

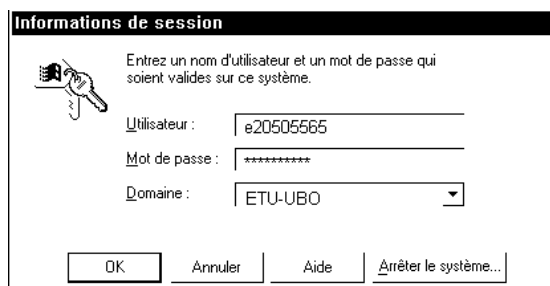
Licence de Psychologie - Semestre N° 5 - TD n° 1

Première partie : Windows XP - Présentation générale

1 Une session de travail sous Windows XP

1.1 Ouverture de session

Une période d'utilisation d'un poste de travail par un utilisateur donné est appelée session. Vous affichez le dialogue d'ouverture de session en appuyant simultanément sur les trois touches Ctrl+Alt+Suppr. Complétez le dialogue en ouvrant la session à l'aide de vos identifiants ENT :



N.B. Pour des raisons de confidentialité, le mot de passe ne s'affiche pas "en clair".

1.2 Si le mot de passe n'est pas reconnu ...

Ouvrez la session en utilisant le compte :

Utilisateur : etudiant

Mot de passe : ubo

Domaine : leta20_p__ (Ce poste).

Vous devez ensuite modifier ou confirmer votre mot de passe pour provoquer une nouvelle synchronisation des différentes bases de données dans lesquelles celui-ci est enregistré. Pour cela :

- Ouvrez votre ENT dans un navigateur Web (Internet Explorer ou Firefox).
- Activez l'onglet "Mon Bureau" puis l'item "Mon compte" et cliquer sur le lien "changer votre mot de passe".
- Indiquez ensuite (deux fois) votre nouveau mot de passe (qui, sans inconvénient, peut être identique à l'ancien).
- Validez en cliquant sur le bouton "Terminer".
- Quittez le navigateur, fermez la session "etudiant" et tentez de vous reconnecter.

Si vos login/mot de passe sont refusés à l'ouverture de l'ENT, vous devez au préalable refaire la procédure de validation de votre compte ENT à l'aide de votre passeport informatique.

1.3 Fermeture de session

Pour terminer la session en cours, vous pouvez, soit choisir le menu Démarrer-Arrêter..., soit appuyer de nouveau sur les touches Ctrl+Alt+Suppr. vous obtenez un dialogue proche de celui-ci (les dialogues diffèrent selon la méthode utilisée) :



En cours de journée, choisissez le premier item. En fin de journée, les appareils peuvent être arrêtés.

Remarque. Pour arrêter l'appareil, un raccourci pratique consiste à appuyer sur l'interrupteur de mise sous tension.

1.4 Accès par Internet

1.4.1 Le serveur Web du Département d'Informatique

Chargez le logiciel Internet Explorer (icône sur le bureau), ou Mozilla Firefox.

Dans la zone d'édition Adresse, saisissez : <http://geai.univ-brest.fr/>

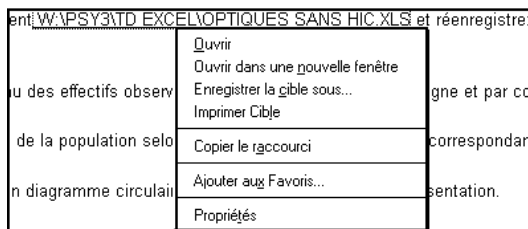
A partir de la page d'accueil, cliquez sur "Les membres du Département" puis sur "---> Documents pédagogiques sous la responsabilité de F.G. Carpentier" ou directement sur "Documents mis en ligne par F.-G. Carpentier".

Cliquez sur ce lien, puis sur le lien vers la fiche de TD que vous voulez consulter.

Vous pouvez aussi télécharger les fichiers cités dans le texte. Pour cela :

Cliquez, avec le bouton droit de la souris, sur le lien vers le fichier convoité

Dans le menu local qui apparaît, choisissez : Enregistrer la cible sous...



Ce serveur abrite également d'autres documents :

- les programmes d'enseignement

(adresse <http://geai.univ-brest.fr/enseignements-lmd/>)

- des "tables statistiques en ligne" dont voici un échantillon :

- Des pages interactives permettant de réaliser des traitements statistiques tels que le test du khi-2 par exemple :

Voici les résultats de votre khi²:

- Le khi² observé est de : 5.95671
- Le degré de liberté est de : 1
- Le niveau de significativité est de : 0.01466138

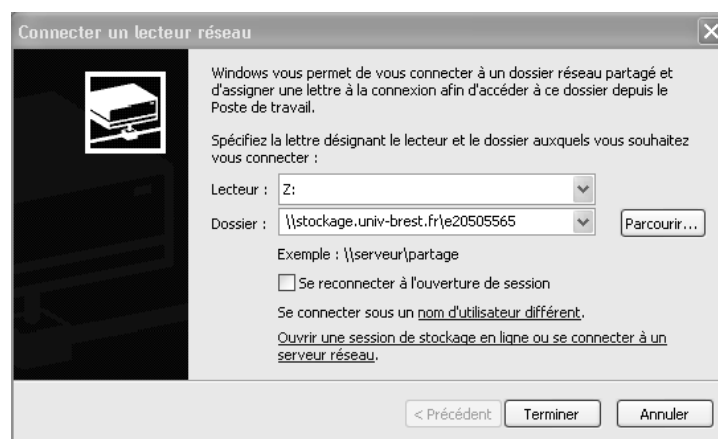
entrer une valeur de seuil:

1.5 L'espace de stockage sur l'ENT

Vous disposez d'un espace de stockage sur votre ENT (espace numérique de travail). Vous pouvez y accéder par l'intermédiaire d'une interface Web, depuis les salles informatiques, ou depuis votre domicile. A partir de nos salles de TD, vous pouvez également "monter" cet espace de stockage comme disque réseau.

Les manipulations sont analogues à celles décrites dans le paragraphe "Utiliser son compte", permettant de monter le volume de votre binôme étudiant.

- Utilisez le menu local Connecter un lecteur réseau, accessible à l'aide d'un clic droit sur l'icône du poste de travail.
- Complétez la fenêtre de connexion en indiquant votre identifiant sur l'ENT



- Cliquez sur l'item "Se connecter sous un nom d'utilisateur différent" et complétez le dialogue en indiquant votre identifiant et votre mot de passe de l'ENT :



Cliquez ensuite sur "OK" puis sur "Terminer".

Pour monter l'espace de stockage d'un 2^{ème} étudiant, on rencontre le même problème que précédemment. On contourne ce problème en utilisant une autre désignation pour le serveur : "stockage" par exemple, ou son adresse IP : 195.83.247.68.

Remarque Importante : Tel qu'il est configuré en cette rentrée 2011, le serveur ne permet pas d'enregistrer directement sur le volume réseau à partir d'un logiciel tel que Word, Excel ou Statistica. Cette limitation ne concerne pas tous les documents, mais s'applique notamment aux classeurs Statistica. En revanche, on peut, en principe :

- Enregistrer son travail sur un disque local.
- Quitter le logiciel.

- Copier le fichier correspondant depuis le disque local sur le volume réseau correspondant à son espace de stockage.

De même, pour reprendre un travail à partir d'un fichier existant sur son espace de stockage :

- Copier le fichier depuis le volume réseau sur un disque local.
- Travailler sur le fichier enregistré localement.
- Mettre à jour le fichier du serveur de stockage.

1.6 Quelques manipulations courantes sous Windows

1.6.1 Le répertoire "Mes Documents" - Où enregistrer ses documents ?

L'interface fournie par les différentes versions de l'explorateur donne un statut particulier au "Bureau" et à l'emplacement "Mes documents". Qu'en est-il exactement ?

Sur un poste de travail isolé, ces deux emplacements ne sont que des répertoires du disque local : le bureau est le répertoire C:\Documents and Settings\ et "Mes Documents" est C:\Documents and Settings\.

Sur des appareils faisant partie d'un domaine, tels que ceux de nos salles, ces deux répertoires sont synchronisés avec un "profil" enregistré sur le serveur. Ainsi :

- A l'ouverture de la session, la totalité des fichiers enregistrés dans ce profil doit être recopiée sur le disque dur local
- A la fermeture de la session, tous les fichiers de ces répertoires, modifiés durant la session, doivent être copiés sur le serveur.

Une conséquence de ce comportement est que l'ouverture/fermeture de session peut devenir très lente, si des fichiers nombreux et volumineux doivent être ainsi copiés. *Par ailleurs, il est important de refermer correctement sa session en fin de travail, faute de quoi, les fichiers sur lesquels vous avez travaillé ne seraient pas copiés sur votre profil du serveur.*

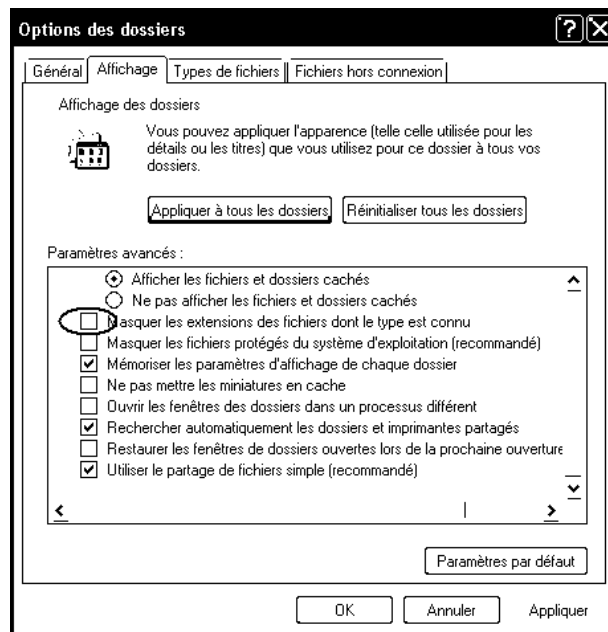
Il est également important de noter que le compte "etudiant" est un compte local verrouillé. Si vous avez ouvert la session à l'aide de ce compte, tout ce que vous avez enregistré dans le répertoire "Mes Documents" et sur le bureau sera enregistré sur le disque de l'appareil et non pas sur un volume réseau.

1.6.1.1 Noms de fichiers

Vous pouvez, a priori, attribuer le nom de votre choix à vos fichiers. Notez qu'un nom de fichier se termine par une extension de 3 lettres, en général caractéristique du format de fichier. Par exemple :

Mon premier document.doc..... Document Word
 Mon 2è document.xls Document Excel
 Mon 3è document.cwk Document ClarisWorks.

Cette extension est ajoutée automatiquement par le logiciel. Elle peut être affichée ou masquée selon les options choisies dans l'Explorateur ou le Poste de Travail. Le menu Outils-Options des dossiers, onglet Affichage, de l'explorateur permet de forcer l'affichage des extensions, ce qui est généralement utile.



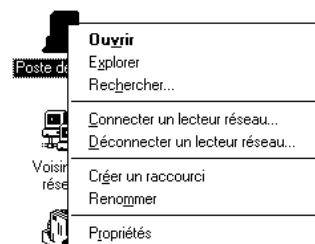
1.6.1.2 Ouvrir un document par double clic

Lorsqu'on fait un double-clic sur un fichier, le système d'exploitation analyse l'extension et charge le logiciel correspondant à cette extension. Toutefois, plusieurs logiciels peuvent être utilisés pour manipuler certains documents. Ainsi, par exemple, un document "texte seul" (extension .txt) pourra être ouvert aussi bien par Word que par Wordpad, ou Notepad.

Pour ouvrir un document en utilisant un autre logiciel que celui prévu par défaut, utilisez le bouton droit de la souris et le menu contextuel Ouvrir avec... au lieu du double clic, ou chargez d'abord le logiciel voulu, puis utilisez son menu Fichier - Ouvrir.

1.6.1.3 La souris et ses boutons

Le bouton gauche s'utilise de façon classique. Le bouton droit donne généralement accès à un *menu contextuel*, propre à l'objet sélectionné. Par exemple, si l'on utilise le bouton droit sur l'icône Poste de travail, on obtient :



1.6.1.4 Copier des fichiers

Pour copier des fichiers, plusieurs méthodes sont possibles. En voici trois :

- Pour copier un (des) fichier(s) d'un endroit quelconque d'un disque dur vers, par exemple, une clé USB montée comme disque F:, on peut utiliser l'Explorateur : il suffit de sélectionner le fichier, puis d'utiliser le menu Fichier-Envoyer vers...- Disque amovible (F:).

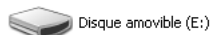
Le Poste de travail offre une fonctionnalité analogue.

- Pour copier un (des) fichier(s) d'une source vers une cible, on peut aussi, à l'aide de l'explorateur :
 - Afficher la source et sélectionner les fichiers ;
 - Utiliser le menu Edition-Copier ;
 - Afficher la cible ;
 - Utiliser le menu Edition-Coller.

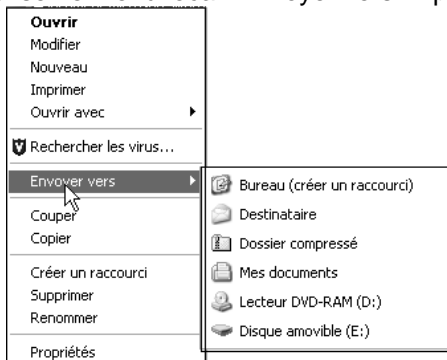
- Enfin il est également possible d'ouvrir deux exemplaires de l'explorateur (ou un Explorateur et un Poste de travail), d'afficher la source et la cible dans deux fenêtres différentes et de faire glisser les fichiers désirés de la première vers la seconde.

Les clés USB

Sur la plupart des postes de travail, une rallonge USB sur la table de travail permet d'insérer la clé sans quitter sa place. La clé est en principe reconnue et montée automatiquement lorsque vous l'insérez dans un port USB de l'appareil. Elle apparaît alors comme "disque amovible" ; la lettre de lecteur dépend de la configuration de l'appareil :



On peut, comme précédemment, utiliser le menu local "Envoyer vers..." pour y copier des fichiers :



Avant de retirer la clé du lecteur, il faut arrêter le périphérique à l'aide de l'icône présente dans la barre de tâches :



Windows affiche alors un dialogue du type représenté ci-dessous :



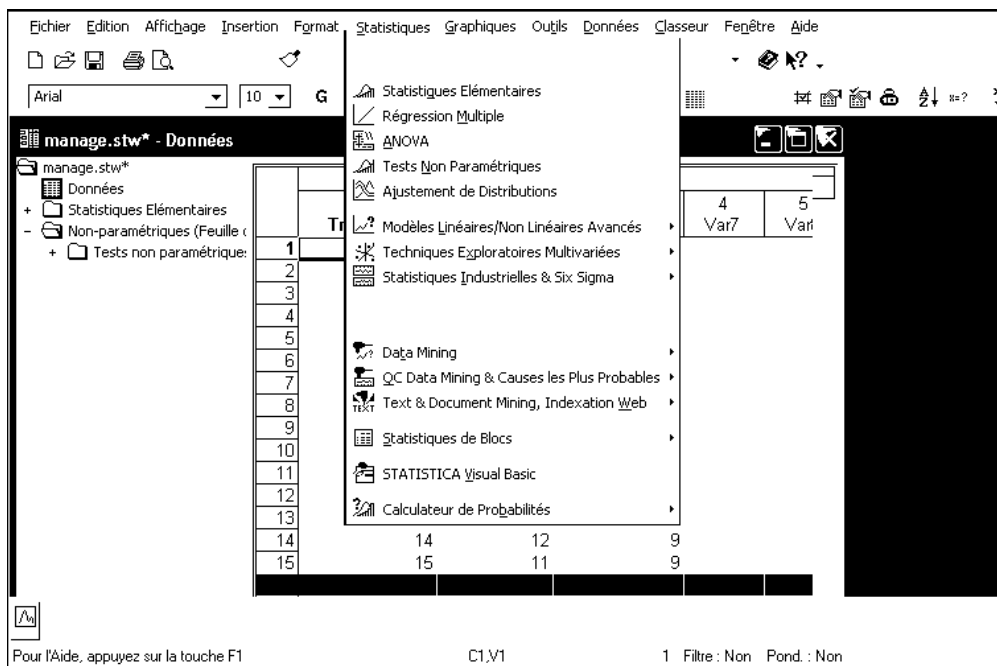
Remarque : Il peut arriver qu'une clé ne soit pas bien reconnue par Windows. Dans ce cas, on peut essayer de refermer la session, et de la ré-ouvrir, après avoir introduit la clé, ou même de redémarrer l'appareil, avec la clé présente dans un port USB de l'appareil. D'autres comportements étonnants peuvent parfois se produire : par exemple, votre clé semble être correctement reconnue, mais l'explorateur affiche comme contenu, celui de la clé du précédent utilisateur du poste - contenu d'ailleurs accessible car il a été copié dans le cache présent sur le disque dur ! Si vous constatez de tels comportements, redémarrez le poste après avoir inséré votre clé.

Deuxième partie : Introduction à STATISTICA

2 Statistica : l'interface utilisateur

2.1 L'écran de travail

Dans le cadre de l'EC PSY54AA du premier semestre, nous utiliserons essentiellement un logiciel de traitements statistiques : Statistica. Mais nous verrons également comment transférer les résultats obtenus sur le logiciel dans un document traitement de textes ou tableur. Nous utilisons actuellement la version 7.1 du logiciel.



2.2 Organiser son espace de travail sous Statistica:

Statistica est un logiciel assez stable et fiable, à condition de respecter quelques règles élémentaires. La première d'entre elles est :

Pour travailler avec Statistica, ne pas utiliser un compte tel que le compte "etudiant" dont le profil est verrouillé.

Ouvrez une session avec votre login sous XP, puis chargez le logiciel Statistica. La configuration par défaut du logiciel n'est pas vraiment satisfaisante. Nous allons donc commencer par régler la configuration à nos besoins.

2.2.1 Le menu Outils - Options

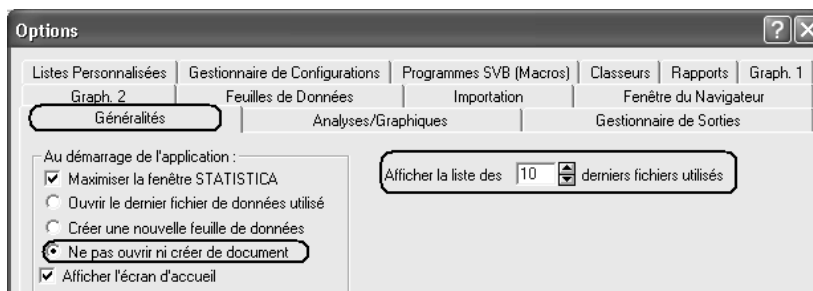
Le menu Outils - Options contient de nombreuses possibilités de paramétrage de Statistica. Heureusement, seules quelques-unes d'entre elles méritent d'être retouchées.

Ouvrez la fenêtre de dialogue accessible par le menu Outils-Options et explorez les différents onglets qui y sont rassemblés.

N.B. En principe, les options ainsi choisies sont enregistrées dans le profil de l'utilisateur lorsque l'on quitte le logiciel. Mais ce n'est pas toujours le cas lorsqu'on ouvre une session dans le domaine ETU-UBO. Par ailleurs, il n'y a aucun enregistrement si le compte est verrouillé ou si Statistica se plante en cours de travail.

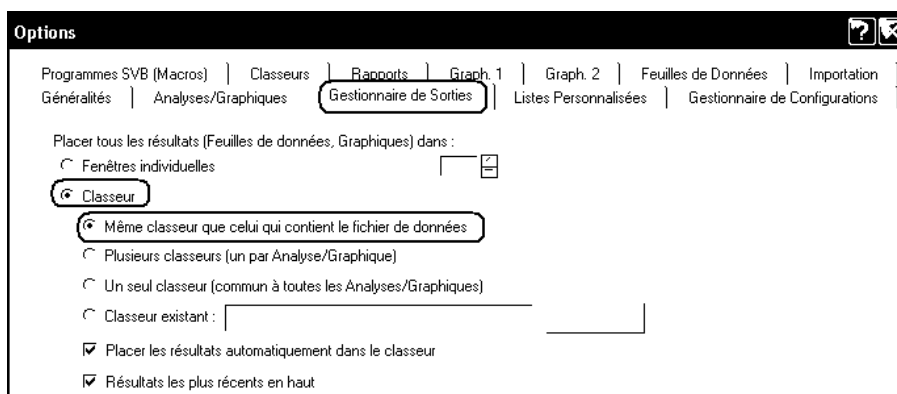
2.2.1.1 Spécifier le répertoire d'enregistrement par défaut

Dans l'onglet **Généralités**, on pourra limiter le nombre d'items figurant dans le menu Fichier (10 au lieu de 16 par exemple) et éviter qu'au démarrage, Statistica crée une nouvelle feuille de données ou cherche à ouvrir le dernier fichier utilisé :

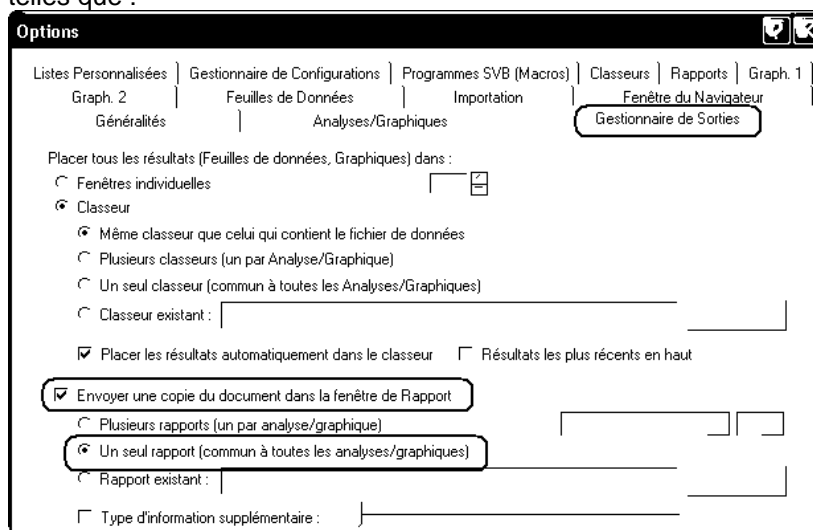


2.2.1.2 Gérer les sorties

La manière la plus commode de gérer nos documents avec Statistica consiste à rassembler dans un même classeur la ou les feuilles de données et les résultats de traitements concernant ces données. Ce comportement sera obtenu à l'aide du réglage suivant, sous l'onglet "Gestionnaire de sorties" :



Il peut également être commode de demander à Statistica de placer une copie des résultats dans un rapport, en utilisant des options telles que :



En effet un rapport peut être enregistré au format .rtf pour être relu sur une autre machine par un logiciel de traitement de textes, même si Statistica n'est pas installé sur l'appareil. *Cependant, cette pratique présente plus d'inconvénients que d'avantages.* En effet :

- Les rapports produisent rapidement des fichiers très volumineux. Un rapport, ou un classeur contenant un ou des rapports devra être compressé (zippé) avant d'être envoyé par mail. Et par ailleurs, un rapport trop volumineux semble provoquer des plantages du logiciel dans certains cas.

- Si plusieurs séances de travail sont nécessaires pour réaliser le traitement, un nouveau rapport sera créé à chaque séance, ce qui est assez peu pratique.
- Si le rapport produit au cours d'une séance de travail est inséré dans le classeur (par exemple à l'aide du menu local Insérer - Toutes les fenêtres du classeur), Statistica se plante de façon systématique, lorsque l'on quitte le logiciel.

Ne pas insérer dans le classeur courant le rapport produit au cours d'une séance de travail à l'aide du menu local "Insérer - Toutes les fenêtres", faute de quoi Statistica se plante lorsque l'on quitte le logiciel.

Il semble que l'on évite le plantage en utilisant le menu : Insérer - Document Statistica - Créer à partir d'une fenêtre.

En revanche, on pourra utiliser un rapport pour y taper de courts commentaires textuels, l'interprétation du résultat d'un traitement par exemple.

2.3 Les objets manipulés par Statistica

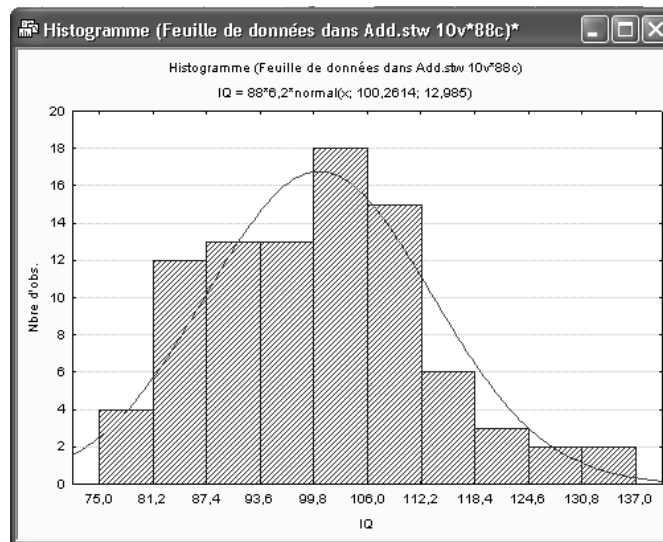
La **feuille de données** est organisée en variables et observations. Les colonnes sont les variables. Chaque ligne représente en règle générale, un individu statistique, appelé *observation*.

	1 Trimestre	2 H_Incident	3 D_Incident	4 Var7
1	1	11	8	
2	2	11	13	
3	3	14	12	
4	4	21	17	
5	5	12	14	
6	6	10	9	
7	7	15	10	
8	8	15	12	
9	9	17	13	
10	10	9	10	
11	11	12	8	
12	12	12	13	
13	13	15	12	
14	14	12	9	
15	15	11	9	

Les feuilles de données peuvent être enregistrées comme fichiers autonomes (fichiers *.sta). Elles contiennent les données d'entrée sur lesquelles s'effectuent les traitements statistiques. Les résultats de ces traitements s'affichent dans un document de sortie. Plusieurs possibilités sont offertes.

Fenêtre de rapport : C'est la méthode traditionnelle pour gérer les résultats produits par le logiciel. Un rapport se comporte plus ou moins comme un document produit par un traitement de textes. On peut insérer des commentaires, modifier la mise en forme, spécifier la mise en page, la numérotation des pages, l'en-tête et le pied de page en vue de l'impression. Les rapports peuvent être enregistrés comme fichiers autonomes (fichiers *.str).

Les résultats de sortie peuvent également être dirigés vers des fenêtres individuelles. Les résultats numériques sont alors affichés dans des fenêtres de données. Les graphiques sont affichés dans des **fenêtres de graphiques** (fichiers *.stg).



Les classeurs : les données d'entrée et de sortie peuvent également être stockées comme onglets dans un classeur. Un classeur est un "conteneur" accueillant d'autres objets, organisés sous forme hiérarchique. Ils correspondent aux fichiers de type *.stw.

Variable	N Actifs	Moyenne
H Incident	15	13,13333
D Incident	15	11,26667

Traitements statistiques

Statistica est organisé en modules, accessibles à partir du menu Statistiques. Chaque module contient un groupe de procédures statistiques reliées entre elles. Par exemple, le module "Statistiques élémentaires" se présente comme suit :



2.4 Copier - coller entre classeurs, entre un classeur et un objet Statistica

Pour déplacer un objet d'un classeur à un autre, il suffit de déplacer son icône depuis le volet gauche du premier classeur dans le volet gauche du second. On peut également utiliser les menus locaux Copier et Coller obtenus à l'aide d'un clic droit dans le volet gauche de chaque classeur.

Le menu local "Insérer" du volet gauche d'un classeur permet également d'insérer dans ce classeur un document contenu dans une fenêtre indépendante. Il suffit de choisir les options : Document Statistica - Créer à partir d'une fenêtre.

L'opération faite par Statistica est soit une copie (l'original de l'objet est conservé) soit un déplacement (l'original de l'objet n'est pas conservé) selon le paramétrage choisi dans le menu Outils - Options - Onglet Classeurs - Item "En cas d'ajout d'un document dans le classeur".

En cas d'ajout d'un document dans le classeur	
<input type="radio"/>	Déplacer le document initial dans le Classeur
<input checked="" type="radio"/>	Envoyer une copie du document initial dans le Classeur

2.5 Supprimer un objet d'un classeur

Il est également possible de supprimer un objet d'un classeur, à l'aide d'un clic droit et de l'item de menu Supprimer. Cela permet notamment de ne garder, pour un traitement donné, que le résultat le plus abouti. Attention cependant : lorsque l'on supprime un objet qui n'est pas une feuille de la hiérarchie, on supprime en même temps tous les objets qui en dépendent.

3 Travailler avec Statistica : saisie de données numériques

Utilisez le menu Fichier pour créer un nouveau classeur, comportant une feuille de données vierge, ou, de manière équivalente, une nouvelle feuille de données, placée dans un nouveau classeur. Si nécessaire, utilisez le menu Données - Observations - Ajouter pour disposer de 15 lignes dans la feuille de données.

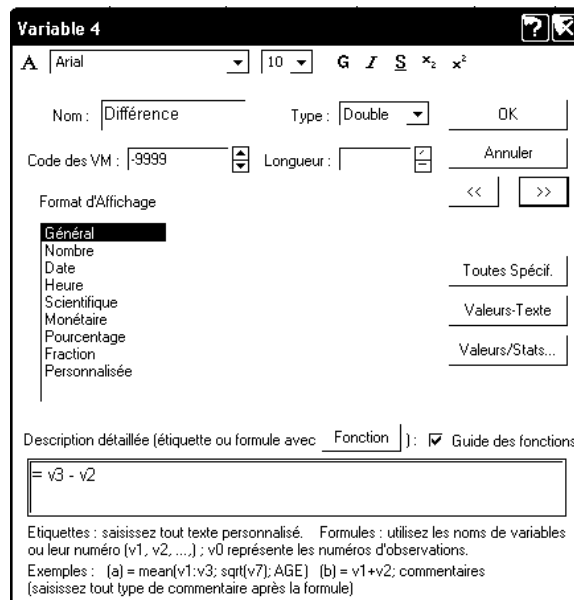
Faites un double-clic sur les têtes de colonnes des variables, ou utilisez le menu Données - Spécifications de la variable... pour attribuer aux variables Var1 à Var3 les noms "Trimestre", "H_Incident" et "D_Incident" Saisissez dans cette feuille les observations des variables "Trimestre", "H_Incident" et "D_Incident" ci-dessous.

Trimestre	H_Incident	D_Incident
1	11	8
2	11	13
3	14	12
4	21	17
5	12	14
6	10	9
7	15	10
8	15	12
9	17	13
10	9	10
11	12	8
12	12	13
13	15	12
14	12	9
15	11	9

3.1.1 Variables calculées : Calcul du protocole dérivé des différences individuelles

On veut calculer la différence $D_Incident - H_Incident$, pour chacun des 15 trimestres.

Faites un double-clic sur la tête de la colonne N°4. Complétez comme suit la fenêtre de dialogue qui s'affiche :



Attention, Statistica n'est pas très souple du point de vue de la syntaxe. Un espace devant le signe "=", par exemple, rend la formule incorrecte.

Le comportement de Statistica sur ces variables calculées est particulier. Ce n'est ni celui d'un tableur tel qu'Excel, ni celui d'un logiciel de statistiques tel que Minitab ou Statgraphics.

Modifiez une valeur observée dans l'une des colonnes H_Incident ou D_Incident. La valeur correspondante de la colonne Différence est-elle modifiée ? Conclusion : par défaut, il n'y a pas de recalcul automatique des formules.

Utilisez le menu Données - Recalculer les formules ... et demandez le recalcul de toutes les formules de la feuille.

De même, modifiez manuellement une valeur de la colonne Différence. Demandez de nouveau le recalcul des formules.

3.1.2 Remarque importante concernant les variables calculées

Comment les variables calculées se comportent-elles lorsque l'on demande un recalcul après avoir modifié la feuille ?

Il faut noter que les formules de calcul ne sont pas remises à jour lorsqu'on modifie la structure de la feuille, par exemple en insérant une variable (colonne). Par exemple, si on insère une variable entre les colonnes 2 et 3, un recalcul de la variable Différence, toujours calculée à partir de la formule $=v3-v2$, prendra en compte la "nouvelle" variable v3 au lieu de la variable D_Incident, conduisant ainsi à des résultats incorrects.

On peut également écrire les formules de calcul en utilisant les noms des variables au lieu de leurs noms génériques v1, v2, etc. Ainsi, pour la variable calculée Différence, la formule de calcul pourrait s'écrire :

`= 'D_Incident' - 'H_Incident'`

Notez bien l'usage de cotes autour des noms de variables, obligatoire si les noms de variables utilisent des caractères "ambigus" tels que : espace, tiret (ambiguïté avec le signe moins), etc.

Cette solution permet de s'affranchir du problème précédent. Il vaut donc mieux utiliser cette syntaxe, même si la saisie des formules est plus fastidieuse. Mais on retrouve un problème analogue si l'une des variables intervenant dans les formules est renommée.

3.1.3 Trier les données selon les valeurs d'une ou plusieurs variables



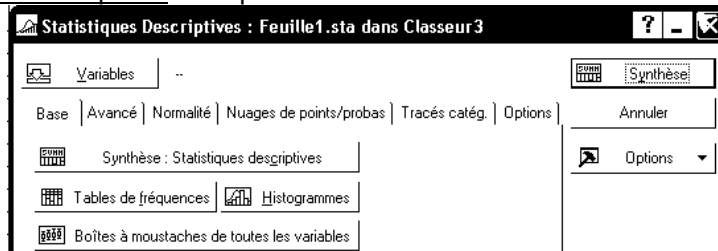
Utilisez le bouton trier . La boîte de dialogue qui s'affiche permet d'effectuer un tri hiérarchisé. Par exemple, trie les données selon les valeurs de la colonne Différence, puis refaites le tri selon les valeurs de la colonne Trimestre.

3.2 Calcul des paramètres de statistiques descriptives

On veut calculer les paramètres relatifs aux variables H_Incident, D_Incident et Différence :

Utilisez le menu Statistiques - Statistiques élémentaires...

Sélectionnez Statistiques Descriptives et cliquez sur le bouton OK.



Sélectionnez les variables H_Incident, D_Incident et Différence.

Sous l'onglet Avancé, sélectionnez les paramètres dont vous souhaitez le calcul, puis cliquez sur le bouton Synthèse. On obtient par exemple :

Variable	Statistiques Descriptives (Feuille de données1)					
	N Actifs	Moyenne	Minimum	Maximum	Variance	Ecart-type
H_Incident	15	13,13	9,00	21,00	9,55	3,09
D_Incident	15	11,27	8,00	17,00	6,50	2,55
Différence	15	-1,87	-5,00	2,00	5,41	2,33

Les résultats s'affichent dans un document de type "feuille de données", contenu dans un classeur.

Remarquez que l'analyse reste active, ce qui permet de demander d'autres traitements accessibles dans le même module sans redéfinir les paramètres.

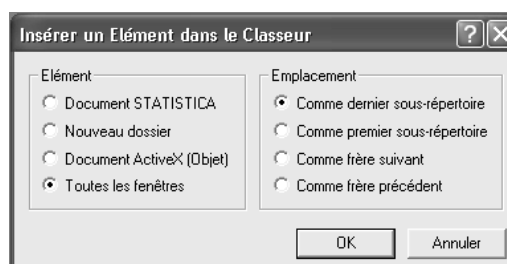
3.3 Enregistrement du document

Si vous avez correctement réglé les options conseillées au début de la séance de travail, votre fichier de données et la feuille de résultats se trouvent rassemblés dans le même classeur. Vous pouvez utiliser le menu Fichier - Enregistrer sous... pour enregistrer ce classeur sous le nom Manage. L'extension caractéristique des classeurs Statistica, .stw sera automatiquement ajoutée par le logiciel.

Une feuille de données affichée dans une fenêtre indépendante peut être enregistrée comme fichier indépendant. Par exemple, activez la feuille contenant les données et enregistrez-la sous le nom : Manage. L'extension .sta sera automatiquement ajoutée par le logiciel.

Mais, avec Statistica, il est souvent préférable d'enregistrer les données et les résultats dans un même classeur. Pour insérer la feuille de données comme élément du classeur :

- Activez la fenêtre du classeur
- Faites un clic-droit dans la partie gauche de cette fenêtre
- Sélectionnez l'item "Insérer", puis le choix "Toutes les fenêtres".



Lorsqu'une feuille est insérée dans un classeur, elle doit être rendue active pour que l'on puisse exécuter des traitements sur les données qu'elle contient. Pour rendre active une feuille, ou changer de feuille active :

- cliquez sur l'icône de la feuille et en utilisez le menu : Classeur - Feuille de données active;
- ou cliquez avec le bouton droit sur l'icône de la feuille et en utilisez l'item "Feuille de données active" du menu local.

Dans un classeur, une feuille active est repérée par un liseré rouge autour de son icône :

Feuille active :  Manage.sta Feuille inactive :  Manage.sta

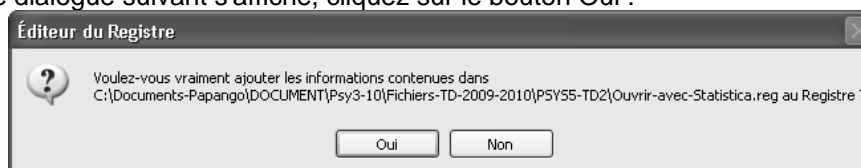
Remarque importante :

L'extension .stw, caractéristique des classeurs Statistica, est aussi utilisée pour certains documents OpenOffice. En conséquence, lorsqu'on fait un double-clic sur un fichier d'extension .stw, c'est généralement OpenOffice qui se charge, et non pas Statistica. C'est notamment le comportement que l'on observe lorsque la session est ouverte sur le compte "etudiant". Pour ouvrir le fichier à l'aide de Statistica, il faut donc charger le logiciel Statistica puis :

- soit utiliser le menu Fichier - Ouvrir ;
- soit faire glisser l'icône du fichier dans la fenêtre de Statistica.

On trouve en de multiples emplacements sur le serveur des salles de TD un fichier nommé "Ouvrir-avec-Statistica.reg" qui permet de redonner la priorité à Statistica pour l'ouverture des fichiers ayant l'extension .stw. Pour l'utiliser :

- Avant de charger Statistica ou l'un de ses fichiers, faites un double-clic sur l'icône "Ouvrir-avec-Statistica.reg".
- Lorsque le dialogue suivant s'affiche, cliquez sur le bouton Oui :



Il semble malheureusement que la modification ne soit pas mémorisée à la fermeture de la session. La manipulation doit donc être refaite lors de chaque session de travail.

4 Travailler avec Statistica : saisie de données catégorielles

A partir de la version 7, Statistica permet d'attribuer à une variable le type "texte". Cependant, lorsqu'on doit saisir une variable nominale comportant un nombre réduit de modalités, il est généralement plus pratique d'utiliser une variable numérique pour laquelle des étiquettes de texte auront été définies.

Exemple :

Les parfums agréables aident-ils un étudiant à apprendre plus efficacement ? Hirsch et Johnston pensent que la présence d'un parfum floral peut améliorer la capacité d'apprentissage d'un sujet dans certaines situations. Dans leur expérience, 21 sujets devaient effectuer à 6 reprises un parcours dans un labyrinthe tracé sur une feuille de papier. Pour chacune des 6 épreuves, les sujets portaient un masque. Pour 3 épreuves, le masque était parfumé. Pour les 3 autres épreuves, le masque ne l'était pas. Pour une moitié des sujets, le masque parfumé correspondait aux 3 premières épreuves, pour l'autre moitié, aux trois dernières. A chaque essai, les expérimentateurs notaient le temps nécessaire au parcours du labyrinthe par le sujet. On demandait également à chaque sujet s'il trouvait le parfum agréable ou non.

Les variables rassemblées dans le fichier de données sont les suivantes :

- ID : identifiant du sujet
- Genre : sexe du sujet
- Fumeur : le sujet est-il fumeur (Y = oui, N = non)
- Opinion : opinion du sujet sur le parfum (positive, négative, indifférente)
- Age : âge du sujet en années
- Ordre : 1 = le masque parfumé est porté aux épreuves 4 à 6 ; 2 = le masque parfumé est porté aux épreuves 1 à 3.
- Sans parfum : moyenne des temps mis par le sujet aux trois épreuves avec masque sans parfum

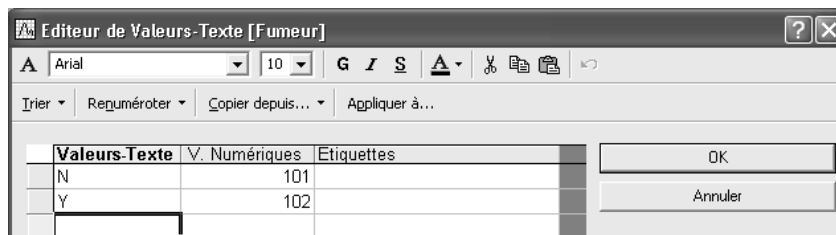
Avec parfum : moyenne des temps mis par le sujet aux trois épreuves avec masque parfumé.

Les données observées sont les suivantes :

	ID	Genre	Fumeur	Opinion	Age	Ordre	Sans parfum	Avec parfum
1	1	M	N	pos	23	1	30,6	38,0
2	2	F	Y	neg	43	2	48,4	51,6
3	3	M	N	pos	43	1	60,8	56,7
4	4	M	N	neg	32	2	36,1	40,5
5	5	M	N	neg	15	1	68,5	49,0
6	6	F	Y	pos	37	2	32,4	43,2
7	7	F	N	pos	26	1	43,7	44,6
8	8	F	N	pos	35	2	37,1	28,4
9	9	M	N	pos	26	1	31,2	28,2
10	10	F	N	indiff	31	2	51,2	68,5
11	11	F	Y	pos	35	1	65,4	51,1
12	12	F	Y	indiff	55	2	58,9	83,5
13	13	F	Y	pos	25	1	54,5	38,3
14	14	M	Y	indiff	39	2	43,5	51,4
15	15	M	N	indiff	25	1	37,9	29,3
16	16	M	N	pos	26	2	43,5	54,3
17	17	M	Y	neg	33	1	87,7	62,7
18	18	M	N	neg	62	2	53,5	58,0
19	19	F	Y	pos	54	1	64,3	52,4
20	20	F	N	neg	38	2	47,4	53,6
21	21	M	N	neg	65	1	53,7	47,0

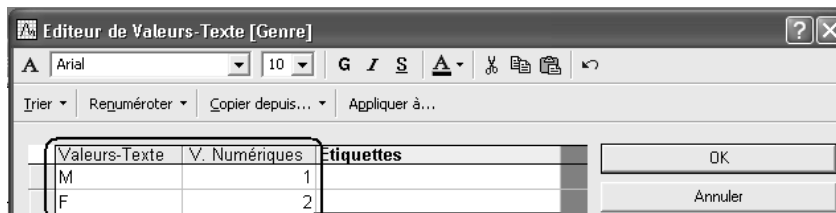
Ouvrez le classeur Statistica Parfums1.stw. Les données y ont été saisies, à l'exception des valeurs de la variable "Genre".

Quel est le codage choisi pour la variable "Fumeur" ? Affichez ses caractéristiques (utilisez le menu Données - Spécifications de la variable... ou faites un double-clic sur la tête de la colonne concernée). On observe qu'il s'agit d'une variable numérique. Cliquez alors sur le bouton "Valeurs-Texte". On obtient :



On voit que les textes "N" et "Y" qui s'affichent dans cette colonne sont des "étiquettes" correspondant aux valeurs numériques 101 et 102. Il s'agit là du choix fait par défaut par Statistica : la première valeur texte saisie dans une variable numérique se voit affecter la valeur 101, la seconde la valeur 102, etc.

On veut coder numériquement et saisir la variable "Genre" : le sexe masculin sera affiché "M" et codé 1 pendant que le sexe féminin sera affiché "F" et codé 2. Pour cela, affichez les spécifications de la variable "Genre" puis cliquez sur le bouton "Valeurs-Texte". Complétez la fenêtre de dialogue de la façon suivante :



Saisissez ensuite les valeurs de cette variable. Vous pouvez aussi bien saisir le nombre (1 ou 2) que la valeur texte correspondante (M ou F).

4.1 Recensement (tri à plat) sur une variable nominale

Pour faire un tri à plat des observations selon les valeurs de la variable "Genre" :

- Utilisez le menu Statistiques - Statistiques élémentaires
- Sélectionnez la méthode "Tables de fréquences" et activez l'onglet "Avancé".
- Veillez à ce que le bouton radio "Toutes les valeurs distinctes" et l'option "avec étiquettes" soient cochés.
- Cliquez sur le bouton Synthèse ou le bouton "Synthèse : tables de fréquences".

On doit obtenir :

Table de fréquences : Genre

	Effectif	Effectifs Cumulés	%age	%age Cumulé
M	11	11	52,38095	52,3810
F	10	21	47,61905	100,0000
VM	0	21	0,00000	100,0000

N.B. Dans ce tableau, "VM" signifie "valeurs manquantes". Nos données ne comportent ici aucune valeur manquante.

4.2 Calculer des paramètres descriptifs (moyennes, variances, etc) "par groupe"

La variable "Genre" permet de définir deux échantillons indépendants de sujets. On veut calculer, par exemple, la moyenne, la variance et l'écart type des variables dépendantes "Avec parfum" et "Sans parfum" dans chacun des deux groupes.

- Utilisez le menu Statistiques - Statistiques élémentaires
- Sélectionnez la méthode "Décompositions ; tableaux non factoriels".
- Sous l'onglet "Base", indiquez les variables "Sans parfum" et "Avec parfum" comme variables dépendantes et "Genre" comme variable de classement.



- Activez l'onglet "Statistiques descriptives" et cochez les options "Ecart type" et "Variance" puis cliquez sur le bouton "Synthèse".

Vous devriez obtenir :

Statistiques Descriptives par Groupes
N=21(pas de VM dans les vars dépendantes)

Genre	Sans parfum	Sans parfum	Sans parfum	Sans parfum	Avec parfum	Avec parfum	Avec parfum	Avec parfum
	Moyenne s	N	Ec- Types	Variance	Moyenn es	N	Ec- Types	Varianc e
M	49,72	11	17,49	305,82	46,82	11	11,52	132,61
F	50,34	10	10,88	118,43	51,52	10	15,43	237,96
TsGrpes	50,01	21	14,36	206,30	49,06	21	13,39	179,16

Remarque. Cette méthode permet également d'indiquer plusieurs variables catégorielles. Les paramètres descriptifs sont alors calculés dans tous les groupes définis par les combinaisons de modalités des différentes variables nominales. Testez-le par exemple, avec les variables "Genre" et "Fumeur".

5 Travailler avec Statistica : saisie des réponses à une question à réponses multiples

Statistica permet également un certain nombre de traitements sur des questions à réponses multiples. Deux types d'encodage sont possibles, appelés par Statistica *réponse multiple* et *dichotomie multiple* :

- Le sujet interrogé peut sélectionner de 0 à k items : on prévoit alors k colonnes, dans lesquelles on saisit les réponses choisies par le sujet (*réponse multiple*) ;

- Pour chaque question, on définit autant de colonnes que de modalités de réponse. On saisit alors 1 ou 0 selon que la modalité a été choisie par le sujet ou non (*dichotomie multiple ou tableau disjonctif*).

Exemple : Un questionnaire comporte deux questions. La question Q1 comporte 5 items : A, B, C, D et E et chaque sujet peut choisir au maximum 3 de ces items. La question Q2 comporte 3 items : U, V, W et chaque sujet peut en choisir 2 au maximum.

5.1 Saisie des données (réponse multiple)

Il n'est pas indispensable de respecter une cohérence dans les associations valeurs numériques / valeurs textes entre les différentes colonnes représentant une même question. C'est cependant préférable, ne serait-ce que pour éviter les erreurs de saisie.

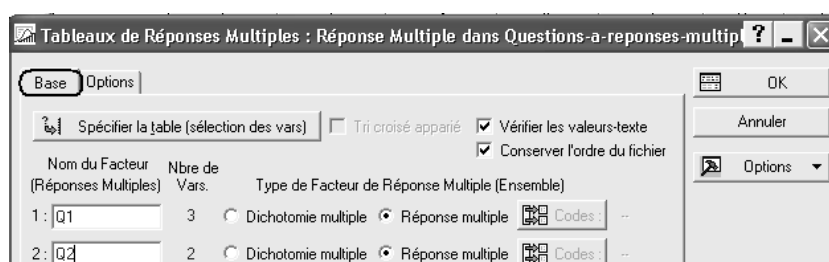
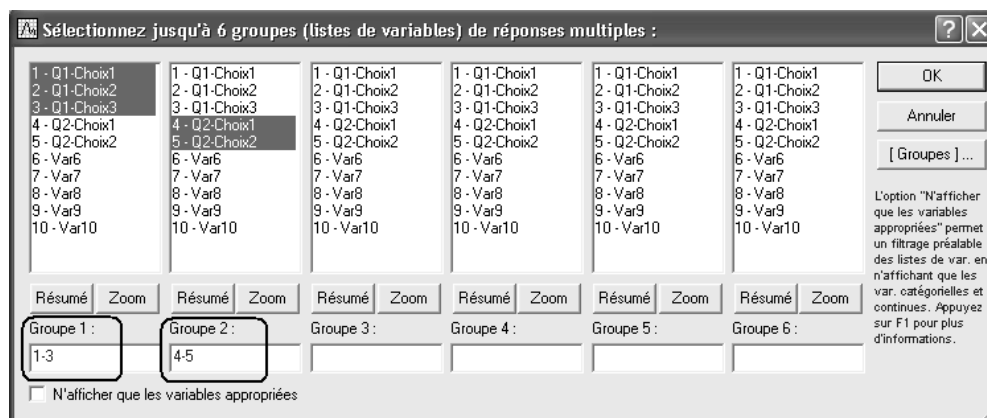
- Ouvrez un nouveau classeur Statistica.
- Nommez la première colonne "Q1-choix1" et affectez les codes numériques 1 à 5 aux items A à E.
- Pour les deux colonnes suivantes : nommez-les respectivement "Q1-choix2" et "Q1-choix3", affichez la spécification de chacune d'elles et affichez la fenêtre de dialogue "Valeurs-texte".
- Utilisez le bouton "Copier depuis..." pour définir pour ces variables les mêmes associations valeur numérique / valeur texte que pour la variable "Q1-choix1".
- Ensuite, procédez de même pour définir les variables "Q2-choix1" et "Q2-choix2".

Saisissez ensuite les données suivantes :

	Q1-Choix1	Q1-Choix2	Q1-Choix3	Q2-Choix1	Q2-Choix2
1	A	B	C	U	V
2	B	D	E	V	W
3	C	D	A	W	U
4	D	E		U	
5	E	B		U	
6	A	B		V	
7	A	C		V	
8	B			U	

Recensement et croisement des variables

- Utilisez le menu Statistiques - Statistiques élémentaires
- Sélectionnez la méthode "Tableaux de réponses multiples".
- Sous l'onglet "Base", spécifiez les variables comme suit :



- Cliquez sur le bouton OK. La fenêtre "Résultats des tableaux de réponses multiples" s'affiche alors.
- Sous l'onglet "Avancé", cliquez sur le bouton "Tables de fréquences". Vous obtenez par exemple :

N=8 Fréquences ; Réponses identiques ignorées

Catégorie	Effectifs	%age de Réponses	%age d'obs.
A	4	22,22	50,00
B	5	27,78	62,50
C	3	16,67	37,50
D	3	16,67	37,50
E	3	16,67	37,50
Total	18	100,00	225,00

Ainsi la réponse "A" par exemple, a été choisie par 50% des sujets (4 sujets sur 8), et représente 22,22% de l'ensemble des 18 réponses recueillies.

On peut également obtenir des tris croisés combinant les réponses aux deux questions. Ces tableaux croisés peuvent être basés soit sur le nombre de répondants, soit sur le nombre de réponses. Ainsi :

- Sous l'onglet "Options", dans la zone d'édition "Totaux/pourcentages basés sur" activez le bouton radio "Nombre de répondants".
- Sous l'onglet "Avancé" cliquez sur le bouton "Tableaux détaillés à double entrée". On obtient :

Tableau de Synthèse 2 x 2				
(Totaux = nb de répondants)				
	Q2 U	Q2 V	Q2 W	Totaux Ligne
A	2	3	1	4
B	3	3	1	5
C	2	2	1	3
D	2	1	2	3
E	2	1	1	3
Total	5	4	2	8

Lecture du tableau : sur les 8 répondants, 2 ont sélectionné à la fois "A" dans leurs réponses à la question 1 et "U" dans leur réponse à la question 2. La réponse "A" a été choisie par 4 répondants et la réponse "U" par 5 répondants.

Sous l'onglet "Options", dans la zone d'édition "Totaux/pourcentages basés sur" activez le bouton radio "Nombre de réponses".

Sous l'onglet "Avancé" cliquez sur le bouton "Tableaux détaillés à double entrée". On obtient :

Tableau de Synthèse 2 x 2				
(Totaux = nb de réponses)				
	Q2 U	Q2 V	Q2 W	Totaux Ligne
A	2	3	1	6
B	3	3	1	7
C	2	2	1	5
D	2	1	2	5
E	2	1	1	4
Total	11	10	6	27

Lecture du tableau : comme précédemment, sur les 8 répondants, 2 ont sélectionné à la fois "A" dans leurs réponses à la question 1 et "U" dans leur réponse à la question 2. Les combinaisons de réponses "A-U", "A-V", "A-W" représentent 6 des 27 combinaisons observées.

Remarques et compléments : on pourra obtenir des tableaux analogues au précédent, mais exprimés en pourcentages (calculés sur les totaux lignes, les totaux colonnes ou le total général des répondants ou des réponses) en utilisant les boîtes à cocher de l'onglet "Options".

5.2 Saisie des données (dichotomie multiple)

De manière sensiblement équivalente, on peut obtenir les mêmes traitements en structurant les données sous la forme "dichotomie multiple".

Créez une nouvelle feuille de données dans le classeur et rendez-la active. Pour cela, par exemple :

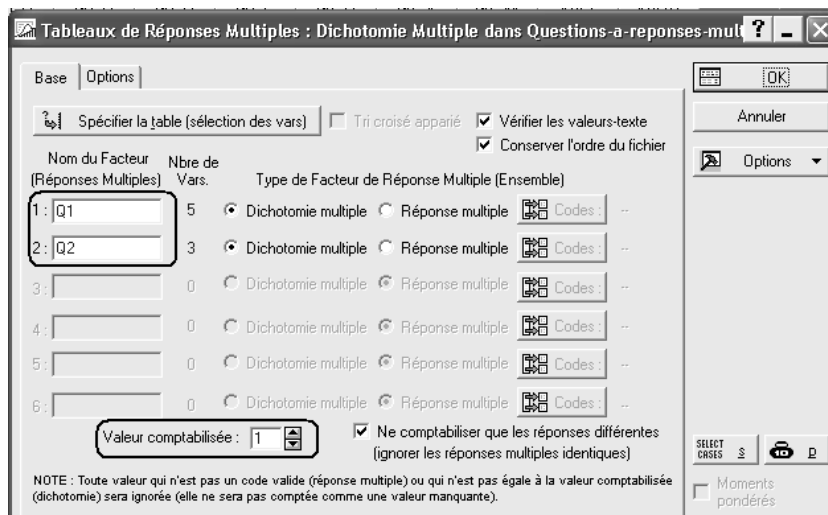
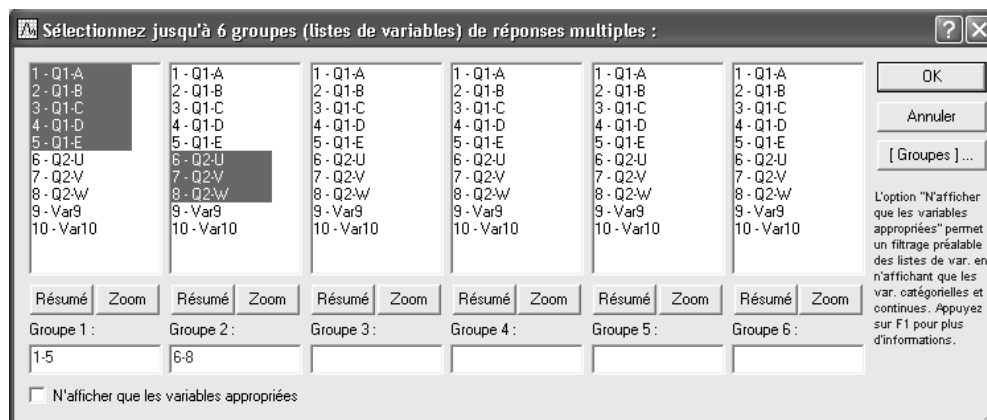
- Cliquez sur l'icône symbolisant le classeur dans la marge de gauche du classeur
- Affichez le menu local à l'aide du bouton droit de la souris
- Sélectionnez "Insérer" puis "Document Statistica" et enfin "Nouveau document" Type : feuille de données.
- Cliquez sur l'icône de cette feuille et à l'aide du menu local, sélectionnez "Feuille de données active".

Définissez cette fois-ci 8 variables nommées Q1-A, Q1-B, Q1-C, Q1-D, Q1-E, Q2-U, Q2-V, Q2-W et saisissez les données observées sous la forme suivante :

	Q1-A	Q1-B	Q1-C	Q1-D	Q1-E	Q2-U	Q2-V	Q2-W
1	1	1	1			1	1	
2		1		1	1		1	1
3	1		1	1		1		1
4				1	1	1		
5		1			1	1		
6	1	1					1	
7	1		1				1	
8		1				1		

N.B. On a choisi ici de coder les modalités non choisies à l'aide de valeurs manquantes. En fait, toute valeur autre que 1 pourrait convenir.

Comme précédemment, utilisez le menu Statistiques - Statistiques élémentaires et sélectionnez la méthode "Tableaux de réponses multiples", mais spécifiez les variables comme suit :



Faites ensuite des traitements analogues aux précédents. Vous devriez retrouver les mêmes résultats.

6 Représentations graphiques d'une variable nominale

6.1 Le cas Fast-food

Vous menez une enquête sur les préférences de consommation des jeunes adultes. Plus particulièrement, vous vous intéressez :

- 1) à leurs préférences pour différents types de fast-food
- 2) à leurs préférences pour différents types de voitures
- 3) à leur comportement (déclaré) par rapport à certains restaurants de fast-food.

De plus, vous notez également le sexe du sujet.

Ces données sont rassemblées dans la feuille de données du classeur [fastfood.stw](#).

Les variables observées sont décrites ci-dessous :

SEXE (variable nominale simple). Le sexe du sujet interrogé est entré comme variable catégorisée dans le fichier de données (i.e., HOMME, FEMME).

Fast-food favori (variable à réponses multiples). Dans le questionnaire utilisé pour cette étude, on demande aux sujets de sélectionner leurs trois types de fast-food préférés dans une liste comportant 8 types. Les 8 types de fast-food proposés sont :

- (1) Hamburger
- (2) Sandwiches
- (3) Poulet
- (4) Pizza
- (5) Fast-food Mexicain
- (6) Fast-food Chinois
- (7) Fruits de mer
- (8) Autres

Les trois choix de chaque sujet sont saisis dans le fichier comme variable à réponse multiple, c'est-à-dire que leur premier choix est saisi dans la variable FASTF_1, leur second choix dans la variable FASTF_2 et leur troisième choix dans la variable FASTF_3.

Voiture favorite (variable à réponses multiples). Dans ce cas, nous avons demandé à chaque individu de renseigner les voitures (marque et modèle) qu'ils aimeraient posséder (l'argent n'étant pas un problème). Ces réponses (particulièrement les marques et les modèles) ont été codées en quatre catégories :

- (1) Voiture de sport, nationale
- (2) Voiture de tourisme, nationale
- (3) Voiture de sport, étrangère
- (4) Voiture de tourisme, étrangère

Comme pour la variable fast food favori (voir ci-dessus), cette variable a été enregistrée comme une variable à réponses multiples, c'est-à-dire que les préférences des personnes interrogées ont été renseignées dans les variables Voitur_1 à Voitur_3. Remarque : dans ce cas, les individus peuvent répéter la même réponse trois fois (par exemple, ils peuvent citer 3 voitures de sports comme leurs voitures les plus convoitées). Dans le cas des fast foods ci-dessus, les réponses multiples identiques n'étaient pas autorisées (c'est-à-dire, ignorées).

Enfin, nous avons demandé aux personnes interrogées d'indiquer les différents restaurants rapides (locaux) qu'elles avaient eu l'occasion de visiter au cours des deux semaines précédant l'enquête, parmi les quatre existants. Les données ont été enregistrées sous la forme d'une variable pour chaque restaurant particulier. Les quatre variables, Burger_1 à Burger_4, représentant les quatre restaurants locaux :

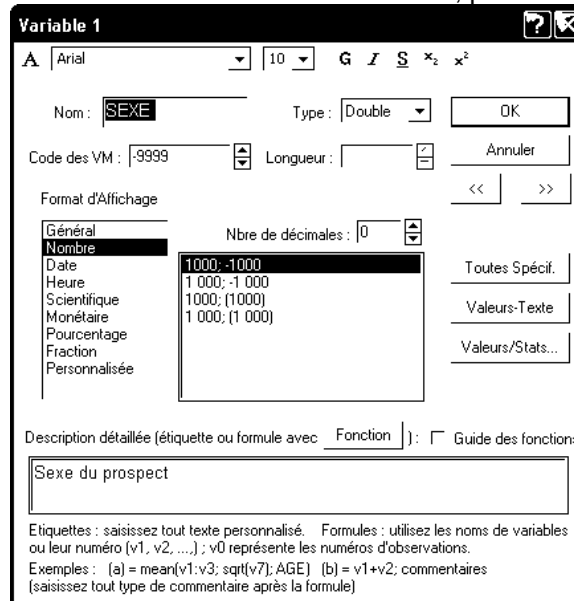
- (1) Burger Meister
- (2) Bill's Best Burgers
- (3) Hamburger Heaven
- (4) Bigger Burger

Si une personne interrogée a déclaré avoir mangé dans au moins un des ces restaurants récemment, un 1 a été entré dans la colonne respective ; sinon, la colonne respective reste vierge. Par conséquent, il s'agit d'une

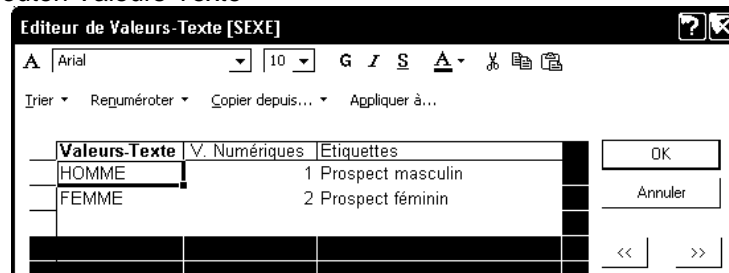
dichotomie multiple, et nous souhaitons tabuler le nombre (ou la proportion) de personnes interrogées ayant déclaré avoir mangé dans chacun des quatre restaurants locaux.

6.2 Examen des types des variables

Faites un double-clic sur une tête de colonne. Pour la variable SEXE, par exemple, on obtient :



Cliquez ensuite sur le bouton Valeurs-Texte



Autrement dit, la variable SEXE est une variable numérique, dont les valeurs 1 et 2 sont étiquetées HOMME et FEMME. Statistica proposera donc tous les traitements propres à une variable numérique pour cette variable. A nous d'éviter les traitements qui n'auraient pas de sens.

Avec Statistica 7.1, on peut également donner le type "texte" à la variable SEXE. Pour cela, faites un double-clic sur la tête de colonne correspondante et attribuez la valeur "texte" au champ Type. Certains traitements propres aux variables numériques ne sont alors plus possibles (par exemple, le tri numérique), mais la plupart le reste encore (par exemple : somme, moyenne, etc).

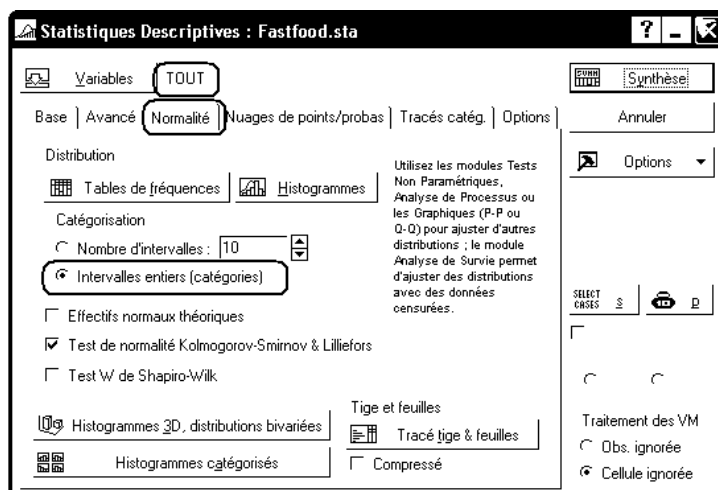
6.3 Tri à plat sur variables nominales

Pour chacune des variables, quelles sont les différentes modalités ? Quels sont les effectifs et fréquences de chacune d'elles.

Utilisez le menu Statistiques - Statistiques élémentaires, puis l'item Statistiques Descriptives.

Cliquez sur le bouton Variation et sélectionnez toutes les variables.

Cliquez sur l'onglet Normalité et activez le bouton Intervalles entiers (Catégories) :



Cliquez ensuite sur le bouton Tables de fréquences. Vous obtenez pour chaque variable un tableau du type suivant :

Table de fréquences : SEXE: Sexe du prospect (Fastfood.sta)							
	Effectifs	Effectifs Cumulés	% Actifs	Indiv.	% Cumulé Ind. Act.	% toutes Observ.	% Cumulé du Total
HOMME: Prospect masculin	164	164	82,00		82,00	82,00	82,00
FEMME: Prospect féminin	36	200	18,00		100,00	18,00	100,00
VM	0	200	0,00			0,00	100,00

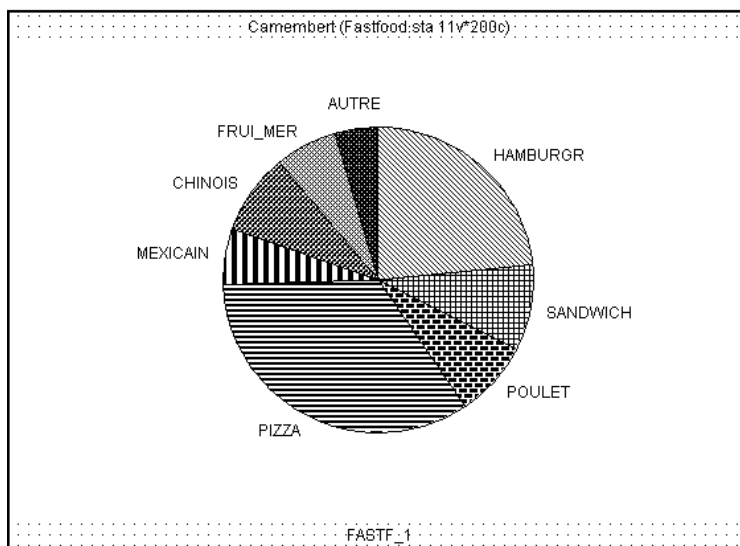
6.4 Diagrammes circulaires

Remarque générale à propos des graphiques. De nombreux graphiques peuvent être réalisés aussi bien à l'aide d'items du menu Graphiques qu'à partir d'items du menu Statistiques. Cependant, les graphiques produits à l'aide du menu Graphiques sont remis à jour lorsque les données sont modifiées, contrairement à ceux produits à l'aide du menu Statistiques.

On veut représenter les distributions de la variable FASTF_1 à l'aide d'un diagramme circulaire.

Utilisez le menu Graphiques - Graphiques en 2D - Camemberts (Diagr. Circulaires)...

Veillez à ce que le paramètre "Intervalle d'effectifs" soit réglé sur Mode entier. Vous devriez obtenir le résultat suivant :



6.5 Graphiques obtenus à partir d'un bloc de données

On veut réaliser un diagramme circulaire pour la variable FASTF_1, pour les sujets de sexe féminin.

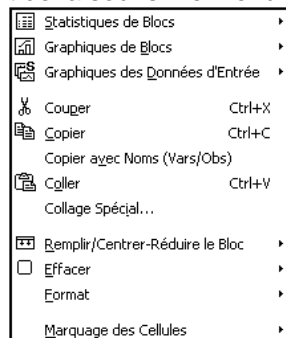
Plusieurs méthodes sont possibles.

Première méthode :

Triez les observations selon les valeurs de la variable SEXE.

Sélectionnez dans la feuille de données, les 36 observations de la variable FASTF_1 relatives aux sujets féminins.

Cliquez sur la sélection avec le bouton droit de la souris. Le menu local suivant s'affiche alors :



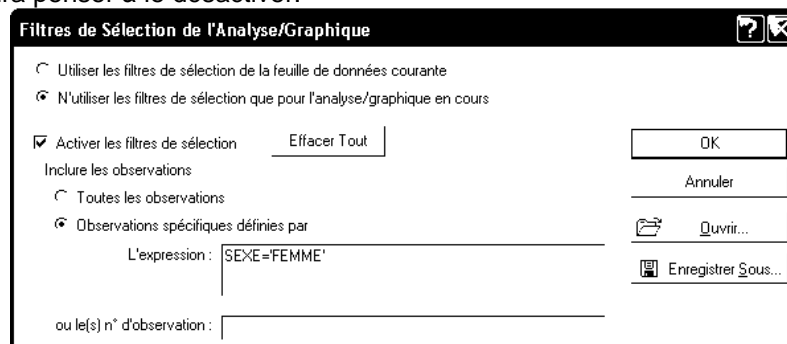
Sélectionnez Graphiques de Blocs - Graphique Perso. issu du Bloc par Colonnes, puis, dans la fenêtre "Sélectionnez un graphique", Graphiques en 2D - Camemberts.

Modifiez ensuite la légende du graphique pour rappeler la manière dont il a été obtenu.

Deuxième méthode :

Utilisez, comme dans le cas des premiers graphiques, le menu Graphiques - Graphiques en 2D - Camemberts (Diagr. Circulaires)...

Dans la fenêtre Camemberts (Diagrammes circulaires), sélectionnez la variable FASTF_1, puis cliquez sur le bouton "Select Cases". Introduisez alors le filtre : SEXE = 'FEMME'. Veillez à cliquer sur le bouton "N'utiliser les filtres de sélection que pour l'analyse/graphique en cours". Sinon, le filtre restera actif pour les autres traitements, et il faudra penser à le désactiver.



6.6 Produire des graphiques pour chacun des groupes définis par une variable nominale.

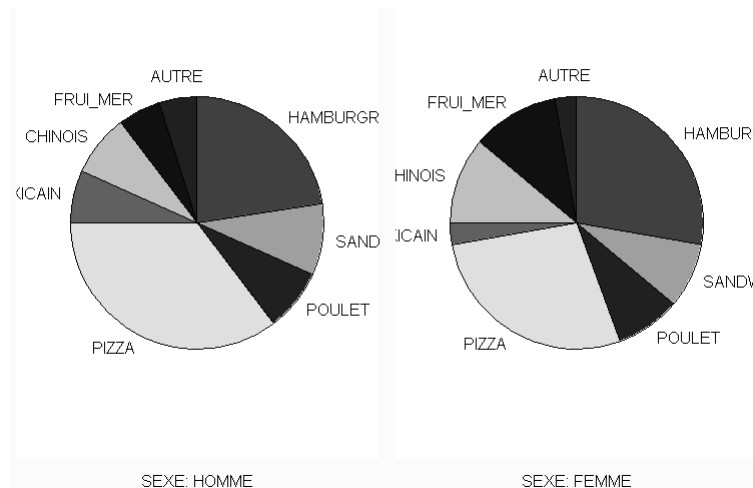
On veut réaliser deux diagrammes circulaires pour la variable FASTF_1, l'un concernant les hommes et le second concernant les femmes.

Utilisez le menu Graphiques - Graphiques catégorisés - Camemberts (Diagr. Circulaires)...

Veillez à ce que le paramètre "Intervalle d'effectifs" soit réglé sur Mode entier.

Indiquez FASTF_1 comme "variable" et SEXE comme "catégorie X".

Vous devriez obtenir :



6.7 Produire des graphiques utilisant l'ensemble des réponses à une question à réponses multiples

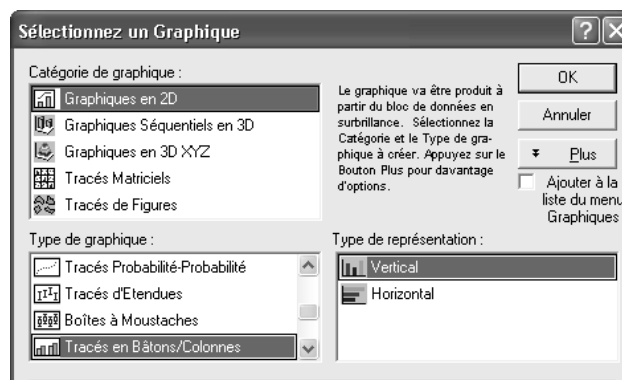
Il n'est pas simple de produire des diagrammes circulaires représentant l'ensemble des réponses à une question à réponses multiples. En revanche, on peut assez facilement produire des diagrammes en bâtons.

On veut par exemple représenter la répartition des quatre types de voitures dans l'ensemble des réponses observées.

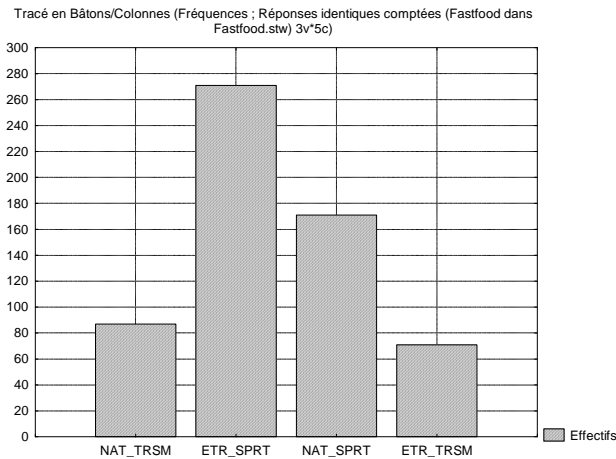
- Utilisez le menu Statistiques - Statistiques élémentaires et la méthode "Tableaux de réponses multiples" pour obtenir le tableau des effectifs concernant les variables VOITUR_1 à VOITUR_3. Veillez à ce que la boîte "Ne comptabiliser que les réponses différentes" ne soit pas cochée. Vous devriez obtenir :

	Effectifs	%age de Réponses	%age d'obs.
NAT_TRSM	87	14,50	43,50
ETR_SPRT	271	45,17	135,50
NAT_SPRT	171	28,50	85,50
ETR_TRSM	71	11,83	35,50
Total	600	100,00	300,00

- Sélectionnez le bloc de données des effectifs des 4 types de voitures, sans sélectionner le total.
- Activez le menu local puis sélectionnez Graphiques de Blocs - Graphique issu du bloc par colonne
- Sélectionnez alors la catégorie de graphique "Graphiques en 2D" et le type de graphique "Tracés en bâtons/colonnes".



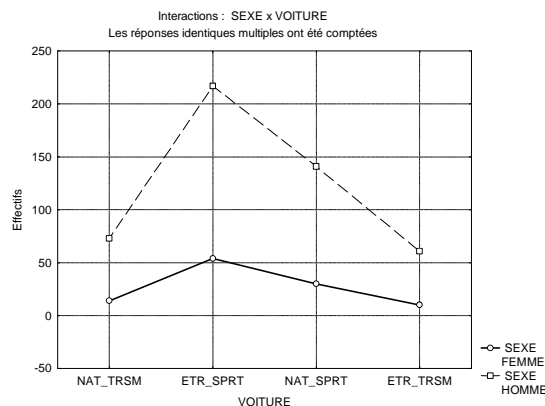
Vous devriez obtenir le résultat suivant :



6.8 Graphique illustrant l'interaction entre deux questions à réponses multiples

Les préférences pour tel ou tel type de voiture dépendent-elles du sexe. On va utiliser les possibilités graphiques de la méthode "Tableaux de réponses multiples" pour illustrer graphiquement la réponse à cette question.

- Utilisez le menu Statistiques - Statistiques élémentaires puis la méthode "Tableaux de réponses multiples".
- Sélectionnez la variable SEXE (qui est en fait une variable ordinaire) et les 3 variables rassemblant les réponses à la question relative aux types de voiture. Comme précédemment, veillez à ce que la boîte "Ne comptabiliser que les réponses différentes" ne soit pas cochée.
- Lorsque la fenêtre de dialogue "Résultats des tableaux de réponses multiples" s'affiche, cliquez sur le bouton "Tracés d'interactions d'effectifs". Vous obtenez le graphique suivant, qui semble montrer que la préférence pour les voitures étrangères sportives est nettement plus marquée chez les hommes que chez les femmes.



6.9 Produire un tableau de contingence

On veut constituer un tableau de contingence à partir des modalités des variables SEXE et FASTF_1.

Utilisez le menu Statistiques - Statistiques élémentaires, puis l'item Tableaux et tris croisés.

Sélectionnez l'onglet Tris croisés.

Spécifiez SEXE dans la première liste de variables et FASTF_1 dans la seconde.

Cliquez sur les boutons OK, puis Synthèse.

Vous devriez obtenir le tableau suivant :

	HAMBUR GR	SANDWICH	POULET	PIZZA	MEXICAIN	CHINOIS	FRUITES	AUTRE	Totaux
HOMME	37	15	13	58	11	13	9	8	164
FEMME	10	3	3	10	1	4	4	1	36
Ts Grpes	47	18	16	68	12	17	13	9	200

7 Graphiques relatifs à des variables numériques

On considère le cas suivant :

Données de Mireault (1990) à propos de l'impact de la mort d'un parent.

Mireault a recueilli des données auprès de 381 étudiants universitaires dont certains avaient perdu un de leurs parents pendant leur enfance, suite à un décès. Les étudiants se répartissaient en trois groupes. Le groupe 1 se composait des sujets dont un des parents étaient décédé. Le groupe 2 comprenait les sujets dont les parents vivaient toujours et étaient restés mariés. Le groupe 3 regroupait les étudiants dont les parents avaient divorcé. Mireault voulait observer les effets de la perte d'un parent sur le niveau actuel de symptomatologie mesuré chez la personne par l'inventaire condensé des symptômes et sur la perception par l'individu de sa propre vulnérabilité vis-à-vis de toute perte future.

Pour toutes les mesures, un score élevé indique que le concept mesuré est davantage présent.

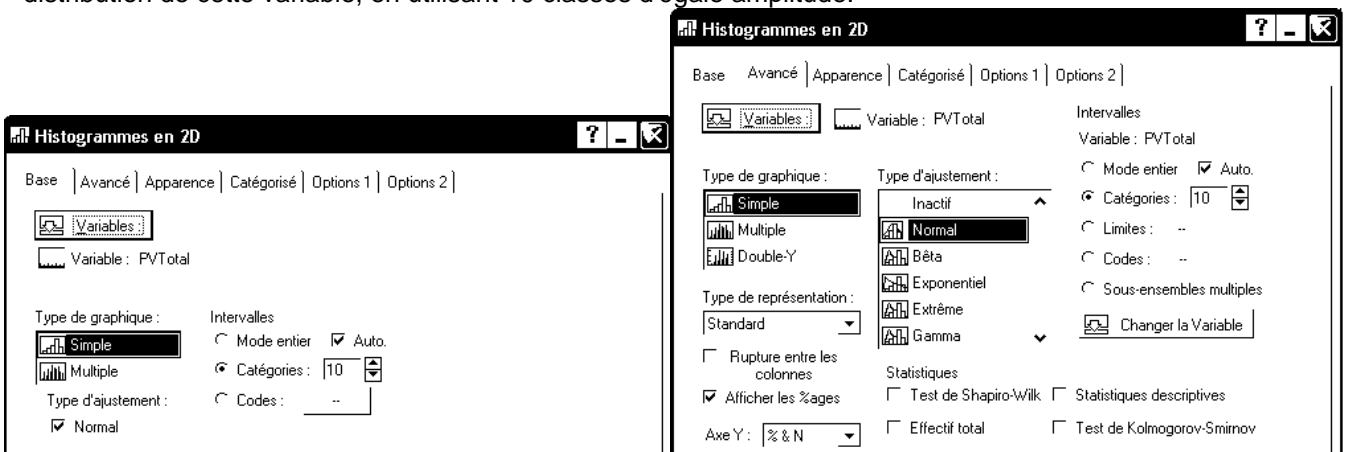
Col 1	ID	= Numéro d'identification du sujet
Col 2	Group	= 1 = décès; 2 = marié; 3 = divorcé
Col 3	Gender	= Sexe. 1 = masculin; 2 = féminin
Col 4	YearColl	= Année d'étude. 1 = première année; 2 = deuxième année; etc.
Col 5	College	= Type d'études. 1 = lettres et sciences; 2 = santé; 3 = ingénieur; 4 = sciences économiques et/ou commerciales; 5 = agriculture
Col 6	GPA	= Moyenne des points obtenus. 4 = très bon; 3 = bon; 2 = satisfaisant; 1 = passable; 0 = insuffisant
Col 7	LostPGen	= Sexe du parent décédé
Col 8	AgeAtLos	= Age de l'enfant lors du décès du parent
Col 9	SomT	= Score T de somatisation
Col 10	ObsessT	= Score T de comportements obsessionnels et compulsifs
Col 11	SensitT	= Score T de sensibilité interpersonnelle
Col 12	DepressT	= Score T de dépression
Col 13	AnxT	= Score T d'anxiété
Col 14	HostT	= Score T d'hostilité
Col 15	PhobT	= Score T d'anxiété phobique
Col 16	ParT	= Score T d'idées paranoïaques
Col 17	PsyT	= Score T de comportements psychotiques
Col 18	GSIT	= Score T sur l'index global des symptômes
Col 19	PVTotal	= Score total de vulnérabilité perçue
Col 20	PVLoss	= Score de vulnérabilité perçue aux pertes ultérieures
Col 21	SuppTotl	= Score T de soutien social

Les données correspondantes se trouvent dans le classeur [Mireault.stw](#).

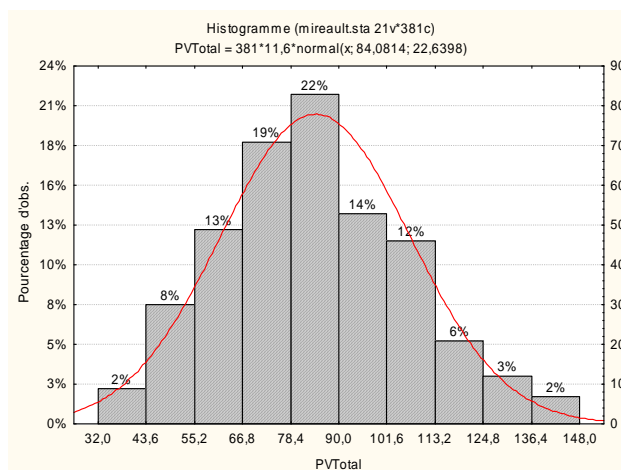
7.1 Histogrammes

On veut réaliser un histogramme représentant la distribution de la variable PVTotal.

Affichez les paramètres descriptifs de cette variable (moyenne, max, min, nombre d'observations, écart type ...) Utilisez ensuite le menu **Graphiques - Graphiques en 2D - Histogrammes...** pour réaliser un histogramme de la distribution de cette variable, en utilisant 10 classes d'égale amplitude.



Vous devriez obtenir un graphique du type suivant :



L'affichage des pourcentages au dessus des barres est obtenu à l'aide de la boîte à cocher "Afficher les %ages", la double graduation, à la fois en pourcentages et en effectifs est obtenue en sélectionnant la valeur "% & N" pour le champ "Axe Y" dans la fenêtre de dialogue "Histogrammes en 2D".

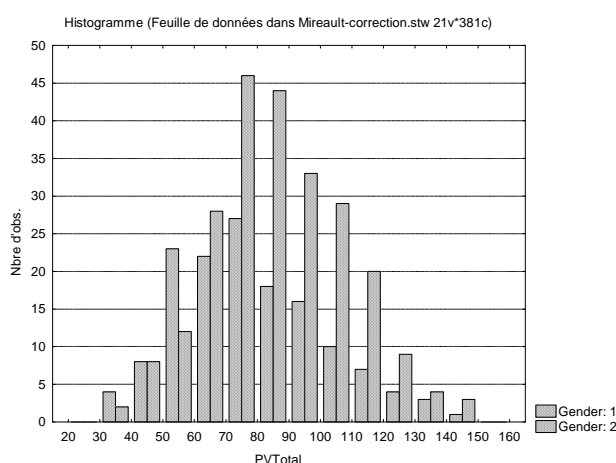
7.2 Histogrammes superposés ou empilés comparant des groupes indépendants

On souhaite réaliser un graphique de type histogramme, faisant apparaître la distribution pour chacun des sexes. On souhaite ne faire qu'un seul graphique rassemblant les deux distributions : c'est ce que Statistica appelle des graphiques *superposés*. On souhaite également que les barres relatives aux deux sexes soient "l'une au dessus de l'autre". Là, pour Statistica, il s'agit d'histogrammes *empilés*.

On obtient facilement un graphique comportant les deux histogrammes juxtaposés en utilisant l'onglet "Catégorisé" du menu Graphiques - Histogramme... Mais, pour obtenir des histogrammes superposés, il faut passer par le menu Graphiques - Graphiques catégorisés.

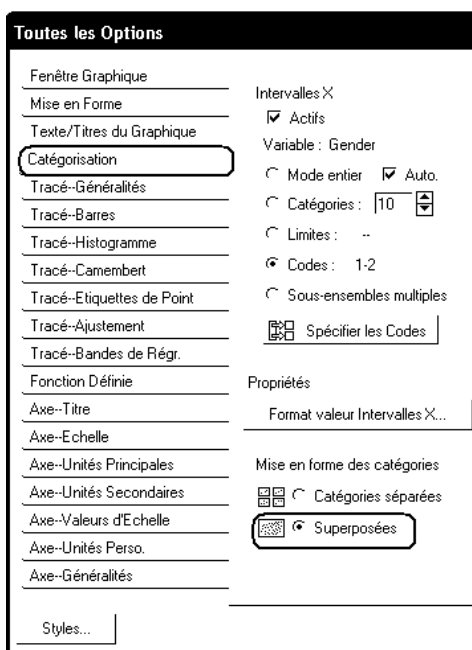
- Utilisez le menu Graphiques - Graphiques catégorisés - Histogrammes
- Sous l'onglet "Base", indiquez PVTotal comme variable et Gender comme "Catégorie X". Sélectionnez "Multiple" comme type de graphique.

On obtient le résultat suivant :

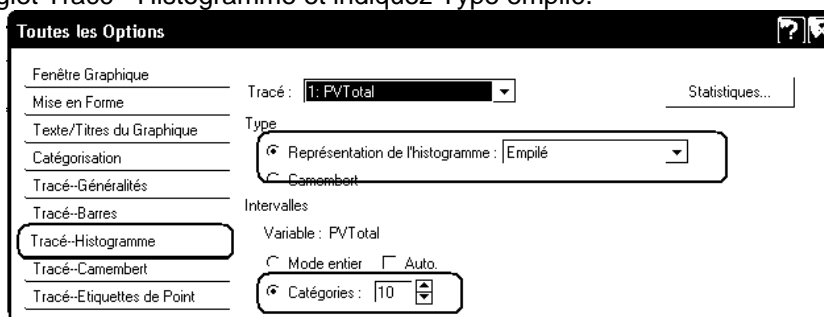


Pour obtenir un histogramme à barres empilées :

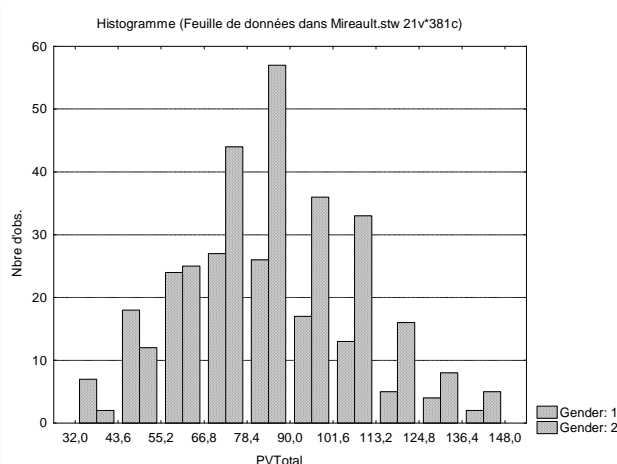
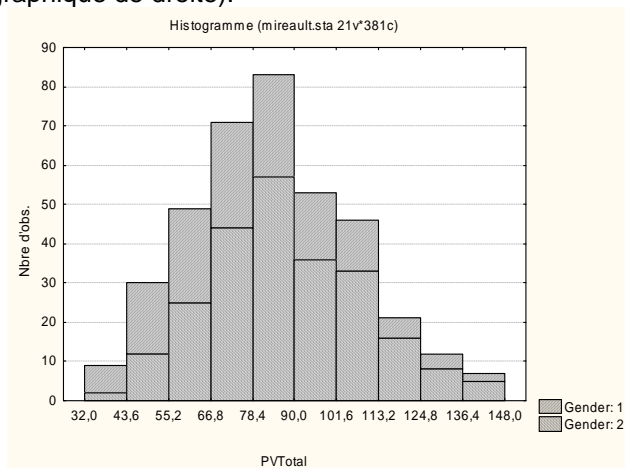
Cliquez sur le graphique avec le bouton droit de la souris, et sélectionnez le menu : Propriétés du graphique (Toutes les options). Activez ensuite l'onglet Catégorisation et sélectionnez "superposées" dans la zone "Mise en forme des catégories".



- Activez l'onglet Tracé - Histogramme et indiquez Type empilé.



Vous devriez obtenir un résultat du type ci-dessous à gauche. Mais, en raison d'un bug de Statistica 7, les rectangles correspondant à Gender=2 ne s'affichent pas. On préférera donc revenir à un graphique "standard" (graphique de droite).



7.3 Graphiques à barres pour une variable ordinale.

La variable YearColl peut être considérée comme une variable ordinale. Pour représenter la distribution de cette variable dans la population étudiée, on peut utiliser les mêmes outils que précédemment, mais en séparant les barres les unes des autres :

Pour cela utilisez le menu local Propriétés du graphique (Toutes options)... , puis l'onglet Tracé - Histogramme, et la zone d'édition "largeur des barres".

7.4 Graphiques de type "boîte à moustache"

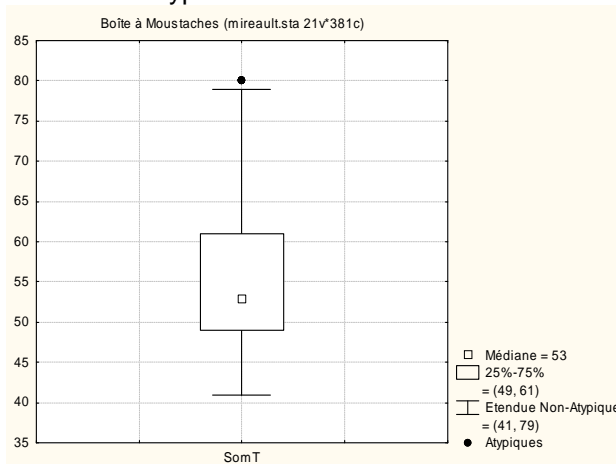
Pour des variables dont la distribution n'est pas aussi régulière que celle de PVTtotal, il peut être intéressant de la représenter à l'aide d'une boîte à moustaches (box and whiskers).

Considérons, par exemple, la variable SomT.

- Calculez d'abord ses valeurs extrêmes, sa médiane et ses quartiles.

	N Actifs	Médiane	Minimum	Maximum	1er Quartile	3ème Quartile
SomT	375	53	41	80	49	61

- Construisez ensuite un graphique de type "boîte à moustaches", avec la médiane comme point central. Vous devriez obtenir un résultat du type suivant :



Règles de construction du graphique : les bases de la boîte représentent les premier et troisième quartiles. Ainsi, la boîte représente les 50% d'observations centrales. On calcule ensuite l'écart inter-quartile (61-49=12) et on le multiplie par 1,5 : $12 * 1,5 = 18$. Pour les valeurs extérieures à l'intervalle délimité par les quartiles, les valeurs qui s'écartent de plus de 18 du quartile correspondant sont considérées comme atypiques.

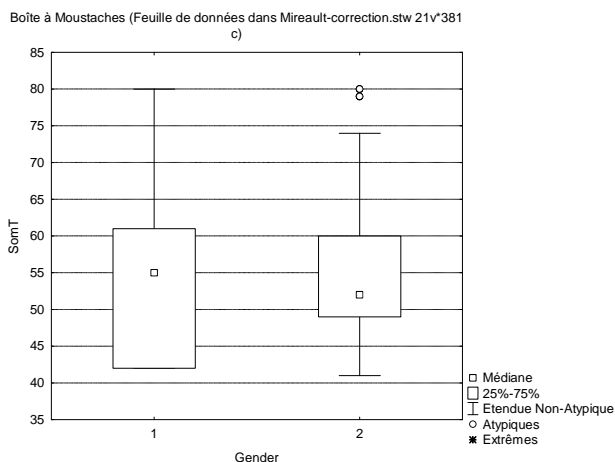
Sur notre exemple, le minimum est de 41. La distance au premier quartile ne dépasse pas 49-41, c'est-à-dire 8. Toutes les valeurs inférieures au premier quartile sont donc typiques et la moustache correspondante va du premier quartile au minimum. En revanche, le maximum est de 80, et la distance de ce maximum au troisième quartile est de 80-61, c'est-à-dire 19. La moustache correspondante s'arrête donc sur la valeur 79, et la valeur 80 est représentée comme valeur atypique.

7.5 Boîtes à moustaches comparant deux groupes indépendants

Ces représentations en boîtes à moustaches sont notamment intéressantes pour comparer les résultats observés sur deux groupes. On veut, par exemple, réaliser, par exemple, des boîtes à moustaches concernant la variable SomT pour les deux groupes définis par la variable SEXE.

- Utilisez le menu Graphiques - Graphiques en 2D - Boîtes à moustaches.
- Indiquez SomT comme variable dépendante et Gender comme variable de classement.

Vous devriez obtenir le résultat suivant :



Remarque 1. On peut aussi utiliser le menu Graphiques - Graphiques catégorisés - Boîtes à moustaches en spécifiant Gender comme "Catégorie X". Le graphique n'est alors pas tout à fait présenté de la même façon.

Remarque 2. On peut également réaliser des boîtes à moustaches pour chacune des conditions obtenues en croisant le sexe et l'année d'études (variable YearColl). Indiquez par exemple l'une des variables comme "Catégorie X" et l'autre comme "Catégorie Y". Refaites le même graphique en indiquant cette fois l'une des variables comme variable de classement et l'autre comme "Catégorie X" et comparez avec le graphique précédent.

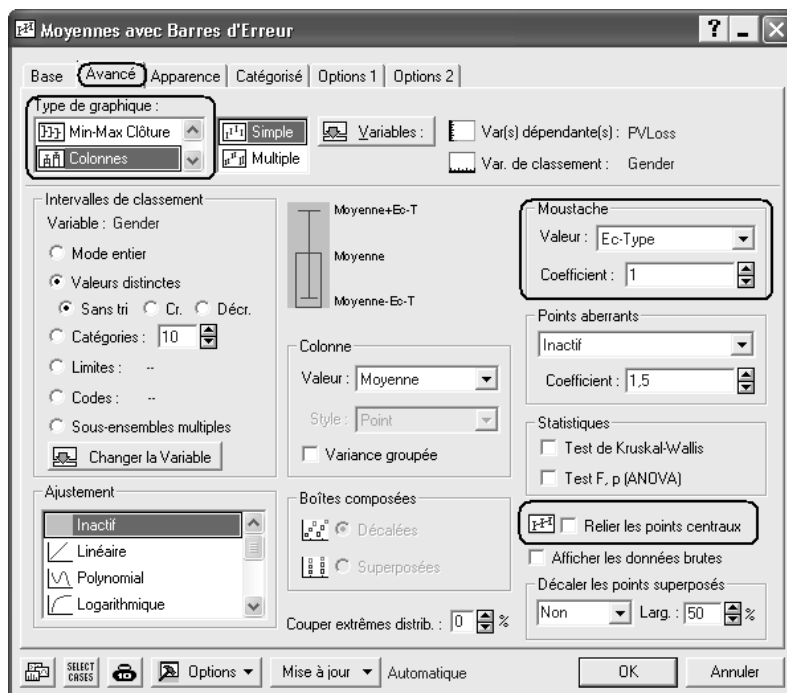
7.6 Graphiques des moyennes par groupe ou des moyennes avec barres d'erreur

On trouve classiquement dans les publications de psychologie des graphiques représentant à l'aide de rectangles, les moyennes dans les différents groupes définis par une variable catégorielle (ou la combinaison de deux variables catégorielles), avec éventuellement une indication relative à l'écart type.

Par exemple, on veut représenter les moyennes de la variable PVLoss dans les deux groupes définis par la variable Gender.

- Utilisez le menu Graphiques - Moyennes avec barres d'erreur...
- Spécifiez PVLoss comme variable dépendante et Gender comme variable de classement.
- Sélectionnez le type de graphique "Colonnes"
- Dans la zone d'édition "Moustache", indiquez : Valeur : Ec-Type et Coefficient = 1.
- Sous l'onglet "Avancé", désactivez la boîte : "relier les points centraux".

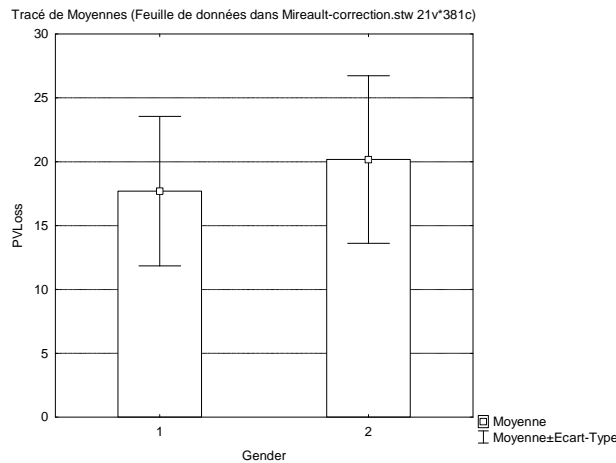
La fenêtre de l'onglet "Avancé" devrait se présenter comme suit :



Dans le graphique produit, l'origine des graduations de l'axe vertical correspond à PVLoss=10. On obtiendra un graphique plus pertinent en rétablissant 0 comme origine des graduations. Pour cela, faites un double-clic sur l'axe vertical, activez l'onglet "Echelle". Puis, dans la zone d'édition "Intervalle", sélectionnez le mode "manuel" et indiquez 0 et 30 comme bornes :



On obtient ainsi le graphique suivant :

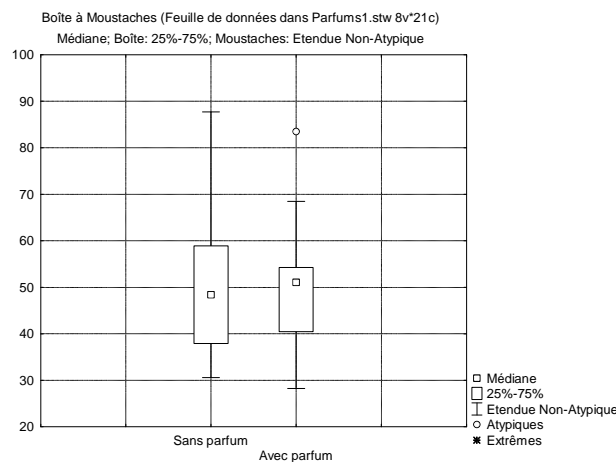


7.7 Graphiques comparant deux groupes appariés

Ouvrez le classeur Parfums1.stw.

Les deux colonnes "Avec parfum" et "Sans parfum" contiennent les valeurs de la variable dépendante dans deux conditions : il s'agit de groupes appariés.

Pour réaliser avec Statistica des graphiques comparant ces deux conditions, il faut sélectionner les colonnes, puis, à l'aide du bouton droit de la souris, afficher le menu local de la sélection et utiliser l'item "Graphiques de blocs". On peut ainsi comparer les deux conditions à l'aide de boîtes à moustaches :



En revanche, ce type de représentation ne peut pas être combiné avec une variable de catégorisation (telle que Genre ou Fumeur sur notre exemple).

8 Nuages de points

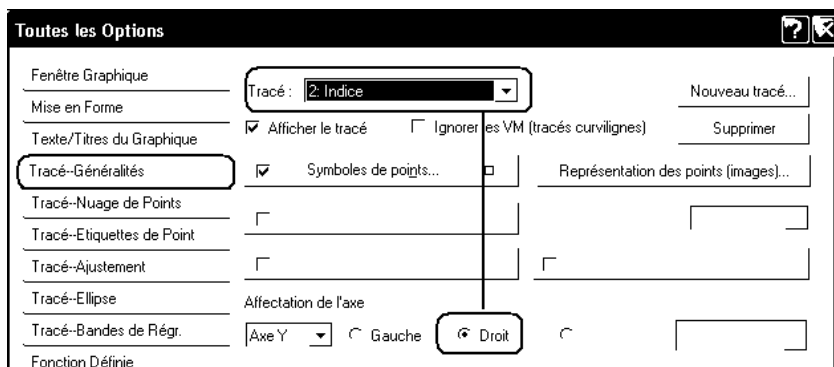
Ouvrez le classeur Variations-Circadiennes.stw et affichez le rapport qui contient une courte présentation des données.

On veut réaliser une première représentation graphique, de type nuage de points, avec en abscisse le moment de la journée et en ordonnée, la vitesse.

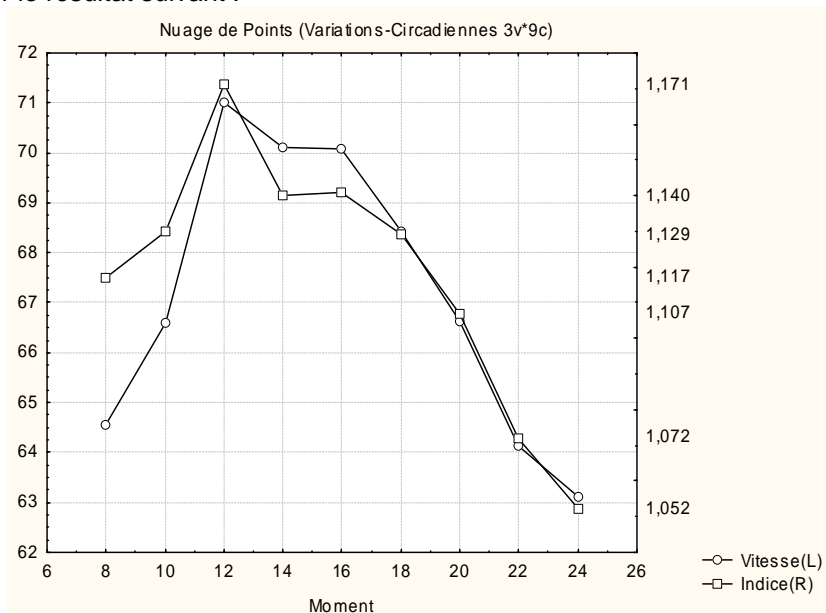
- Utilisez le menu Graphiques - Nuages de points...
- Spécifiez Moment comme variable X et vitesse comme variable Y.
- Désactivez la boîte à cocher Type d'ajustement : linéaire.
- Cliquez sur le bouton OK.

Vous obtenez ainsi un graphique dont les points ne sont pas reliés. Pour le modifier, cliquez avec le bouton droit de la souris sur le graphique et utilisez le menu Propriétés du graphique (toutes options), puis l'onglet Tracé - généralités.

On souhaiterait réunir sur un même graphique les nuages de points (Moment, Vitesse) et (Moment, Indice). Reprenez la manipulation précédente, mais sélectionnez Vitesse et Indice comme variables Y. Indiquez Type de graphique : multiple. On obtient alors les deux nuages de points, mais avec une seule graduation sur l'axe des ordonnées, ce qui ne donne pas un résultat satisfaisant. Pour modifier ceci : cliquez avec le bouton droit de la souris sur le graphique et utilisez le menu Propriétés du graphique (toutes options)



Affichez ensuite l'onglet Tracé - Généralités et associez la variable Vitesse à l'axe Y de gauche et la variable Indice à l'axe Y de droite (cf. copie d'écran ci-dessus) et, indiquez un motif pour joindre les points du nuage. Vous devriez obtenir le résultat suivant :



Par défaut, Statistica n'indique pas de valeurs sur l'axe de droite. Pour afficher les valeurs, sélectionnez l'axe, et à l'aide d'un clic droit, sélectionnez le menu local "Valeurs d'échelle". Cochez l'option "Automatique - aux graduations principales".

Nous souhaitons maintenant étudier la corrélation éventuelle entre les deux variables Vitesse et Indice. Réalisez un nuage de points en spécifiant Vitesse comme variable X, et Indice comme variable Y. Vous pouvez, également, indiquer les moments : 8 h, 10 h, etc comme étiquettes, à l'aide de l'onglet Tracé - Etiquettes de points.

Nous souhaitons, de plus, connaître le coefficient de corrélation entre les deux séries de données. Utilisez le menu Statistiques Élémentaires - Matrices de corrélation. Plusieurs méthodes sont possibles pour indiquer les paramètres. Par exemple : Cliquez sur le bouton "Deux listes". Indiquez Vitesse comme première liste, et Indice comme seconde liste, puis cliquez sur le bouton Synthèse. Le coefficient de corrélation entre les deux séries est : $r = 0,90$.