

محاضرات في مادة تحليل قواعد المعطيات | التسويقية |



GHICHI ALI

السنة الدراسية 2017 – 2018

المحور رقم (03)

الإحصاء الوصفي عبر مجموعة برامج (SPSS)

الوصف الإحصائي للبيانات

• مقدمة

• وصف البيانات

• اختبار توزيع البيانات (الطبيعية أو Normality)

أوامر وصف البيانات عبر برنامج (SPSS)

• الأمر التكرارات (Frequencies)

• الأمر وصف (Descriptive)

• مقارنة بين الأمر (Frequencies) و الأمر (Descriptive)

• الأمر (Explore)

• الأوامر المختلفة لاختبار مدى خضوع البيانات للتوزيع الطبيعي (Normality Test)

الوصف الإحصائي للبيانات

• مقدمة

- بعد الخوض في كيفية التعامل مع البيانات وإجراء بعض التحويلات عليها سنتعرض في هذا الدرس كيفية استخدام حزمة SPSS في عمليات العرض والتحليل الإحصائي.
- بصفة عامة عمليات العرض والتحليل الإحصائي سوف لا تحتاج لمجهود كبير او معلومات اكثر مما تعلمناه وستتم ببسر ويمكن لأي مستخدم مهما كانت خلفيته الإحصائية استدعاء الأوامر وتنفيذها بسهولة.
- لكن يجب الحذر قبل تنفيذ أي عملية من حيث هل هذه العملية لها شروط يجب ان تتحقق قبل اجرائها ام لا.
- وسوف نولى هذا الموضوع عناية كبيرة.

الوصف الإحصائي للبيانات

• وصف البيانات

- بعد جمع البيانات وادخالها الى شاشة المحرر ننتقل بعد ذلك الى عرض البيانات احصائيا.
- ويتضمن العرض الاحصائي او الاحصاء الوصفي للبيانات تنظيم البيانات ووصفها بطريقة تجعلها مفهومة اكثر للمستخدم.
- فنوع البيانات ووحدة قياسها تحدد الطريقة المستخدمة لوصف البيانات ويمكن تلخيص طرق العرض الاحصائي في الآتى:
- **جدولة البيانات (Tabulation)** : وهو وضع البيانات فى جدول احصائي ويختلف شكل الجدول الاحصائي حسب نوع البيانات وصفية أو كميّة متقطعة أو مستمرة.
- **تمثيل البيانات بيانيا (Graphical Representation)** : وهو التعبير عن البيانات بشكل بياني يتلائم مع طبيعة البيانات. على سبيل المثال:

الوصف الإحصائي للبيانات

• وصف البيانات

- أ- البيانات التي لها وحدة قياس اسمية أو ترتيبية: ترسم باستخدام الأعمدة البيانية (Bar Chart) أو رسمة الدائرة (Pie Chart) بشرط ان تكون تقسيمات المتغير ليست كبيرة (اقل من 10 تقسيمات).
- ب- البيانات المستمرة: (وحدة قياسها فتره او نسبة) الموضوعه فى جدول تكرارى بعد الترميز ترسم بالمدرج التكرارى (Histogram) والمنحنى التكرارى (Frequency Curve) ويستخدم المدرج التكرارى ايضا فى حالة البيانات الكمية المتقطعة.
- ج- يمكن تمثيل البيانات المستمرة ايضا باستخدام رسمة الصندوق (Box Plot) والتي تعتمد على الوسيط والربيعين.
- د- يمكن تمثيل البيانات الكمية (متقطعه أو مستمرة) باستخدام رسمة الساق والأوراق (Stem and leaf)
- حساب بعض المقاييس الإحصائية (Statistical Measures) : يمكن التعبير عن البيانات الإحصائية بحساب مقياس احصائى أو اكثر لها وتوجد عدة مقاييس احصائية هى:

الوصف الإحصائي للبيانات

• وصف البيانات

- **(أ) مقاييس النزعة المركزية (Measures of central Tendency) :** يعرف مقياس النزعة المركزية للبيانات بأنه الرقم الذى تتمركز حوله البيانات. وتوجد عدة مقاييس للنزعة المركزية منها:
 - **الوسط الحسابى (Arithmetic Mean)** ويصلح للبيانات الكمية فقط وهو وحيد ولكنه يتأثر بالقيم الشاذة.
 - **الوسيط (Median):** وهى القيمة التى تقع فى منتصف البيانات بعد الترتيب وهو وحيد ولا يتأثر بالقيم الشاذة.
 - **المنوال (Mode) :** وهى القيمة الأكثر شيوعا ولا يتأثر بالقيم الشاذة لكنه قد لا يكون وحيد او قد لا يكون موجودا.
- مقاييس النزعة المركزية كلها مقاييس مطلقة أى لها نفس تمييز البيانات الأصلية وبذلك لا تصلح للمقارنة بين مجموعتين أو أكثر اذا اختلفت وحدة القياس.
- **(ب) مقاييس التشتت المطلق والنسبى:** ويقاس مدى البعد بين مفردات المتغير أى أن مقياس التشتت يقاس مدى التباعد والتنافر بين قيم المتغير أو قد يقاس مدى بعد البيانات عن مقياس النزعة المركزية الخاص بها وتوجد مقاييس مطلقة للتشتت ومقاييس نسبيه ومن مقاييس التشتت:

الوصف الإحصائي للبيانات

• وصف البيانات

1. **المدى (Range):** وهو البعد بين أكبر وأقل قيمة وهو يتأثر بالقيم الشاذة
2. **الانحراف المعياري (Standard deviation):** وهو الجذر التربيعي الموجب للتباين وهو وحيد وقياس تشتت البيانات عن الوسط.
3. **معامل الاختلاف (Coefficient of Variation):** وهو من المقاييس النسبية للتشتت وهو من أحسن مقاييس التشتت النسبية. ويستخدم معامل الاختلاف أحيانا للتعرف على القيم الشاذة للبيانات.
4. **مقياس الالتواء (Measure of Skewness):** وهو مقياس تشتت نسبي يحدد هل البيانات متماثلة أم ملتوية.
5. **مقياس التفطح (Measure of Kurtosis):** وهو مقياس نسبي يقيس ارتفاع قمة المنحنى. إن كبر قيمة مقياس التشتت يعنى وجود تشتت واختلاف أكثر ويطلق على كبر التشتت أحيانا عدم التجانس، مقاييس التشتت المطلقة كلها موجبه لكن المقاييس النسبيه قد تكون موجبه أو سالبه.

الوصف الإحصائي للبيانات

• اختبار توزيع البيانات (الطبيعية أو Normality)

- من الفروض الهامة في الاختبارات الإحصائية المعلمية ان يكون التوزيع الاحتمالي للبيانات المستخدمة هو التوزيع الطبيعي. بدون ذلك الشرط لا يمكن تطبيق الاختبارات المعلمية من الناحية العلمية.

للتأكد من أن البيانات تتوزع حسب التوزيع الطبيعي توجد عدة طرق منها:

1. طرق تعتمد على الرسم البياني

2. طرق تعتمد على حساب مقياس احصائي للبيانات

3. طرق تعتمد على اجراء اختبار احصائي

ويمكن استخدام احد تلك الطرق للتأكد من أن البيانات لها التوزيع الطبيعي وسنتعرض الطرق السابقة الذكر كل على حده.

الوصف الإحصائي للبيانات

• اختبار توزيع البيانات (الطبيعية أو Normality)

1. اولا الاعتماد على الأشكال البيانية:

حيث ان منحني التوزيع الطبيعي متماثل حول الوسط الحسابي لذا ستعتمد فكرة الأشكال البيانية على مفهوم التماثل عن طريق اسقاط عمود من قمة المنحنى وبحث الجزئين المقسم لهما الشكل هل متساويان (متماثل حول العمود) فتكون للبيانات التوزيع الطبيعي أم غير متساوى فتكون البيانات ليس لها التوزيع الطبيعي.

• وسوف نستخدم الأشكال التالية

• المدرج التكرارى (Histogram)

• رسمة الساق والاوراق (Stem and Leaf Plot)

• رسمة الصندوق (Box Plot): تكون البيانات متماثله اذا كان البعد بين الربيع الأدنى والوسيط يساوى البعد بين الربيع الأعلى والوسيط

• رسمة الاحتمال الطبيعي (Normal Probability Plot): نقوم برسم البيانات المشاهدة والقيم المتوقعة المناظرة اذا كانت البيانات لها التوزيع الطبيعي ستقع النقاط فى الشكل على شكل خط مستقيم.

• رسمة الاتجاه للمنحنى الطبيعي (De-trended Normal Plot): ونحصل عليها برسم الانحراف الحقيقى للنقاط على الخط المستقيم فاذا كانت النقاط على الشكل المرسوم ليس لها نمط حول الخط المرسوم حول الصفر فان هذا يعنى انها تتوزع حسب التوزيع الطبيعي.

الوصف الإحصائي للبيانات

• اختبار توزيع البيانات (الطبيعية أو Normality)

2. ثانيا الاعتماد على مقياس احصائي:

• بحساب معامل الالتواء فإذا كان مسويا الصفر كانت البيانات متماثله وإذا كان معامل التفرطح مساويا الصفر أو كانت البيانات معتدله التفرطح وبالتالي فإن البيانات تتوزع حسب التوزيع الطبيعي.

3. ثالثا استخدام اختبار احصائي

• يوجد ايضا العديد من الاختبارات الاحصائية التي تختبر هل البيانات تتوزع حسب التوزيع الطبيعي ام لا؟

• وسوف نهتم هنا بوصف البيانات الاحصائية لمتغير واحد فقط وايضا اختبار طبيعية البيانات هل البيانات سيكون لها التوزيع الطبيعي ام لا؟

أوامر وصف البيانات عبر برنامج (SPSS)

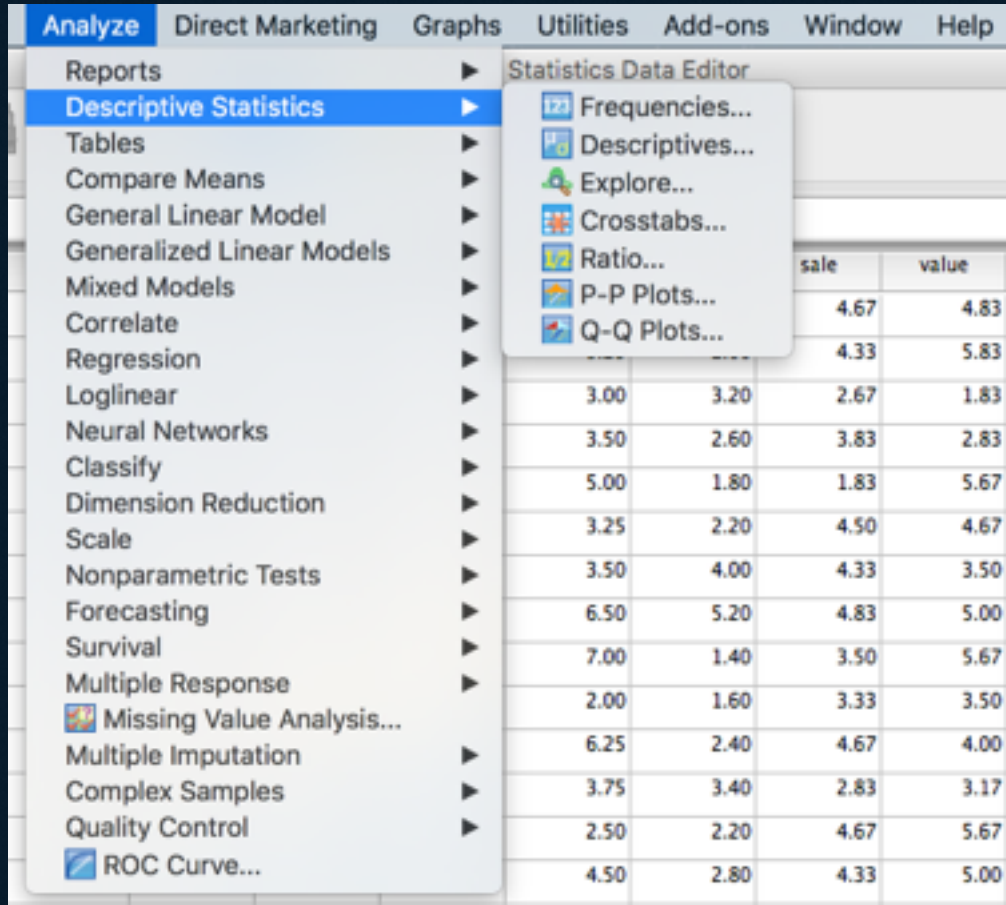
سنعتمد في تقديم هذه الأوامر الثلاثة (Frequency, Descriptive, Explore) على قاعدة البيانات في الملف المسمى (Seniors.sav)، ويشتمل هذا الملف على العديد من المتغيرات كما يبين الشكل في الأسفل:

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
18	free5	Numeric	8	2	studying	None	None	8	Right	Scale	Input
19	free6	Numeric	8	2	television	None	None	8	Right	Scale	Input
20	free7	Numeric	8	2	going out	None	None	8	Right	Scale	Input
21	free8	Numeric	8	2	cultural activiti...	None	None	8	Right	Scale	Input
22	free9	Numeric	8	2	walking	None	None	8	Right	Scale	Input
23	gender	Numeric	8	2	gender	{1.00, male}...	None	8	Right	Nominal	Input
24	mrhp	Numeric	8	2	main responsi...	{1.00, no mr...	None	8	Right	Nominal	Input
25	place	Numeric	8	2	dwelling	{1.00, city}...	None	8	Right	Ordinal	Input
26	numfamily	Numeric	8	2	number of per...	{1.00, 1}...	None	8	Right	Ordinal	Input
27	age	Numeric	8	2	age	{1.00, 25-3...	.00	8	Right	Nominal	Input
28	education	Numeric	8	2	education	{1.00, elem...	99.00	8	Right	Ordinal	Input
29	income	Numeric	8	2	income	{1.00, <125...	-99.00	8	Right	Ordinal	Input
30	rankA	Numeric	8	0	ranking of drin...	{1, most pr...	None	8	Right	Ordinal	Input
31	rankAA	Numeric	8	0	ranking of drin...	{1, most pr...	None	8	Right	Ordinal	Input
32											

تمثل هذه القاعدة من البيانات قياسات للعديد من السلوكيات الشرائية لعينة مكونة من 310 مستهلك. كما تقدم أيضا خياراتهم في أوقات الراحة والفراغ، أين تم استخدام سلم ليكرت السباعي في تقييم مواقفهم.

أوامر وصف البيانات عبر برنامج (SPSS)

تشتمل القائمة (Descriptive Statistics) على قائمة الأوامر التالية:



يتم الوصول إلى القائمة من خلال المسار التالي:

Analyze/Descriptive Statistics/
ثم ننتقي الأمر المناسب

1. الأمر: (Frequencies)

2. الأمر: (Descriptive)

3. الأمر: (Explore)

4. الأمر: (Crosstabs)

5. الأمر: (Ratio)

6. الأمر: (P-P Plots)

7. الأمر: (Q-Q plots)

أوامر وصف البيانات عبر برنامج (SPSS)

• الأمر التكرارات (Frequencies)

يتم الوصول إلى الأمر (Frequencies) من خلال المسار التالي:

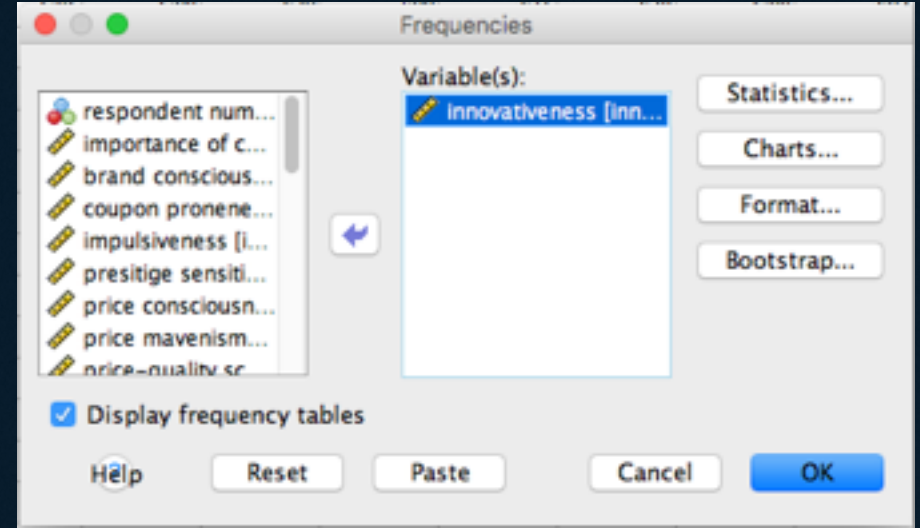
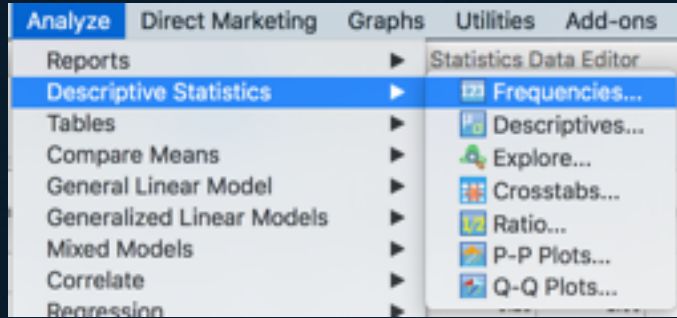
Analyse/Descriptive Statistics/Frequencies

The screenshot shows the SPSS Statistics interface. The 'Analyze' menu is open, and the path 'Analyze > Descriptive Statistics > Frequencies...' is highlighted. The background shows a data editor window with the following data:

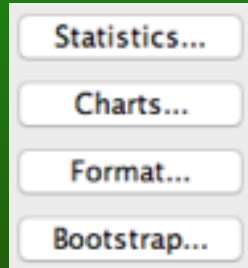
	respr	conv	brand	coup	impuls
1	144.00	3.33	4.00	3.71	1.33
2	235.00	2.33	3.50	5.14	2.67
3	286.00	5.67	3.00	2.57	2.33
4	234.00	7.00	1.00	2.00	2.67
5	285.00	3.00	4.50	3.57	1.67
6	248.00	3.00	3.00	5.29	1.67
7	212.00	5.33	3.50	4.86	3.33
8	205.00	4.00	4.50	4.29	3.00
9	226.00	4.00	4.50	4.86	2.00
10	186.00	3.67	2.50	5.29	2.33
11	241.00	5.00	5.50	3.71	2.33
12	126.00	3.33	3.50	2.71	4.00
13	207.00	3.33	1.00	6.29	2.33
14	127.00	4.67	3.50	3.43	2.33
15	153.00	3.00	1.00	1.00	1.00

أوامر وصف البيانات عبر برنامج (SPSS)

1. الأمر: (Frequencies) عند اختيار هذا الأمر يظهر مربع الحوار التالي:



بعد إدخال المتغير يتم انتقاء المقاييس
الملائمة من خلال هذه القائمة بالضغط
على كل على حدى وملء ما يلائم الباحث



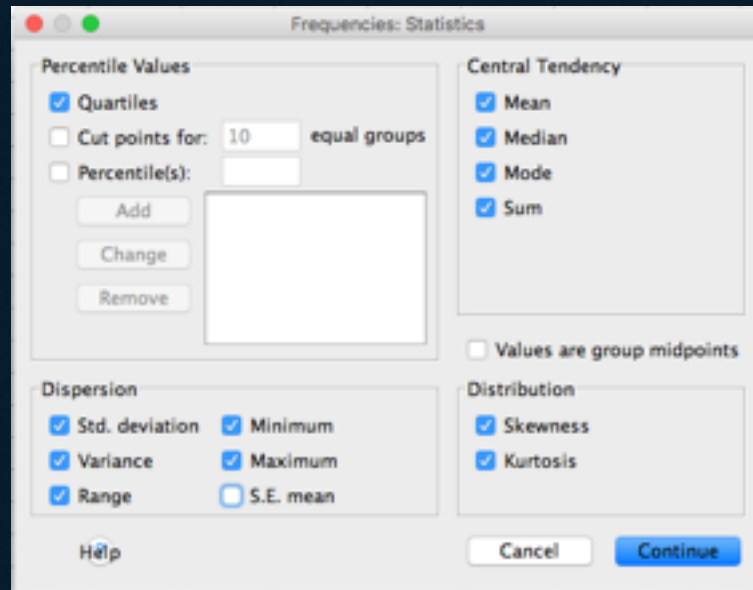
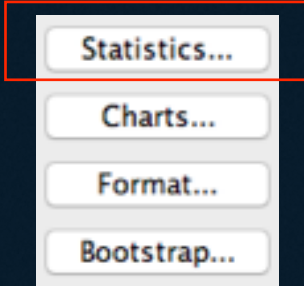
ثم ندخل المتغير المراد تحليله إلى مربع التحليل المسمى
(Variables) من قائمة المتغيرات عبر السهم التالي:

مثلا في هذه الحالة تم إدخال المتغير (Innovativeness)

أوامر وصف البيانات عبر برنامج (SPSS)

1. الأمر: (Frequencies):

من خلال الضغط على إيقونة (Statistics) يمكنك الحصول على كل من مقاييس النزعة المركزية (1) ومقاييس التشتت (2) و الشكل (3)، ويتم ذلك بتأشير الباحث في الخانات الملائمة كما يظهر في مربع الحوار أسفله:



نضغط على



(Continue) ثم

على (OK)

أنظر المخرجات

في الصفحة الموالية

أوامر وصف البيانات عبر برنامج (SPSS)

1. الأمر: (Frequencies):

بالتطبيق على المثال السابق للملف (Seniors.sav) والمتغير (Innovativeness) نتحصل على الجدولين التاليين:

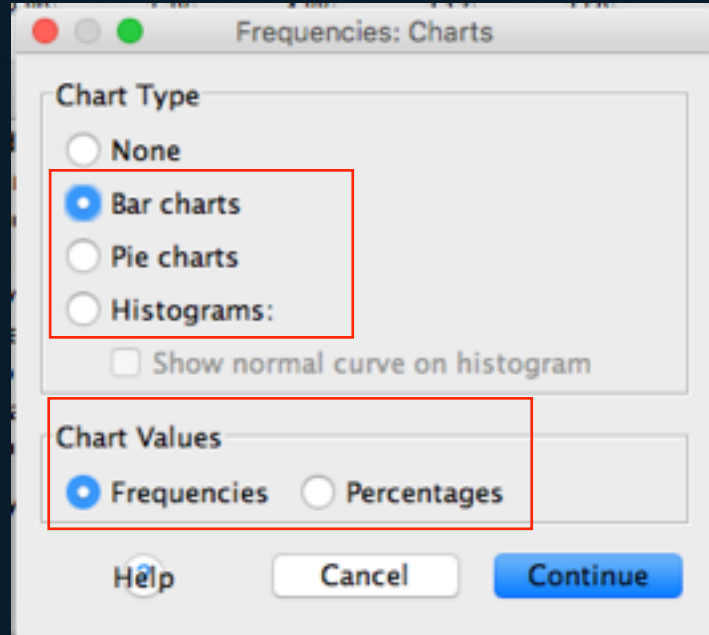
innovativeness		
N	Valid	310
	Missing	0
Mean		3.6826
Median		3.6000
Mode		4.00
Std. Deviation		.90315
Variance		.816
Skewness		.068
Std. Error of Skewness		.138
Kurtosis		.266
Std. Error of Kurtosis		.276
Range		5.60
Minimum		1.00
Maximum		6.60
Sum		1141.60
Percentiles	25	3.0000
	50	3.6000
	75	4.2000

innovativeness					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	1	.3	.3	.3
	1.20	1	.3	.3	.6
	1.40	1	.3	.3	1.0
	1.60	3	1.0	1.0	1.9
	1.80	1	.3	.3	2.3
	2.20	12	3.9	3.9	6.1
	2.40	12	3.9	3.9	10.0
	2.60	9	2.9	2.9	12.9
	2.80	22	7.1	7.1	20.0
	3.00	16	5.2	5.2	25.2
	3.20	24	7.7	7.7	32.9
	3.40	30	9.7	9.7	42.6
	3.60	25	8.1	8.1	50.6
	3.80	22	7.1	7.1	57.7
	4.00	31	10.0	10.0	67.7
	4.20	30	9.7	9.7	77.4
	4.40	21	6.8	6.8	84.2
	4.60	13	4.2	4.2	88.4
	4.80	11	3.5	3.5	91.9
	5.00	6	1.9	1.9	93.9
	5.20	5	1.6	1.6	95.5
	5.40	6	1.9	1.9	97.4
	5.60	2	.6	.6	98.1
	5.80	2	.6	.6	98.7
	6.00	3	1.0	1.0	99.7
	6.60	1	.3	.3	100.0
Total		310	100.0	100.0	

أوامر وصف البيانات عبر برنامج (SPSS)

1. الأمر: (Frequencies):

من خلال الضغط على إيقونة (Chart) يمكنك الحصول على كل من التمثيل البياني في شكل أعمدة أو دوائر نسبية أو مدرج تكراري (1)، وكما يمكن إختيار أساس عرض القيم في الأشكال إما عبر التكرارات أو النسب المئوية (2)، ويتم ذلك بتأشير الباحث في الخانات الملائمة كما يظهر في مربع الحوار أسفله:



نضغط على

ثم (Continue)

على (OK)

أنظر المخرجات

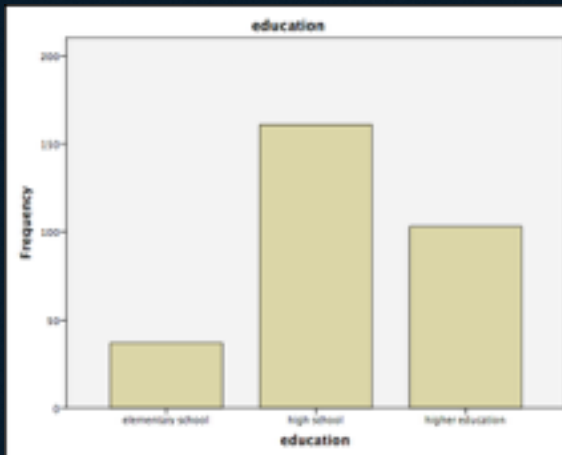
في الصفحة الموالية

أوامر وصف البيانات عبر برنامج (SPSS)

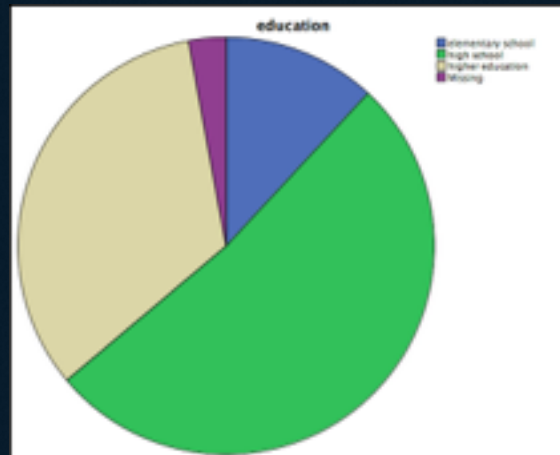
1. الأمر: (Frequencies):

بالتطبيق على المثال السابق للملف (Seniors.sav) والمتغيرين (Education, value) نتحصل على الرسومات البيانية التالية:

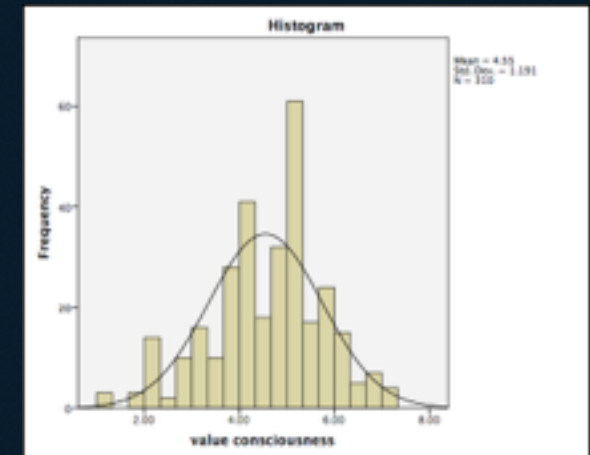
Bar charts



Pie charts



Histograms:

 Show normal curve on histogram


لإدخال المزيد من التحسينات على الشكل نضغط مرتين على الشكل ونتعامل مع مربع الحوار لمزيد من العرض في جوانب المعلومات والشكل على حد سواء.

محاضرات في مادة تحليل قواعد المعطيات | التسويقية |

