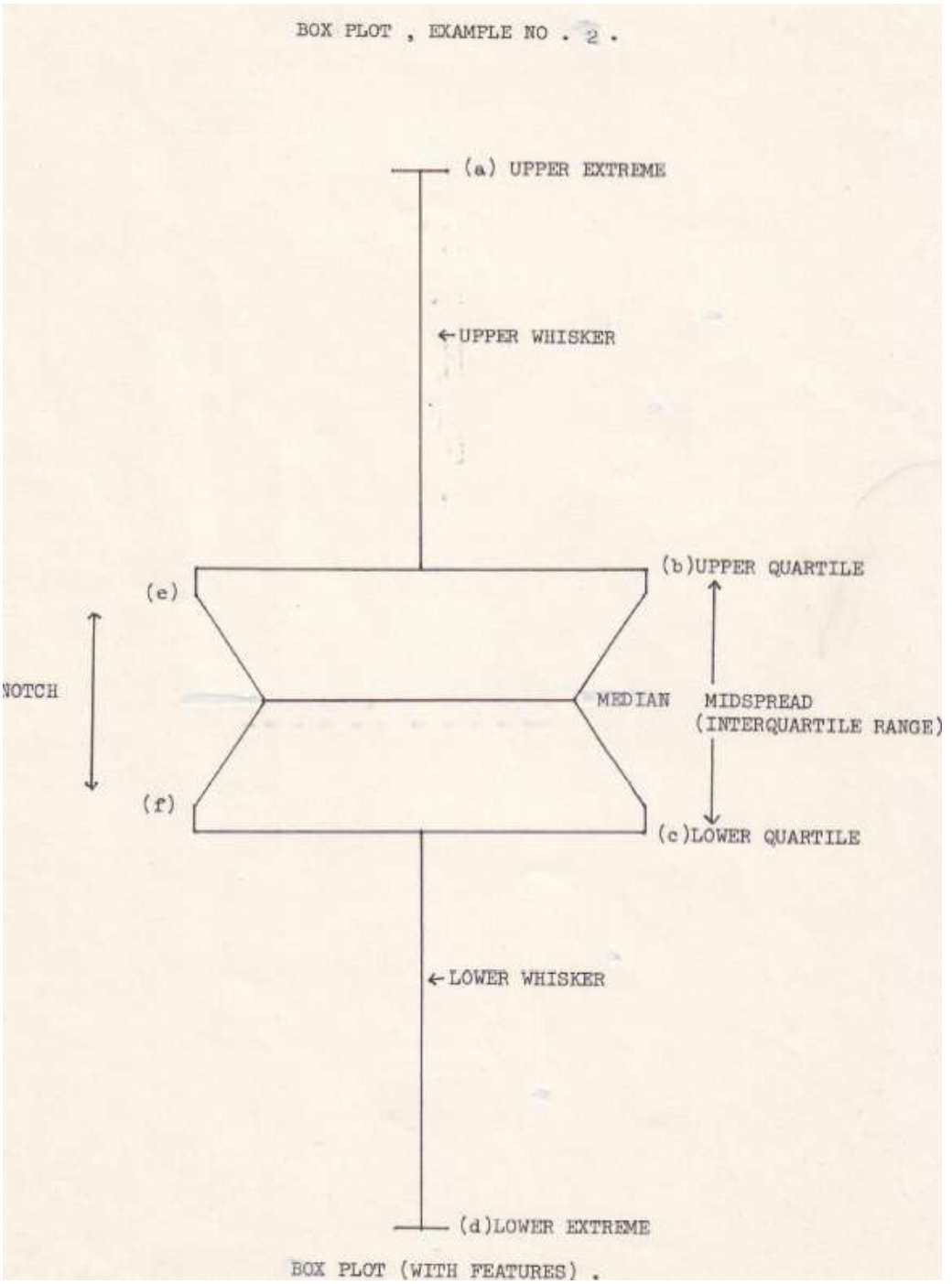
Plot no.3 box length&outlier values



Plot no.4 box length& box width


Estimated Marginal Means of MEASURE_1
Estimated Marginal Means
25
20
15
10
1
2
3
4
length
width
1
2
3
4

شكل رقم (1A)


BOX PLOT , EXAMPLE NO . 2 .
(a) UPPER EXTREME
←UPPER WHISKER
(b)UPPER QUARTILE
(e)
NOTCH
MEDIAN
MIDSPREAD
(INTERQUARTILE RANGE)
(f)
(c)LOWER QUARTILE
←LOWER WHISKER
(d)LOWER EXTREME
BOX PLOT (WITH FEATURES) .