برنامج التحليل الإحصائي SPSS

مقدمة إلى بيئة SPSS: استكشاف البيانات EXPLORE

رابط الدورة على الإنترنت : https://library.wpu.edu.sy/?page_id=7028&lang=ar

لمزيد من الدورات التدريبية: https://library.wpu.edu.sy/?page_id=7019&lang=ar

لأي استفسار: Waseem.ramadan@wpu.edu.sy

جدول المحتويات

جدول الأشكال 2

1 استكشاف البيانات Explore في SPSS 3

1 - 1 شرح مربع الحوار الخاص باستكشاف البيانات Explore 3

1 - 2 فحص التوزيع الطبيعي للمتغير Normality plots with tests 7

1 - 2 - 1 كولموجوروف-سميرنوف (KS) 7

1 - 2 - 2 شابيرو-ويلك 8

1 - 3 أنواع مخططات boxplot في إجراء استكشاف البيانات Explore 10

1 - 3 - 1 اثنين من المتغيرات الرقمية، عدم وجود عوامل 11

1 - 3 - 2 اثنين من المتغيرات الرقمية، وعامل واحد متغير 13

Dr. Wassim Ramadan For SPSS Learning @ WPU University

جدول الأشكال

شكل 1 : مسار القائمة لأمر Explore.....3
شكل 2: مربع حوار الإجراء Explore الخاص باستكشاف المتغيرات المستمرة.....4
شكل 3 :الاحصاءات الوصفية المتاحة مع إجراء استكشاف البيانات Explore.....5
شكل 4: المخططات البيانية المتاحة مع إجراء استكشاف البيانات Explore.....6
شكل 5: الخيارات Options المتاحة مع إجراء استكشاف البيانات Explore.....7
شكل 6: مخطط Normal Q-Q لفحص التوزيع الطبيعي للبيانات.....9
شكل 7: مخطط Deterended Normal QQ للتوزيع الطبيعي.....10
شكل 8 :مخطط Boxplot لاثنين من المتغيرات الرقمية بدون عوامل وفي حال استخام مستويات المعامل معاً ..11
شكل 9: مخطط Boxplot لاثنين من المتغيرات الرقمية بدون عوامل وفي حال استخدام المتغيرات المستقلة معاً 12
شكل 10: مخطط Boxplot لاثنين من المتغيرات الرقمية مع عوامل وفي حال استخام مستويات المعامل معاً ...14
شكل 11: مخطط Boxplot لاثنين من المتغيرات الرقمية مع عوامل وفي حال استخدام المتغيرات المستقلة معاً.15

Dr. Wassim Ramadan For SPSS Learning © VPU University

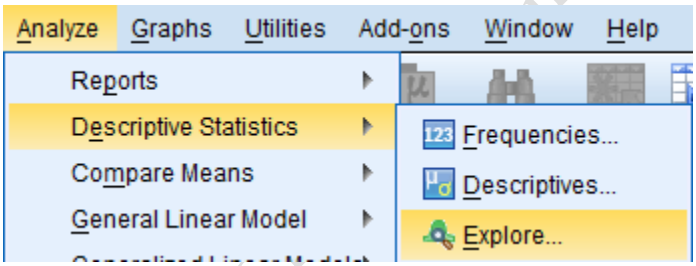
1 استكشاف البيانات Explore في SPSS

يستخدم الإجراء Explore لإعطاء إحصاءات ورسوم بيانية لعينة بأكملها، أو مجموعات فرعية منها. يتم اعتماد المتغيرات الرقمية المستمرة والمعرفة على أنها من النوع مقياس Scale. يمكن أيضا استخدام الإجراء Explore لتقييم التوزيع الطبيعي لمتغير رقمي مع بعض الإحصاءات الاستقرائية (احصاء استدلالي) الخاصة والمخططات البيانية التشخيصية المفصلة.

1 - 1 شرح مربع الحوار الخاص باستكشاف البيانات Explore

لتنفيذ إجراء استكشاف البيانات Explore، انقر فوق **تحليل <الاحصائيات الوصفية <استكشاف** . شكل 1

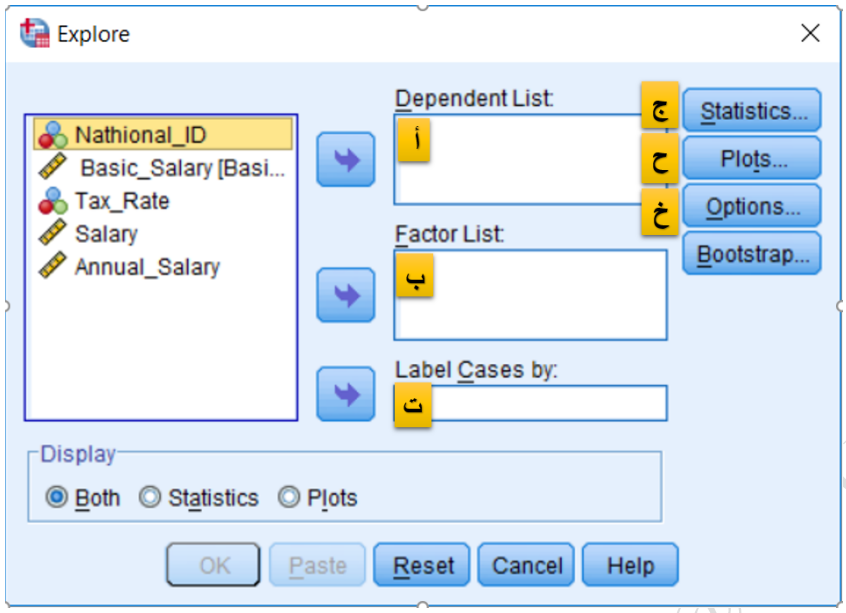
Analyze > Descriptive Statistics > Explore.



شكل 1: مسار القائمة لأمر Explore

فيظهر مربع الحوار الخاص بالإجراء Explore. شكل 2

Dr. Wassim Ramadan For Statistics @ VPU University



شكل 2: مربع حوار الإجراء Explore الخاص باستكشاف المتغيرات المستمرة

أ- **قائمة المتغيرات التابعة :** المتغيرات الرقمية المستمرة التي ترغب في تحليلها.

ب- **قائمة العوامل** (اختياري): متغير أو مجموعة متغيرات فئوية تستخدم لتجزيء البيانات بواسطتها إلى مجموعات جزئية .إذا تم تحديد متغير في هذا حقل المتغيرات الفئوية، سيقوم إجراء Explore بإنتاج ملخصات إحصائية فردية للمتغير الرقمي بمقابل كل فئة من فئات المتغير الفئوي.

ت- **تسمية الحالات بواسطة** (اختياري): متغير يعبر عن ترقيم أو تسمية لكل حالة. تستخدم هذه الأسماء في تسمية الحالات المتطرفة في البيانات. .إذا لم يتم تحديد هذا الحقل، سوف يستخدم SPSS رقم الصف لتسمية هذه الحالات المتطرفة.

ج- **الاحصائيات :**خيارات اختيارية تتيح الحصول على تقرير عن الإحصاءات الوصفية للمتغير المدروس. يوضح الشكل 3الخيارات الممكنة وهي:

a. **Descriptives** الإحصاءات الوصفية (مفعلة افتراضيا). وهي
   i. المتوسط Mean
   ii. 5% trimmed mean
   iii. الوسط Median
   iv. التباين Variance
   v. الانحراف المعياري Standard deviation
   vi. القيمة العظمى والصغرى Minimum And Maximum
   vii. المدى Range

Dr. Wassim Ramadan For SPSS Learning @ U University

viii. المجال بين الشرائح الربعية Interquartile range (IQR)

ix. الالتواء والتفرطح Skewness, And Kurtosis

x. الأخطاء المعيارية للمتوسط، الالتواء والتفرطح Standard errors for the mean skewness and kurtosis

b. **M-estimators**

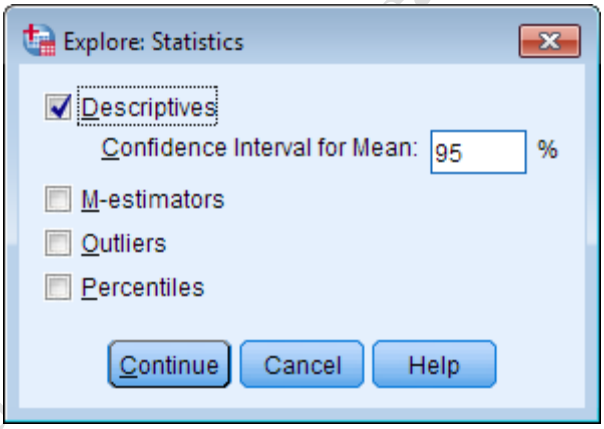
تتيح الحصول على بدائل عن المتوسط والوسط

c. Outliers القيم المتطرفة

تتيح الحصول على الخمس قيم المتطرفة العليا والخمسة الدنيا. إذا تم تعبئة الحقل الاختياري المرمز ب **تسمية الحالات بواسطة** label cases by ستستخدم تسمية هذه القيم المتطرفة بواسطته.

d. **Percentiles** النسب المئوية

يتيح هذا الخيار الحصول على حدود النسبة المئوية التالية: 5، 10، 25، 50، 75، 90، و95 للبيانات.

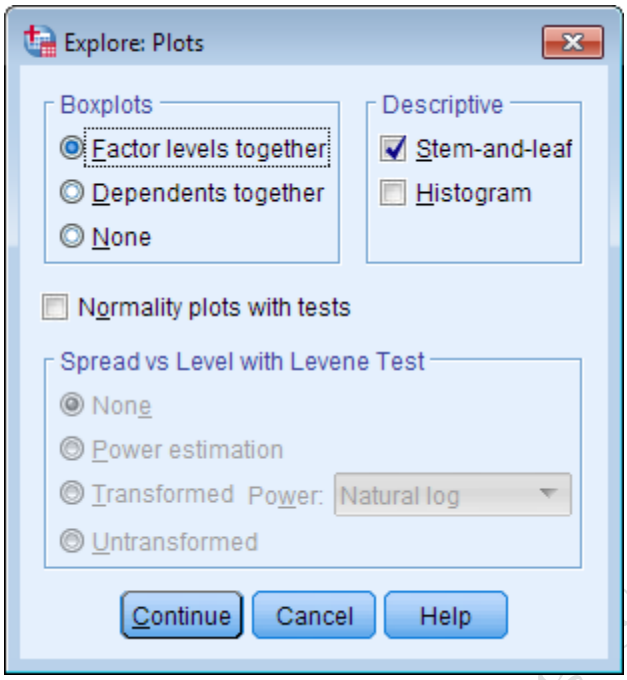


شكل 3: الاحصاءات الوصفية المتاحة مع إجراء استكشاف البيانات Explore

ح- **الرسوم البيانية** خيارات اختيارية تتيح الحصول على الرسوم والمخططات البيانية للمتغير المدروس. يوضح الشكل 4شكل 3الخيارات الممكنة.

يسمح هذا الخيار أيضا بالحصول على اختبارالتوزيع الطبيعي للمتغير **Normality plots with tests** . تشمل الخيارات الأخرى إنشاء المخطط البياني boxplots والمخططات الجذعية والورقية Stem-and-leaf، ومخطط المدرج التكراري Histogram. افتراضيا، يُنشأ إجراء الاستكشاف مخطط boxplots والمخططات الجذعية والورقية لكل متغير رقمي مستمر.

Dr. Wassim Ramadan Learning @ WPU University



شكل 4: المخططات البيانية المتاحة مع إجراء استكشاف البيانات Explore

تحمل الخيارات في منطقة Boxplots معنى فقط إذا كنت قد حددت أكثر من متغير مستمرلدراسته ، أو إذا كنت قد حددت متغير عامل واحد على الأقل. وهي :

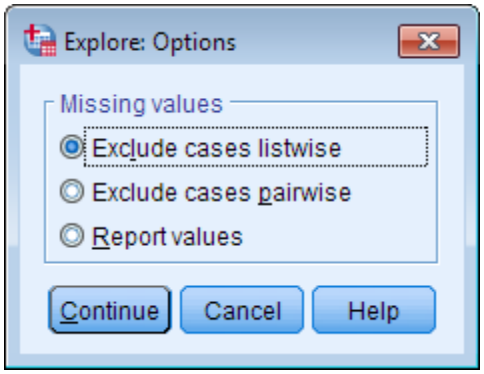
- **مستويات العامل معا Factor levels together** : سوف تُنشأ رسوم بيانية منفصلة لكل متغير مستمر .إذا تم تحديد متغير عامل، سوف تظهر مجموعات أو فئات المتغير العامل معاً في كل مخطط boxplot .
- **المتغيرات المستقلة معا Dependents together:** سيضع boxplots كل المتغيرات المستمرة على نفس الرسم البياني .إذا تم تحديد متغير عامل، سوف تظهر مجموعات أو فئات المتغير العامل معاً في كل مخطط boxplot .

خيار فحص التوزيع الطبيعي للمتغير **Normality plots with tests:** يتيح الخيار الحصول على كل من الاختبارات الإحصائية الاستدلالية (الاستقرائية) للتوزيع الطبيعي مع إنشاء مخطط التوزيع الطبيعي للمتغير أيضاً. لاحظ أنه ليس من الممكن طلب الإحصاءات دون إنشاء المخطط، أو العكس بالعكس . على وجه التحديد، فإنه سوف يجري اختبار كولموغوروف-سميرنوف Kolmogorov-Smirnov ، واختبار شابيرو ويلك Shapiro-Wilk، ومخططات QQ.

إذا تم تحديد متغير عامل، يمكنك استخدام اختبار **انتشار مقابل مستوى مع اختبار ليفين Spread vs Level with Levene Test** : اختبار ليفين لتجانس التباين (أي التباين المستمر عبر مستويات المتغير العامل).

Dr. Wasim Ramadan For SPSS Learning © WPU University

خ- **خيارات :**التحكم في كيفية التعامل مع **القيم المفقودة Missing Values:** افتراضيا، يتم استخدام أسلوب الإقصاء listwise. انظر الشكل 5



شكل 5: الخيارات Options المتاحة مع إجراء استكشاف البيانات Explore

- **استبعاد Listwise** سوف يستبعد من التحليل كل الحالات التي تتضمن قيم مفقودة لأي من المتغيرات المدروسة. في هذه الحالة فإن عدد الحالات الصالحة لكل متغير هي نفسها لباقي المتغيرات.
- **استبعاد Pairwise** سوف يستبعد من التحليل الحالات التي تتضمن قيم مفقودة لكل متغير من المتغيرات المدروسة على حدا وبشكل مستقل عن باقي المتغيرات. في هذه الحالة فإن عدد الحالات الصالحة لكل متغير ليست بالضرورة متساوية مع باقي المتغيرات (بمعنى أخر فهي قد تكون مختلفة).
- **تقرير القيم Report Values**: يؤثر فقط على التحليلات التي تشـــمل على متغير عامل .إذا تم تحديد هذا الخيار، سيتم التعامل مع الحالات التي تتضمن قيم مفقودة للمتغير العامل على أنها فئة بحد ذاتها.

1 - 2 فحص التوزيع الطبيعي للمتغير Normality plots with tests

عندما يتم تحديد خيار فحص التوزيع الطبيعي للمتغير **Normality plots with tests** ، يتم إضافة جدول التوزيع الطبيعي ومخططي **QQ** إلى صفحة النتائج.

يحوي جدول **اختبارات التوزيع الطبيعي** على اثنين من الاختبارات الفرضـــية المختلفة للتوزيع الطبيعي : اختبار كولموغوروف-سميرنوف Kolmogorov-Smirnov ، واختبار شابيرو ويلك Shapiro-Wilk،

1 - 2 - 1 كولموجوروف-سميرنوف (KS)

هو اختبار لابارامتري nonparametric test . ومن الناحية التقنية يمكن اســـتخدامه لاختبار فيما إذا كان للبيانات توزيع محدد معروف أو لا (وليس فقط التوزيع الطبيعي)

Dr. Wassim Ramadan For SPSS Learning @WPU University

- الفرضية الأساسية (العدم) هي أن البيانات تأتي من توزيع محدد؛
- أما الفرضية البديلة هي أن البيانات لا تأتي من التوزيع المحدد .

يقوم SPSS بتحديد واستكشاف التوزيع الطبيعي تلقائيا ولاداعي للتفكير بتحديد نوع التوزيع المطلوب اختباره.

1 - 2 - 2 شابيرو-ويلك

هو اختبار بارامتري parametric test.

- الفرضية الأساسية (العدم) هي أن العينة تم استخلاصها من التوزيع الطبيعي
- الفروضية البديلة هي أن العينة لا تتبع التوزيع الطبيعي.

المعايير المستخدمة لرفض أو عدم رفض فرضية العدم هو نفسه لكل الاختبارات

- إذا كانت القيمة p أصغر من α (مستوى المعنوية، عادة = 0.05) نرفض فرضية العدم .حيث أن هناك أدلة كافية على أن البيانات لا تتبع التوزيع الطبيعي.
- إذا كانت القيمة p أكبر من α، لا نرفض فرضية العدم .و نقول أنه لا توجد أدلة كافية لاستنتاج أن البيانات لاتتبع التوزيع الطبيعي.

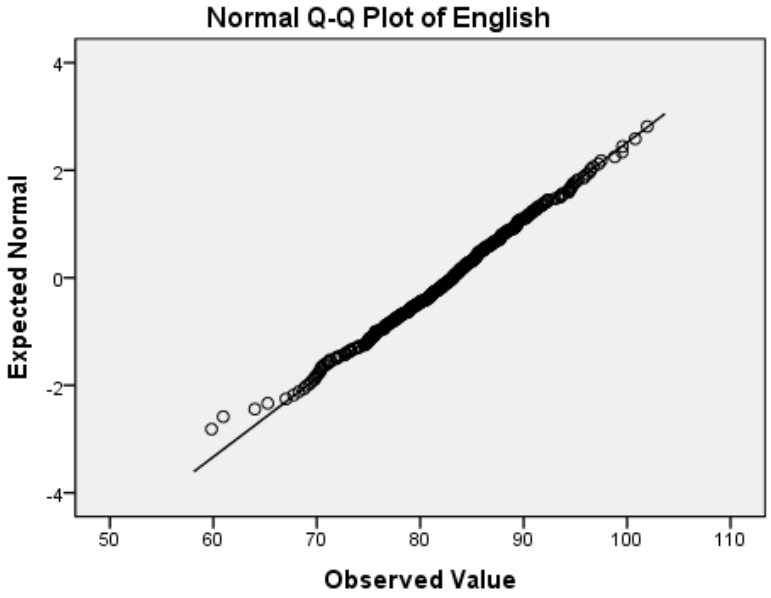
لاحظ أن هناك قيم *p*-values منفصلة لاختبار KS مقابل اختبار شابيرو-ويلك .في الواقع، من الممكن لهذه الاختبارات أن تتوصل إلى نتائج مختلفة.

وللتأكيد، **القيمة الاحتمالية الصغيرة بما فيه الكفاية لـــ P-value تنطوي على ، ولكن لا تثبت، أن البيانات لا تتبع التوزيع الطبيعي** .وفي المقابل، **القيمة الاحتمالية الكبيرة لـــ P-value لا تثبت أن البيانات لاتأتي من توزيع طبيعي .**ونقول أنه ليس هناك أدلة كافية لإقناعنا أن البيانات غير موزعة بشكل طبيعي.

بشكل عام، يجب أن لا تعتمد على نتائج هذه الاختبارات الفرضية وحدها .يجب عليك دائما استخدام مصادر أخرى (وخاصة الرسوم البيانية) لدعم أو دحض ادعاء التوزيع الطبيعي .يجب عليك أن تنظر أيضا إلى حجم العينة وإلى أي خلفية علمية عن مشكلتك. ضع في اعتبارك أن اختبار شابيرو-ويلك حساس جدا للانحرافات حتى التافهة في التوزيع الطبيعي عندما يكون حجم العينة كبير .بمعنى أخر، إذا كان حجم العينة كبيرة جدا، فمن الممكن تماما أن تظهر قيمة P-Value كبيرة جداً في اختبار شابيرو-ويلك، حتى لو كانت الانحرافات عن الوضع الطبيعي صغيرة جداً.

Dr. Wassim Ramadan For SPSS Learning @WPU University

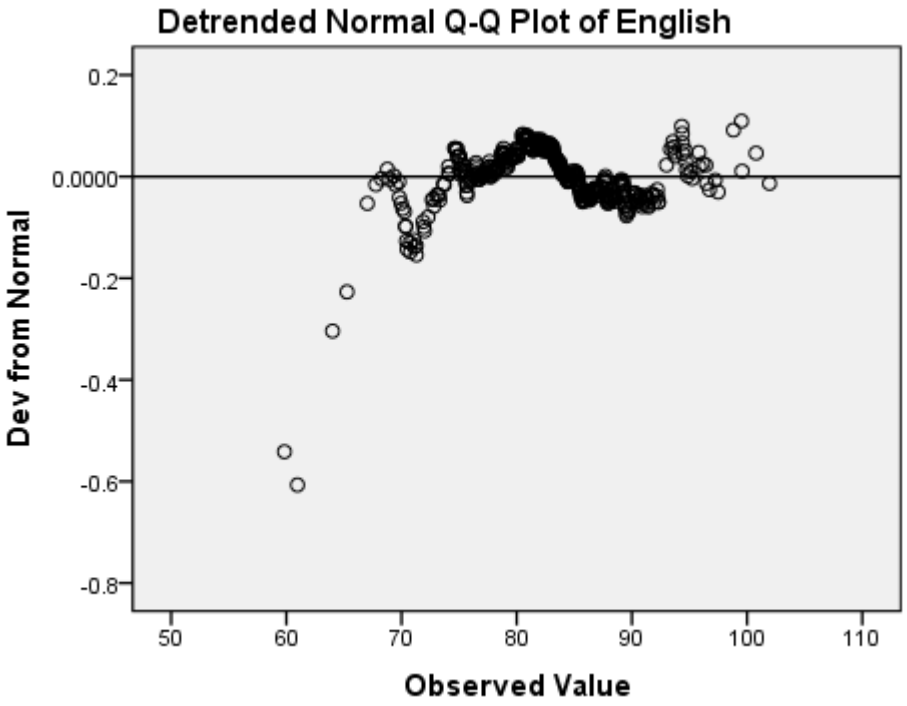
بعد إجرار **اختبارات التوزيع الطبيعي** ، تظهر مخططات **Normal Q-Q (or Quantile-Quantile).** شكل 6



شكل 6: مخطط Normal Q-Q لفحص التوزيع الطبيعي للبيانات

يبين المخطط QQ العادي أو Quantile-Quantile (كما هو مبين في شكل 6) مقارنة بين التجزيئات الملاحظة للبيانات (على شـــكل نقط ودوائر) مع التجزيئات المتوقعة في حال كانت البيانات تتبع التوزيع الطبيعي (على شـــكل خط مستقيم). إذا كان توزيع البيانات طبيعياً فإن النقط ستكون على أو بقرب الخط المستقيم. عند النظر إلى مخطط QQ ، يجب أن نبحث عن النقاط التي تبتعد عن خط القيم المتوقعة، فضلا عن الاتجاهات في القيم الملاحظة.

Dr. Wassim Ramadan For SPSS Learning
PU University



شكل 7: مخطط Deterended Normal QQ للتوزيع الطبيعي

يظهر مخطط **Detrended Normal QQ** نفس المعلومات (شكل 7) كالمخطط السابق، ولكن بطريقة مختلفة . يمثل الخط الأفقي ما نتوقع أن نرى فيما اذا كانت البيانات تتبع التوزيع الطبيعي. تمثل النقاط مقدار *واتجاه* الانحراف في الملاحظات عن التوزيع الطبيعي.

بالإضافة إلى الاختبارات الفرضية ومخططات QQ ، من الجيد لنظر في مخطط **الرسم البياني boxplot** للبيانات الخاصة بك. تعطي مخططات Boxplots نظرة أفضل على القيم المتطرفة ومواقع التجزيئات الخاص بك؛ ويمكن لكل من هذه الأدوات أن تساعدك في اتخاذ القرار فيما إذا كان هناك خروج عن التوزيع الطبيعي في البيانات الخاصة بك، أو إذا كانت كبيرة بما فيه الكفاية لتبرير قلقك من اعتبارها تتبع التوزيع الطبيعي.

1 - 3 أنواع مخططات boxplot في إجراء استكشاف البيانات Explore

في مربع حوار استكشاف: البيانات Explore وبالتحديد في نافذة المخططات (راجع الحرف ح في الشكل 2)، هناك ثلاثة خيارات عرض مختلفة لـ boxplots. – لمزيد من الشرح فقرة ح صفحة 5.

- **مستويات العامل معا Factor levels**
- **المتغيرات المستقلة معا Dependents together**
- **لا يوجد None**

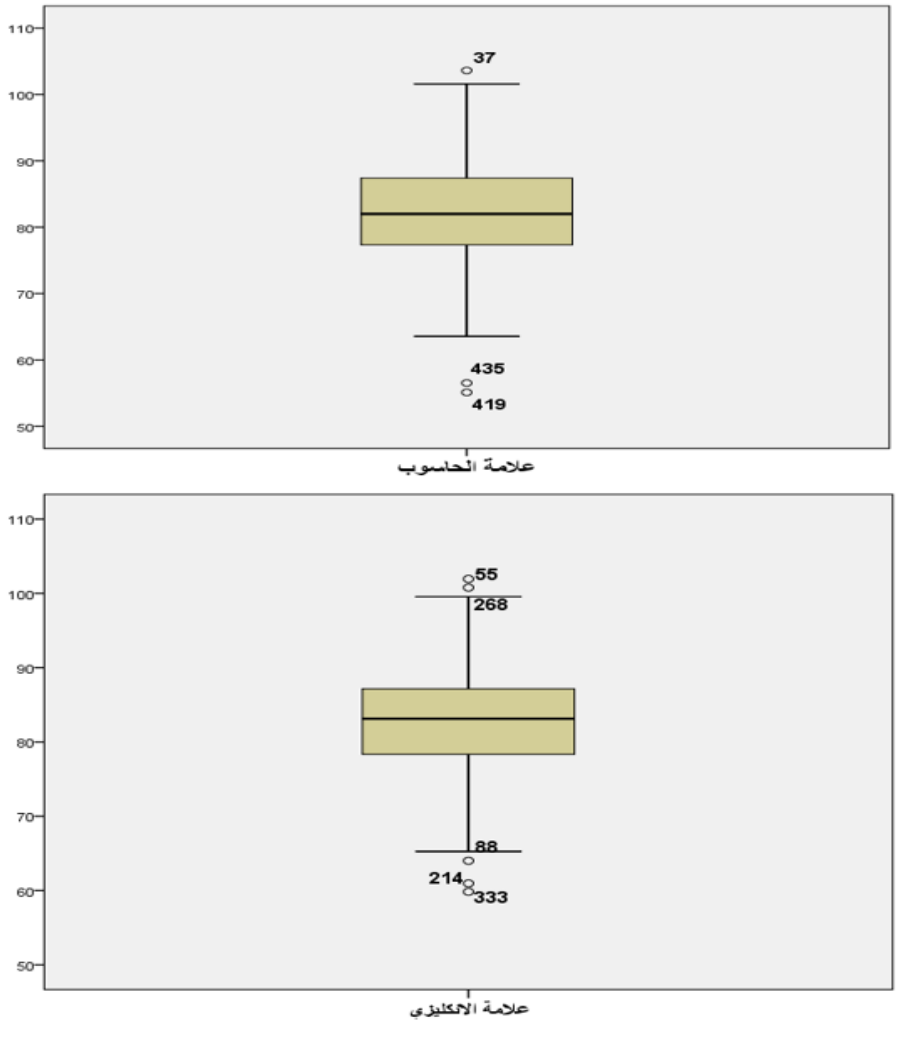
Dr. Wasim Ramadan for SPSS Learning, University

تؤثر هذه الإعدادات فقط على التحليلات التي تشمل اثنين أو أكثر من المتغيرات الرقمية

1 - 3 - 1 اثنين من المتغيرات الرقمية، عدم وجود عوامل

أولاً: مستويات العامل معا Factor levels

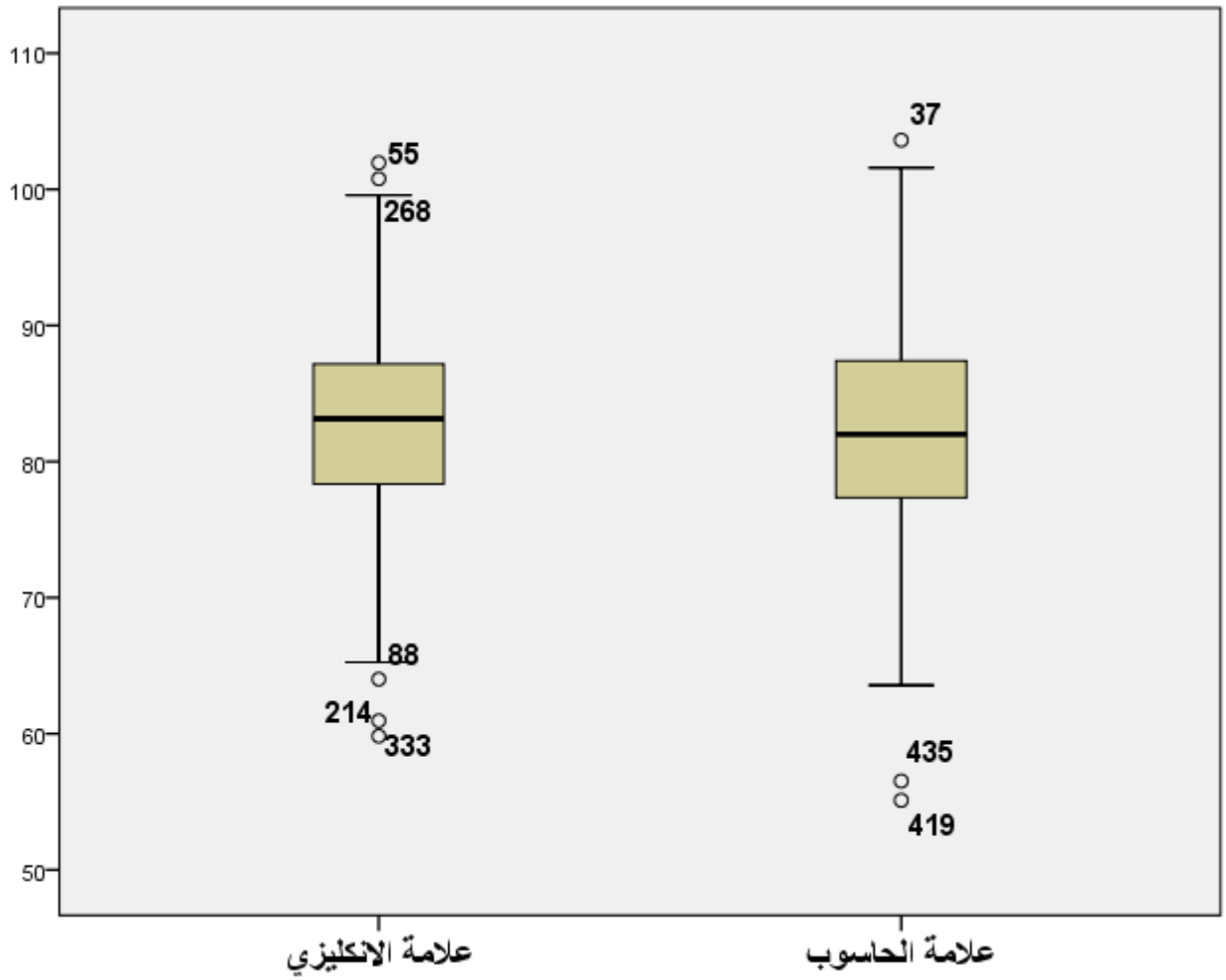
إنشاء رسوم بيانية boxplots منفصلة لكل متغير رقمي. يفيد هذ الخيار إذا كنت لا تحتاج لمقارنة المتغيرات الرقمية الخاصة بك مع بعضها البعض، أو إذا كانت وحدة القياس مختلفة لكل منها.



شكل 8: مخطط Boxplot لاثنين من المتغيرات الرقمية بدون عوامل وفي حال استخدام مستويات المعامل معاً

ثانياً: المتغيرات المستقلة معا Dependents together

إنشاء رسم بياني boxplot واحد لكل المتغير الرقمية لمقارنتها جنباً إلى جنب (على سبيل المثال، قبل الاختبار وبعد الاختبار للمتغيرات) .يجب أن تكون وحاد القياس للمتغيرات الرقمية المستمرة قريبة من بعضها البعض لتحقيق أفضل النتائج . من الناحية المثالية، ينبغي استخدام نفس وحدات القياس.



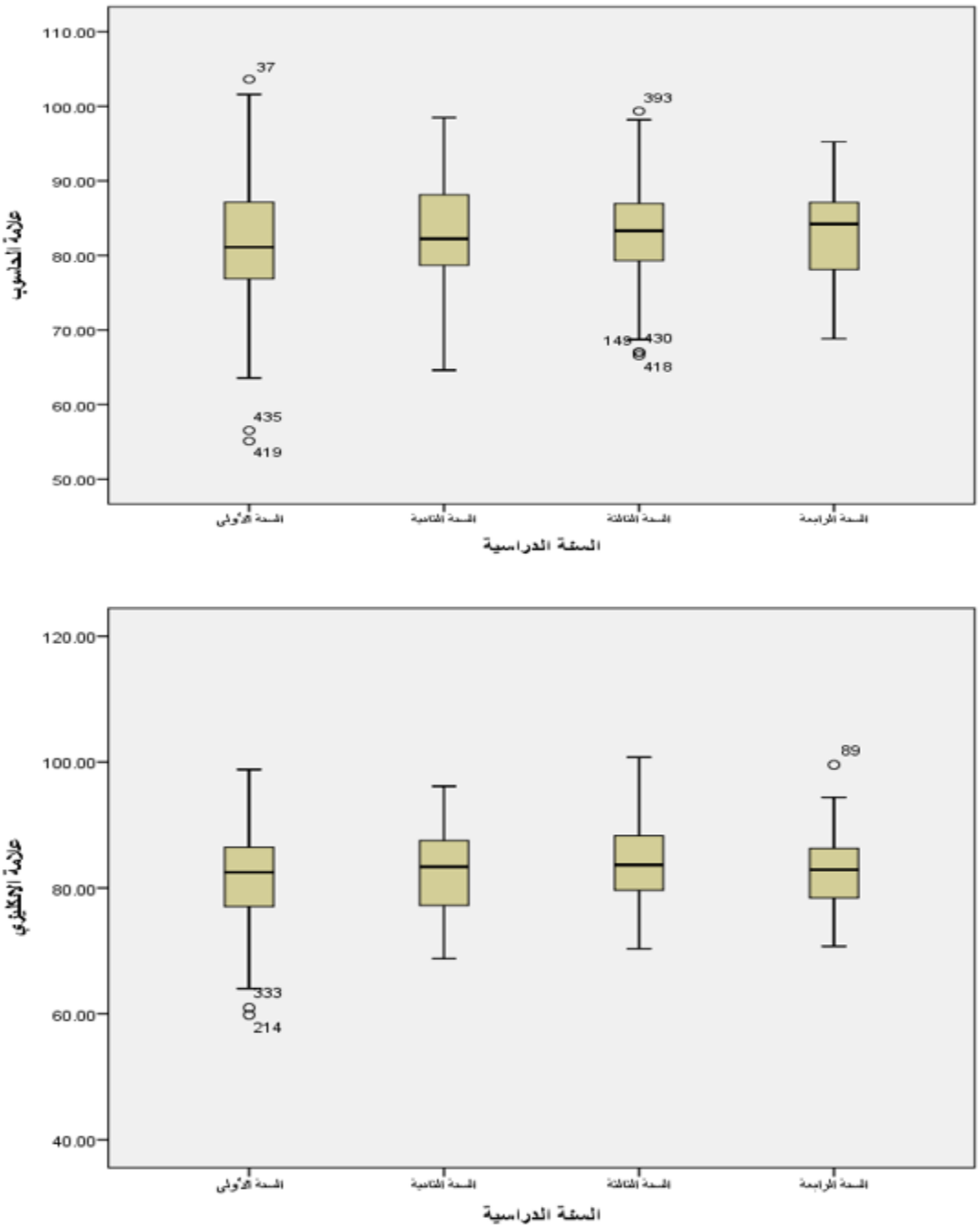
شكل :9 مخطط Boxplot لاثنين من المتغيرات الرقمية بدون عوامل وفي حال استخدام المتغيرات المستقلة معاً

1 - 3 - 2 اثنين من المتغيرات الرقمية، وعامل واحد متغير

أولاً: مستويات العامل معا Factor levels

إنشاء رسوم بيانية منفصلة لكل متغير رقمي .داخل كل رسم بياني، سيكون هناك boxplot لكل مستوى من مستويات العامل .والرسوم البيانية الفردية تظهرمقارنة boxplots جنباً إلى جنب لمستويات العامل .وهذا مفيد في حالات دراسة التباين ANOVA حيث تريد أنتننظر إلى الفروق ضمن المتغير الرقمي فيما يتعلق بالمجموعات الجزئية للعامل.

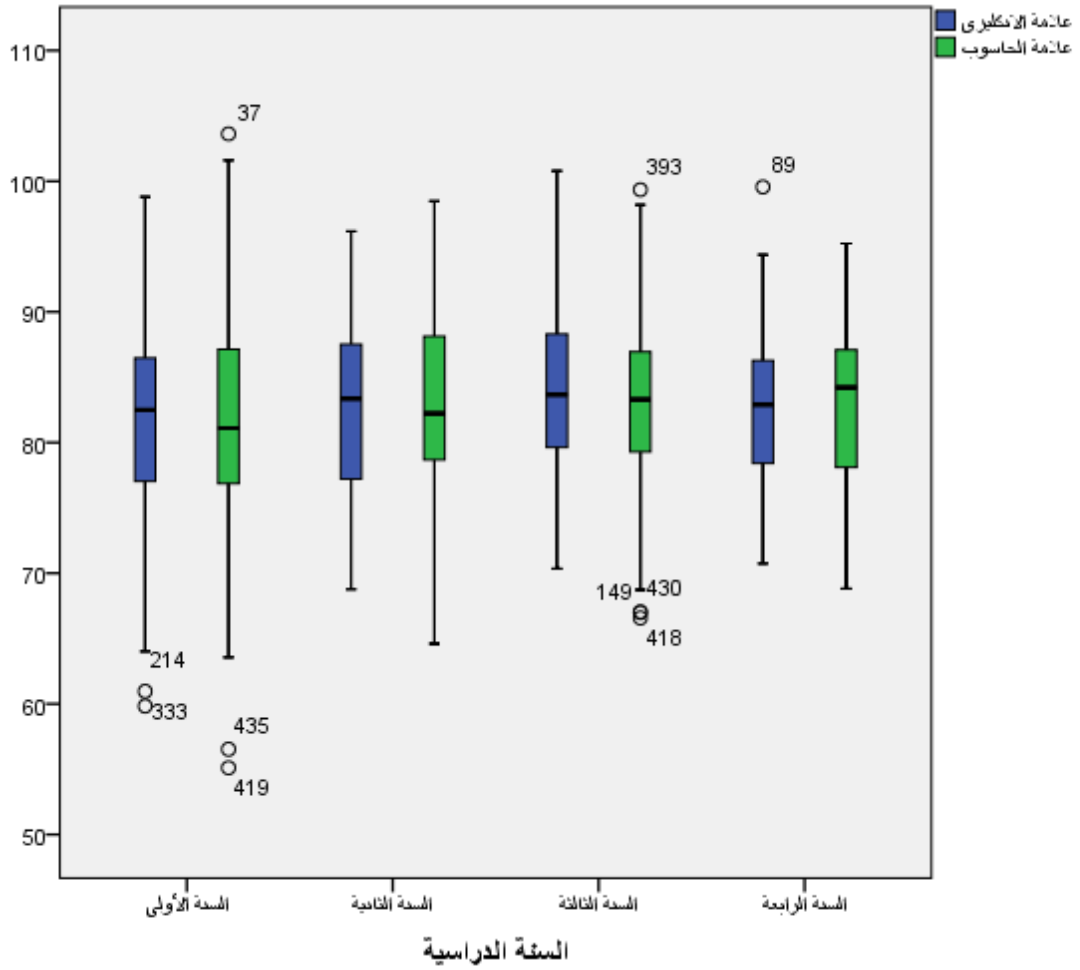
Dr. Wassim Ramadan For SPSS Learning @ WPU University



شكل 10: مخطط Boxplot لاثنين من المتغيرات الرقمية مع عوامل وفي حال استخدام مستويات المعامل معاً

ثانياً: المتغيرات المستقلة معا Dependents together

إنشـاء رسـم بياني boxplot واحد لكل المتغير الرقمية لمقارنتها جنباً إلى جنب وتجميعها من قبل مسـتويات المتغير العامل.



شكل :11 مخطط Boxplot لاثنين من المتغيرات الرقمية مع عوامل وفي حال استخدام المتغيرات المستقلة معاً

Dr. Wassim Ramadan For ... University