

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

Première partie : Windows XP - Configuration des salles - Présentation générale

1 Les salles de TD et le réseau

1.1 Nos appareils ...

1.1.1 ... d'un point de vue matériel

Les appareils que nous utilisons sont des PC. Leur configuration physique est la suivante :

- processeur : AMD Athlon 64
- mémoire vive : 1 Go
- périphériques d'entrée/sortie : clavier, écran, souris, port USB
- périphériques de stockage : lecteur de disquettes, lecteur de DVD, disque dur de 160 Go.
- connexion réseau : reliés par l'intermédiaire d'un réseau Ethernet., permettant le partage des imprimantes, l'accès à un volume partagé sur le serveur, et l'accès à un volume personnel.

1.1.2 ... d'un point de vue logiciel

Le système d'exploitation est Windows XP Professionnel.

Le réseau qui relie les différents postes comporte également un serveur ; fonctionnant sous Linux, il est utilisé pour gérer les connexions aux appareils des salles de PC, et l'impression dans ces salles. Il contient les fichiers sur lesquels nous travaillerons en TD, qui se trouveront dans le répertoire **W:\PSY3**. Il contiendra aussi des "comptes" individuels pour les étudiants, afin d'éviter les manipulations de disquettes.

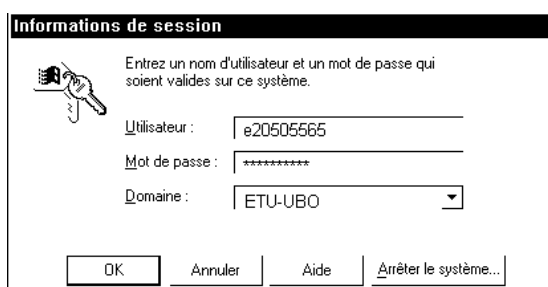
1.1.3 Les logiciels que nous utiliserons cette année.

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.

1.2 Une session de travail sous Windows XP

1.2.1 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".

1.2.2 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.

2 Accès à et par Internet

2.1 Sécurité ... sécurité

Dans une organisation telle qu'une université, dont le réseau reste connecté en permanence au "réseau internet", la sécurité est un souci majeur. En première approche, on peut ranger les problèmes qui se posent en trois catégories.

2.1.1 Au niveau des postes de travail...

Il s'agit de faire en sorte que les postes de travail restent disponibles, avec une configuration stable, et d'éviter que les utilisateurs ne "personnalisent" leur fonctionnement. Aussi, de nombreuses manipulations, effectuées sans problème particulier sur un appareil personnel, sont impossibles à réaliser dans nos salles. Et, dans bien des cas, votre enseignant n'a pas plus de "droits" ou de "pouvoirs" que vous pour les réaliser. En particulier, seul l'administrateur (le technicien chargé des salles) peut :

- installer des logiciels (par exemple les logiciels nécessaires à l'exécution de certains CD-ROM)
- installer de nouveaux périphériques (inutile par exemple, d'amener son scanner, ou d'essayer de déplacer les imprimantes de la salle)
- installer les extensions et autres plug-ins proposés par des sites Web sur Internet.

2.1.2 Sécurité du réseau local

Chaque utilisateur possède des droits étendus sur son propre compte, et très peu de droits sur ceux des autres utilisateurs.

Pour les volumes partagés (le volume "TD sur letsamba" par exemple), les droits sont différents pour les étudiants (lecture seule en général) et pour les enseignants (lecture/écriture). Mais, par exemple, les enseignants ne peuvent pas écrire sur votre compte ni retrouver un mot de passe que vous auriez oublié.

2.1.3 Sécurité vis à vis d'Internet

Cette sécurité est à double visage.

D'une part, l'Université veut couvrir sa responsabilité vis-à-vis d'"initiatives" que pourraient prendre certains étudiants, et vos droits sont très limités.

D'autre part, l'Université, dont le réseau est connecté en permanence à l'Internet, doit se protéger contre les attaques extérieures (virus, piraterie, etc). C'est pourquoi le téléchargement de fichiers à partir de serveurs ftp, par exemple, est (presque) impossible.

2.2 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/>

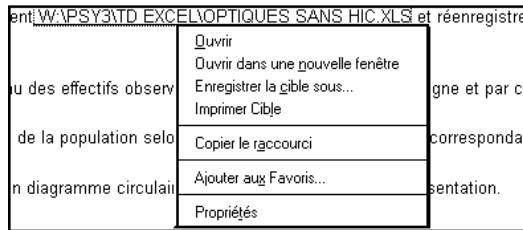
A partir de la page d'accueil, cliquez sur "Les membres du Département" puis sur "---> Documents pédagogiques sous la responsabilité de F.G. Carpentier".

|| 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 :

Calcul de Z critique :

Alpha :

Nature du test :

Test unilatéral

Test bilatéral

Z-critique : 1.959964

- 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:

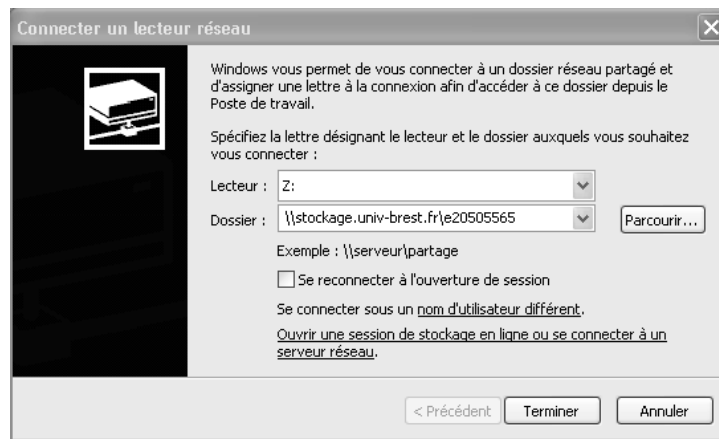
3 L'espace de stockage sur l'ENT

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.

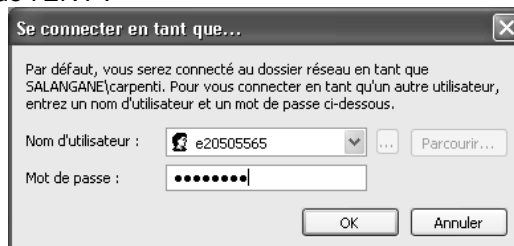
Les manipulations sont analogues à celles décrites dans le paragraphe "Utiliser son compte", permettant de monter le volume de votre binôme étudiant.

- 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 2009, 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 (le disque D: du poste de travail par exemple).
- Quitter le logiciel.
- Copier le fichier correspondant depuis le disque local sur le volume réseau correspondant à son espace de stockage.

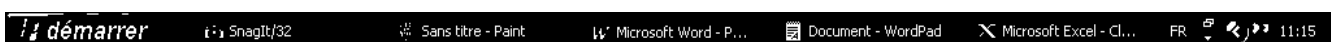
De même, pour reprendre un travail à partir d'un fichier existant sur son espace de stockage :

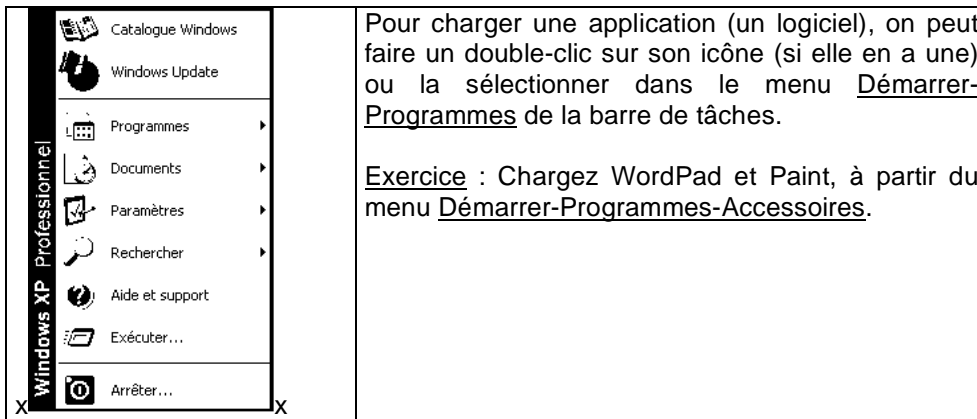
- Copier le fichier depuis le volume réseau sur un disque local.
- Travailler sur le fichier enregistré localement.
- Mettre à jour le fichier du serveur de stockage.

4 Quelques manipulations courantes sous Windows

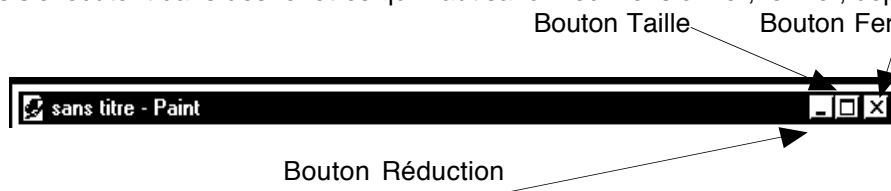
4.1 L'environnement de travail

L'écran affiché après l'ouverture de la session est le bureau. On y trouve diverses icônes et, dans le bas de l'écran la barre de tâches :

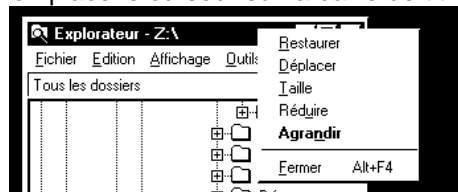




Les applications s'exécutent dans des fenêtres qu'il faut savoir redimensionner, fermer, déplacer, etc



Dans de nombreux cas, plusieurs alternatives sont possibles pour obtenir un résultat donné. Ainsi, pour fermer une fenêtre, on peut aussi sélectionner l'item correspondant dans le menu local que l'on fait apparaître à l'aide du bouton droit de la souris lorsque l'on place le curseur sur la barre de titre



ou appuyer sur la combinaison de touches Alt+F4.

L'ajustement de la taille d'une fenêtre s'effectue à l'aide de la case de taille, en bas à droite de la fenêtre.



4.2 Disques et Fichiers. Manipulations courantes sur les fichiers

4.2.1 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 recopiés 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 recopiés. 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 recopiés sur votre profil du serveur.

Il est également important de noter que le compte "etudiant" est verrouillé (ou devrait l'être - il ne l'est pas au moment où sont écrites ces lignes). 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 perdu à la fermeture de la session.

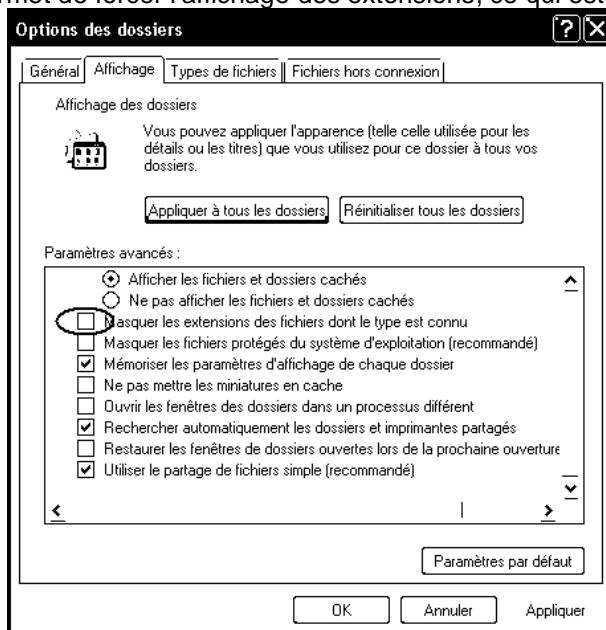
Notez que certains logiciels vous font des propositions plus farfelues, telles que C:\Windows\System32. Evitez à tout prix ces suggestions absurdes.

4.2.2 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.docDocument Word
Mon 2è document.xlsDocument Excel
Mon 3è document.cwkDocument 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.



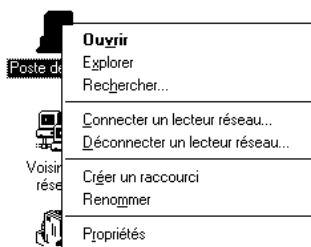
4.2.3 Ouvrir un document par double clic

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.

4.2.4 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 :



4.2.5 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 ;

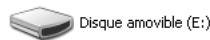
- Afficher la cible ;

- Utiliser le menu Edition-Coller.

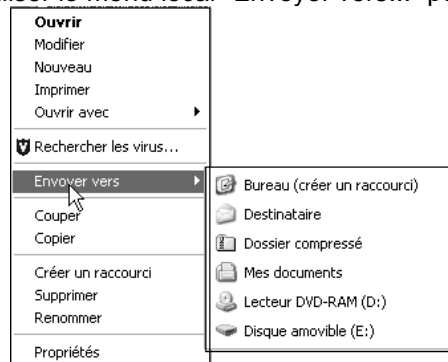
- Enfin il est également possible d'ouvrir deux exemplaires de l'explorateur (ou un Explorateur et un Poste de travail), d'afficher la source et la cible dans deux fenêtres différentes et de faire glisser les fichiers désirés de la première vers la seconde.

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 :



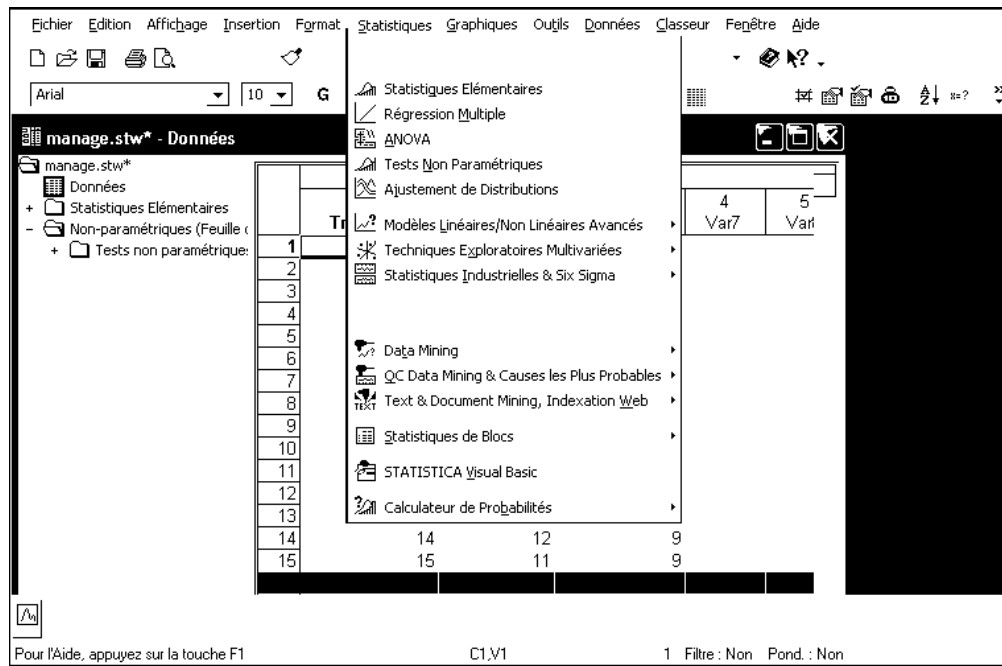
Remarque : Il peut arriver qu'une clé ne soit pas bien reconnue par Windows. Dans ce cas, on peut essayer de refermer la session, et de la ré-ouvrir, après avoir introduit la clé, ou même de redémarrer l'appareil, avec la clé présente dans un port USB de l'appareil. D'autres comportements étonnants peuvent 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

5 Statistica : l'interface utilisateur

5.1 L'écran de travail

Statistica 7.1 est un logiciel dédié aux traitements statistiques. C'est également la "brique" de base des logiciels proposés par Statsoft, et ses possibilités d'interaction avec d'autres logiciels (tableurs, systèmes de gestion de bases de données, traitements de textes, ...) sont nombreuses. En revanche, l'interface utilisateur pourra sembler un peu déconcertante au premier abord.



5.2 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.

5.2.1 Le menu Outils - Options

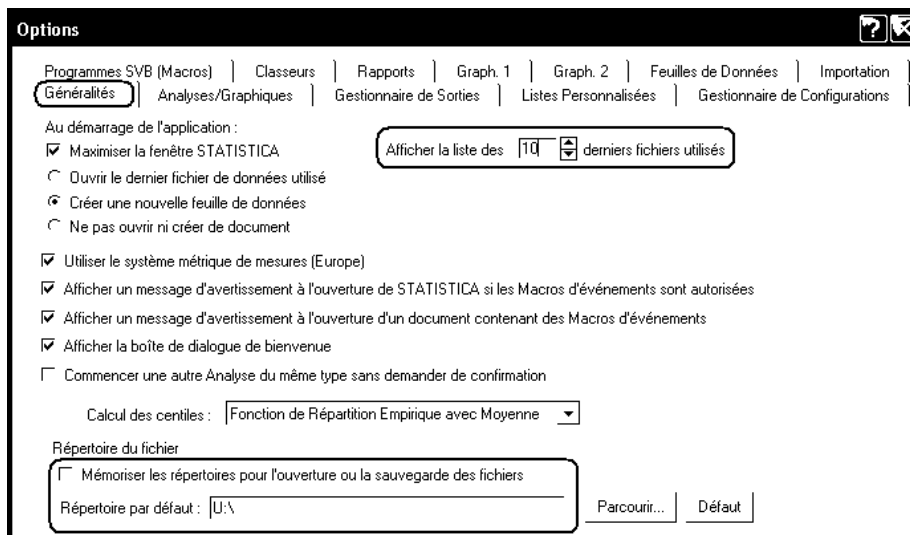
Le menu Outils - Options contient de nombreuses possibilités de paramétrage de Statistica. Heureusement, seules quelques-unes d'entre elles méritent d'être retouchées.

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

N.B. Les options ainsi choisies sont enregistrées dans le profil de l'utilisateur lorsque l'on quitte le logiciel. *Il n'y a aucun enregistrement si le compte est verrouillé ou si Statistica se plante en cours de travail.*

5.2.1.1 Spécifier le répertoire d'enregistrement par défaut

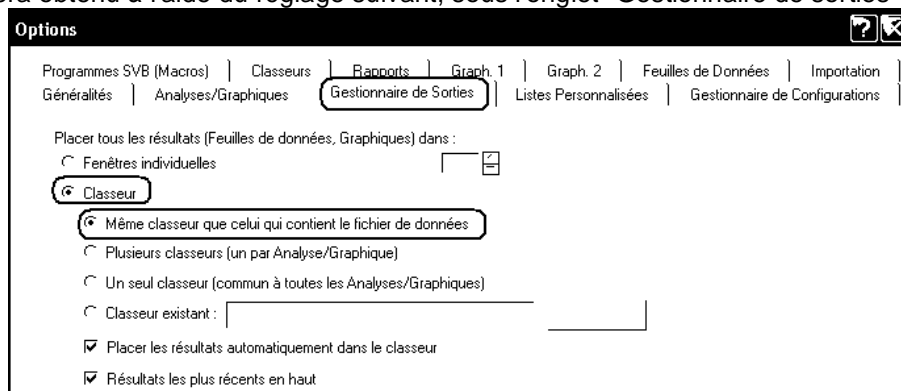
Par défaut, Statistica propose le répertoire "Mes Documents" pour l'enregistrement des nouveaux documents. On peut modifier ce comportement en utilisant l'onglet Généralités :



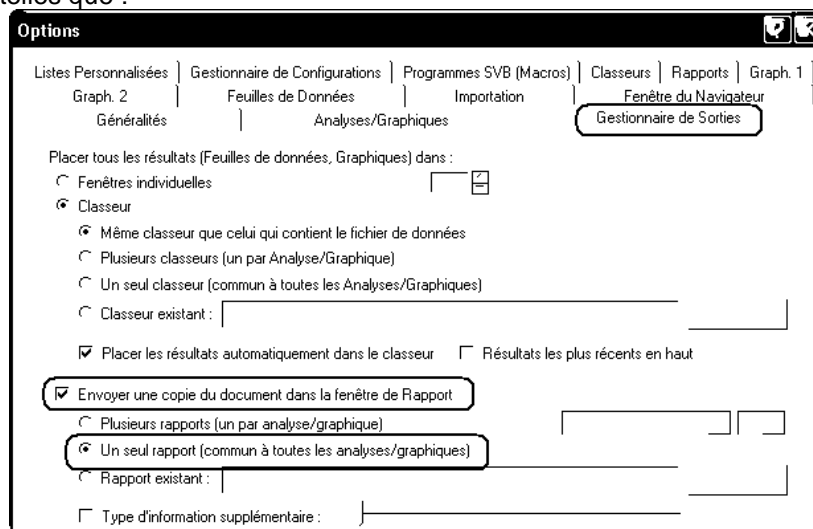
On peut aussi, sans inconvénient, réduire la longueur de la liste des derniers fichiers utilisés (indiquer 10 au lieu de 16 par exemple).

5.2.1.2 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.

5.3 Les objets manipulés par Statistica

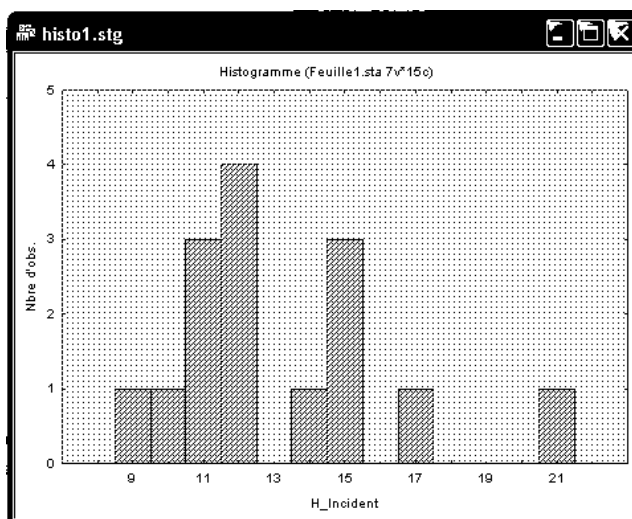
La **feuille de données** est organisée en variables et observations. Les colonnes sont les variables. Chaque ligne représente en règle générale, un individu statistique, appelé *observation*.

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

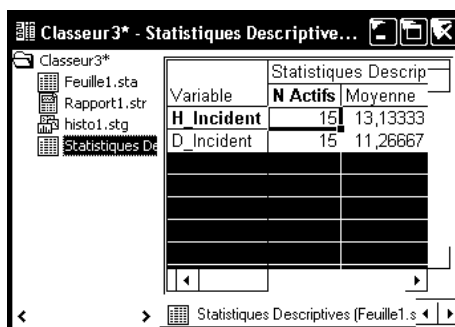
Les feuilles de données peuvent être enregistrées comme fichiers autonomes (fichiers *.sta). Elles contiennent les données d'entrée sur lesquelles s'effectuent les traitements statistiques. Les résultats de ces traitements s'affichent dans un document de sortie. Plusieurs possibilités sont offertes.

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).

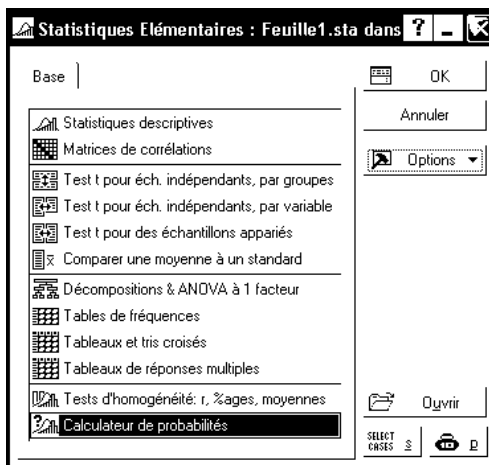


Les classeurs : les données d'entrée et de sortie peuvent également être stockées comme onglets dans un classeur. Un classeur est un "container" accueillant d'autres objets, organisés sous forme hiérarchique. Ils correspondent aux fichiers de type *.stw.



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 :



5.4 Saisie de données

Ouvrez une nouvelle feuille de données. Si nécessaire, utilisez le menu Données - Observations - Ajouter pour disposer de 15 lignes dans la feuille de données. 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

5.5 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 :

The dialog box 'Variable 4' shows the following configuration:

- Font: Arial, Size: 10
- Nom: Différence
- Type: Double
- Code des VM: -9999
- Longueur: [empty]
- Format d'Affichage: Général (selected), Nombre, Date, Heure, Scientifique, Monétaire, Pourcentage, Fraction, Personnalisée
- Description détaillée (étiquette ou formule avec Fonction): Guide des fonctions
- Formulaire: = v3 - v2

Etiquettes : saisissez tout texte personnalisé. Formules : utilisez les noms de variables ou leur numéro (v1, v2, ...); v0 représente les numéros d'observations.
Exemples : (a) = mean(v1:v3; sqrt(v7); AGE) (b) = v1+v2; commentaires (saisissez tout type de commentaire après la formule)

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.

5.6 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 Difference, 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 Difference, 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.

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

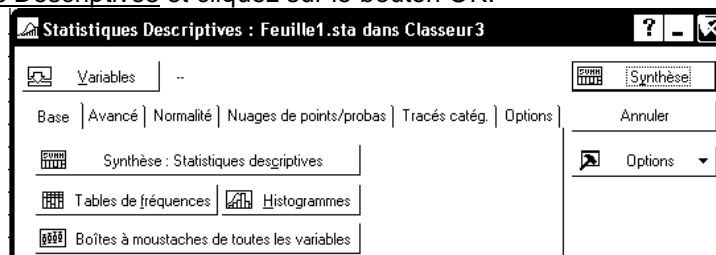
Utilisez le bouton trier . La boîte de dialogue qui s'affiche permet d'effectuer un tri hiérarchisé. Par exemple, trie les données selon les valeurs de la colonne Différence, puis refaites le tri selon les valeurs de la colonne Trimestre.

5.8 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.

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

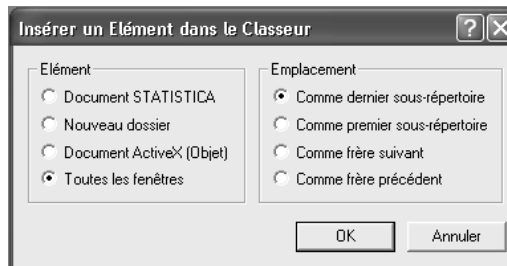
Remarquez que l'analyse reste active, ce qui permet de demander d'autres traitements accessibles dans le même module sans redéfinir les paramètres.

5.9 Enregistrement du document

La feuille de données 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**

Rendez active la feuille Manage.sta du classeur et enregistrez ce classeur en indiquant encore Manage comme nom. Comme l'extension associée à un classeur, .stw, diffère de la précédente, un deuxième fichier sera créé sur le disque.

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.

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 le fichier fastfood.sta.

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 :

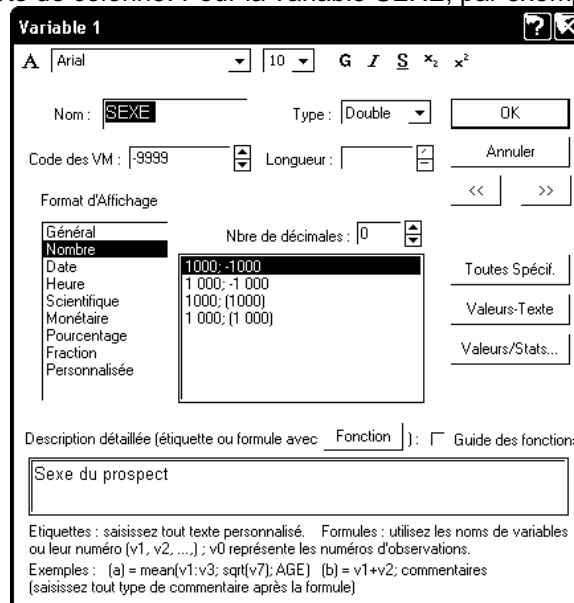
- (1) Hamburger
- (2) Sandwiches
- (3) Poulet
- (4) Pizza
- (5) Fast-food Mexicain
- (6) Fast-food Chinois
- (7) Fruits de mer
- (8) Autres

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.

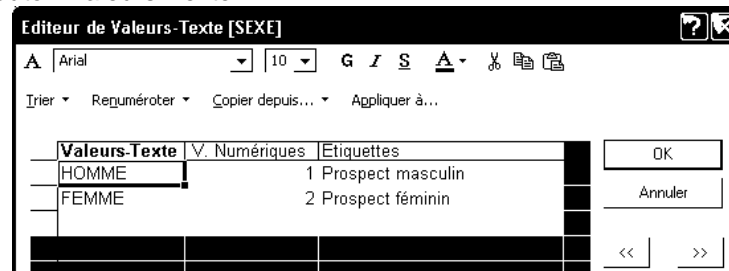
Chargez la feuille de données Fastfood.sta.

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).

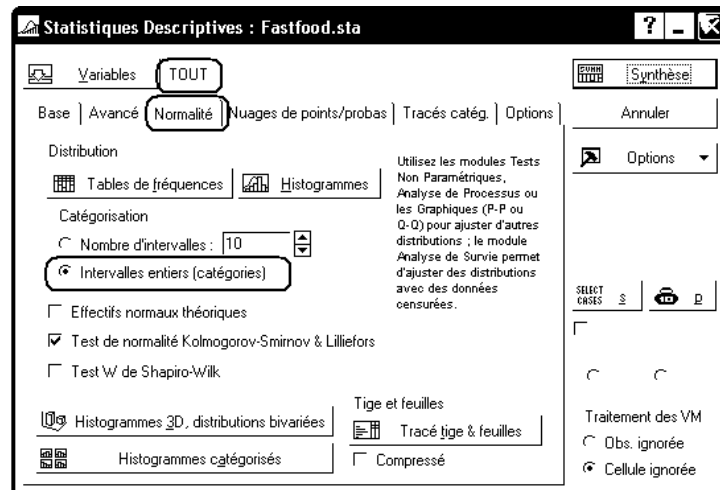
6.3 Tri à plat sur variables nominales

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 Variables 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 :

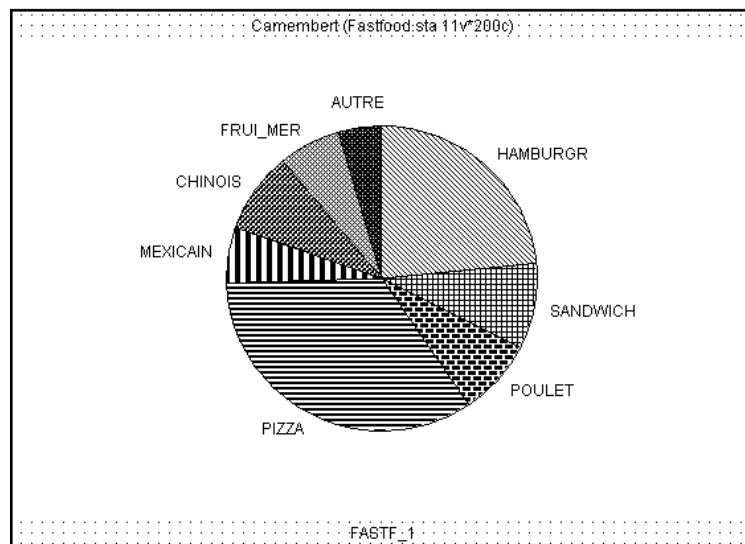
Table de fréquences : SEXE: Sexe du prospect (Fastfood.sta)							
	Effectifs	Effectifs Cumulés	% Indiv. Actifs	% Cumulé Ind. Act.	% toutes Observ.	% Cumulé du Total	
HOMME: Prospect masculin	164	164	82,00	82,00	82,00	82,00	
FEMME: Prospect féminin	36	200	18,00	100,00	18,00	100,00	
VM	0	200	0,00		0,00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

6.4 Diagrammes circulaires

On veut représenter les distributions des variables SEXE et 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

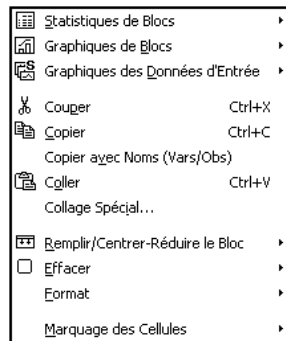
On veut réaliser un diagramme circulaire pour la variable FASTF_1, pour les sujets de sexe féminin. Plusieurs méthodes sont possibles.

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 :

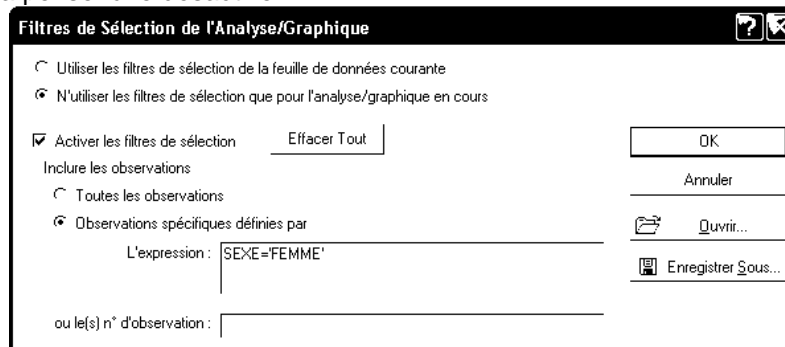


Sélectionnez Graphiques de Blocs - Graphique Perso. issu du Bloc par Colonne, puis, dans la fenêtre "Sélectionnez un graphique", Graphiques en 2D - Camemberts. Modifiez ensuite la légende du graphique pour rappeler la manière dont il a été obtenu.

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 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	SANDWI CH	POULET	PIZZA	MEXICAI N	CHINOIS	FRUI_ME R	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.

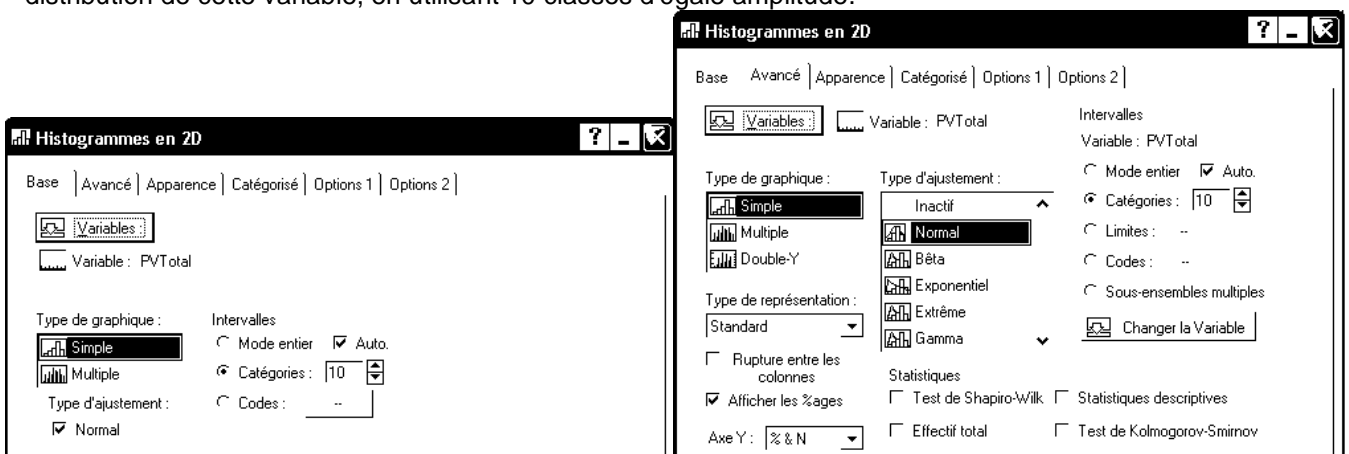
Pour toutes les mesures, un score élevé indique que le concept mesuré est davantage présent.

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

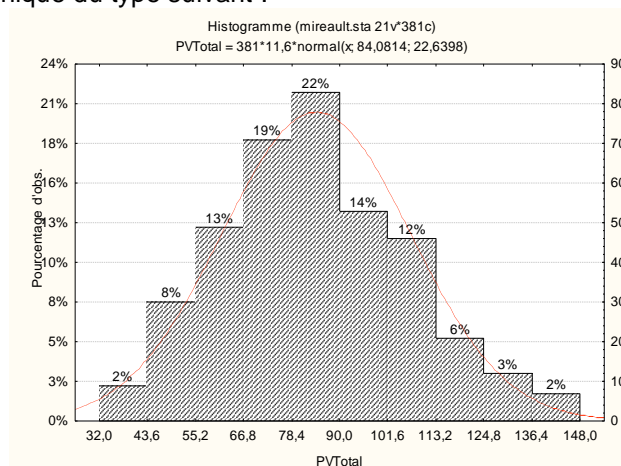
Les données correspondantes se trouvent dans le classeur Mireault.stw.

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".

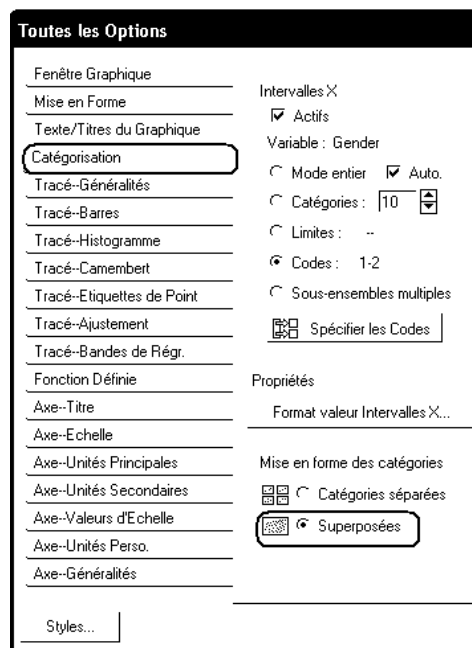
On aurait souhaité réaliser un histogramme avec un découpage en classes d'amplitudes inégales, avec comme bornes : 32, 50, 60, 70, 80, 90, 100, 110, 120, 148. Essayez de le faire en utilisant l'onglet "Avancé" et le bouton "limites" dans la zone d'édition "Intervalles". Le résultat est-il satisfaisant ?

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.

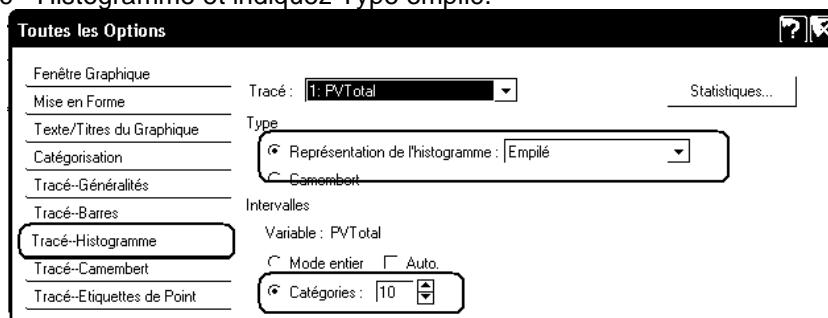
On obtient facilement un graphique comportant les deux histogrammes juxtaposés en utilisant le menu local obtenu en cliquant, avec le bouton droit de la souris sur la colonne correspondant à la variable PVTotale. Sélectionnez l'item Graphiques des données d'entrée - Histogramme en 2D par... et sélectionnez Gender comme seconde variable.

Pour obtenir un histogramme à barres empilées :

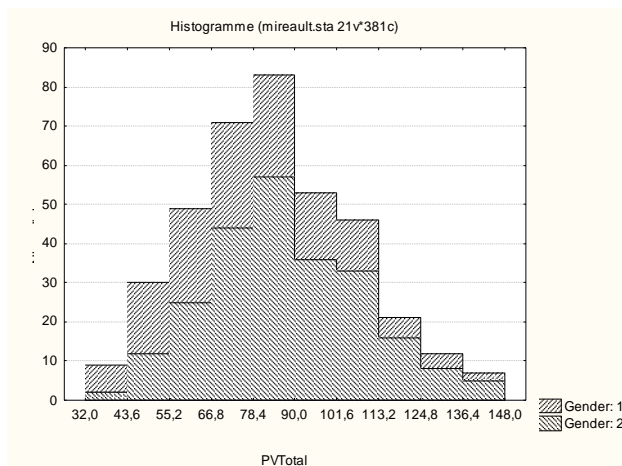
Cliquez sur le graphique avec le bouton droit de la souris, et sélectionnez le menu : Propriétés du graphique (Toutes les options). Activez ensuite l'onglet Catégorisation et sélectionnez "superposées" dans la zone "Mise en forme des catégories".



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



Vous devriez obtenir un résultat du type suivant :



Exercice : Refaites le même graphique en utilisant le menu Graphiques - Histogramme et l'onglet "Catégorisé". De même, refaites le graphique en utilisant le menu Graphiques - Graphiques catégorisés - Histogramme.

7.1 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".

7.2 Graphiques de type "boîte à moustache"

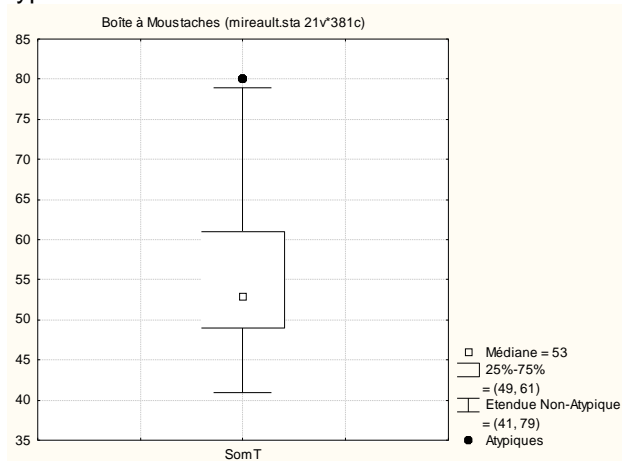
Pour des variables dont la distribution n'est pas aussi régulière que celle de PVTotale, il peut être intéressant de la représenter à l'aide d'une boîte à moustaches (box and whiskers).

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.

Ces représentations en boîtes à moustaches sont notamment intéressantes pour comparer les résultats observés sur deux groupes. Réaliser, par exemple, des boîtes à moustaches concernant la variable SomT pour les deux groupes définis par la variable SEXE. Utilisez par exemple le menu Graphiques - Graphiques

catégorisés - Boîtes à moustaches en spécifiant Gender comme "Catégorie X". 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.

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...

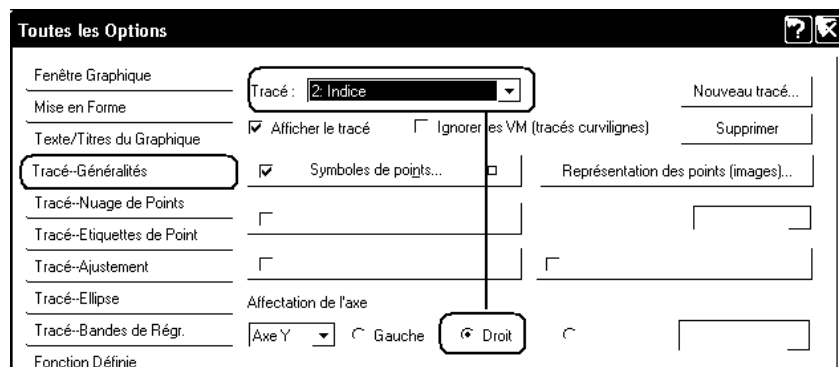
Spécifiez Moment comme variable X et vitesse comme variable Y.

Désactivez la boîte à cocher Type d'ajustement : linéaire.

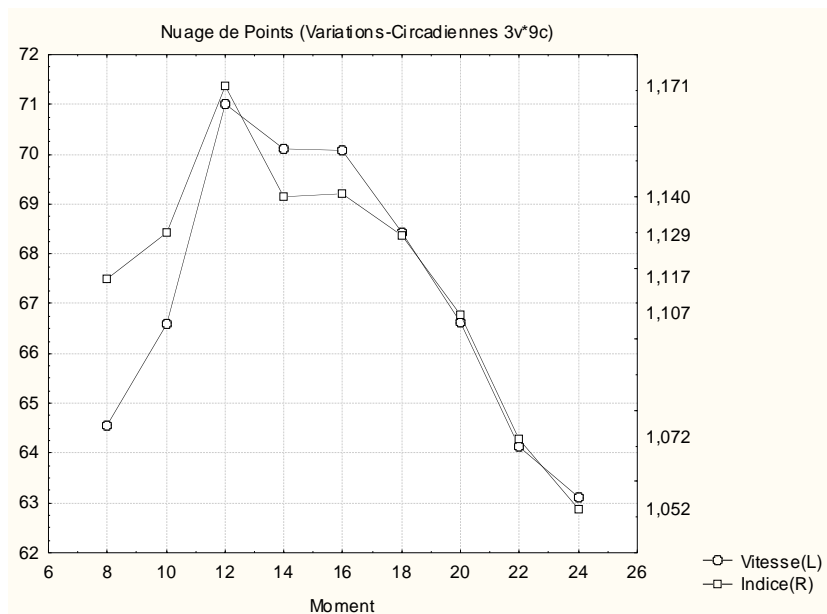
Cliquez sur le bouton OK.

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 : $r = 0,90$.