1

المحاضرة الأولى

ما هو برنامج SAS (Statistical Analysis System)؟

- SAS هو نظام تحليل إحصائي للبيانات والبحوث العلمية المختلفة وإيجاد نتائجها مثل القيم الوصفية (المعدل والمتوسط والتباين والانحراف القياسي والمدى ... الخ) وجداول التحليل الإحصائي للمعاملات المستخدمة في البحوث واختبار المتوسطات لمعرفة الاختلافات المعنوية فيما بينها وإيجاد تحليل الانحدار الخطي بمختلف أنواعه وتحليل الارتباط فضلاً عن رسم العلاقة بين المتغيرات المختلفة على شكل خطوط بيانية. وتجري عملية التحليل عن طريق كتابة البيانات واستخدام جمل برمجية مبسطة لإجراء التحليلات المختلفة.

ويعد من أحد البرامج السهلة الاستخدام، فهو مجهز بجميع الأدوات اللازمة لتحليل البيانات منها:

- خزن المعلومات واستعادتها.
- تحوير البيانات وبرمجتها.
- كتابة التقارير.
- التحليل الإحصائي.
- معالجة الملفات

1- خزن المعلومات واستعادتها Information storage and retrieval :

يستطيع SAS قراءة البيانات المكتوبة بأية صيغة ومن أي شكل من الوسائط مثل الكارتات أو الأقراص (HD، Floppy، CD، أو Flash) أو الشرائط ؛ ومن أي برنامج آخر منها البرامج الخاصة بشركة Microsoft (Word، Excel، V.Basic، Access، ...) أو برامج التحليل الإحصائي الأخرى التي تتعامل مع نظام Microsoft Windows بأنواعه (XP، NT أو MILLINIUM أو غيرها) ومن هذه البرامج برنامج SPSS ؛ بحيث تنظم البيانات والقيم بالشكل الذي يتعامل معها برنامج SAS . يمكن دمج مجموعة من البيانات مع مجموعة أخرى ضمن برنامج SAS باستخدام صيغ حفظ السجلات والتي سوف نأتي على شرحها لاحقاً إذ يمكن تحليلها إحصائياً ؛ وبإمكانه سرد تقارير حول محتوى هذه البيانات. مجاميع البيانات في برنامج SAS تكون ذاتية التوثيق، إذ أنها تحتوي على كل من البيانات وتوصيفها، والتركيبة الخاصة لقاعدة البيانات المستخدمة في برنامج SAS بحيث تقلل من مساحة حفظ البيانات.

2- تحوير البيانات وبرمجتها Data modification and programming :

- عندما يراد عمل تحوير بالبيانات، فإنه يوجد صيغ ودالات جاهزة ومتوفرة من خلال استخدام برنامج SAS . بعض من هذه الصيغ تشكل عمليات جاهزة وقياسية مثل استحداث متغير جديد أو المجموع التراكمي أو البحث عن الأخطاء أو النقص في مجموعة البيانات، في حين يوجد صيغ أخرى تكون
-

- أكثر فاعلية من غيرها مثل صيغ (Do / End أو IF- THEN / ELSE) أو غيرها من الصيغ الأخرى.

3- كتابة التقارير Report writing :

- كما أن برنامج SAS يقرأ على الأغلب معظم البيانات كيفما كتبت، كذلك يستطيع كتابة البيانات في أي صيغة يرتئيها، بالإضافة إلى إمكانية إعطائه تقارير مسبقة التشكيل والصيغ (Procedures) من خلال عملية التنفيذ (RUN) لخطوات البرنامج المختلفة، وكذلك بإمكان ملف مستخدم برنامج SAS (SAS USER) أن يصمم وينتج تقارير مكتوبة بأشكال مختلفة إما أن تظهر مرئية على الشاشة أو مطبوعة على الورق.

4- التحليل الإحصائي Statistical analysis :

- تعد صيغ (Procedures) التحليل الإحصائي في برنامج SAS من أروع ما هو متوفر في هذا البرنامج . فإنها تتراوح ما بين صيغ الإحصاء الوصفي البسيط (Simple Descriptive Statistics) لغاية استخدام تقنات المتغيرات المتعددة والمعقدة، فهي أساسية في تصميمها ومقنعة لمستخدمها وبنتائج تحاليل إحصائية تامة وكاملة متلازمة بسهولة الاستخدام لهذه الصيغ أو الأوامر .

5- معالجة الملفات File handling :

- تركيب القيم للمشاهدات للعديد من مجاميع البيانات والتي تعتبر ضرورية لعملية تحليل البيانات، إذ أن SAS يحتوي على أدوات للكتابة والترتيب وتسلسل ودمج وتحديث مجموعة البيانات. فالعديد من الملفات المدخلة يمكن أن تحفظ بصور متشابهة بإعطائها تسميات للامتدادات تدل على أن هذا الملف يحوي على البيانات الخام قبل برمجتها أو بعد برمجتها في حين ملفات أخرى تحمل تسميات لامتدادات أخرى لحفظها على أنها تمثل نتائج التحليل الإحصائي لملفات البيانات تلك. كما يمكن حفظ مجموعة الملفات لشخص ما على شكل مجلد يحمل اسم ذلك الشخص. بالإمكان حفظ هذه الملفات والمجلدات داخل مجلد برنامج SAS أو خارجه.

لماذا SAS ؟

- توجيه البيانات لنظام واحد من التحليل والذي يمكن أن يفي بجميع احتياجات التحليل الإحصائي هو الهدف الأساس من استخدام برنامج SAS وذلك منذ بداية استخدامه عام 1966. عندما يفي محلل البيانات حاجته من صيغ التحليل فإنه سيكون حر في التركيز على النتائج أكثر مما هو على ميكانيكية الحصول على هذه النتائج. وبدلاً من تعلم لغات برمجة أو عدة حزم تحليلية (Statistical Packages)، وبرامج مفيدة أخرى، فإن SAS يفي بجميع هذه الأغراض لما يحويه من صيغ تحليل واسعة وسهلة الاستخدام مما يعود بالفائدة الكلية على المستخدم وإشباع منافعه من خلال هذا البرنامج .

1

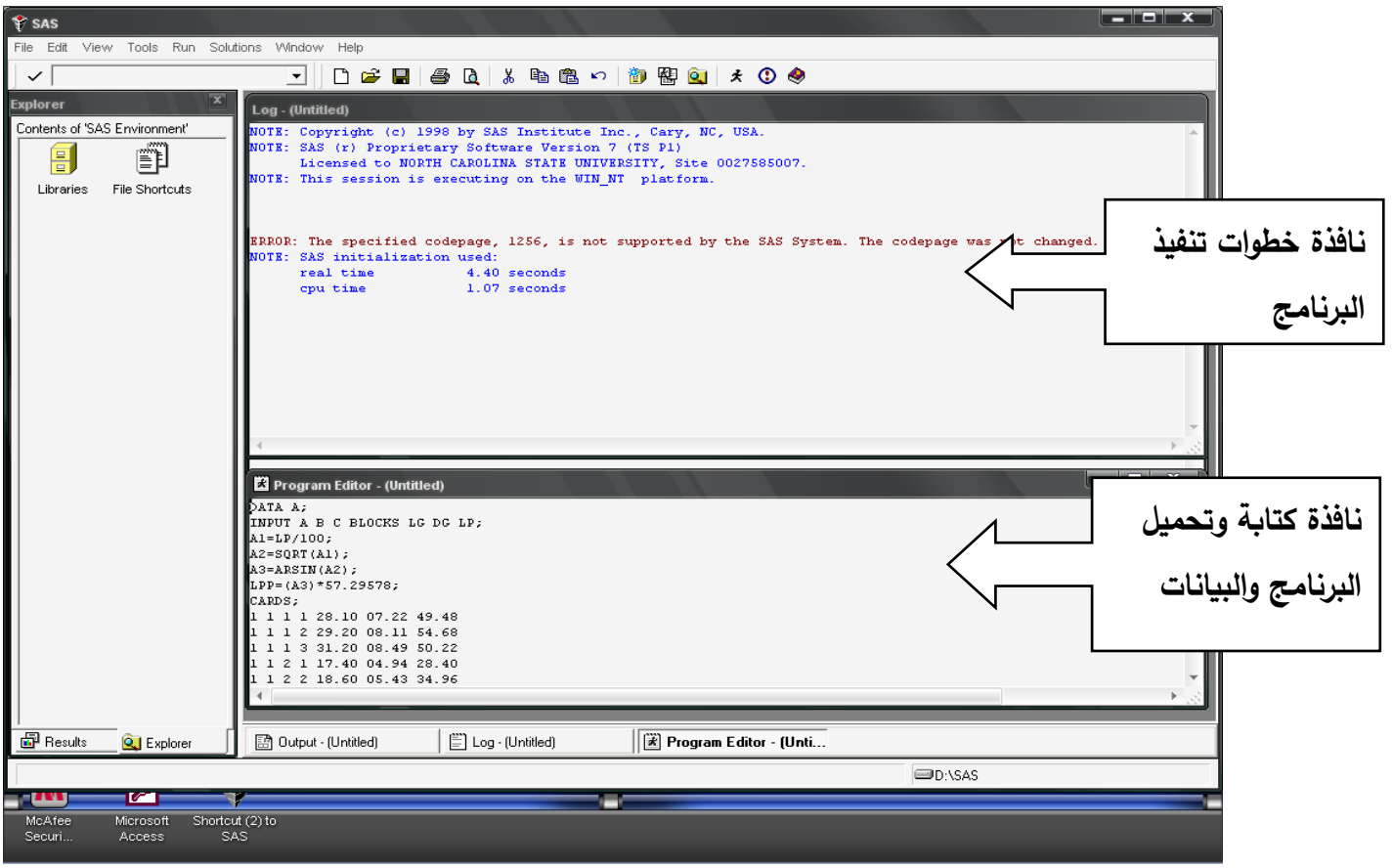
المحاضرة الثانية

نوافذ SAS (SAS Windows)

- ميزة برنامج SAS أنه يحتوي على ثلاث أنواع من النوافذ بالإمكان عرضها بصورة مفردة أو أفقياً أو عمودياً وبأي حجم أو لون يرغبه مستخدم هذا البرنامج كما يمكن الدخول إلى كل نافذة من هذه النوافذ إما عن طريق شريط عنوان النوافذ الموجود أسفل النافذة وفوق شريط المهام أو عن طريق مفاتيح (F) الموجودة في لوحة المفاتيح .

نافذة كتابة وتحميل البرنامج والبيانات (Program Editor Window):

- وتمثل النافذة السفلى من نوافذ البرنامج وتستخدم لكتابة البرنامج وإدخال المتغيرات والبيانات أو استدعاء البيانات المخزونة مسبقاً في ملف بواسطة الأمر (Include) أو استيراد البيانات من برامج أخرى بواسطة الأمر (Import)، أو عن طريق ملف مدخل فيه البرنامج والبيانات ومخزون تحت مسميات عن طريق الأمر (Open) أو عن طريق الأمر (Paste) بالنسبة للبيانات المستنسخة تحت الأمر (Copy) من ملفات معلومة، والذي سوف نقوم بشرح ذلك لاحقاً. مفتاح الدخول لها هو (F5) ومفتاح استعادة البيانات (F4).

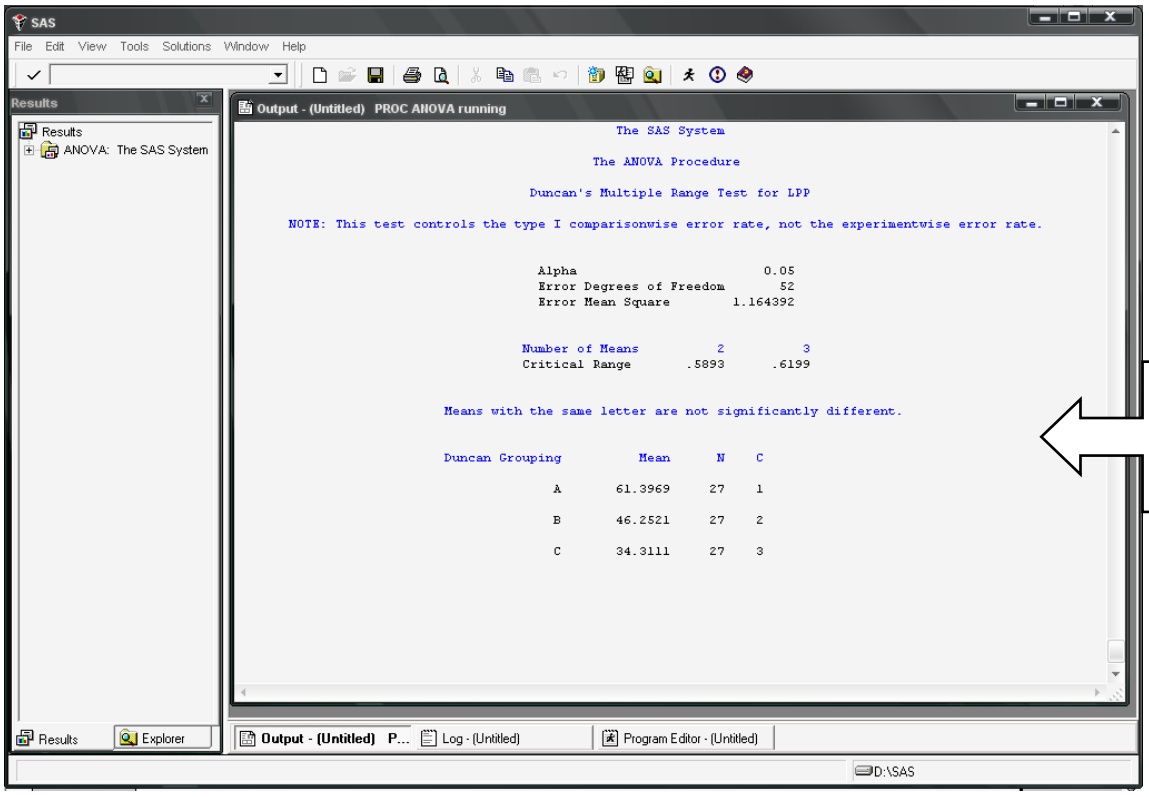


نافذة خطوات تنفيذ البرنامج (Log Window):

- وتتمثل بالنافذة العليا وفيها تعرض التقارير عن تنفيذ خطوات البرنامج وتظهر بصورة مكتوبة ومنه يعرف فيما إذا كانت صيغ المستخدمة بذلك البرنامج مكتوبة بصورة صحيحة ويعرضها بالخط ذو اللون الأسود أو يقترح باستخدام صيغ أخرى أقرب إلى الواقع وحسب طبيعة البيانات المدخلة إذ يعرضها بالخط ذو الأخضر وفي حالة حصول بعض الأخطاء الإملائية البسيطة عند كتابة جمل البرنامج يقوم بتصحيحها ذاتياً على فرض أنها الصيغة الصحيحة للجملة وبدون أي تعطيل في تنفيذ خطوات البرنامج . كما يحدد الأخطاء ونوعيتها بالخط الأحمر حال وجود أي خطأ جوهري في الصيغ المستخدمة أو الجمل المستخدمة في البرنامج . وفي حالة حصول أي نوع من هذه الأخطاء الجوهرية يقوم البرنامج بتعطيل خطوات التنفيذ جميعها أو يقوم بتعطيل تلك الجملة التي حصل فيها خطأ جوهري فيما يستمر اعتياديا بتنفيذ الجمل الصحيحة والتي ليس لها علاقة بذلك الخطأ الجوهري. مفتاح الدخول إلى هذه النافذة هو (F6).

نافذة المخرجات (النتائج)(Window Output):

- وفيها تعرض نتائج التحليل الإحصائي بصورتها التفصيلية وحسب الصيغ المكتوبة في جمل البرنامج عبر نافذة (Program Editor window) والمطلوب الحصول على نتائجها عبر هذا البرنامج . مفتاح الدخول إلى هذه النافذة باستخدام (F7) .



من الذي يستخدم برنامج SAS ؟ Who uses SAS program

- هدف كل مشروع هو الحصول على بيانات، وهذه البيانات يجب أن يتم تحليلها إحصائياً عن طريق إدخالها عبر برامج مختلفة ومن ثم استدعائها وطبعها ثم تدقق من الأخطاء أثناء عملية الإدخال، وأحياناً قد تحتاج إلى إجراء بعض التحويرات في هذه البيانات ولاسيما إذا كانت تمثل نسب مئوية متفاوتة لتحويلها إلى قيم مقلوب جيب الزاوية أو ما يسمى بالتحويل الزاوي للقيم (arcsine) أو تحويلها إلى قيم لوغاريتمية (log10 or lin) أو خلق متغيرا جديد مستوحى من المتغيرات الموجودة في البيانات وغيرها من العمليات الحسابية الأخرى وذلك قبل البدء بكتابة برنامج التحليل الإحصائي المطلوب، كل ذلك متوفرة ضمن فقرات وصيغ برنامج SAS . إذاً برنامج SAS متوفر لجميع الذين يحتاجون إجراء تحاليل إحصائية لبياناتهم. وغالبية مستخدمو البرنامج هم من شريحة العلماء ونخص منهم العاملون في مجالات البحوث الإحصائية و بحوث التسويق، والباحثون في مجال العلوم الطبية وعلوم الحياة، والعلوم الزراعية وعلوم الثروة الحيوانية، كما يستخدم من قبل الأعضاء التنفيذيون في مجال التجارة وعلوم الرياضة والعلوم التربوية والنفسية والجهات البحثية العلمية الأخرى ومحللي برامج الكومبيوتر، بالإضافة إلى طلاب الدراسات العليا للاختصاصات المذكورة أعلاه وطلاب الكليات والمدارس وكل من يحتاج هذا البرنامج لتطبيقه في حياتهم العملية.

إن الأشخاص الذين يريدون استخدام برنامج SAS لغرض التحاليل الإحصائية عليهم أن يكون لديهم بعض الأمور العامة منها:

- عليهم جمع البيانات وتبويبها وترتيبها وإدخالها في برنامج.
- تكوين عدة أسئلة والتي بالإمكان الإجابة عليها عبر تنفيذ خطوات التحليل الإحصائي.
- ج- كتابة برنامج التحليل الإحصائي بخطواته المتسلسلة والصحيحة.
- د – تنفيذ البرنامج.
- هـ- الحصول على النتائج الصحيحة وتلخيصها وعرضها بهدف مناقشتها.

المحاضرة الثالثة

الخطوات العامة لكتابة برنامج Steps of SAS programs

- **هنالك بعض الملاحظات المهمة قبل البدء بشرح خطوات البرنامج منها:**
- **كل جملة من جمل برنامج SAS يجب أن تنتهي بفاصلة منقوطة (;) semicolon للدلالة على نهاية الجملة .**
- **يجب ترك مسافة space بين كلمة وأخرى داخل كل جملة من جمل SAS .**

والآن ندرج الخطوات اللازمة عند كتابة أي برنامج باستخدام SAS

1- جملة تسمية البيانات المعدة للتحليل The DATA step

- **قبل أن نستطيع أن نخبر SAS بالإجابة على الأسئلة التي وضعناها فإنه أولاً يجب علينا تسمية مجموعة البيانات المطلوب إجراء التحليل الإحصائي لها .**

DATA A ; أو DATA FITNESS ;

- **والتسمية هنا إما أن تدل على إسم معين معلوم لدى المحلل أو يدل على نوعية البيانات على أن لا تتعدى عدد أحرفه عن ثمانية أحرف، أو تكون التسمية على شكل حرف واحد أو حرفين أو حرف ورقم أو حرفين ورقم .**
- **وعلى أساس هذه التسمية سوف يتم تنفيذ العمل للخطوات اللاحقة أو إجراء التحويرات أو الإضافة على هذه البيانات أو دمجها مع بيانات أخرى ذات مسميات أخرى أو حفظها في الملفات وكتابة التقارير عنها.**

2- جملة مدخلات البرنامج The INPUT Step :

Example:

INPUT ID NAME $ SEX $ HEIGHT WEIGHT ;

- **عند كتابة جملة المدخلات INPUTفإنه يجب أن يكون لدينا أسماء ورموز للمتغيرات variables المدخلة والمتغيرات تقسم إلى قسمين:**
- **متغيرات عددية أو رقمية**

- متغيرات حرفية

- متغيرات عددية أو رقمية Numerical variables وتتمثل بإطلاق أسماء أو أحرف لهذه المتغيرات وأيضاً يجب أن لا تتعدى عدد الأحرف عن ثمانية أحرف لكل متغير كما HEIGHT أو WEIGHT أو حرف واحد أو حرفين أو حرف ورقم أو حرفين ورقم كما في IDأو Aأو A10 إلخ...، ويقابل هذه المتغيرات في جملة البيانات اللاحقة ثوابت عددية Numerical constants.
- متغيرات حرفية (Character variables): نتبع كما هو الحال في المتغيرات العددية إلا أنها تتبع بالرمز $، كما هو الحال في NAME $ أو gender $ على أن يتم ترك مسافة واحدة بين اسم المتغير والرمز $ للدلالة على المتغير ثابت حرفي Character constant
- أقصى عدد من المتغيرات يمكن أن يتحمله برنامج SAS هو 4000 متغير لمجموعة واحدة من البيانات. كما يجب أن لا ننسى أن نختم جملة INPUTبفاصلة منقوطة (;) كدلالة على انتهاء جملة INPUT .

2- جملة بدأ البيانات The CARDS Step :

CARDS ;

- عندما تكتب جملة CARDS وبعدها الفاصلة المنقوطة (;) دلالة على بدأ بالبيانات، وتكتب البيانات بالترتيب والتسلسل ألموقعي في جملة INPUT
- ولا يجوز استبدال بيانات لمتغير ما مكان متغير آخر وتكون البيانات مسطرة على شكل أعمدة
- إذ يقوم برنامج SAS بقراءة المتغير الأول من جملة INPUT يتبعها قراءة عمود البيانات الأول لحين أخر رقم مدخل في ذلك العمود ثم يعود لقراءة المتغير الثاني ثم العمود الثاني من البيانات وهكذا دواليك إلى أن ينتهي من قراءة المتغيرات والبيانات جميعاً
- عندها سوف ينتقل إلى الجمل الأخرى اللاحقة . كذلك يجب وضع الفاصلة المنقوطة بعد آخر رقم من البيانات للدلالة على الانتهاء من البيانات .

4- جملة الاستثناءات The OPTIONS Step :

OPTIONS pagesize=500 nodate nonumber linesize=120 center ;

- Pagesize : Minimum (أقل حجم) = 20 Maximum (أعلى حجم) = 500,
- Default (عدد الأسطر الاعتيادي) = 60 for number of lines in one page)
- Linesize : for line length the default (عدد الخلايا أفقياً) = 80 , more will be 120-132 .

- (: Center يكون ناتج عملية الطباعة يحتل وسط الورقة المطبوعة (The printout will

- Nodate : عدم طباعة التاريخ في أعلى كل صفحة في النتائج
- Nonumber : عدم طباعة ارقام الصفحات في أعلى كل صفحة في النتائج

5- جملة عنوان التحليل : The TITLE Step

- وفيها يتم طباعة عنوان التحليل في بداية كل عملية تنفيذ لصيغة من صيغ التحليل.
- وتكون داخل علامتي التنصيص المفردة
- وتمثل في هذه الحالة ثابت سلسلة حرفية
- (Character Constant)
- وتقرأ الكلمات مهما كان عدد أحرفها مادامت داخل علامتي التنصيص هذه.

6- جملة صيغ التحليل : The PROC Step

- جملة PROC هي ملخص لكلمة PROCEDURE (صيغة) :
- ذكرنا في الخطوات السابقة كيفية معاملة مجموعة البيانات وإدخالها ضمن حدود الجمل السابقة
- في حين جملة الصيغ (PROC) هي عبارة عن برامج متكاملة للتحليل الإحصائي المتعدد والذي يقوم بقراءة مجموعة البيانات وتنفيذ خطوات التحليل الإحصائي اللازمة ثم طباعة نتائج هذا التحليل.
- بعض من صيغ التحليل تعمل على استحداث مجاميع بيانات جديدة ناتجة عن البيانات الأصلية.
- لدى استخدامك برنامج SAS فإنك تستطيع أن تكتب الجمل والخطوات التي ترغب في إستخدامها في التحليل الإحصائي والحصول على نتائجها بالصورة التي ترغبها خلافاً عن باقي برامج التحليل الإحصائي الجاهزة الأخرى.

- أحياناً تتطلب صيغ SAS جمل إضافية لكل تعطي الصيغة المستخدمة معلومات ونتائج أوسع لما أنت تريده. على سبيل المثال عند استخدام جملة تحليل التباين:

PROC ANOVA ;

- فإنها سوف تتطلب إلى جملة أخرى وهي جملة تصنيف المعاملات مثلاً :

CLASSES TRT;

- **وهنا TRT تمثل التأثير المستقل (Independent effect).**
- **ومن ثم جملة المعادلة المطلوب التحليل وفقها:**

MODEL weight length = TRT;

- **وهنا weight و length يمثلان المتغيرات المعتمدة (Dependent Variables).**
- **وإذا طلبنا اختبار المتوسطات للمعاملات (TRT) باستخدام اختبـار دنكن المتعدد المدى نتبعها بالجملة الآتية:**

Means TRT / Duncan;

إليك بعض جمل الصيغ المستخدمة في برنامج SAS :

PROC ANOVA; , PROC AUTOREG ; , PROC CHART ;

PROC CLUSTER ; , PROC CORR ; , PROC EDITOR ;

PROC FACTOR ; , PROC FREQ ; , PROC GLM ;

PROC MATRIX; , PROC MEANS ; , PROC NESTED ;

PROC NLIN ; , PROC PRINT ; , PROC REG ;

PROC SORT ; , PROC STEPWISE ; , PROC TTEST ;

7- جملة التنفيذ RUN

- **ولا بد من اختتام جملة الصيغ (PROC) بجملة التنفيذ (RUN ;) إذ بدونها لا يتم التحليل الإحصائي أو ظهور النتائج.**
- بالطبع هنالك العشرات من الصيغ الأخرى التي يتعامل بها برنامج SAS، كما أنه لكل صيغة من الصيغ المذكورة أعلاه يوجد جمل أخرى تتبعها أو استثناءات (Options) نود من البرنامج أن يستخرجها وأحياناً أخرى لا نريد أن يستخرجها لنا.
- وسوف نشرح لاحقاً عن كيفية استخدام بعض من هذه الصيغ مشفوعة بأمثلة.

الأمر Proc print;

- **يطبع إجراء PRINT المشاهدات في مجموعة بيانات SAS باستخدام كل أو بعض المتغيرات.**
- **قد ترفق معها جملة VAR لتحديد المتغيرات المراد طباعتها**

تمرين1

لديك البيانات التالية والتي تمثل اسم الطالب وجنسه ودرجته

المطلوب: اكتب برنامج بلغة SAS لتعريف المتغيرات قراءة البيانات في الجدول وطباعتها

name	gender	grade
Ahmad	m	76.5
khalid	m	80.3
sabah	f	83.75
majed	m	80.2
nabil	m	78.4
zainab	f	78.0

الحل تمرين 1:

```
DATA AGRIC;
INPUT NAME $ GENDER $ GRADE;
CARDS;
AHMED   M   76.5
KHALID  M   80.3
SAABAH  F   83.75
MAJED   M   80.2
NABIL   M   78.4
ZAINAB  F   78.0
;
OPTIONS PAGESIZE=500 NODATE NONUMBER;
TILTLE'UNIVERSITY OF MOSUL';
PROC PRINT;
VAR NAME;
RUN;
```

المحاضرة الرابعة

الدوال Functions

هنالك العديد من الدوال التي يمكن استخدامها في برنامج SAS ومن هذه الدوال:

1- الدوال الرياضية Arithmetic Functions :

ABS : دالة القيم المطلقة وصيغتها " ABS (قيمة رقمية) " : هذه الدالة سوف ينتج عنها قيم موجبة مناظرة للقيم المدخلة . أمثلة :

ABS (2.4) = 2.4

ABS (-3)= 3

INT: عرض الأعداد الرقمية الصحية أو بترها وصيغتها " INT(قيمة رقمية) ": في هذه الدالة سوف ينتج الجزء الصحيح من القيمة المدخلة. ففي حالة الرقم المدخل موجب فإن القيمة الرقمية المناظرة أيضاً تكون موجبة، وكذلك إذا كانت القيمة المدخلة سالبة فإن القيمة الرقمية المناظرة أيضاً تكون سالبة.

أمثلة :

INT(2.1) = 2

INT(-2.4) = -2

INT(3) = 3

INT(-1.6) = -1

MAX : أكبر قيمة وصيغتها " MAX(مجموعة من القيم) ": هذه الدالة يمكن أن تكون من أو أكثر من القيم. نتيجتها سوف تتضمن القيمة الكبرى من بين هذه القيم.

أمثلة:

MAX(2,6) = 6

MAX(2,-3) = 2

MAX(3,-3,MISSING) = 3

MIN : أقل قيمة وصيغتها " MIN(مجموعة من القيم) ": هذه الدالة يمكن أن تكون من أو أكثر من القيم. نتيجتها سوف تتضمن القيمة الصغرى من بين هذه القيم.

أمثلة:

MIN(2,6) = 2

MIN(2,-3) = -3

MIN(0,4,MISSING) = 0

2

MOD : باقي القسمة لقيمتين وصيغتها " MOD(القيمة الأولى، القيمة الثانية) ": نتيجة هذه الدالة سوف يكون باقي القسمة للقيمة الرقمية الأولى على القيمة الرقمية الثانية.

أمثلة:

MOD(6 , 3) = 0

MOD(10 , 3) = 1

MOD(11 , 3.5) = 0.5

MOD(10 , -3) = 1

SQRT : الجذر التربيعي وصيغته " SQRT(قيمة رقمية) ": نتيجة هذه الدالة هو قيمة الجذر التربيعي للقيمة الرقمية وهذه القيمة يجب أن لا تكون سالبة.

مثال :

SQRT(9) = 3

ARSIN : مقلوب جيب الزاوية وصيغتها " ARCSIN (قيمة رقمية) : هذه الدالة سوف تنتج عن بما يسمى Arc sine (مقلوب جيب الزاوية أو $sine^{-1}$) لتلك القيمة الرقمية. النتيجة سوف تكون زاوية نصف قطرية (radians) . القيمة الرقمية يجب أن تأخذ قيمة مابين -1 إلى +1 .

أمثلة:

ARSIN(0) = 0

ARSIN(1) = 1.57079

COS : جيب تمام المثلثاني وصيغتها " COS(قيمة رقمية) ": نتيجة هذه الدالة سوف تكون جيب تمام القيمة الرقمية. يفترض بقيمة هذه الدالة أن تمثل زاوية نصف قطرية.

أمثلة :

COS(0.5) = 0.877582

COS(0) = 1

SIN : الجيب المثلثاتي وصيغته " SIN(قيمة رقمية) ": هذه الدالة سوف ينتج عنها جيب الزاوية للقيمة الرقمية. يفترض بقيمة هذه الدالة أن تمثل زاوية نصف قطرية.

أمثلة:

SIN(0.5) = 0.47925

SIN(0) = 0

LOG : اللوغاريتم الطبيعي (natural logarithm) وصيغتها " LOG (قيمة رقمية) ": نتيجة هذه الدالة سوف تكون اللوغاريتم الطبيعي (Naperian) للقيمة الرقمية. القيمة الرقمية يجب أن تكون قيمة موجبة. الأساس للوغاريتم هو e، وتقريبا قيمته = 2.71828

أمثلة:

LOG(1) = 0

LOG(10) = 2.30259

LOG10 : اللوغاريتم الاعتيادي (common logarithm) وصيغتها " LOG10 (قيمة رقمية) ": هذه الدالة سوف ينتج عنها في لوغاريتم اعتيادي (لوغاريتم للأساس 10) للقيمة الرقمية. القيمة الرقمية يجب أن تكون قيمة موجبة.

أمثلة :

LOG10(1) = 0

LOG10(10) = 1

LOG10(100) = 2

وهناك بعض الدوال الإحصائية منها:

الدالة	الإحصاء	الأمثلـــــة
N	عدد القيم الرقمية وبدون فقد	N(1,0,MISSING) = 2
SS	عدد القيم المفقودة	N(1,0,MISSING)= 1
SUM	المجموع	SUM(2,6) = 8
MEAN	الوسط الحسابي (المعدل)	MEAN(2,6) = 4
MIN	أقل قيمة	MIN(2,6) = 2
MAX	أعلى قيمة	MAX(2,6) = 6
RANGE	المدى	RANGE(2,6,3) = 4
STD	الإنحراف القياسي	STD(2,6) = 2.83

4

VAR	التباين	VAR(2,6) = 8
CSS	مجموع المربعات المصححة	CSS(2,6) = 8
USS	مجموع المربعات غير المصححة	USS(2,6) = 40
STDERR	الخطأ القياسي للمتوسط	STDERR(2,6) = 2

هنالك العديد من الدوال الأخرى إلا أننا سوف نكتفي بذكر بهذا القدر من الدوال.

1

المحاضرة الخامسة

استحداث بيانات جديدة من مجموعة البيانات المدخلة

Creating SAS Data Sets from Input Data :

أحيانا يتطلب من المحلل الإحصائي إلى استحداث بيانات جديدة من البيانات المدخلة، فجملة INPUT تصف ترتيب المتغيرات المدخلة في حين جملة CARDS تعتبر كإشارة لبدء البيانات عند إدخالها عبر البرنامج وبحسب تسلسلها في جملة المدخلات INPUT . ألا أنه بالإمكان إدخال جمل أخرى موقعها مابين جملة INPUT وجملة CARDS وهذه الجمل هي التي تحمل صيغ للمتغيرات الجديدة المستحدثة لكي تصطف ضمن جملة INPUT وإن لم تكن بين صفوفها . وهنالك عدة صيغ لاستحداث بيانات جديدة منها:

1- استخدام العمليات الحسابية أو الدوال : ويقصد بالعمليات الحسابية استخدام عمليات الجمع (+) والطرح (-) والضرب (*) والقسمة (/) وإيجاد النسب المئوية (%) من خلال معادلة تستخدم بهدف إيجاد قيم متغير معين، وكذلك استخدام دالة الرفع للقوى 2 (**) إي مربع الأعداد الواقعة ضمن المتغير المطلوب تربيع مفردات أعداده أو استخدام دالة ARCSIN (مقلوب جيب الزاوية) لتعديل النسب المئوية التي تم استخراجها في المعادلة السابقة أو إيجاد الجذر التربيعي لمجموعة من القيم الواقعة ضمن متغير معين (SQRT).

مثال:

```
DATA  FITNESS;
INPUT  A  B ;
C = A + B ;
D = A – B ;
E = A * B ;
F = A / B ;
G = ((A – B)/(0.5*(A+B))*100;
H = G/100;
I =ARSIN(H);
J = A**2 ;
K = SQRT(A);
CARDS ;
6  2
```

2

8 4

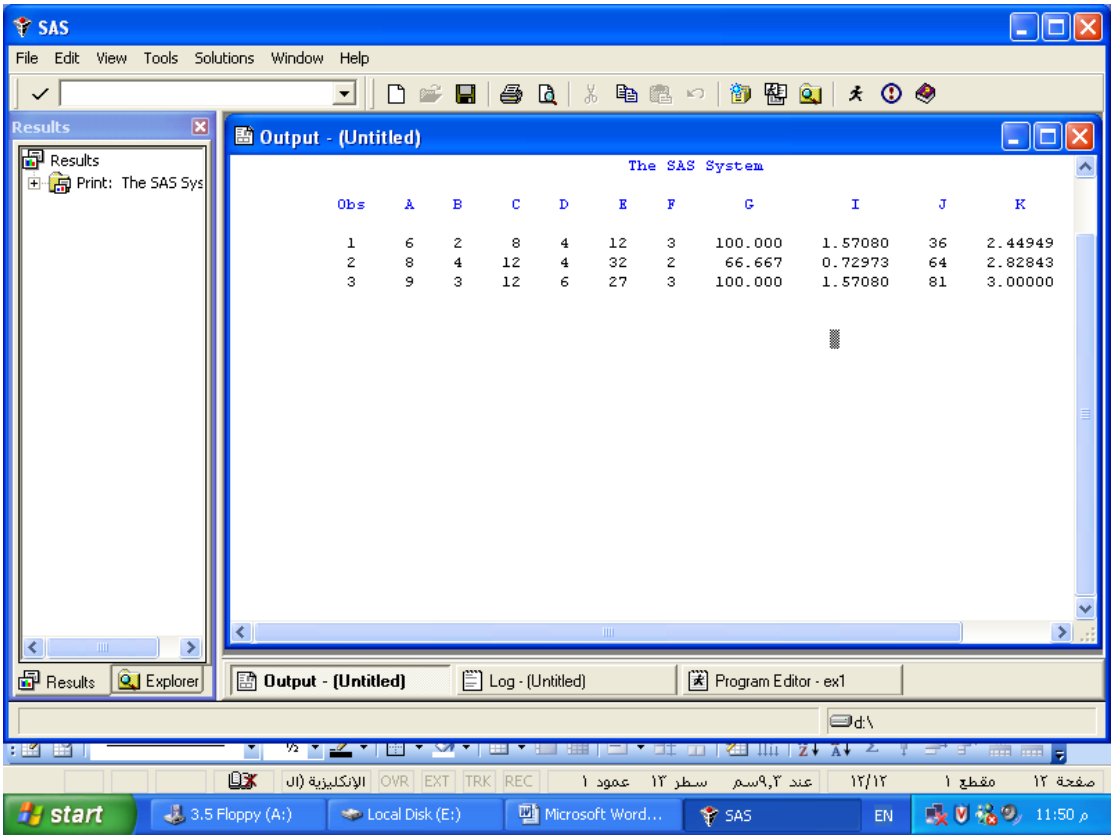
9 3

;

OPTIONS PAGESIZE=500 NODATE NONUMBER;

PROC PRINT DATA=FITNESS; VAR A B C D E F G I J K ; RUN ;

نلاحظ أنه تم استحداث متغيرات جديدة (C، D، E، F، G، I، J و K) من المتغيرين الأصليين (A و B).
ونتيجة تنفيذ البرنامج سوف تكون كما يلي:



3

تمرين:

A	B
6	2
8	4
9	3

- لديك البيانات التالية: اكتب برنامج بلغة SAS لإيجاد ما يلي:
- جمع المتغيرين
- طرح المتغيرين
- ضرب المتغيرين
- قسمة المتغيرين
- $\frac{A-B}{\sqrt{A+B}} * 100$
- مقلوب جيب الزاوية للمتغير B
- الجدر التربيعي للمتغير A
- اللوغاريتم الطبيعي للمتغير B

الحل:

```
DATA A;
INPUC=A+B;
T A B;
D=A-B;
E=A*B;
F=A/B;
G((A-B)/(A+B)**0.5)*100;
H=SIN(G);
I=SQRT(A);
J=LOG(B);
CARD;
6  2
8  4
9  3
```

```
;
OPTIONS PAGESIZE=500 NODATE NONUMBER;
TITLE 'CREATING NEW VARIABLES' ;
PROC PRINT;
VAR A B C D E F G H I J;
RUN;
```

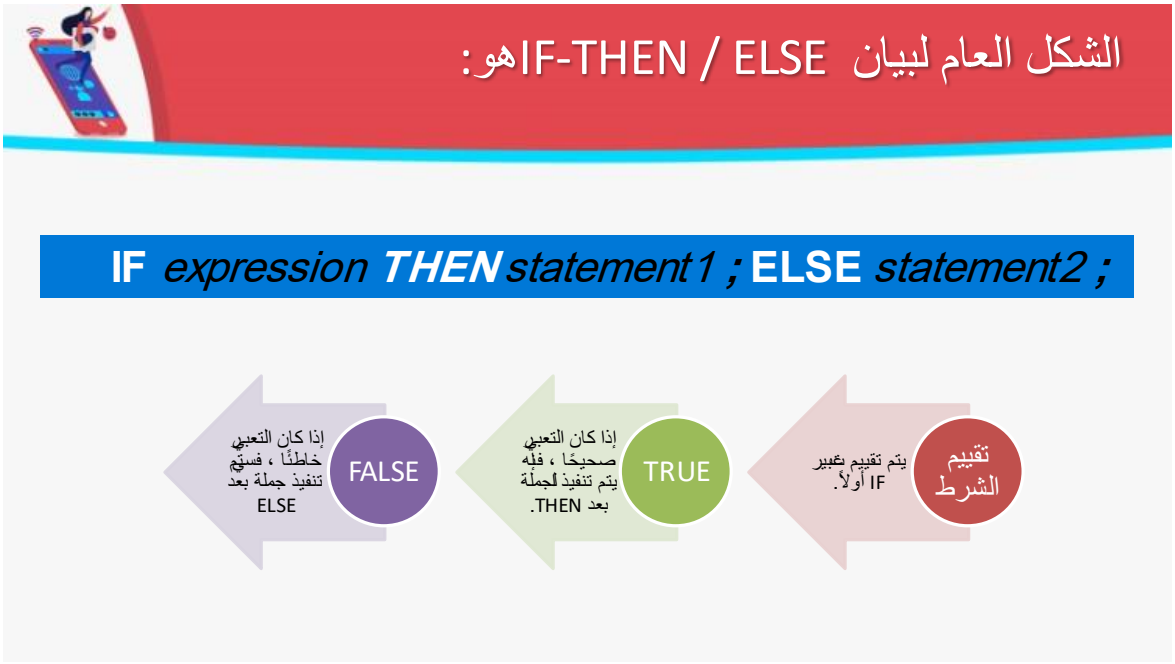
1

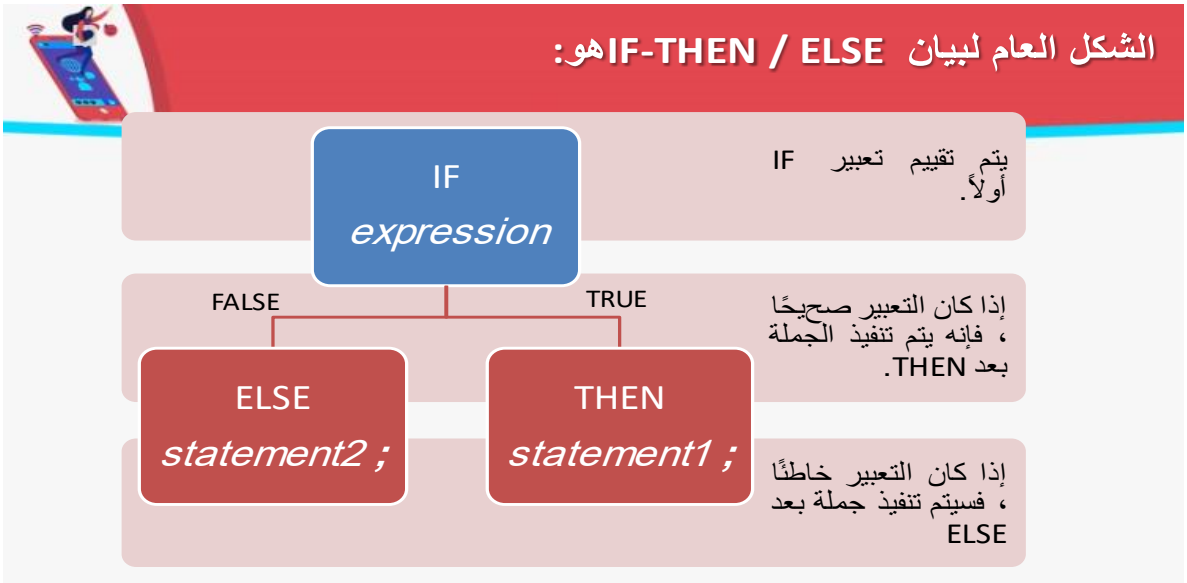
المحاضرة السادسة

- استحداث بيانات باستخدام الجمل الشرطية:

أحيانا يتطلب تحوير بعض المتغيرات من متغير حرفي إلى متغير عددي أو بالعكس أو اســـتحداث مجموعة من المتغيرات وذلك عن طريق استخدام جملة الشرط IF- THEN-ELSE :

استخدم عبارة IF لإجراء عملية مشروطة ، لاختبار تعبير. إذا كان التعبير صحيحًا ، فإنه يتم تنفيذ الجملة بعد THEN. إذا كان التعبير خاطئًا ، فسيتم تنفيذ جملة بعد ELSE





2

أمثلة:

- مثال1
  - if x=y then w=-1; else w=0;
- مثال2
  - if x>y then w=-1; else w=0;
- مثال3
  - if x<y then w=-1; else w=0;
- مثال4
  - if x>=y then w=-1; else w=0;
- مثال5
  - if x<=y then w=-1; else w=0;
- مثال6
  - if x!=y then w=-1; else w=0;

تمرين 1

- لديك البيانات التالية:

year
1960
1976
1982

- اكتب برنامج بلغة sas لاستحداث متغير حرفي جديد باسم COLOR تكون قيمته BLUE اذا كانت السنة 1976 والا فقيمته تكون RED .

```
DATA A;
INPUT YEAR;
IF YEAR=1976 THEN COLOR="BLUE"; ELSE COLOR="RED";
CARDS;
1960
1976
```

```
1982
;
OPTIONS PAGESIZE=500 NODATE NONUMBER;
PROC PRINT;
RUN;
```

فإن نتيجة تنفيذ البرنامج سوف يكون باستحداث بيانات جديدة تحت إسم COLOR وموزعاً حسب جملة الشرط :

The SAS System

OBS	YEAR	COLOR
1	1960	RED
2	1976	BLUE
3	1982	RED

ملاحظة : يمكن أن نستخدم العديد من الرموز الحسابية مع جملة الشرط ومنها: ≥، ≤، <>، > و ≠ .

تمرين : استخدام جملة IF لحذف اسطر بيانات

لديك البيانات التالية: اكتب برنامج بلغة SAS لحذف كل سطر بيانات تكون فيه قيمة المتغير A اكبر من قيمة المتغير B .

A	B
12	10
10	12
06	04
03	08

```
DATA A;
INPUT A B;
IF A > B THEN DELETE  ;
CARDS;
12 10
```

4

```
10 12
06 04
03 08
;
OPTIONS PAGESIZE=500 NODATE NONUMBER;
PROC PRINT;
RUN;
```

ناتج تنفيذ البرنامج بحذف القيم التي تفي بالشرط وهو عندما تكون قيم A أكبر من قيم B .

A	B
10	12
3	8

1

المحاضرة السابعة

إستخدام جملة PROC SORT : وتستخدم هذه الصيغة في عملية تصنيف البيانات وترتيبها حسب الطلب أو الحاجة أثناء التحليل الإحصائي كما يجب أن يستخدم مع هذه الجملة جملة أخرى وهي جملة BY .

تمرين

AGE	WEIGHT	LENGTH
18	65	167
19	66	164
16	57	159
14	48	145

اكتب برنامج بلغة SAS لترتيب البيانات حسب العمر

الحل:

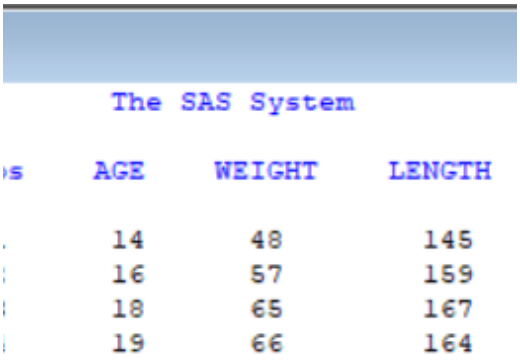
```
DATA  AGR;
INPUT  AGE  WEIGHT   LENGTH;
CARDS;
18   65   167
19   66   164
16   57   159
14   48   145
;
options pagesize=500 nodate nonumber ;
PROC  SORT ;  BY   AGE ;
PROC  PRINT ;  RUN;
```

نتيجة تنفيذ البرنامج سوف يكون:

AGE	WEIGHT	LENGTH
14	48	145
16	57	159
18	65	167
19	66	164

2

نتيجة تنفيذ البرنامج سوف يكون:



2

تمرين

مثال: في دراسة تضمنت الجرعة السمية dose وعدد الفئران الكلي number وعدد الفئران الهالكة killed .

dose number killed
1.69 59 6
1.72 60 13
1.75 62 18
1.78 56 28
1.81 63 52
1.83 59 53
1.86 62 61
1.88 60 60

المطلوب: استحداث متغير آخر يمثل نسبة الهلاكات الى العدد الكلي وترتيب البيانات حسب عدد الفئران؟

```
data t;
  input dose number killed;
 prop = killed/number;
cards;
1.69 59 6
1.72 60 13
1.75 62 18
1.78 56 28
1.81 63 52
1.83 59 53
1.86 62 61
1.88 60 60
Option nodate nonumber;
PROC SORT;
BY NUMBER;
PROC PRINT;
RUN;
```

3

المحاضرة الثامنة

تطبيقات في الإحصاء الوصفي (Descriptive Statistics) :

جدول التوزيع التكراري Frequency Tables وتستخدم فيها الصيغة الآتية :

: PROC FREQ

1- One-way frequency : إذا طلب بتكوين جدول توزيع تكراري ذات اتجاه واحد فإن صيغة هذه الجملة سوف تكون :

```
PROC FREQ ;
TABLES A ;
```

تمرين :

لديك البيانات التالية والتي تمثل اعمار وجنس مجموعة من الطلبة

AGE	GENDER
12	M
16	F
28	F
20	M
22	M
18	F
20	M

المطلوب: اكتب برنامج بلغة sas لإيجاد جدول التوزيع التكراري ذات اتجاه واحد لكل من العمر وجنس الطالب

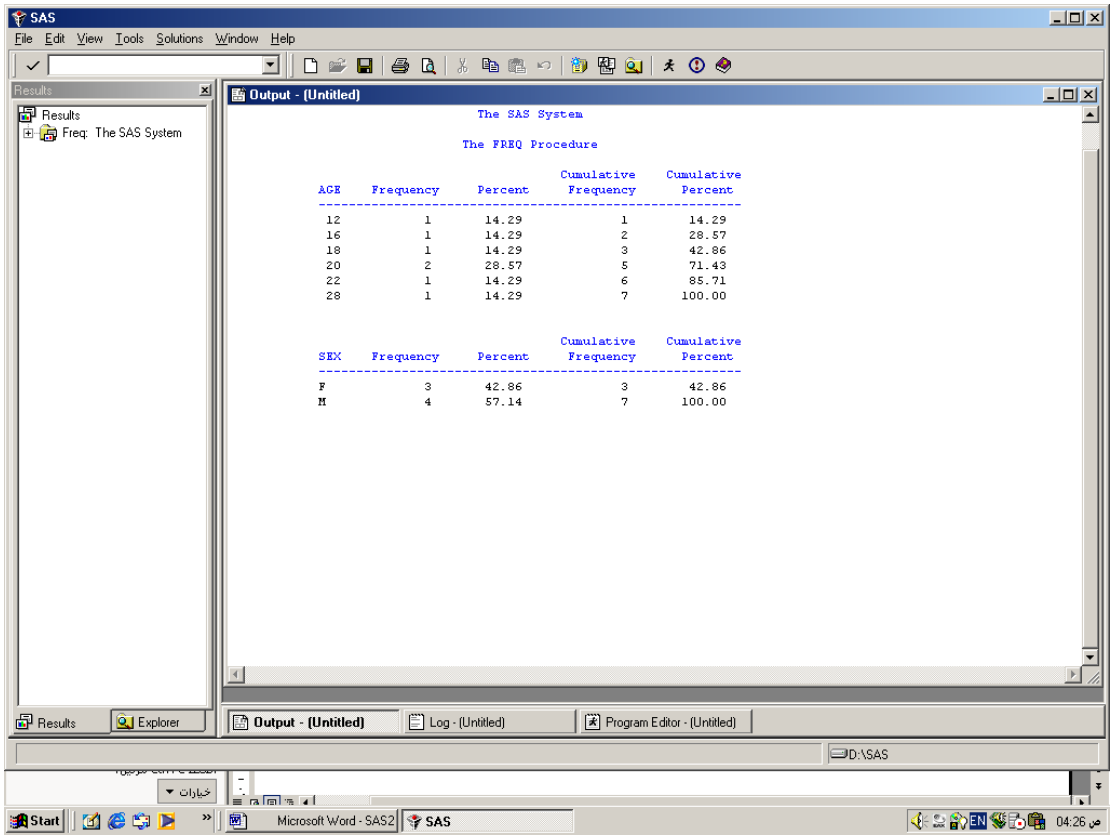
الحل:

```
DATA FREQ;
INPUT AGE SEX $;
DATA FREQ;
INPUT AGE SEX $;
CARDS;
12 M
16 F
28 F
20 M
22 M
18 F
20 M
;
options pagesize=500 nodate nonumber;
```

4

```
PROC FREQ;
TABLES AGE;
TABLES SEX;
RUN;
```

النتائج:



المحاضرة التاسعة

2- TWO –way crosstabulation tables : إذا طلب تحليل جدول تكراري باتجاهين متعامدين، يشترط أن يوجد متغيرين، أحدهما سوف يتمثل بالصفوف والآخر سوف يتمثل بالأعمدة . كما هو في المثال السابق :

تمرين :

لديك البيانات التالية والتي تمثل اعمار وجنس مجموعة من الطلبة

AGE	GENDER
12	M
16	F
28	F
20	M
22	M
18	F
20	M

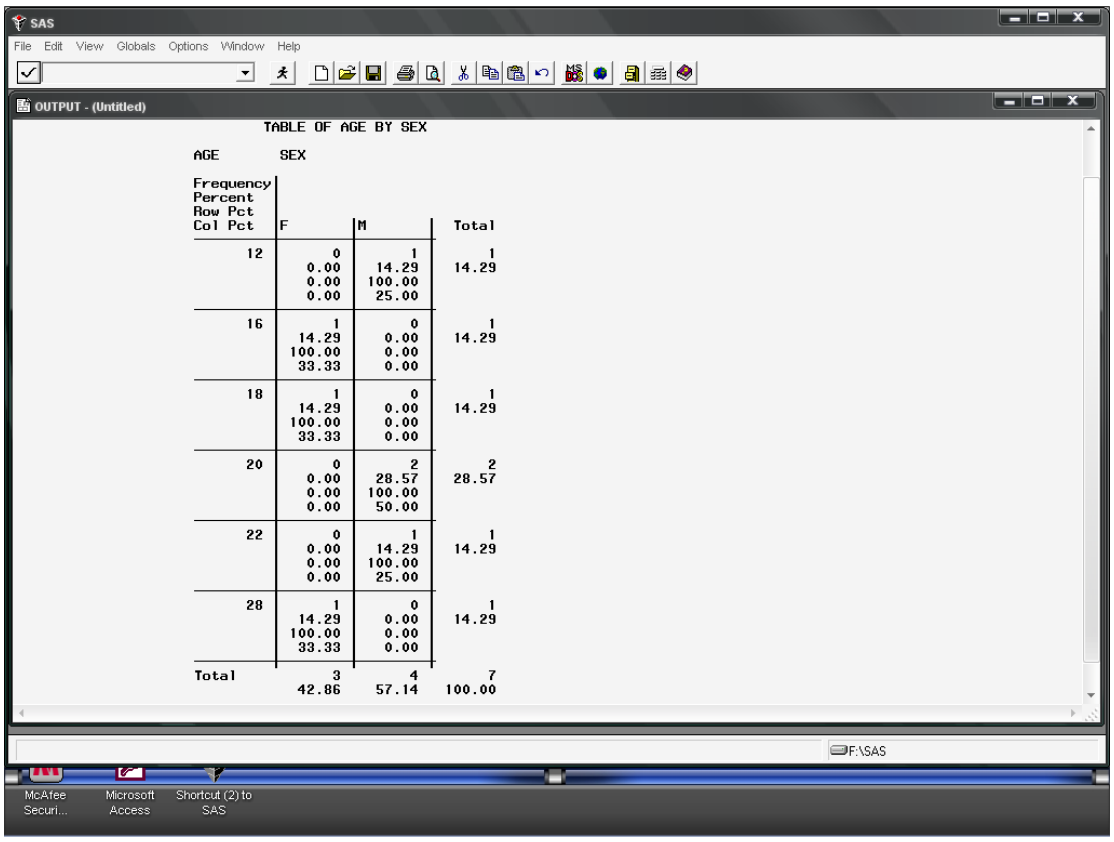
المطلوب: اكتب برنامج بلغة sas لإيجاد جدول التوزيع التكراري ذات اتجاهين متعامدين لكل من العمر وجنس الطالبز

الحل:

```
DATA   FREQ;
INPUT  AGE    SEX  $;
CARDS;
12  M
16  F
28  F
20  M
22  M
18  F
20  M
;
options pagesize=500 nodate nonumber;
PROC  FREQ;
TABLES  AGE*SEX / CHISQ;  RUN;
```

لذلك سوف ينتج عن تنفيذ التحليل الجدول المبين أدناه:

2



يوضح الجدول التوزيع التكراري للجنس لكل من الذكور M والإناث F وذلك حسب الأعمار، كما تظهر لنا النسب المئوية Percent في السطر الثاني من النتائج لكل جنس ولكل عمر على حدا فيما يبين السطر الثالث نسبة كل صف Row Percent والسطر الرابع نسبة كل عمود Column Percent، وأسفل الجدول يظهر لنا المجموع الكلي للنسب التي يجب أن تساوي 100 % .

هنالك بعض الاستخدامات للمطاليب الإضافية OPTIONS ومنها استخدام اختبار مربع كاي (χ^2) (Chisequare Test) لتعطي النتيجة المبينة أدناه :

Statistics for Table of AGE by SEX

Statistic	DF	Value	Prob
Chi-Square	5	7.0000	0.2206
Likelihood Ratio Chi-Square	5	9.5607	0.0887
Mantel-Haenszel Chi-Square	1	0.3225	0.5701
Phi Coefficient		1.0000	
Contingency Coefficient		0.7071	
Cramer's V		1.0000	

WARNING: 100% of the cells have expected counts less than 5. Chi-Square may not be a valid test.

Sample Size = 7

المحاضرة العاشرة

مقاييس التوسط ومقاييس التشتت

Central Tendency and Dispersion Measurements

يمكن استخراج جميع قيم مقاييس التوسط و التشتت عبر صيغة PROC MEANS متبوعة بالمطالب المختلفة من هذه المقاييس .

مثال :

PROC MEANS <option-list>; قائمة الإختيارات

وإليك بعض من هذه الإختيارات للمقاييس الإحصائية حسب تسلسلها :

عدد المشاهدات، بدون مشاهدة مفقودة، الوسط الحسابي، الإنحراف القياسي، أدنى رقم، أعلى رقم، المدى، المجموع، التباين، مجموع المربعات غير المصحح، مجموع المربعات المصحح.

- N NMISS MEAN STD MIN MAX RANGE SUM VAR USS CSS CV STDERR

- وظيفتها:
- عدد المشاهدات، بدون مشاهدة مفقودة، الوسط الحسابي، الانحراف القياسي، أدنى رقم، أعلى رقم، المدى، المجموع، التباين، مجموع المربعات غير المصحح، مجموع المربعات المصحح، معامل الاختلاف، الخطأ القياسي.

بالإمكان استخدام واحد أو أكثر من هذه المقاييس بكتابتها بعد صيغة PROC MEANS وقبل الفاصلة المنقوطة (;) ، ويمكن تقديم أو تأخير أحد المقاييس عن الآخر وحسب الترتيب الذي ترغبه . إذا لم تذكر أي مقياس إحصائي مع صيغة PROC MEANS فإن صيغة PROC MEANS سوف تظهر أسماء المتغيرات المذكور في الجملة التي تعقبها وهي جملة VAR ثم أعدادها والوسط الحسابي والإنحراف القياسي وأدنى قيمة وأعلى قيمة :

Name of the variable N MEAN STD MIN MAX

بالإضافة عند طلب التصنيف (CLASS) فإنه سوف يطبع عدد المشاهدات N Obs .

تمرين:

أخذت عينة من ستة أشخاص وسجلت أعمارهم وكانت كما يلي :

Yi = 10 18 22 17 13 18

المطلوب : أوجد باستخدام برنامج SAS مقاييس التوسط التشتت التالية: متوسط الأعمار Yi، التباين S2 الانحراف القياسي S، الخطأ القياسي Sy، معامل الاختلاف CV%

الحل:

```
DATA A;
INPUT yi ;
CARDS;
10
18
```

2

22
17
13
18
;
OPTIONS PAGESIZE=500 NODATE NONUMBER;
PROC MEANS MEAN VAR STD STDERR CV;
VAR yi ;
RUN ;

النتائج:

The SAS System

The MEANS Procedure

Analysis Variable : AGE

Mean	Variance	Std Dev	Std Error	Coeff of Variation
16.3333333	17.8666667	4.2268980	1.7256239	25.8789673

1

المحاضرة الحادية عشر

اختبار المتوسطات وتحليل التباين Test of Means and Analysis of Variance

أولاً – اختبار t (t-test) :

يستخدم إختبار (t) لفحص الفرضية بعدم وجود فروق معنوية بين المتوسطات المطلوب اختبارها. ويعتبر هذا الإختبار أحدى طرق الإختبار ذات الإتجاه الواحد في تحليل التباين لمقارنة مستويين من التصنيفات. PROC TTEST يحسب مقاييس t المبنية على أساس الفرضية بأن التباينين لمجموعتين من البيانات متساوية. وفي نفس الوقت تحسب المقاييس على أساس الفرضية البديلة بوجود أختلافات بين تباينات المجموعتين . إذ يظهر لكل قيمة من قيم t، درجات الحرية (degrees of freedom) ومستوى الإحتمالية. ويلاحظ أن الفرضيات المخطط لها في اختبار t بأن هذه المتغيرات تتوزع توزيعاً طبيعياً و مستقلاً داخل كل مجموعة من مجاميع المقارنة . والصيغة المستخدمة في هذا الإختبار هي:

```
PROC TTEST DATA= SAS-data-set ;
    CLASS variable;                    /* required */
    VAR variables;
    BY variables;
```

الأمر Proc ttest يستعمل لغرض اجراء اختبار t لاختبار الاختلافات بين متوسطي مجموعتينز

CLASS variable ; : إعطاء اسم متغير التصنيف (أو التجميع) ويجب أن يصاحب عبارة PROC TTEST في اختبار T لعينتين مستقلتين.

- **يجب أن يحتوي متغير الفئة على مستويين ، ومستويين فقط.**
- **إذا تم استخدامه بدون عبارة VAR، فإن جميع المتغيرات الرقمية في بيانات الإدخال باستثناء تلك التي تظهر في عبارة CLASSسيتم تضمينها في التحليل.**
- **يمكنك استخدام إما متغير رقمي أو حرف في جملة .CLASS**

تمرين:

نفرض لدينا سلالتين معينتين من الدجاج و أخذنا عينتين كل منهما مؤلفة من 5 دجاجات وسجل متوسط أوزانها فهل يوجد فرق بين السلالتين أم لا ؟ استخدم برنامج SAS في الحل.

السلالة الأولى = 280 ، 300 ، 310 ، 260 ، 280 ، والسلالة الثانية = 260 ، 255 ، 245 ، 220 ، 230

الحل :

```
DATA HENS;
INPUT BREED EGG;
```

2

```
CARDS;
280   1
300   1
310   1
260   1
280   1
260   2
255   2
245   2
220   2
230   2
;
OPTIONS PAGESIZE=500 NODATE NONUMBER;
PROC TTEST   ; CLASS  BREED;  VAR  EGG ; RUN;
```

النتائج

Statistics

Variable	BREED	N	Lower CL Mean	Mean	Upper CL Mean	Lower CL Std Dev	Std Dev	Upper CL Std Dev	Std Err	Minimum	Maximum
EGG	1	5	261.8	286	310.2	11.679	19.494	56.016	8.7178	260	310
EGG	2	5	221.13	242	262.87	10.07	16.808	48.298	7.5166	220	260
EGG	Diff (1-2)		17.456	44	70.544	12.294	18.2	34.868	11.511		

T-Tests

Variable	Method	Variances	DF	t Value	Pr > |t|
EGG	Pooled	Equal	8	3.82	0.0051
EGG	Satterthwaite	Unequal	7.83	3.82	0.0053

Equality of Variances

Variable	Method	Num DF	Den DF	F Value	Pr > F
EGG	Folded F	4	4	1.35	0.7808

المحاضرة الثانية عشر

صيغة تحليل التباين PROC ANOVA : وهي واحدة من الصيغ العديدة في المستخدمة في نظام التحليل الإحصائي SAS بهدف تحليل التباين . ANOVA تعمل على تحليل التباين لقيم المتغيرات المستمرة كما هي مقياس للعديد من المتغيرات المصنفة(في جملة CLASS)
صيغة PROC ANOVA مصممة للعمل على البيانات المتوازنة (عدم وجود قيم مفقودة) إي أن كل متغير يمتلك إلى نفس العدد من قيم المشاهدات المستجيبة لكل متغير من المتغيرات المعروضة في جملة التصنيف (CLASS) . فإذا كانت غير متوازنة فإن صيغة ANOVA سوف لا تعطي النتائج المطلوبة وعليه يجب استخدام صيغة (PROC GLM) والتي صممت من أجل تحليل التباين غير المتوازن .(كما سوف نذكرها لاحقاً) .
والصيغة المستخدمة في PROC ANOVA والجمل الملحقة بها كما يلي :

```
PROC ANOVA ;
CLASS variables;                    /* required */
    MODEL dependents=effects / options;          /* required */
    MEANS effects / options;
```

جملة CLASS وهي من الجمل الأساسية والتي يذكر فيها المتغيرات أسماء المتغيرات المستقلة أو المتغيرات المؤثرة والمراد إختبارها، مثل أنواع المعاملات أو العوامل المؤثرة وذلك حسب التجارب المستخدمة وتصاميمها مثل : التصميم العشوائي الكامل CRD ، أو تصميم القطاعات العشوائية الكاملة RCBD، أو التجارب العاملية Factorial experiments … ألخ.
أما في جملة المعادلة MODEL وهي أيضاً من الجمل الأساسية فيجب وضع أسماء المتغيرات المعتمدة (dependents) في الجهة اليسرى من المعادلة والمتغيرات المؤثرة (independents) في الجهة اليمنى من المعادلة . أما الإختيارات (options) فسوف نتجاهلها الآن وهي تستخدم من قبل المستويات المتقدمة في التحليلات الإحصائية.
أما جملة MEANS وهي جملة اختبار المتوسطات إذ يذكر فيها أسماء العوامل المؤثرة والمذكورة في جملة CLASS . اما الإخيارات (options) المذكورة فيها فهي تمثل أنواع الإختبارات للمتوسطات ومن أشهرها هو إختبار TTEST لأختبار متوسطين فقط و اختبار DUNCAN المتعدد المدى لإختبار عدد غير محدد من المتوسطات واختبارLSD أقل فرق معنوي والذي يختبر أيضاً العديد من المتوسطات، وكذلك اختبار WALLERS والذ هو اختبار دنكن المعدل وهناك اختبارات أخرى متعددة تستخدم حسب الهدف من التحليل والتخصص المعين للشخص المحلل وحاجته من أنواع الإختبارات.

تمرين
مثال : بيانات حاصل الفردي من فول الصويا مسجلة في مخطط تجربة لدراسة تأثير أربعة مستويات من التسميد النتروجيني بثلاث مكررات :

t_2	15	t_4	13	t_3	20	t_2	9
t_3	17	t_1	15	t_1	17	t_4	17
t_4	12	t_2	12	t_3	23	t_1	22

المطلوب : هل يوجد فرق بين مستويات التسميد ؟ استخدم برنامج SAS في الحل واختبر الفروق بين المتوسطات باستخدام اختبار DUNCAN .

2

الحل :

```
DATA BEAN;
INPUT  TRT   REP  PROD;
CARDS;
1  1  17
1  2  15
1  3  22
2  1  09
2  2  15
2  3  12
3  1  20
3  2  17
3  3  23
4  1  13
4  2  17
4  3  12
;
OPTIONS PAGESIZE=500 NODATE NONUMBER;
PROC ANOVA;
CLASS  TRT;
MODEL PROD = TRT ;
MEANS  TRT / DUNCAN;
RUN ;
```

المحاضرة الثالثة عشر

صيغة PROC GLM :

وتسمى تحليل التباين باستخدام معادلة خط الإنحدار العام (General Linear Model) ويمكن لهذه الصيغة من تحليل البيانات المتقطعة أو المصنفة (Discrete variables) ذات المقياس النوعي كما أنها تعمل على تحليل البيانات المستمرة (Continuous variables) ذات المقياس الكمي وعليه فإن صيغة GLM يمكنها أن تشمل في طياتها العديد من أنواع التحاليل منها :

- simple regression الإنحدار البسيط
- multiple regression الإنحدار المتعدد

تحليل التباين (ANOVA) وبالأخص للبيانات غير المتوازنة .

- analysis of variance (ANOVA), especially for unbalanced data
- analysis of covariance تحليل التباين المشترك
- response-surface models معادلات الإستجابة السطحية
- weighted regression الإنحدار الموزون
- polynomial regression الإنحدار متعدد الحدود
- partial correlation معامل الإرتباط الجزئي

تحليل التباين متعدد المتغيرات

- multivariate analysis of variance (MANOVA)

تحليل التباين متكرر المقاييس

- repeated measures analysis of variance.

والصيغة المستخدمة له والجمل التابعة لها كما يلي :

```
PROC GLM options ;
    CLASS variable-list;
    MODEL dependents= independents / options;      /* required */
    MEANS effects / options;
```

كما هو الحال في صيغة PROC ANOVA ماعدا أن البيانات هنا غير متزنة .

استخدام صيغة GLM للوحدات التجريبية غير المتوازنة:

مثال : في تجربة لدراسة أربعة مستويات من التسميد النتروجيني (صفر ، 10 ، 20 و 30 كغم / دونم) على حاصل القطن ، سجلت البيانات المبينة في أدناه والتي تمثل حاصل القطن للوحدة التجريبية :

t_1	24	t_4	140	t_2	92		98 t_3
t_2	115	t_1	52	t_1	64	t_3	100
t_4	172	t_2	66	t_3	45	t_1	45

المطلوب : هل يوجد فروق معنوية بين معاملات التسميد ، استخدم برنامج SAS في حل المسألة واختبر معنوية الفروق بين المتوسطات باستخدام اختبار DUNCAN المتعدد المدى.

2

الحل :

```
DATA COTTON ;
INPUT  TRT  REP  PROD;
CARDS;
1  1  024
1  2  052
1  3  045
2  1  092
2  2  115
2  3  064
2  4  066
3  1  098
3  2  100
3  3  045
4  1  140
4  2  172
;
OPTIONS PAGESIZE=500 NODATE NONUMBER ;
PROC GLM ;
CLASS  TRT ;
MODEL  PROD = TRT;
MEANS  TRT / DUNCAN ;
 RUN;
```

1

المحاضرة الرابعة عشر

استخدام صيغ معامل الارتباط PROC CORR :

- هو مقياس لدرجة العلاقة بين المتغيرين Y, X ويرمز له بالرمز r، ويحقق معامل الارتباط الخطي المتباينة:
- أي أن قيمة معامل الارتباط محصورة بين +1 و 1-، وتدل قيمته على درجة العلاقة بين المتغيرين أو المتغيرات موضع الدراسة من حيث أنها قوية، متوسطة، أو ضعيفة،
- وأما الإشارة فإنها تصف نوعية العلاقة هل هي عكسية أم طردية، فالإشارة السالبة تدل على وجود علاقة عكسية أما الموجبة فتدل على وجود علاقة طردية بين المتغيرين موضع الدراسة.
  - صيغة معامل الارتباط تقيس قوة العلاقة الخطية بين متغيرين.
  - أي أن المتغير x يمكن أن يعبَّر عنه بالضبط كدالة خطية للمتغير y
  - وقيمة معامل الارتباط تتراوح بين +1 و -1
  - معتمدة فيما إذا كان هنالك علاقة مباشرة سواءً طردية أم عكسية بين المتغيرين x و y
- فعندما يوجد علاقة قوية بين المتغيرين فإن قيم x و y سوف تكون قريبة أو تقع على خط التوافق.
- أما إذا كانت العلاقة صفر بين المتغيرين فذلك يعني بأن المتغير المعين ليس على القابلية التنبؤية الخطية التي يمكن أن يعبر عنها للمتغير الآخر.
- وإذا كانت تتوزع توزيعاً طبيعياً فإن علاقة صفر تعني بأن المتغيرات مستقلة بعضها عن البعض.

الصيغة المستخدمة للارتباط الخطي البسيط والجمل الملحقة بها كما يلي:

```
PROC CORR <option-list>;
    VAR variable-list;
```

- قائمة options تشمل اختيار نوع الارتباط المطلوب وكما ذكرنا منه Pearson وSpearman
- أو غيرهما وعدم كتابتها فالبرنامج يختار ارتباط Person
- أما قائمة المتغيرات variables فتشمل جميع المتغيرات المطلوب إيجاد معاملات الارتباط بينها.

2

تمرين 1:

- لدراسة العلاقة بين الدخل والاستهلاك بالدنانير في مدينة غزة، أخذت عينة مكونة من عشرة أسر فأعطت النتائج التالية:

الدخل	300	350	500	600	900	1000	900	1200	1050	250
الاستهلاك	280	340	500	550	800	750	850	1050	1000	250

المطلوب: اكتب برنامج بلغة sas لإيجاد معامل الارتباط بين المتغيرين الدخل والاستهلاك

الحل:

```
DATA  REG;
INPUT  income  consump;
CARDS;
300     280
350     340
500      500
600      550
900      800
1000    750
900      850
1200    1050
1050    1000
250      250
;
OPTIONS PS=500 NODATE NONUMBER;
 PROC CORR ;
 VAR income  consump;
RUN ;
```

النتائج

3

The SAS System

The CORR Procedure

2 Variables: income consump

Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
income	10	705.00000	345.96885	7050	250.00000	1200
consump	10	637.00000	294.16737	6370	250.00000	1050

Pearson Correlation Coefficients, N = 10
Prob > |r| under H0: Rho=0

	income	consump
income	1.00000	0.98220 <.0001
consump	0.98220 <.0001	1.00000

من النتائج السابقة يمكن استنتاج ما يلي:

- **2- معامل الارتباط بين الدخل والاستهلاك = 0.982 وهو يدل على وجود ارتباط طردي قوي بينهما**
- معنوية الارتباط : **Prob < 0.0001 لذا فان الارتباط معنوي**

المحاضرة الخامسة عشر

إستخدام صيغ معامل الإنحدار PROC REG :

الانحدار الخطي البسيط صيغة PROC REG :
الهدف من دراسة الانحدار هو إيجاد دالة العلاقة بين المتغيرين المستقل والتابع والتي تساعد في تفسير التغير الذي قد يطرأ على المتغير التابع (y) تبعاً لتغير في قيم المتغير المستقل (χ).
معادلة الانحدار:

$$y = a + bx$$

- **صيغة الانحدار هي صيغة عامة للانحدار.**
- **بينما طرق الانحدار الأخرى في نظام SAS تمثل حالة متقدمة من التطبيق .**
- **PROC REG تزودك بتسعة طرق مختارة للمعادلات، اختبار المعادلة الخطية و فرضية المتعددة المتغيرات، تولد رسم الأشكال الموزعة للبيانات ومختلف المقاييس الإحصائية و غيرها من حالات الإحصاء المتقدم .**

والصيغة المستخدمة للانحدار الخطي البسيط والجمل التابعة لها كما يلي :

```
PROC REG;
 MODEL dependents= regressors / <options>;
```

- **تمثل options هنا الخيارات لنوع معادلة الانحدار المطلوبة . وناتج المخرجات output المطلوبة.**

تمرين 1:

- لدراسة العلاقة بين الدخل والاستهلاك بالدنانير في مدينة غزة، أخذت عينة مكونة من عشرة أسر فأعطت النتائج التالية:

الدخل	300	350	500	600	900	1000	900	1200	1050	250
الاستهلاك	280	340	500	550	800	750	850	1050	1000	250

المطلوب: اكتب برنامج بلغة sas لإيجاد نموذج انحدار الاستهلاك على الدخل.

الحل:

```
DATA  REG;
INPUT  income  consump;
CARDS;
300    280
```

2

```
350    340
500    500
600    550
900    800
1000   750
900    850
1200   1050
1050   1000
250    250
;
OPTIONS PS=500 NODATE NONUMBER;
PROC REG ;
 MODEL consump = income;
 RUN;
```

النتائج:

The SAS System

The REG Procedure
Model: MODEL1
Dependent Variable: consump

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	751330	751330	218.73	<.0001
Error	8	27480	3435.01392		
Corrected Total	9	778810			

Root MSE	58.60899	R-Square	0.9647
Dependent Mean	637.00000	Adj R-Sq	0.9603
Coeff Var	9.20078		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	48.22929	43.91308	1.10	0.3040
income	1	0.83514	0.05647	14.79	<.0001

من النتائج السابقة يمكن استنتاج ما يلي:

1- نموذج انحدار الاستهلاك على الدخل هو:

Consump = 48.229 + 0.835 * Income

2- معامل الارتباط بين الدخل والاستهلاك = 0.982 وهو يدل على وجود ارتباط طردي قوي بينهما

$$Sig. = 0.000 < \alpha = 0.05$$