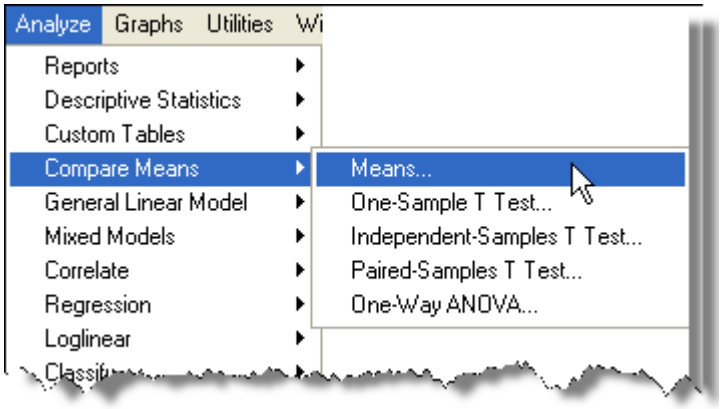
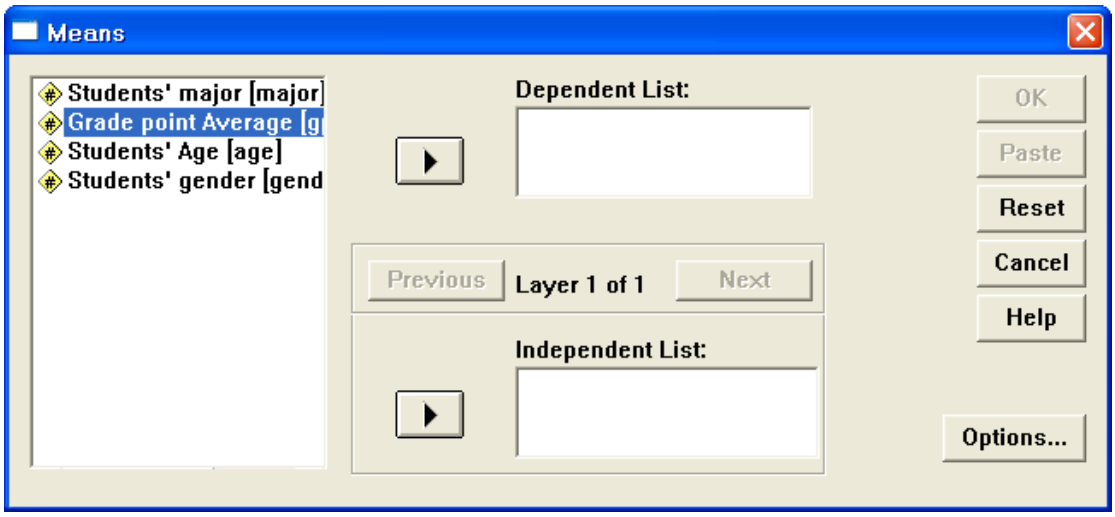
تحليل إحصائي مبسط على بيانات الطلاب.

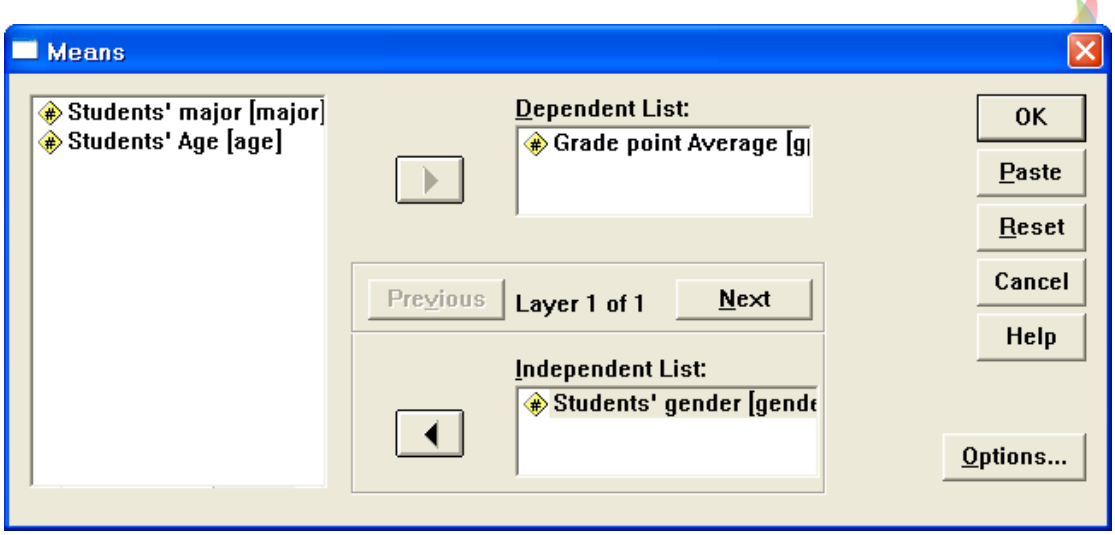
لنفرض أن لدينا الرغبة في معرفة متوسط المعدل التراكمي للطلاب والطالبات مع عرض النتائج بيانياً. ويتم ذلك بالنقر على النحو التالي:



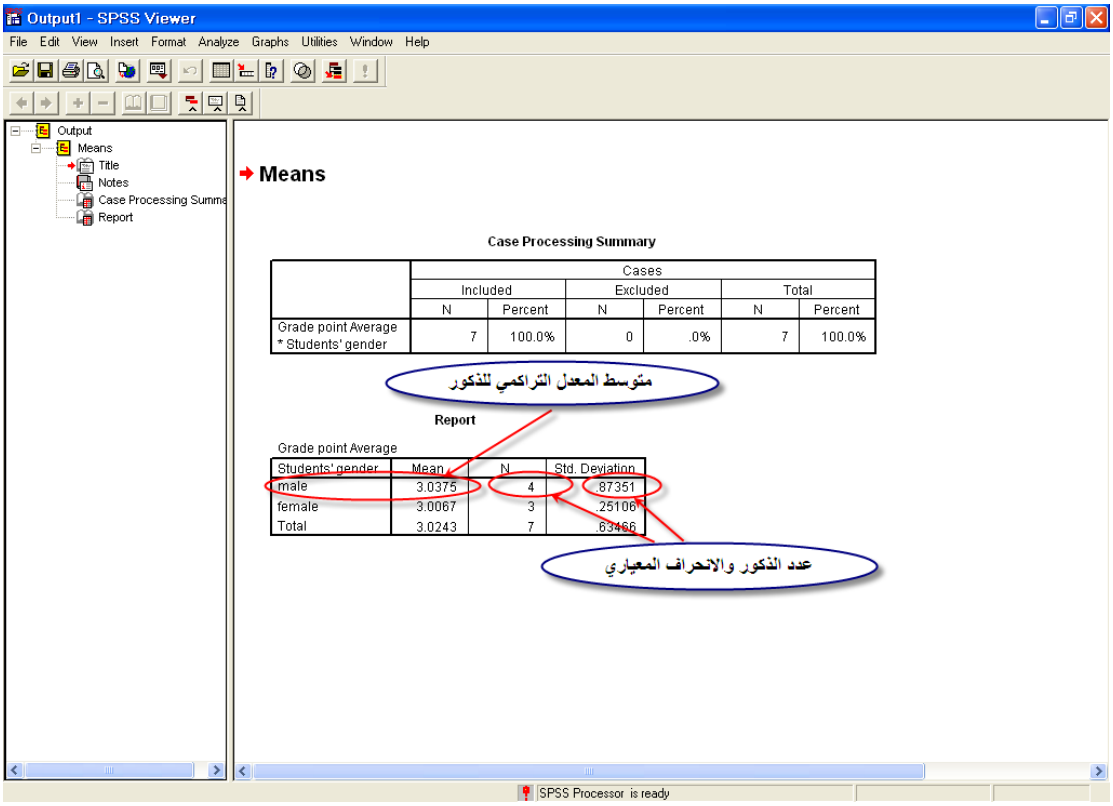
وبذلك يظهر مربع الحوار التالي:



ونظراً لان الهدف هو معرفة متوسط المعدل التراكمي وذلك حسب جنس الطالب (ذكر أو أنثى)، نضع المتغير gpa في قائمة المتغيرات التابعة Dependent List ونضع gender في قائمة المغيرات المستقلة Independent List وذلك بالنقر على ▶ المجاورة لكل قائمة.



وبالنقر على OK تظهر لنا النتائج التالية في شاشة عارض النتائج.



التعامل والتحكم في البيانات
Data Manipulation

مقدمة

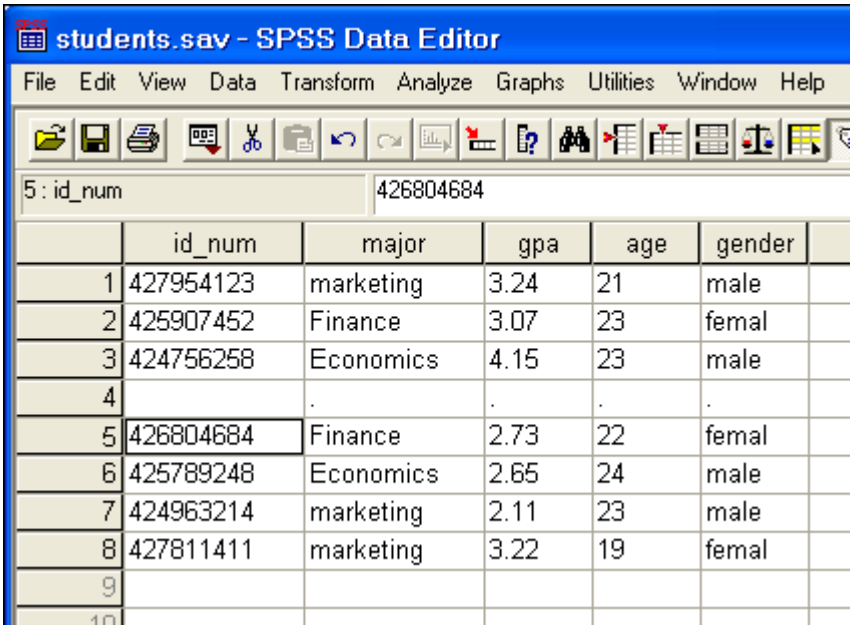
يتطلب التحليل الإحصائي المقدرة على التعامل مع البيانات وذلك لوضعها بالصورة المناسبة للتحليل. وسيتم خلال هذا الفصل القيام بعمليات تهدف إلى تعديل البيانات لتلائم التحليل، وتشمل عمليات إضافة أو حذف متغيرات أو حالات، اختيار مجموعة جزئية من البيانات أو مجموعة جزئية من المتغيرات، البحث عن متغيرات أو حالات، إعادة ترميز قيم متغيرات أو تكوين متغيرات جديدة باستخدام متغيرات متوفرة حالياً.

إضافة وحذف متغيرات وحالات Insert and Delete

تبرز الحاجة دائما إلى إضافة متغيرات جديدة إلى ملف بيانات أو إضافة حالات جديدة، ففي بعض الحالات يتطلب التحليل إضافة متغيرات جديدة أو يتم الاستغناء عن متغيرات غير مفيدة. كذلك فإنه قد تحتاج إلى إضافة مشاهدات جديد للمتغيرات الموجودة حالياً في ملف البيانات وذلك بسبب نسيانها سهواً أو بسبب توفر بيانات جديدة.

إضافة وحذف حالات

يتم إضافة حالة جديد، بالنقر على أيقونة إضافة حالة من شريط الأدوات، أو باختيار Insert Cases من قائمة البيانات.

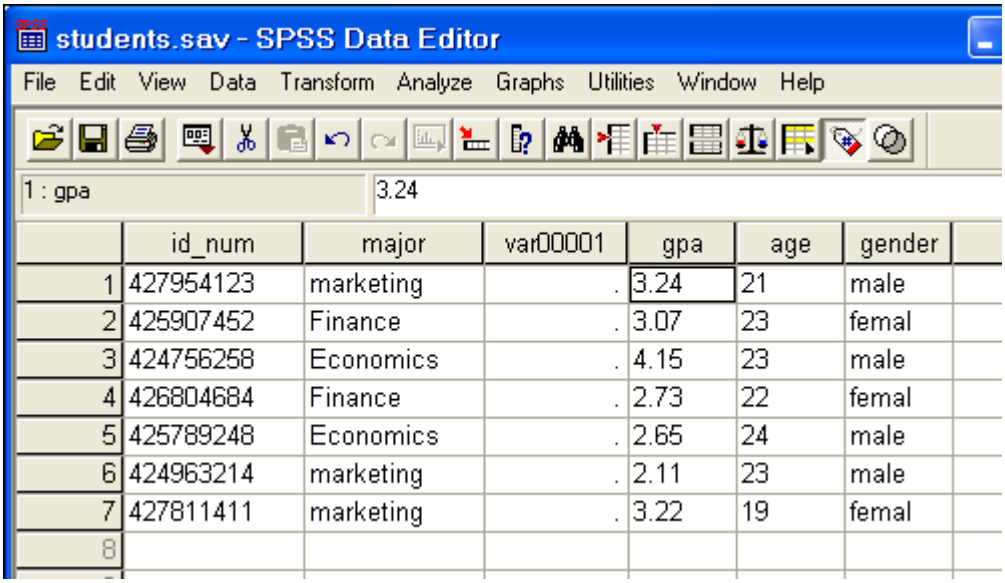


	id_num	major	gpa	age	gender
1	427954123	marketing	3.24	21	male
2	425907452	Finance	3.07	23	femal
3	424756258	Economics	4.15	23	male
4		.	.	.	.
5	426804684	Finance	2.73	22	femal
6	425789248	Economics	2.65	24	male
7	424963214	marketing	2.11	23	male
8	427811411	marketing	3.22	19	femal
9					

وبذلك يتم إضافة حالة جديدة تقع إلى الأعلى من الخلية المختارة. ويمكن حذف حالة ما وذلك بالنقر بواسطة زر الفارة الأيمن على الحالة المراد حذفها ثم اختيار Clear أو باختيار الصف التي تقع فيه الحالة ثم الضغط على زر Delete من لوحة المفاتيح.

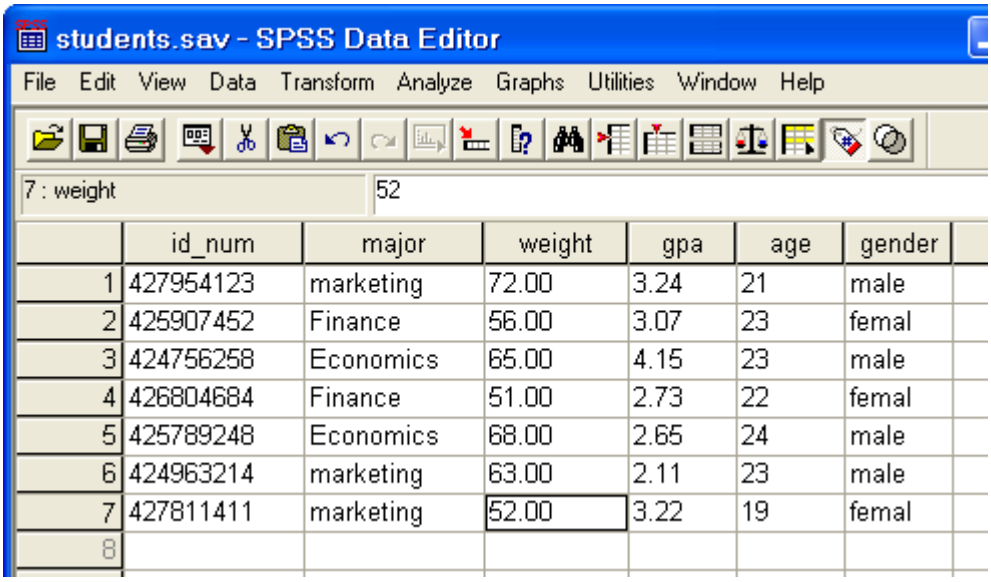
إضافة وحذف متغيرات

لنفرض أن لديك الرغبة في إضافة متغير جديد يمثل وزن الطالب في ملف البيانات، ويتم ذلك بالنقر على أيقونة إضافة متغير من شريط الأدوات، أو باختيار Variable Insert من قائمة البيانات. ويقوم البرنامج بإضافة عمود جديد باسم VAR0001 بحيث يقع إلى اليسار من الخلية المختارة.



	id_num	major	var00001	gpa	age	gender
1	427954123	marketing	.	3.24	21	male
2	425907452	Finance	.	3.07	23	femal
3	424756258	Economics	.	4.15	23	male
4	426804684	Finance	.	2.73	22	femal
5	425789248	Economics	.	2.65	24	male
6	424963214	marketing	.	2.11	23	male
7	427811411	marketing	.	3.22	19	femal
8						

وبعد إضافة المتغير الجديد يتطلب الأمر إجراء العديد من التعديلات على ملف البيانات مثل تعديل خصائص المتغير وإضافة البيانات المقابلة لكل حالة.



	id_num	major	weight	gpa	age	gender
1	427954123	marketing	72.00	3.24	21	male
2	425907452	Finance	56.00	3.07	23	femal
3	424756258	Economics	65.00	4.15	23	male
4	426804684	Finance	51.00	2.73	22	femal
5	425789248	Economics	68.00	2.65	24	male
6	424963214	marketing	63.00	2.11	23	male
7	427811411	marketing	52.00	3.22	19	femal
8						

ويمكن حذف أي متغير وذلك بالنقر بواسطة زر الفارة الأيمن على المتغير المراد حذفه ثم اختيار Clear أو باختيار العمود الذي يقع فيه المتغير ثم الضغط على زر Delete من لوحة المفاتيح.

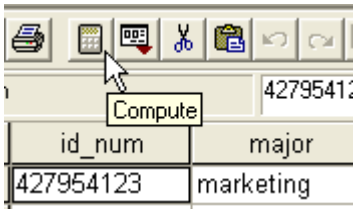
حساب قيم جديدة لمتغير أو إعادة ترميز متغير

تبرز الحاجة في الكثير من الأحيان إلى تحويل بعض المتغير إلى متغيرات جديد تحمل قيم مرتبطة بالمتغير السابق أو استخدام عدة متغيرات لإنشاء متغير جديد. كذلك قد نحتاج إلى إعادة ترميز متغير بحيث تكون القيم التي أقل من 10 في متغير ما تحمل الرقم 1 في المتغير الجديد والقيم التي بين 10 وأقل من 20 تحمل الرقم 2 في المتغير الجديد وهكذا.

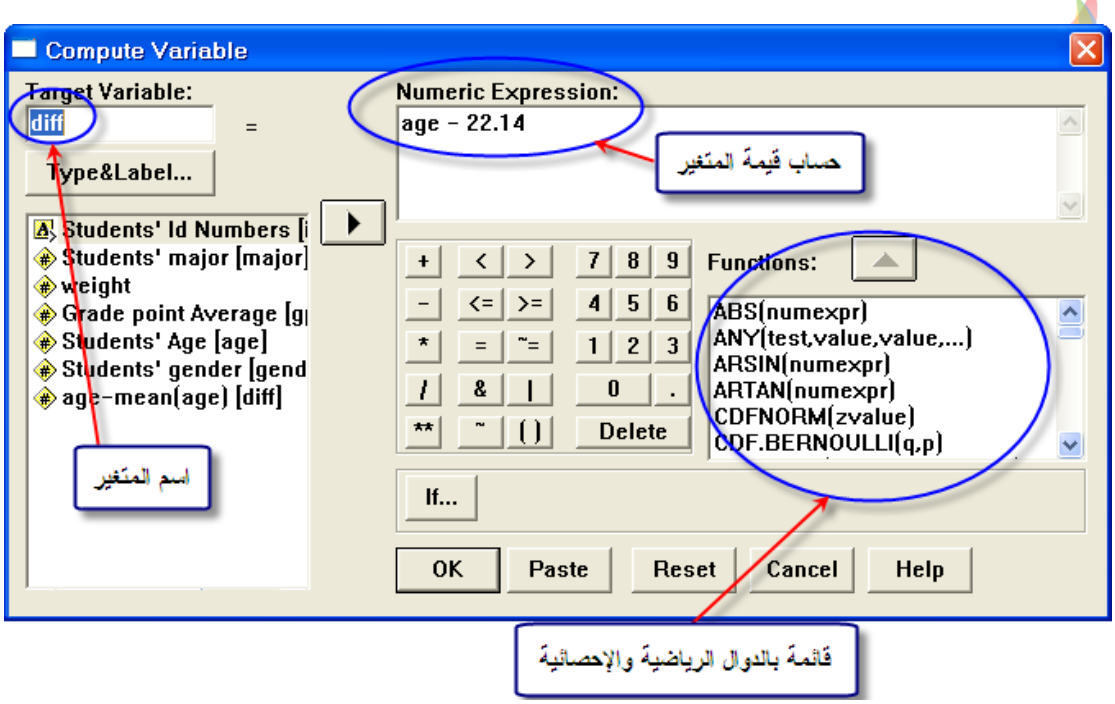
حساب قيم جديدة لمتغير

تتم عملية حساب قيم جديد لمتغير بناء على قيم متوفرة لمتغير أو متغيرات أخرى باستخدام الأمر Compute والذي يمكن الباحث من استخدام العديد من الدوال الرياضية والإحصائية لإجراء الحسابات اللازمة.

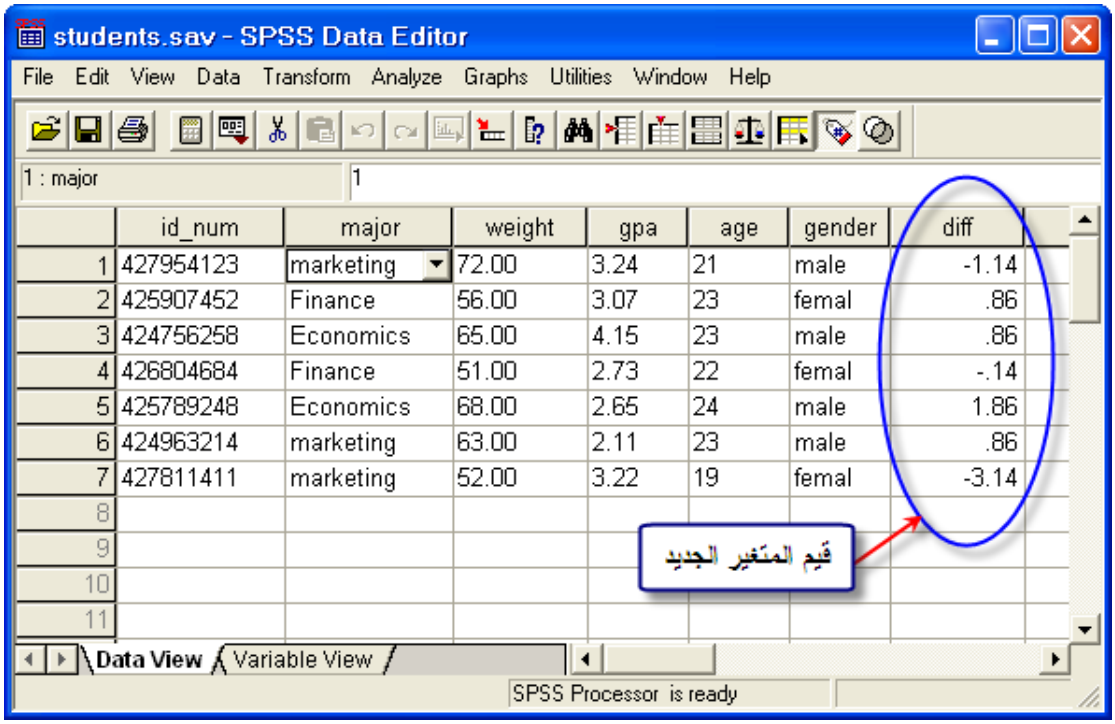
لنفرض أن لدى الباحث الرغبة في تكوين متغير جديد يحمل الإسم "diff" والذي يمثل الفروقات بين عمر الطالب ومتوسط الأعمار لجميع الطلاب. وللقيام بذلك نستخدم الأمر Compute من قائمة Transform أو بالنقر على أيقونة Compute بعد إضافتها إلى شريط الأدوات.



وعند ذلك يظهر لنا مربع الحوار التالي.

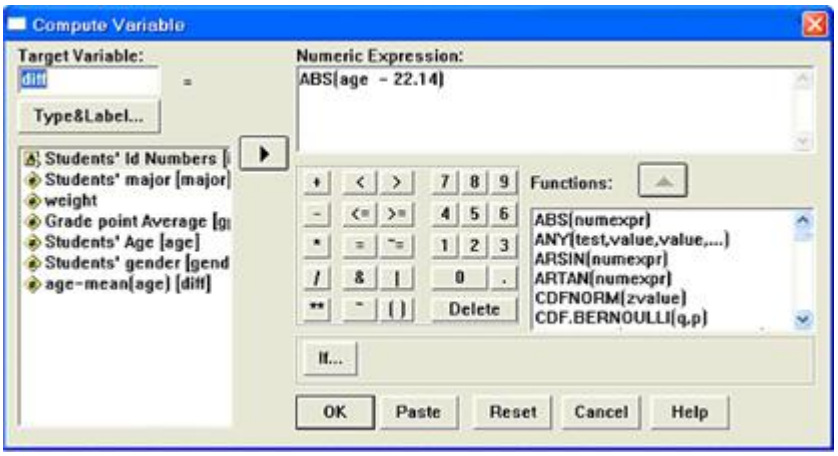


ولحساب القيم الجديدة للمتغير "diff" نقوم بطرح متوسط عمر الطلاب من عمر الطالب. ويوضح الشكل السابق قائمة الدوال الإحصائية المتوفر والتي يمكن استخدامها لحساب قيم المتغير الجديد. ويتم وضع اسم المتغير الجديد في المربع الذي يقع أسفل كلمة Target Variable، كذلك بتم تحديد بعض خصائص المتغير بالضغط على زر Type&Label، ثم يتم وضع العبارة المستخدمة لحساب قيم المتغير الجديد في مربع Numeric Expression. وبالضغط على زر OK يظهر لنا المتغير الجديد كما في الشكل التالي.

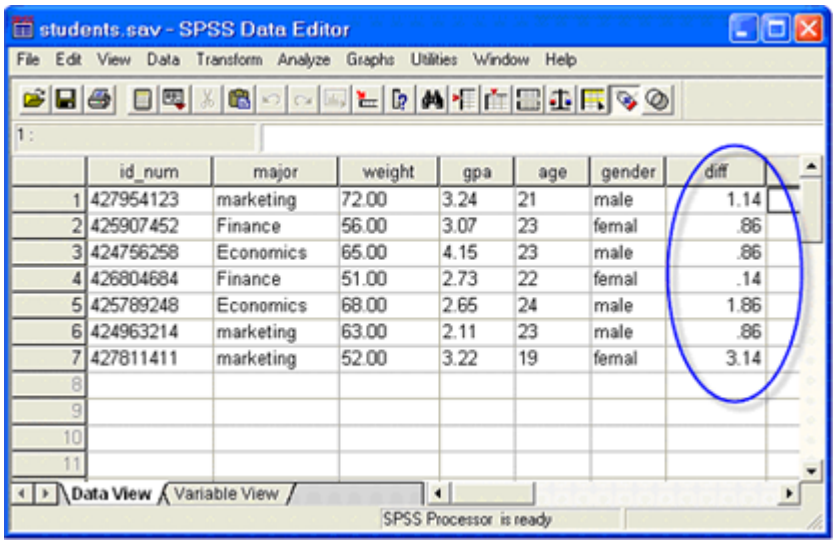


استخدام الدوال الرياضية والإحصائية

يوجد في قائمة الدوال مجموعة كبيرة من الدوال والتي تتجاوز 70 دالة رياضية وإحصائية وبعض الدوال الأخرى التي تساعد على حساب قيم متغيرات جديدة, وتعتبر هذه الدوال أداة لحساب قيمة المتغير الجديد بالاعتماد على القيم المقابلة للمتغير المتوفر حالياً. فعلى سبيل المثال فإنه استخدام دالة ABS(numexpr) لحساب القيمة المطلقة للفرق بين عمر الطالبة ومتوسط أعمار الطلاب، ويتم ذلك باختيار الدالة ثم الضغط على زر وذلك لرفع الدالة المختارة إلى مربع Numeric Expression ثم إضافة العبارة السابقة لحساب قيمة المتغير الجديد كما هو موضح في الشكل التالي.



وبعد الضغط على زر OK، نشاهد أن جميع قيم المتغير الجديد "diff" أصبحت موجبة.



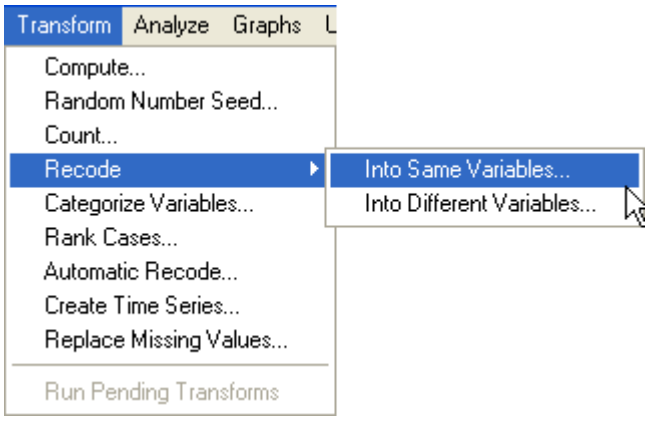
	id_num	major	weight	gpa	age	gender	diff
1	427954123	marketing	72.00	3.24	21	male	1.14
2	425907452	Finance	56.00	3.07	23	femal	.86
3	424756258	Economics	65.00	4.15	23	male	.86
4	426804684	Finance	51.00	2.73	22	femal	.14
5	425789248	Economics	68.00	2.65	24	male	1.86
6	424963214	marketing	63.00	2.11	23	male	.86
7	427811411	marketing	52.00	3.22	19	femal	3.14
8							
9							
10							
11							

إعادة ترميز وتصنيف المتغيرات

إعادة ترميز المتغيرات هي عملية تكوين متغير جديد مرتبط بمتغير موجود حالياً في ملف البيانات بحيث يتم تحويل متغيرات كمية أي قابلة للقياس Scale Variable إلى متغيرات وصفية Nominal Variable. أما تصنيف المتغيرات فهي عملية مشابهة لعملية إعادة ترميز المتغيرات إلا أن المتغير الجديد يعكس عدد مستويات التصنيف ويتم تصنيف المتغير على فترات متساوية الطول.

إعادة ترميز المتغيرات Recode

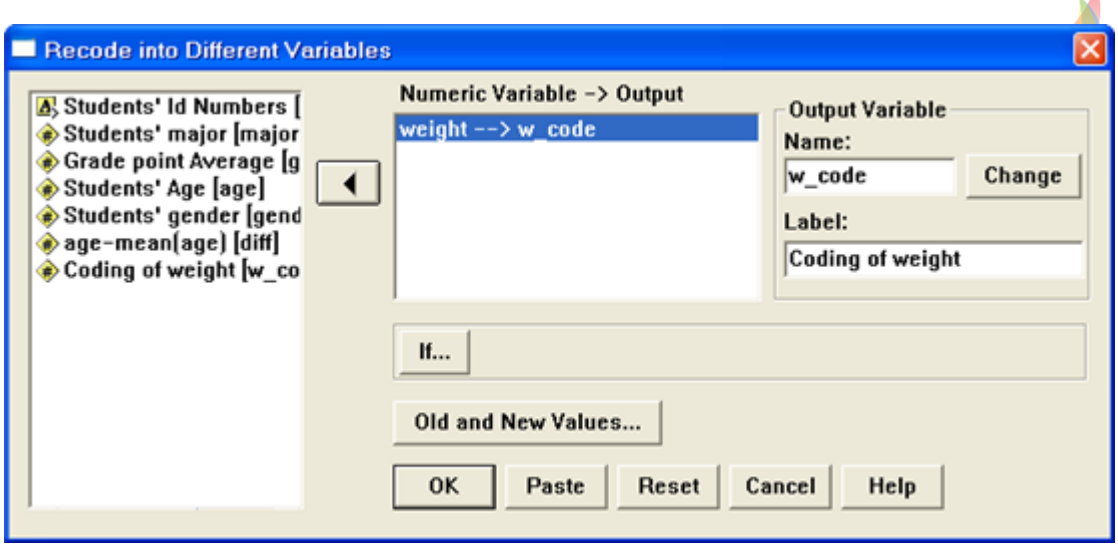
يستخدم الأمر Recode لتكوين متغير جديد بحيث تكون قيم المتغير عبارة عن مستويات للتصنيف أو عبارة عن قيم محدد ليتم تحويل المتغير الكمي إلى متغير وصفي. ويتم ذلك باختيار الأمر Recode من قائمة Transform كما في الشكل التالي.



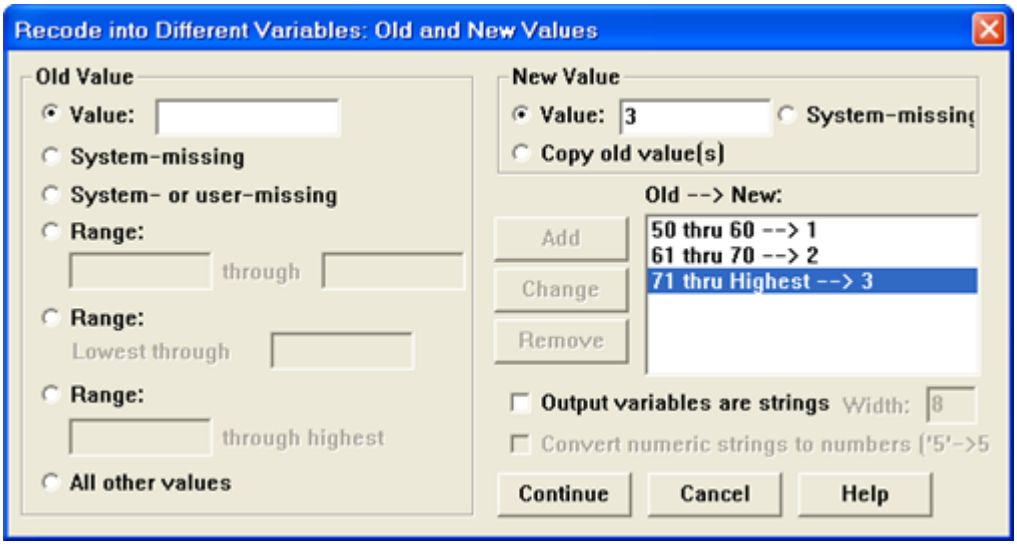
ويشمل أمر إعادة الترميز إلى خيارين، حيث يمكن بواسطة الخيار الأول وضع الترميز الجديد في نفس المتغير السابق، في حين يتم وضع قيم الترميز الجديدة باسم متغير جديد عند استخدام الخيار الثاني. لنفرض أن لدينا الرغبة في تكوين متغير تصنيفي لوزن الطلاب بحيث تكون قيمة المتغير الجديد بحسب الفئات التالية

قيم المتغير الجديد	قيم المتغير السابق
1	50−60
2	61−70
3	≥71

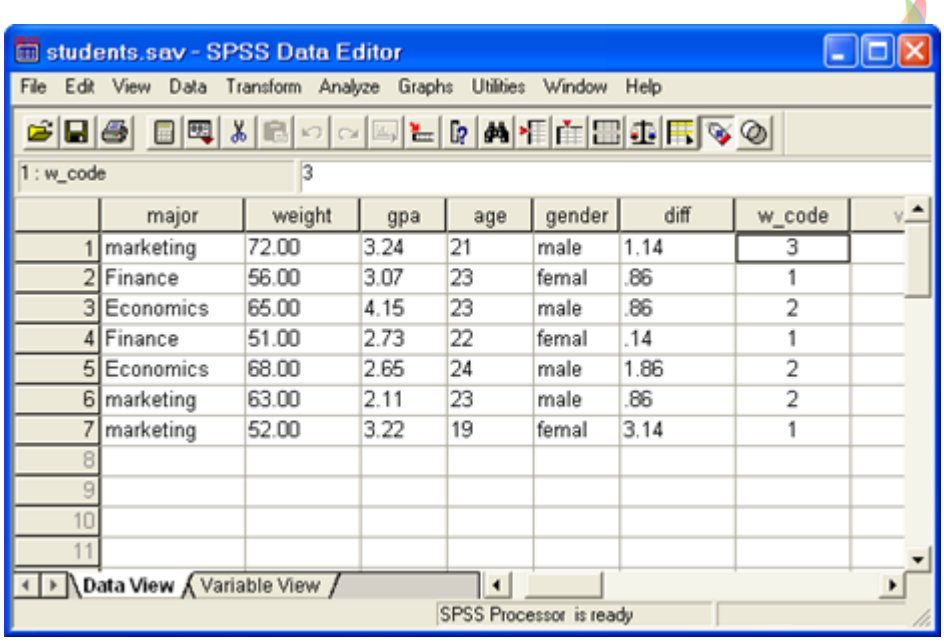
وسيكون اسم المتغير الجديد "w_code". وبالنقر على Recode من قائمة Transform وأختيار Into Different Variable يظهر مربع الحوار التالي.



ويتم من خلال مربع الحوار السابق تحديد المتغير المستهدف في عملية إعادة الترميز وتحديد اسم المتغير الجديد ووضع وصف للمتغير الجديد. ولتحديد قيم المتغير الجديد، يتم النقر على زر Old and New Values، وبذلك يظهر لنا مربع الحوار التالي.



ويتم من خلال مربع الحوار السابق تحديد قيم محدد أو فترات محددة للمتغير السابق وكذلك تحديد قيم الترميز الجديد للمتغير الجديد. ويمكن إضافة العدد المناسب من الفترات بأطوال مختلفة وذلك حسب رغبة الباحث، وعند الانتهاء من تحديد الفترة أو القيمة السابقة وتحديد قيمة الترميز الجديد يتم النقر على زر Add إلى قائمة Old→New وذلك لقبول الترميز الجديد. وعند الانتهاء من إضافة جميع الفترات، يتم النقر على زر Continue ثم OK ليتم بذلك تكوين المتغير الجديد ومشاهدة المتغير الجديد في ورقة عارض البيانات.



	major	weight	gpa	age	gender	diff	w_code
1	marketing	72.00	3.24	21	male	1.14	3
2	Finance	56.00	3.07	23	femal	.86	1
3	Economics	65.00	4.15	23	male	.86	2
4	Finance	51.00	2.73	22	femal	.14	1
5	Economics	68.00	2.65	24	male	1.86	2
6	marketing	63.00	2.11	23	male	.86	2
7	marketing	52.00	3.22	19	femal	3.14	1
8							
9							
10							
11							