



الفصل السادس: تحليل التباين

العنوان	رقم الصفحة
مقدمة	3
1. المتغيرات	3
1.1. المتغير المستقل	3
2.1. المتغير التابع	4
2. الفرضيات	4
3. شروط التحليل	4
4. حساب المؤشرات الإحصائية في ANOVA	5
5. الاختبار باستخدام SPSS	6
المراجع العربية	14
المراجع الأجنبية	14
مقترحات وتمارين للفصل السادس	15

الكلمات المفتاحية:

تحليل التباين، ANOVA، معامل F.

ملخص:

يتناول الفصل كيفية اختبار فرضية تساوي المتوسطات بين أكثر من مجموعتين من خلال اللجوء إلى تحليل التباين أو ما يعرف اختصاراً بـ ANOVA. ويشرح الفصل كيفية اختبار أثر المتغير المستقل (العامل) في المتغير التابع (الكمي) باستخدام ANOVA. كما يبين شروط تحليل ANOVA وفرضياته وكيفية تطبيقه ضمن SPSS.

أهداف تعليمية:

بعد اطلاع الطالب على مضمون ومحتوى هذا الفصل، ستتوافر لديه القدرة على تحقيق الأهداف التالية:

- إدراك شروط وكيفية تطبيق تحليل التباين ANOVA
- فهم طبيعة المتغيرات في تحليل التباين
- التعرف على كيفية حساب المؤشرات الإحصائية ضمن ANOVA
- إدراك كيفية تفسير مخرجات ANOVA

المخطط:

1. المتغيرات Variables
 - 1.1 المتغير المستقل Independent variable
 - 2.1 المتغير التابع dependent variable
2. الفرضيات Hypothesis
3. شروط التحليل Assumptions of the analysis
4. حساب المؤشرات الإحصائية في ANOVA Calculating statistical coefficients in ANOVA
5. الاختبار باستخدام SPSS Test using SPSS

مقدمة

استخدمنا في الفصل السابق تحليل t للعينات المستقلة Independent Samples t -test لاختبار فرضية تساوي المتوسطات بين مجموعتين. ولكن قد نحتاج في الكثير من الحالات إلى اختبار فرضية تساوي المتوسطات بين أكثر من مجموعتين (ثلاثة أو أكثر)، يمكننا في هذه الحالة اللجوء إلى تحليل التباين Analysis of Variance أو ما يعرف اختصاراً بـ ANOVA. يعتبر تحليل التباين إذاً امتداداً لاختبار t للعينات المستقلة ويمكن استخدامه لمقارنة متوسطات مجموعتين أو أكثر.

1. المتغيرات

يسمى تحليل التباين بالأحادي (أو أحادي الاتجاه) One-Way ANOVA إذا كان لكل مفردة من مفردات العينة علامة أو إجابة على متغيرين. يسمى المتغير الأول المتغير العامل Factor أو المتغير المستقل Independent Variable. أما الثاني فهو المتغير التابع Dependent Variable.

1.1. المتغير المستقل

يكون المتغير المستقل أو العامل عادةً متغيراً غير كمي أو غير قياسي Non-metric Variable أي أنه يكون متغيراً من النوع الاسمي Nominal أو الترتيبي Ordinal. يكون لهذا المتغير عدداً محدداً من الفئات أو المستويات. وهو المتغير الذي سيقسم العينة الكلية إلى عدد من المجموعات التي يراد مقارنة متوسطاتها الحسابية. ومن الأمثلة على هذا المتغير نذكر:

- **مستوى التعليم:** أقل من بكالوريا، بكالوريا، شهادة جامعية، ماجستير، دكتوراه
- **المحافظة:** دمشق، حلب، اللاذقية
- **الحالة الاجتماعية:** عازب، متزوج، مطلق
- **الجامعة:** جامعة دمشق، جامعة تشرين، جامعة حلب، الجامعة الافتراضية

ويمكن أن يتضمن تحليل التباين أكثر من متغير مستقل غير كمي. ففي حال كان لدينا متغير مستقل واحد نكون كما ذكرنا في حالة تحليل التباين أحادي الاتجاه، أما إذا كان لدينا متغيران مستقلان (عاملان) فنطلق على التحليل اسم تحليل التباين ثنائي الاتجاه Two-Way ANOVA، وهكذا في حالة n متغير مستقل نكون أمام حالة n -way ANOVA. وسنقتصر في هذا المقرر على دراسة تحليل التباين أحادي الاتجاه.

2.1. المتغير التابع

يكون المتغير التابع متغيراً من النوع الكمي أي من النوع المدرج Interval أو النسب Ratio. وهو المتغير الذي سيتم فحص مساواة متوسطه لكل فئة من فئات المتغير العامل.

2. الفرضيات

الهدف الأساسي من تحليل التباين كما ذكرنا سابقاً هو مقارنة متوسطات متغير كمي يسمى المتغير التابع بين فئات المتغير العامل Factor أو بين المجموعات التي يعرفها المتغير المستقل. تنص فرضية العدم على أن المتوسطات التي تتم مقارنتها متساوية أي أن المتغير المستقل لا يؤثر في المتغير التابع.

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \dots = \mu_c$$

أما الفرضية البديلة فتتص على وجود اختلاف بين متوسطين على الأقل من متوسطات المجموعات التي تتم مقارنتها أي أنها تنص على وجود أثر للمتغير المستقل في المتغير. فإذا رفضت الفرضية التي تقول إن متوسطات هذه الفئات متساوية فأى هذه المتوسطات متساوية وأياها غير متساوية؟ للإجابة على هذا السؤال تستخدم المقارنات البعدية Post Hoc Comparisons لمقارنة متوسطات المتغير التابع لكل زوجين من الفئات أو المجموعات على حدة. فإذا كان عدد الفئات الكلية ثلاثة، فإن عدد المقارنات البعدية سيكون ثلاث مقارنات. وبالتحديد ستكون هذه المقارنات بين المجموعتين الأولى والثانية، وبين المجموعتين الأولى والثالثة، وفي النهاية بين المجموعتين الثانية والثالثة.

3. شروط التحليل

يوجد شرطان أساسيان لاستخدام اختبار ANOVA:

- **التوزيع الطبيعي:** يجب أن يكون توزيع المتغير التابع (المتغير الكمي) طبيعياً لكل مجتمع من مجتمعات (مجموعات) المتغير العامل Factor. وقد وجد من خلال الأبحاث أن عدم تحقق هذا الشرط لا يؤثر بشكل كبير في نتيجة تحليل التباين إذا كان عدد أفراد المجموعات التي تتم مقارنتها جيداً. وبهذا قد تكون نتيجة تحليل التباين دقيقة إلى حد ما حتى لو كان توزيع المتغير التابع غير طبيعي
 - **تجانس التباين (تساوي تباينات المتغير التابع):** يجب أن لا يختلف تباين المتغير التابع بين المجموعات التي تتم مقارنتها. وكما في اختبار t للعينات المستقلة يتم التأكد من تحقق هذا الشرط من خلال اختبار ليفين Leven's test. وقد وجد أيضاً أنه يمكن الاعتماد على نتيجة ANOVA حتى في حال عدم تحقق هذا الشرط بشكل تام
- ولا بد من الإشارة إلى أنه ينصح بأن تكون المجموعات أو العينات التي تتم مقارنتها متقاربة من حيث عدد أفرادها.

4. حساب المؤشرات الإحصائية في ANOVA

تكمن الفكرة الأساسية في ANOVA في اختبار تغيرات المتغير التابع بين المجموعات وبناءً على هذه التغيرات تحديد ما إذا كان هناك من سبب وجيه للاعتقاد بأن متوسطات مجتمعات المجموعات (أو مستويات العامل) تختلف بشكل ذي دلالة إحصائية.

يتم في ANOVA تجزئة التباين إلى جزئين: تباين بين المجموعات Between-Groups Variance وتباين ضمن المجموعات Within-Groups Variance. ويعتمد الاختبار على حساب المعامل F من الصيغة التالية:

$$\frac{\text{متوسط مربعات الانحرافات بين المجموعات}}{\text{متوسط مربعات الانحرافات بين المجموعات}} = \frac{\text{Mean Square}_{\text{between}}}{\text{Mean Square}_{\text{within}}} = F = \frac{\text{التباين بين المجموعات}}{\text{التباين ضمن المجموعات}}$$

$$\text{Mean Square}_{\text{between}} = \frac{\text{Sum of Squares}_{\text{between}}}{\text{degrees of freedom}_{\text{between}}} = \frac{\sum_{j=1}^k n_j (\bar{x}_j - \bar{x})^2}{k - 1}$$

حيث:

\bar{x} : المتوسط الكلي

\bar{x}_j : متوسط المجموعة j

n_j : حجم المجموعة j

k : عدد المجموعات

$$\text{Mean Square}_{\text{within}} = \frac{\text{Sum of Squares}_{\text{within}}}{\text{degrees of freedom}_{\text{within}}} = \frac{\sum_{j=1}^k \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2}{n - k}$$

يعتمد معامل F إذاً على درجتي حرية. تتعلق الأولى بالتباين بين المجموعات $k - 1$ أو عدد فئات المتغير المستقل ناقصاً واحداً. وتعلق الأخرى بالتباين ضمن المجموعات $n - k$ أو حجم العينة ناقص عدد فئات المتغير المستقل.

إذا كانت قيمة F أصغر أو تساوي واحد فإن ذلك يدل على عدم وجود فروقات بين متوسطات المجتمعات الإحصائية وبالتالي عدم إمكانية رفض فرضية العدم. أما إذا كانت قيمة F أكبر من الواحد فلا بد من العودة إلى الدلالة الإحصائية لاختبار ANOVA لمعرفة إمكانية رفض فرضية العدم (p أو sig أصغر من α).

5. الاختبار باستخدام SPSS

لتوضيح كيفية تطبيق اختبار ANOVA أحادي الاتجاه لنأخذ المثال التالي: قامت إحدى الشركات التي تمتلك سلسلة متاجر عبر البلاد بإجراء تجربة لاختبار أثر مستوى الترويج داخل المتجر in-store promotion في مستوى المبيعات. تم اختبار ثلاثة مستويات من الترويج داخل المتجر:

1. مستوى ترويج مرتفع داخل المحل

2. مستوى ترويج متوسط داخل المحل

3. مستوى ترويج منخفض داخل المحل

تم اختيار 30 متجراً بشكل عشوائي لإجراء التجربة، وتم إخضاع كل 10 متاجر لإحدى الحالات الثلاثة السابقة. تم إجراء الاختبار لمدة شهرين وتم قياس المبيعات sales في كل متجر ومن ثم تحويل أرقام المبيعات إلى مقياس من 10 علامات. كما تم قياس معدل حضور الزبائن في المتاجر clientele rating على مقياس من 10 درجات. يوضح الجدول التالي البيانات التي تم جمعها للتحليل في نهاية التجربة.

Store Number	In-Store Promotion	Sales	Clientele Rating
1	1.00	10.00	9.00
2	1.00	9.00	10.00
3	1.00	10.00	8.00
4	1.00	8.00	4.00
5	1.00	9.00	6.00
6	2.00	8.00	8.00
7	2.00	8.00	4.00
8	2.00	7.00	10.00
9	2.00	9.00	6.00
10	2.00	6.00	9.00
11	3.00	5.00	8.00
12	3.00	7.00	9.00
13	3.00	6.00	6.00
14	3.00	4.00	10.00
15	3.00	5.00	4.00
16	1.00	8.00	10.00
17	1.00	9.00	6.00
18	1.00	7.00	8.00
19	1.00	7.00	4.00
20	1.00	6.00	9.00
21	2.00	4.00	6.00
22	2.00	5.00	8.00
23	2.00	5.00	10.00
24	2.00	6.00	4.00
25	2.00	4.00	9.00
26	3.00	2.00	4.00
27	3.00	3.00	6.00
28	3.00	2.00	10.00
29	3.00	1.00	9.00
30	3.00	2.00	8.00.

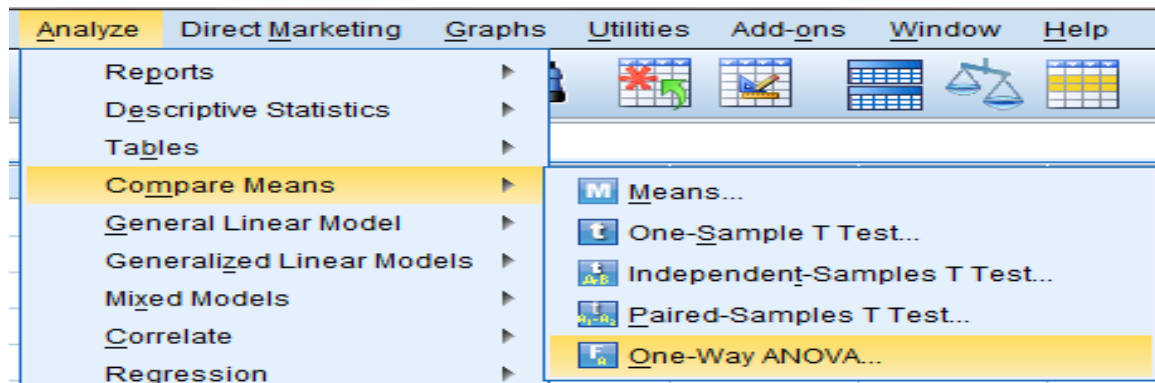
جدول 1.6. بيانات تجربة أثر الترويج داخل المتجر في مستوى المبيعات

تسعى الشركة كما ذكرنا إلى اختبار أثر الترويج داخل المتجر في مستوى الترويج. يتمثل المتغير المستقل في هذا المثال في الترويج داخل المتجر وهو متغير اسمي ذو ثلاث فئات (مرتفع، متوسط، منخفض) أي أنه يقسم العينة إلى ثلاث مجموعات. أما المتغير التابع فهو مستوى المبيعات ضمن المحلات. تشير فرضية العدم إلى عدم وجود أثر لمستوى الترويج داخل المتجر في مستوى المبيعات أي أنها تشير إلى عدم وجود اختلاف بين متوسطات المبيعات بين مجموعات المتاجر الثلاثة التي تم تعريفها من خلال مستوى الترويج المطبق في المتاجر.

$$H_0 : \mu_1 = \mu_2 = \mu_3$$

أما الفرضية البديلة فتشير إلى وجود أثر لمستوى الترويج داخل المتجر في مستوى المبيعات أي أنها تشير إلى وجود اختلاف بين متوسطات المبيعات بين مستويي ترويج على الأقل. ولتطبيق الاختبار في SPSS يمكننا اتباع المسار التالي (كما في الشكل):

Analyze ← Compare Means ← One-Way ANOVA



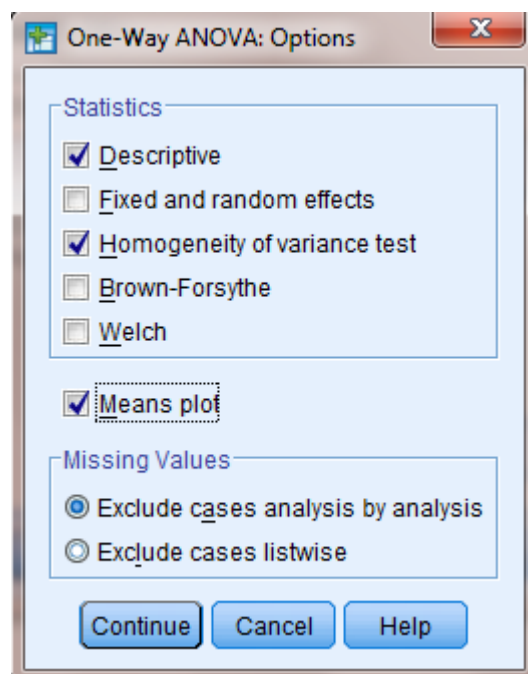
شكل 1.6. كيفية تطبيق ANOVA في SPSS

في النافذة One-Way ANOVA نقوم بنقل المتغير المستقل "الترويج داخل المتجر" إلى الصندوق Factor وننقل المتغير التابع "المبيعات" إلى الصندوق Dependent List.



شكل 2.6. النافذة One-Way ANOVA

وننقر فوق الزر Options وعند ظهور النافذة One-Way ANOVA نختار Descriptive لإظهار المتوسطات الحسابية للمبيعات في كل مجموعة ونختار Homogeneity of Variance Test لإظهار نتيجة اختبار تجانس التباين أو اختبار ليفين. كما يمكننا اختيار Means Plot للحصول على تخطيط أو تمثيل بياني لأثر الترويج داخل المتجر في مستوى المبيعات.



شكل 3.6. النافذة One-Way ANOVA: Options

يبين الجدول الأول Descriptives في نتيجة الاختبار مجموعة من الإحصاءات الوصفية المتعلقة بمبيعات المتجر وفق مستويات الترويج الثلاثة حيث نلاحظ أن المبيعات ارتفعت بشكل ملحوظ مع ارتفاع مستوى الترويج (M منخفض = 3.70 \leftarrow متوسط = 6.20 \leftarrow مرتفع = 8.30) واختبار الدلالة الإحصائية لهذه الفروقات لا بد من العودة إلى نتيجة ANOVA.

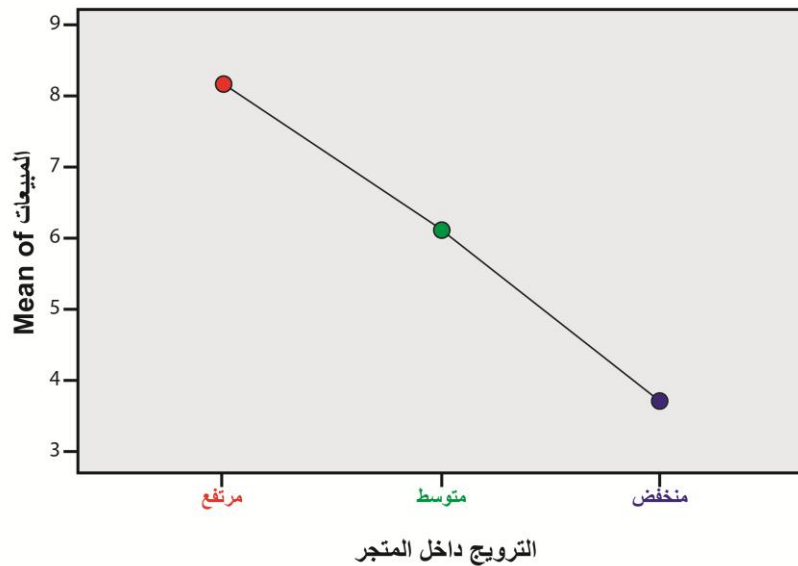
Descriptives

المبيعات

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
مرتفع	10	8.30	1.337	.423	7.34	9.26	6	10
متوسط	10	6.20	1.751	.554	4.95	7.45	4	9
منخفض	10	3.70	2.003	.633	2.27	5.13	1	7
Total	30	6.07	2.532	.462	5.12	7.01	1	10

جدول 2.6. جدول الإحصاءات الوصفية Descriptives

Means Plots



شكل 4.6. الوسط الحسابي لمستوى المبيعات وفق مستوى الترويج داخل المتجر

تظهر نتيجة اختبار Levene في الجدول Test of Homogeneity of Variances أن $\text{sig} = 0.275$ وهي أكبر من $\alpha (0.05)$ وبالتالي فإن فرضية العدم في اختبار ليفين القائلة بتجانس أو تساوي تباين المبيعات بين المجموعات التي تتم مقارنتها لا يمكن رفضها مما يدل على تحقق شرط تجانس التباين.

Test of Homogeneity of Variances

المبيعات

Levene Statistic	Df1	Df2	Sig
1.353	2	27	.275

جدول 3.6. اختبار تجانس التباينات

يظهر جدول ANOVA أن قيمة $P\text{-value}$ (أو sig) المصاحبة لإحصائية F أقل من 0.05

$$F(2,27)=17,944 ; \text{sig} < 0.001$$

ولهذا نستطيع رفض فرضية العدم وقبول الفرضية البديلة القائلة بوجود فروق معنوية بين متوسطات المبيعات في المتاجر تبعاً لنوع الترويج المستخدم.

ANOVA

المبيعات

	Sum of squares	df	Mean Square	F	Sig.
Between Groups	106.067	2	53.033	17.944	.000
Within Groups	79.800	27	2.956		
Total	185.867	29			

جدول 4.6. نتيجة اختبار ANOVA

ولكن عند رفض فرضية العدم فإنه لا يوجد دليل واضح على وجود فروق ذات دلالة إحصائية بين كافة متوسطات المبيعات في مستويات الترويج الثلاثة فكل ما نستطيع استنتاجه من جدول ANOVA أن هناك فرقاً معنوياً بين متوسطي المبيعات وفق مستويي ترويج على الأقل.

ونظراً لأهمية الإجابة على هذا التساؤل وضع الإحصائيون مجموعة من الطرق التي تسمح باختبار الفروق بين متوسطات المجموعات المقارنة ومن بين هذه الطرق نجد طريقة Bonferroni.

حيث تسمح هذه الطريقة بإجراء مقارنات متعددة Multiple Comparisons لاختبار معنوية الفرق لكل زوج من فئات أو حالات المتغير المستقل (العامل).

ويمكن إظهار واختيار اختبار المقارنات البعدية بالنقر فوق الزر Post Hoc في النافذة One-Way ANOVA.

Post Hoc Tests

Multiple Comparisons

المبيعات

Bonferroni

الترويج داخل المتجر (I)	الترويج داخل المتجر (J)	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
مرتفع	متوسط	2.100*	.769	.033	.14	4.06
	منخفض	4.600*	.769	.000	2.64	6.65
متوسط	مرتفع	-2.100*	.769	.033	-4.06	-.14
	منخفض	2.500	.769	.009	.54	4.46
منخفض	مرتفع	-4.600*	.769	.000	-6.56	-2.64
	متوسط	-2.500*	.769	.009	-4.46	-.54

*. The mean difference is significant at the 0.05 Level

جدول 5.6. المقارنات المتعددة وفق طريقة Bonferroni

يظهر اختبار Bonferroni وجود فروق ذات دلالة إحصائية بين مبيعات المتاجر بحسب مستوى الترويج داخل المتجر فجميع المقارنات تظهر فروقاً ذات دلالة معنوية بين متوسطي المجموعتين المقارنتين. وبالعودة إلى الجدول الأول الذي يظهر متوسط المبيعات للمتاجر بحسب طريقة الترويج المتبعة نلاحظ أن المتاجر التي قامت بأعلى مستوى ترويج قد حققت أعلى مستوى مبيعات ($M = 8.3$) تلتها المتاجر ذات الترويج المتوسط ($M = 6.3$) فيما جاءت المتاجر التي قامت بأقل جهد ترويجي ضمن المتجر في المرتبة الأخيرة من حيث المبيعات ($M = 3.7$). تدل هذه النتائج أن على المتاجر التي ترغب بزيادة مبيعاتها أن تقوم بجهود أكبر في مجال الترويج ضمن المتجر.

ولاختبار أثر الترويج داخل المتجر في معدل حضور الزبائن أعاد الباحث اختبار ANOVA مع استبدال المتغير التابع أي أنه استبدل المبيعات بمعدل حضور الزبائن فحصل على النتيجة التالية:

Descriptives

الزبائن

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
مرتفع	10	7.40	2.271	.718	5.78	9.02	4	10
متوسط	10	7.40	2.271	.718	5.78	9.02	4	10
منخفض	10	7.40	2.271	.718	5.78	9.02	4	10
Total	30	7.40	2.191	.400	6.58	8.22	4	10

Test of Homogeneity of Variances

الزبائن

Levene Statistic	df1	df2	Sig.
.000	2	27	1.000

ANOVA

الزبائن

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.000	2	.000	.000	1.000
Within Groups	139.200	27	5.156		
Total	139.200	29			

جدول 6.6. اختبار أثر الترويج داخل المتجر في معدل حضور الزبائن

تشير النتيجة بوضوح إلى عدم وجود أثر للترويج داخل المتجر في معدل حضور الزبائن. تعتبر هذه النتيجة منطقية على اعتبار أن الترويج داخل المتجر لن يؤثر إلا في الزبائن المتواجدين في المتجر أما خارجه فإن الزبائن لن يكونوا معرضين لهذا الترويج وبالتالي لن يؤثر هذا النوع من الترويج في اجتذاب زبائن جدد إلى المتجر.

المراجع العربية:

- البلداوي، عبد الحميد عبد المجيد (2007)، أساليب البحث العلمي والتحليل الإحصائي: التخطيط للبحث وجمع البيانات يدوياً وباستخدام SPSS، الطبعة الثالثة، دار الشروق، عمان، الأردن.
- الطويل، ليلى (2014)، منهجية البحث العلمي، كلية الاقتصاد جامعة تشرين، سورية.
- نجيب، حسين علي؛ الرفاعي، غالب عوض صالح (2006)، تحليل ونمذجة البيانات باستخدام الحاسوب: تطبيق شامل للحزمة SPSS، الطبعة الأولى، الأهلية للنشر والتوزيع، عمان، الأردن.

المراجع الأجنبية:

- Blumberg B., Cooper D.R., & Schindler P.S. (2005), Business Research Methods, Mcgraw–Hill, Berkshire.
- Coakes S.J. (2005), SPSS for Windows: Analysis without Anguish, John Wiley, Australia.
- Ho R. (2006), Handbook of Univariate and Multivariate Data Analysis and Interpretation with SPSS, Chapman & Hall/CRC, USA.
- Landau S. & Everitt B.S. (2004), A Handbook of Statistical Analysis Using SPSS, Chapman & Hall/CRC Press, USA.
- Malhotra N.K. (2010), Marketing Research: An Applied Orientation, 6th Edition, Pearson, USA.
- Malhotra N.K. & Briks D.F. (2007), Marketing Research: An Applied Approach, 3rd European Edition, Pearson Education Limited, Italy.
- Mooi E. & Sarstedt M. (2011), A Concise Guide to Market Research: The Process, Data, and Methods Using IBM SPSS Statistics, Springer, Germany.
- Morgan G.A., Leech N.L., Gloeckner G.W., & Barrett K.C. (2004), SPSS for Introductory Statistics: Use and Interpretation, 2nd Edition, Lawrence Erlbaum Associates, USA.
- Pallant J. (2007), SPSS Survival Manual: A Step by Step Guide to Data Analysis Using SPSS for Windows, 3rd Edition, McGraw Hill, USA.

مقترحات وتمارين للفصل السادس

بهدف مساعدة الطالب على مراجعة هذا الفصل وتثبيت الأفكار الأساسية، يمكنه محاولة الإجابة على الأسئلة التالية.

1. ما هي طبيعة المتغير المستقل في تحليل التباين؟
(الحل في الفقرة: 1.1.)
2. ما هي طبيعة المتغير التابع في تحليل التباين؟
(الحل في الفقرة: 2.1.)
3. ما هي فرضية العدم والفرضية البديلة في تحليل التباين؟
(الحل في الفقرة: 2.)
4. ما هي شروط تحليل التباين؟
(الحل في الفقرة: 3.)
5. قم بإنشاء ملف SPSS جديد وعرف المتغيرات وأدخل فيه البيانات الواردة في الجدول رقم 1.6. ثم قم بإعادة اختبار أثر مستوى الترويج داخل المتجر في مستوى المبيعات.