

Modalisa 4

Création, analyse de questionnaires et d'entretiens pour MAC & PC



ODBC & SQL

Vous trouverez dans cette section les informations qui vous permettront d'utiliser Modalisa comme outil de data mining. Le module client ODBC permet de lire et de copier des données dans Modalisa.

Aucune fonction de modification des données originales (dans la base cible) n'est accessible. Modalisa vous permet donc de consulter les données tout en garantissant totalement leur intégrité.

Ce module est fourni en standard dans Modalisa 4.

Modalisa a été avant tout conçu comme un logiciel destiné aux chargés d'études et de recherche. Tout particulièrement à l'intention de ceux et celles dont les notions de statistiques étaient anciennes ou incomplètes.

Dans cet esprit, les fonctions «client ODBC» de Modalisa ont pour objectif de fournir aux chargés d'études un moyen simple d'accéder aux bases de données via le module ODBC.

Cette fonction est tout particulièrement utile dans les entreprises au sein desquelles le Directeur Informatique ou ses affidés n'est nullement disposé à laisser les chargés d'études ou de recherche "farfouiller" sur leurs précieuses bases.

Notons ici le gouffre culturel séparant l'informaticien du chargé d'études. Un informaticien est formé pour être la vestale des données de l'entreprise, sa hantise est le crash, la corruption de données...

Pour le chargé d'études un gisement de données est une source dans laquelle il va sélectionner des éléments nécessaires à l'édification de ses hypothèses. Le chargé d'études doit faire parler les données, donc les manipuler, les recoder, les transformer. Toutes opérations qui ne peuvent que hérisser les cheveux d'un informaticien gestionnaire.

On constate que les outils d'analyse de données ou de data mining présents sur les sites informatiques n'offrent pas la souplesse de recodage présente traditionnellement dans les outils de traitement de questionnaires et d'analyse de données.

L'auteur de Modalisa a donc décidé d'intégrer un module d'accès aux données à Modalisa plutôt que de créer un nouveau programme dit de data mining.

Ceci confère à Modalisa un avantage supplémentaire en renforçant son caractère d'outil du chargé d'études.

Tout comme le traitement des questions textes ou le serveur http Modalisa intègre en standard ce module ODBC.

Caractéristiques fonctionnelles du module

Modalisa ODBC a trois caractéristiques principales :

- 1- Respect total les données de la base (pas de commande de modification intégrée).
- 2- Copie des données sélectionnées sur l'ordinateur du chargé d'étude. Ceci afin de permettre de réaliser les traitements «off line»
- 3- Langage SQL volontairement limité aux commandes d'import et de sélection des données, commandes d'interrogation.

Principe de fonctionnement

Le module ODBC de Modalisa fonctionne à l'identique sur Windows 95, NT et MacOS.

Il permet aux utilisateurs d'une plate-forme de récupérer des données présentes sur d'autres plate-forme.

Un Macintosh pourra à travers le réseau récupérer des données sur une base de données de Microsoft Access présente sur un PC équipé de Windows95, 98 ou NT.

Modalisa ODBC permet à un ordinateur portable de récupérer des données sur un serveur voire un mainframe.

La procédure du choix du pilote (driver) ODBC dépend de l'environnement matériel de l'utilisateur.

Toutefois, cette procédure est peu différente qu'on soit utilisateur d'un PC ou d'un Macintosh.

INSTALLATION SUR WINDOWS

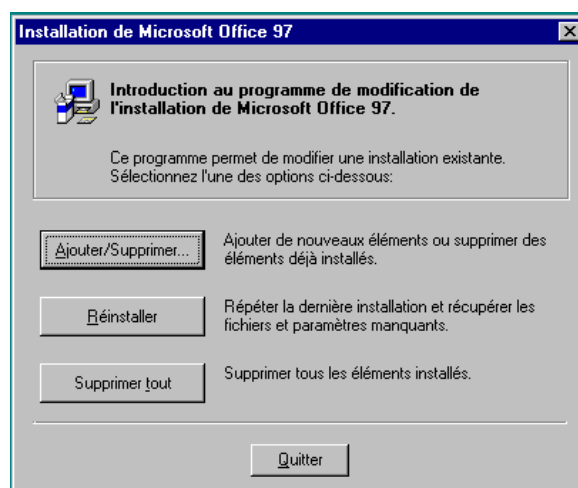
Installation des pilotes ODBC

Pour installer le support d'ODBC et le pilote du serveur SQL de Microsoft ou le pilote d'Oracle sur votre PC, utiliser le CD-ROM d'installation de Microsoft Office.

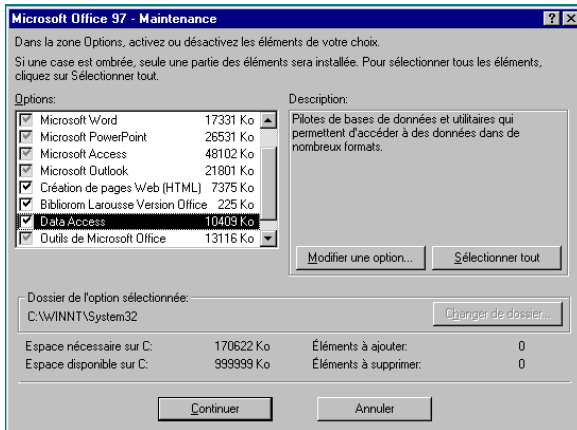
Une installation typique de Microsoft Office n'installe pas le support ODBC.

Installation de pilotes de base de données intégrés

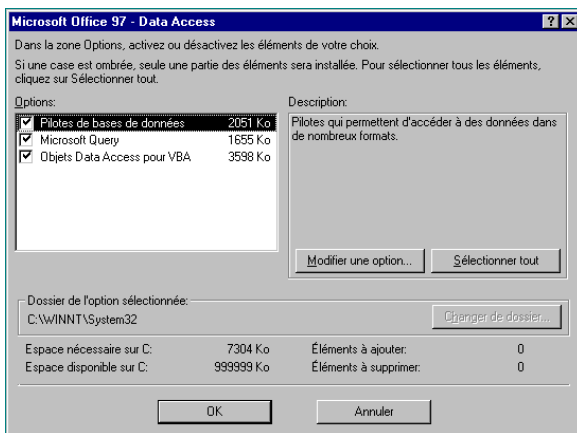
1- Relancer le programme d'installation de Microsoft Office, puis cliquer sur Ajouter/Supprimer dans la fenêtre ci-dessous.



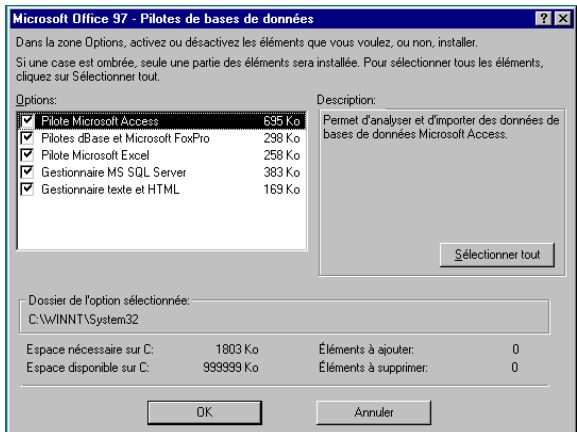
2- Sélectionner Data Access, puis cliquer sur Modifier une option dans la fenêtre Maintenance.



3- Sélectionner Pilotes de bases de données, puis cliquer sur Modifier une option.



4- Sélectionner les pilotes à installer, puis double-cliquer sur OK.



5- Cliquer sur Continuer, puis suivre les instructions des boîtes de dialogue restantes du programme d'installation.

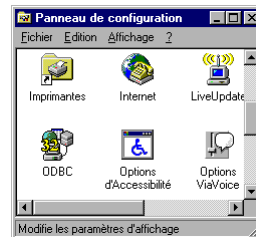
Remarque :

Les pilotes pour Lotus 1-2-3, Paradox et Microsoft Exchange/Outlook ne sont pas inclus dans le programme d'installation, mais sont disponibles dans le Complément de CD-ROM Office 97.

Installation des sources de données ODBC

Lors de l'exécution du programme d'installation du pilote d'ODBC du serveur SQL de Microsoft, du pilote d'Oracle, ou tout autre pilote d'ODBC, il est possible d'installer de nouvelles sources de données ou de modifier les sources de données existantes.

Pour les utilisateurs de Microsoft Windows 95 ou NT, cliquer sur le bouton Démarrer, pointer sur Paramètres puis cliquer sur Panneau de configuration.

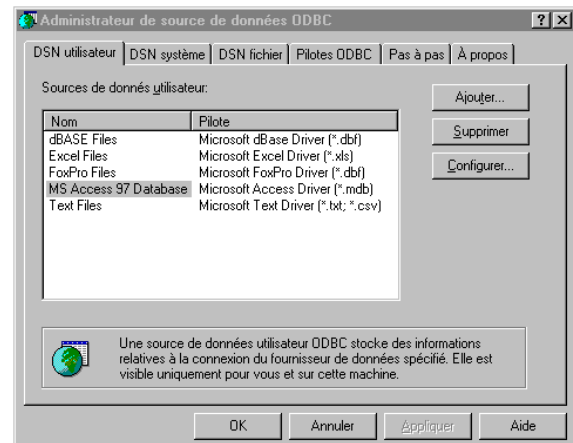


Dans la fenêtre Panneau de configuration, double cliquer sur l'icône "ODBC" pour accéder à l'Administrateur de source de données ODBC (bases de données).

Le gestionnaire ODBC est utilisé pour spécifier les sources de données.

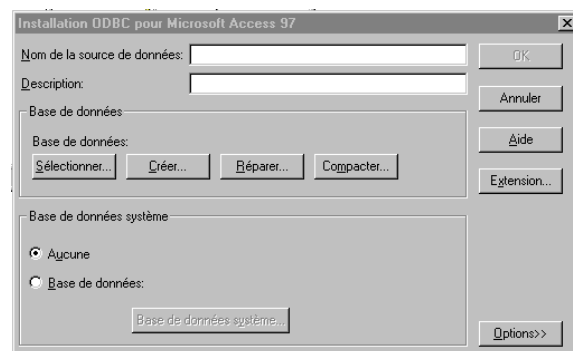
Source de données : "Data Source".

Une source de données utilisateur ODBC stocke des informations relatives à la connexion du fournisseur de données spécifié. Elle est visible uniquement pour l'utilisateur et sur sa machine.



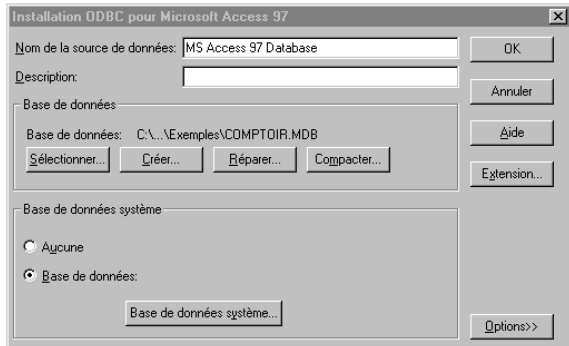
L'Administrateur ODBC offre le choix de source de données (DSN) utilisateur, l'ajout, la suppression et la configuration d'une nouvelle source de données.

Pour définir une nouvelle source de données pour le pilote actuellement installé, cliquer sur Ajouter.



Pour modifier la définition d'une source de données existante, cliquez sur un nom dans la liste des Sources de données, puis cliquez sur Installer.

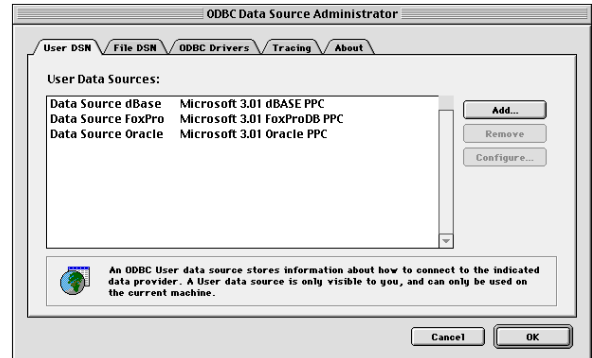
Le bouton Configurer permet de choisir la version de la base de données et de sélectionner un répertoire (dossier) ou d'utiliser le répertoire en cours par défaut.



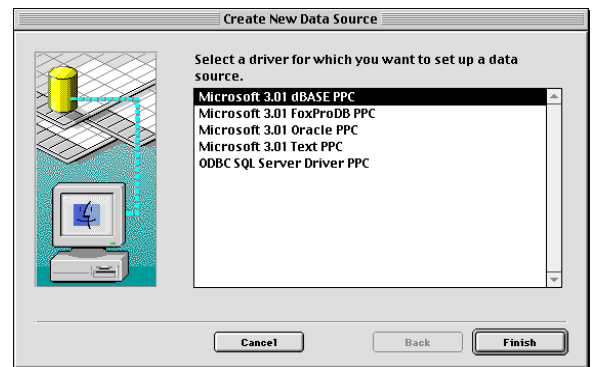
L'onglet ODBC drivers affiche les pilotes ODBC disponibles.

Dans l'onglet User DSN (nom du source de données utilisateur), on peut ajouter, configurer ou supprimer des sources de donnée.

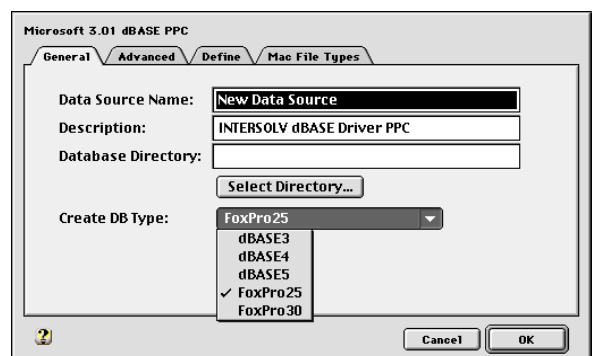
Pour ajouter une nouvelle source de données, il suffit de cliquer sur le bouton Add



Dans la fenêtre Create New Data Source, sélectionner le pilote ODBC puis cliquer sur Finish.



Enfin, inscrire le nom de la nouvelle source de données, sélectionner le répertoire (dossier), le type et/ou la version de la base de données dans la fenêtre suivante :



Remarque:

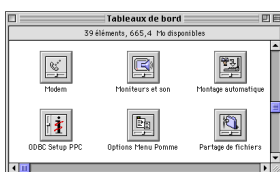
Si l'option ODBC est absente dans les Tableaux de bord du système, il est recommandé d'installer les pilotes ODBC à partir du CD-ROM Office 98.

Important : Sur MacOS le module de connexion est à acquérir auprès de la société éditrice mentionnée lors du lancement. Il n'est pas fourni en standard par Microsoft.

INSTALLATION SUR MAC OS

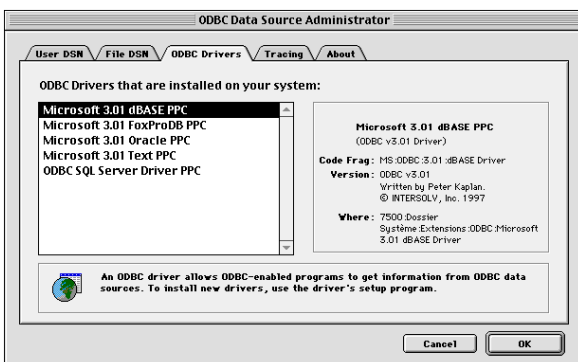
La présence dans la suite Office 98 des driver ODBC pour la plus part des bases de données présentes sur l'environnement dominant offre aux utilisateurs de Macintosh une réelle opportunité de rupture de leur isolement.

Il est à noter que la suite Office 98 ne fonctionne que sur power Mac. Nous ne saurons trop conseiller l'usage de Mac disposant d'un processeur rapide (100 MHZ et +) et largement dotés en mémoire vive (minimum de 64 MO).



Sur Mac (Power PC), dans le menu pomme, sélectionner Tableaux de bord, puis cliquer sur l'icône ODBC Setup PPC.

Une fenêtre de dialogue nommée "ODBC Data Source Administrator" sera affichée, les différents onglets de cette fenêtre permettent de choisir, de rajouter, de supprimer et de configurer les pilotes d'ODBC.

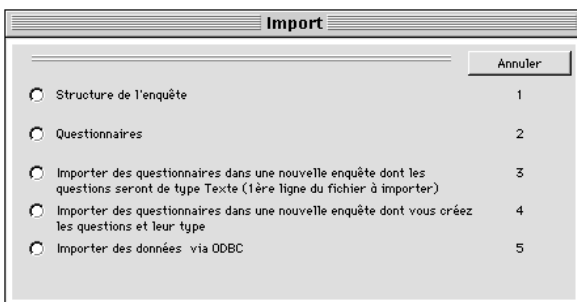


UTILISATION DU CLIENT ODBC



Sélectionner l'article Import ASCII du menu Fichier qui ouvrira la fenêtre Import.

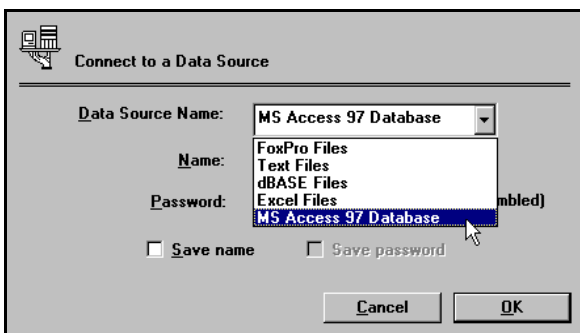
Pour Modalisa, le mode ODBC est l'un des modes d'import de données en format ASCII (texte seul).



Le cinquième choix ouvre la fenêtre "Login", celle-ci dépend de système d'exploitation sur lequel s'exécute Modalisa.

Login sur Windows

Sous un environnement PC & Compatibles, le choix du driver ODBC est possible aussi au niveau de la fenêtre de dialogue "Login".

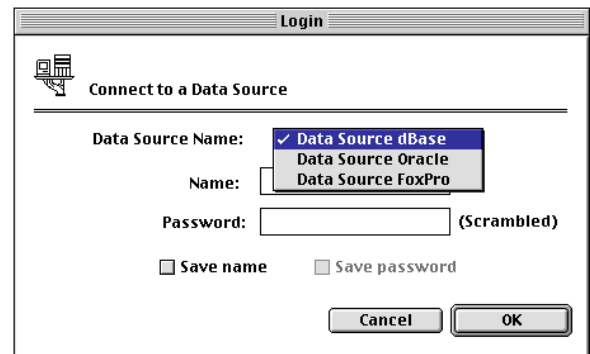


Login sur MacOS

Sur Macintosh, sélectionner le nom de la source de données (Data Source Name) ce qui correspond au pilote ODBC choisi lors de la configuration de "ODBC Setup PPC" du tableaux de bord.

Les source de données utilisateurs affichées dans l'onglet User DSN de la fenêtre ODBC Data Source

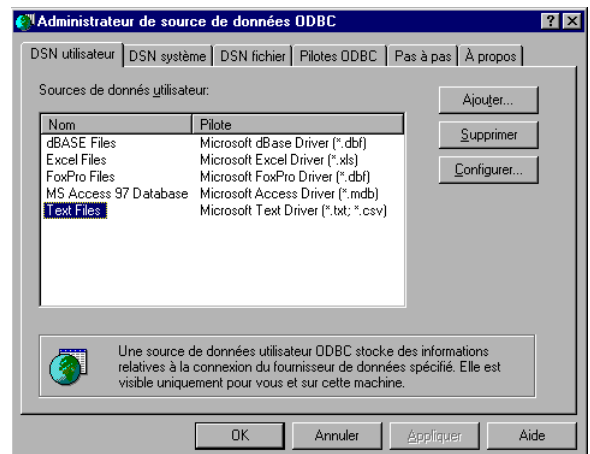
Administrator seront disponibles sur la fenêtre Login (pour plus de détails voir § Installation sur MacOS)



Exemple d'utilisation de sources de données de type Texte.

Pour importer des fichiers de type texte dans Modalisa via le module ODBC, il est indispensable de configurer sous Windows l'ODBC32 et sous MacOS l'ODBC Setup PPC.

Configuration de l'ODBC sous Windows 95 & NT

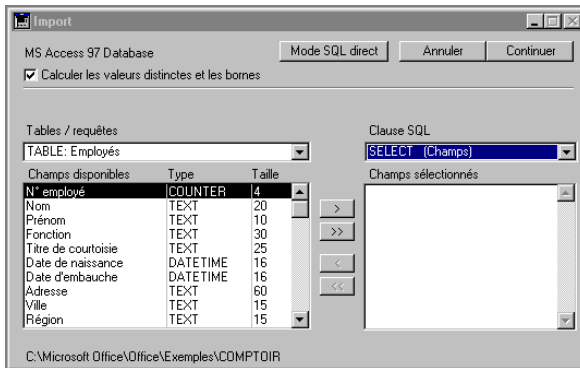


Sélectionner Text Files dans la fenêtre ci-dessus, puis cliquer sur le bouton "Configurer"

Sélection des données

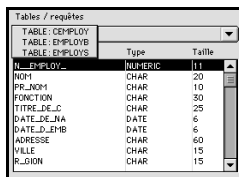
Mode SQL Assistant

Dans la fenêtre Import, Modalisa charge les tables/requêtes disponibles sur la base de données



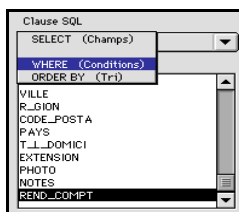
Modalisa fait un inventaire sur toutes les données (Calcul des valeurs distinctes et les bornes) des fichiers (calcul de minimum, de maximum et de l'étendue) puis montre les 100 premières fiches (enregistrements) de chaque fichier (tables / requêtes).

Un menu déroulant permet de sélectionner la table dont on désire afficher les champs.



Modalisa affiche les champs disponibles, leur type ainsi que leur taille.

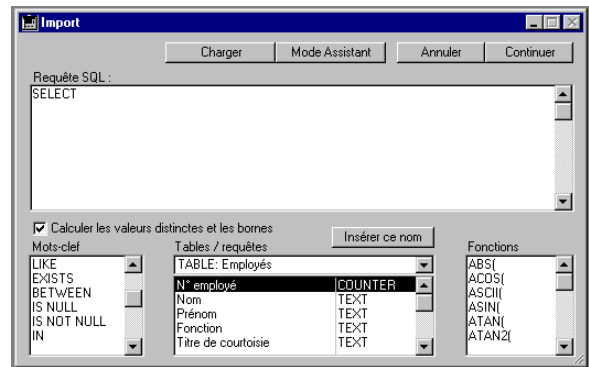
Il est alors possible de sélectionner un ou plusieurs de ces champs, d'établir des conditions de sélection ainsi que d'opérer des tris.



Ceci peut être réalisé sur des champs de diverses Tables/requêtes disponibles en utilisant les trois clauses SQL du mode assisté : SELECT (Champs), WHERE (Conditions), ORDER BY (Tri).

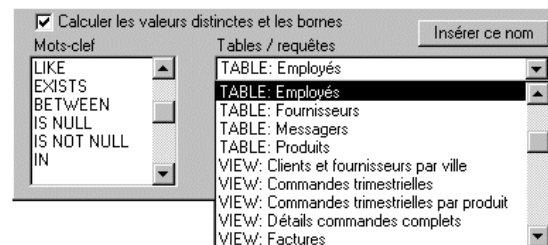
Mode SQL direct

Le bouton Mode SQL direct permet d'activer une nouvelle fenêtre qui autorise l'écriture de requêtes en langage SQL.



Pour écrire une requête SQL, l'utilisateur peut se servir des trois listes : Mots-clé, Fonctions, Tables/Requêtes.

Sélectionner un mot-clé, une fonction et/ou un champ, utiliser le double clic pour l'inscrire dans la zone de rédaction de la requête SQL.



Utiliser le bouton Insérer ce nom pour inscrire dans le cadre Requête SQL le nom du fichier actif dans la liste Table/Requête.

Ainsi pour les utilisateurs qui maîtrisent le langage SQL, il leurs est possible d'écrire des requêtes simples comme pour visualiser les données d'une table : `SELECT * FROM Nom_Table`, ou compliquées comme pour l'exemple qui suit (liste des commandes trimestrielles de la BD comptoir de Microsoft Access)

```
SELECT DISTINCTROW Catégories.[Code catégorie],  
Catégories.[Nom de catégorie], Produits.[Nom du produit],  
Sum([Détails commandes complets].PrixTotal) AS  
VentesProduit
```

```
FROM Catégories INNER JOIN (Produits INNER JOIN  
(Commandes INNER JOIN [Détails commandes complets] ON  
Commandes.[N° commande] = [Détails commandes  
complets].[N° commande]) ON Produits.[Réf produit] =  
[Détails commandes complets].[Réf produit]) ON  
Catégories.[Code catégorie] = Produits.[Code catégorie]
```

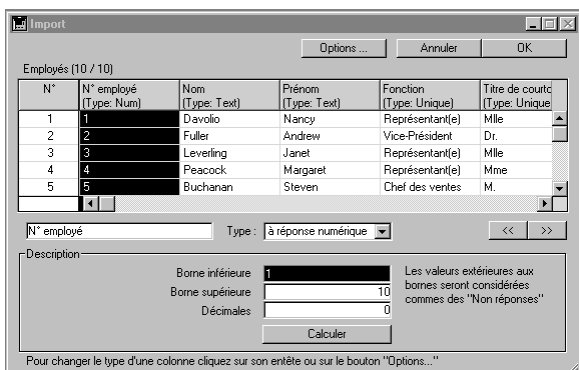
```
WHERE (((Commandes.[Date commande]) Between #1/1/95#  
And #12/31/95#))
```

```
GROUP BY Catégories.[Code catégorie], Catégories.[Nom de  
catégorie], Produits.[Nom du produit]
```

```
ORDER BY Produits.[Nom du produit];
```

TYPAGE DES DONNEES

Le bouton Continuer exécute la requête SQL.



Modalisa permet alors d'adapter le type des données présentes dans la base aux besoins des traitements et analyses ultérieures.

On peut ainsi décider qu'un champ texte devienne dans Modalisa une variable numérique, qu'un champ texte devienne une question fermée...

Changement du nom des variables

L'utilisateur peut renommer les champs des tables de la base de données.

En effet il est courant que les informaticiens donnent des noms abrégés à leurs champs, ceux-ci étant réservés à un usage strictement interne.

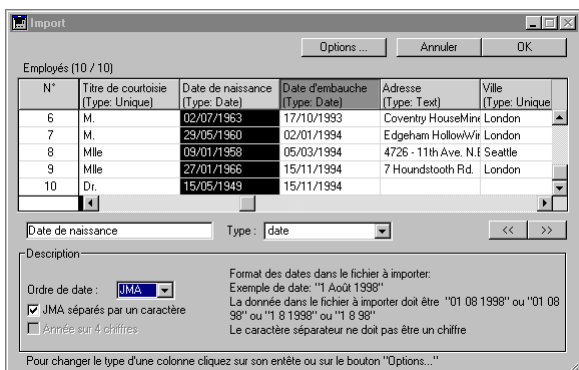
Par ailleurs des traditions du milieu font que l'on tend à choisir les noms les plus courts tout en restant le plus significatifs possibles.

Toutefois ces noms ne sont pas utilisables en tant que tels pour la réalisation d'une étude.

Les destinataires finaux d'une étude attendent des intitulés limpides. Modalisa permet donc de renommer CP_Clt en Code Postal du client.

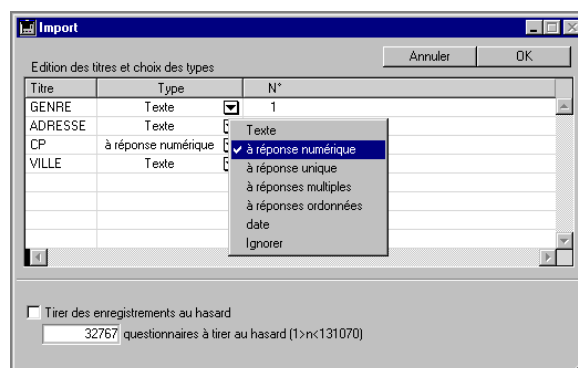
Changement de type de variables

Pour changer le type d'une colonne, cliquer sur l'entête de la colonne, sur l'option "type" ou sur le bouton "option".



On trouvera sur la zone Description un certain nombre d'annotations dynamiques qui serviront de repère à l'utilisateur débutant.

Le bouton Option active une fenêtre qui présente les différentes variables et leurs types.



Un simple clic sur le bouton à droite de la cellule type, déroule un menu local proposant le choix entre les types de variable utilisables.

Le recodage des variables (questions) lors de l'import des données via ODBC est d'une grande utilité, qui peut rendre l'enquête exploitable de suite.

En effet, on peut parler de l'anti-chambre de l'analyse de données tout simplement.

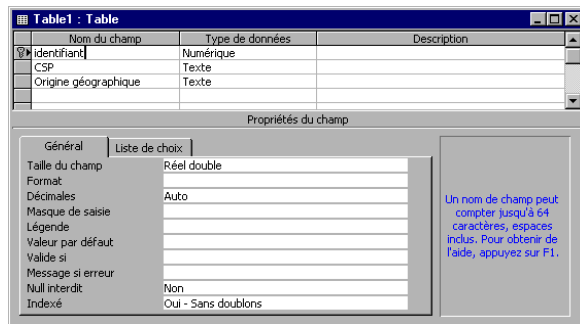
Ainsi les questions de type texte seront transformées en question à réponse unique ou multiple ou même numérique.

Exemple 1 : jointure de variables et typage de données

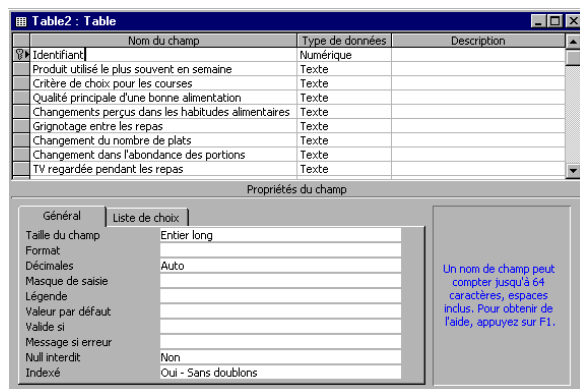
Sur la base de l'enquête habitudes alimentaires, on a construit deux sous enquêtes (table1, table2) ayant des structures (questions) différentes et des individus (questionnaires) en commun.

Le code identifiant des deux tables est unique.

Structure de la Table1



Structure de la table 2



L'écriture d'une requête en mode SQL direct de Modalisa permet de fusionner les deux sous enquêtes (tables).

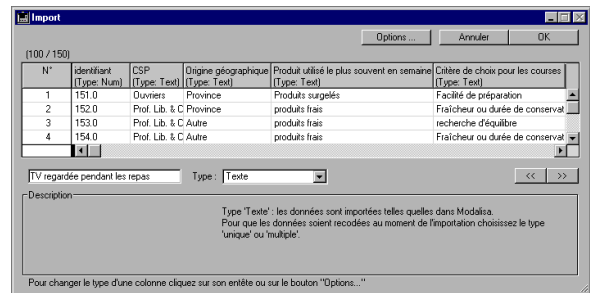
```
SELECT Table1.identifiant, Table1.CSP,
Table1.[Origine géographique], Table2.[Produit
utilisé le plus souvent en semaine], Table2.[Critère
de choix pour les courses], Table2.[Qualité principale
d'une bonne alimentation], Table2.[Changements
perçus dans les habitudes alimentaires],
Table2.[Grignotage entre les repas],
Table2.[Changement du nombre de plats],
Table2.[Changement dans l'abondance des portions],
Table2.[TV regardée pendant les repas]
```

FROM Table1, Table2

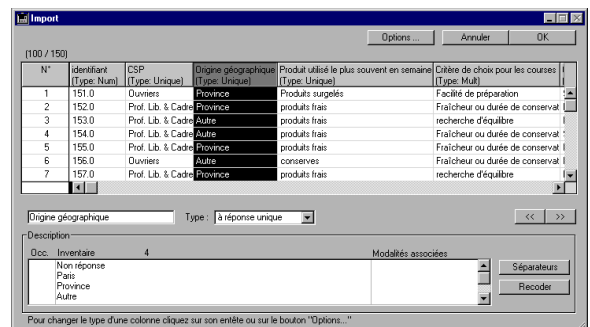
WHERE Table1.identifiant = Table2.Identifiant

ORDER BY Table1.identifiant

Lors de l'import, et avant l'intégration de l'enquête, Modalisa offre la possibilité de travailler sur la structure (typage, recodage, changement de libellés...).



Dans cet exemple, l'utilisateur peut constater que toutes les variables hormis l'identifiant sont de type texte. Si on intègre l'enquête dans Modalisa dans l'état, il n'est possible de réaliser des traitements sans recoder les variables.



Un des avantages de cette fusion via ODBC est la mise à jour. Toute modification des données (questionnaires) sur la base de données est susceptible d'être réactualisée sous Modalisa par un simple clic sur le bouton Maj ODBC de l'écran base.

Exemple 2 : jointure de variables et intersection de 2 tables

Supposons qu'on a deux tables (fichiers) :

PParis (le personnel résidant à Paris).

code	Nom	Prénom	sexe	Age
a1	Van Impe	J. Luc	Homme	50
a2	Ranjard	Sophie	Femme	45
a256	Kerkeni	Ridha	Homme	35

Personnels (fichier principal du personnels)

N° employé	Nom	Prénom	Fonction	Titre de	Date de nais.	Adresse	Ville	Notes
a15	Davolio	Nancy	Représentant	Mlle	08/12/48	507 - 20th Ave. E	Seattle	Diplôme universitaire
a16	Fuller	Andrew	Vice-Président	Dr.	18/02/52	308 W. Capital M	Tacoma	Diplôme universitaire
a17	Leverling	Janet	Représentant	Mlle	30/08/63	722 Moss Bay E	Kirkland	Diplôme d'ingénieur
a18	Peacock	Margaret	Représentant	Mme	19/09/37	4110 Old Redmond	Fremont	Diplôme universitaire
a19	Buchanan	Steven	Chef des ventes	M	04/03/55	14 Garrett Hill	London	Diplôme de l'Univer
a20	Sugama	Michael	Représentant	M	02/07/63	Coventry House	London	Diplôme de l'Univer
a21	King	Robert	Représentant	M	29/05/60	Edgeham Hollow	London	A beaucoup voyagé
a22	Callahan	Laura	Assistante co	Mlle	08/01/58	4726 - 11th Ave. E	Seattle	Diplôme universitaire
a23	Oodsworth	Anne	Représentant	Mlle	27/01/66	7 Houndstooth	London	Diplôme universitaire
a1	Van Impe	J. Luc	Vice-Président	M	25/10/55	3 rue des Montils	Paris	Diplôme universitaire
a2	Ranjard	Sophie	Vice-Président	Mme	31/12/51	4 rue des Montils	Paris	Diplôme universitaire
a256	Kerkeni	Ridha	Technicien	M	02/10/60	5 rue des Montils	Paris	Diplôme universitaire

Le code identifiant de la table Personnels est le N° employé, celui de la table PParis est code.

Pour faire une jointure des deux tables sur le code identifiant afin d'avoir le restant des variables du fichier principal (Personnels) concernant le personnels résidant à Paris :

Modalisa, avec son module Import via ODBC permet d'écrire en mode SQL direct la requête suivante :

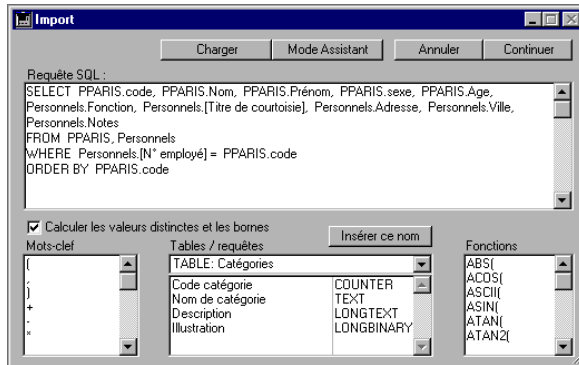
```
SELECT PPARIS.code, PPARIS.Nom,
PPARIS.Prénom, PPARIS.sexe, PPARIS.Age,
```

Personnels.Fonction, Personnels.[Titre de courtoisie],
Personnels.Adresse, Personnels.Ville,
Personnels.Notes

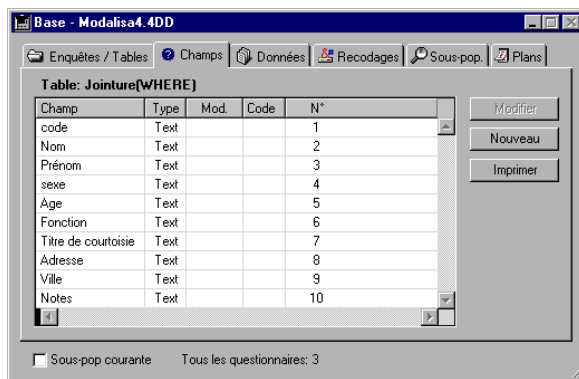
FROM PPARIS, Personnels

WHERE Personnels.[N° employé] = PPARIS.code

ORDER BY PPARIS.code



On peut distinguer que les champs de l'enquête/Table intitulée Jointure(WHERE) sont le résultat d'une concaténation des deux tables mentionnées ci-dessus.



La jointure des tables et l'usage de la clause WHERE pour les conditions sont d'une grande importance dans les divers travaux d'analyse et de traitement d'enquêtes.

Prenons le cas d'une enquête ayant plus de 300 questions (variables).

L'utilisateur ne peut introduire les données de l'enquête sur une seule feuille Excel qui ne dispose que de 256 colonnes.

La solution est de scinder les questions en deux feuilles, en rajoutant un code identifiant sur chacune des deux feuilles qui permettra par la suite de faire la jointure et de traiter l'enquête sous Modalisa.

Le mode SQL direct de Modalisa permet la liaison de plusieurs tables. L'opération de jointure réalise une liaison entre deux tables, en se basant sur l'égalité entre l'une des variables de chaque table.

Exemple 3 (union de 2 tables)

L'opérateur UNION effectue l'union des résultats de deux requêtes SQL.

Si l'opérateur UNION est suivi du mot clé ALL, les doublons (enregistrement égaux) seront conservés.

Le type des variables sélectionnées des deux tables doit être compatible. Si les tables unies sont des sous-ensembles de la même table, l'emploi de l'opérateur UNION est similaire à celui du OR placé entre deux conditions (WHERE).

On dispose de la table clients (100 individus) et la table fournisseurs (20 individus) ayant des champs compatibles dans des structures différentes. Si on veut connaître les clients et les fournisseurs par ville, il suffit d'écrire la requête SQL suivante :

SELECT Clients.Ville, Clients.Société,
Clients.Contact, Clients.Notes

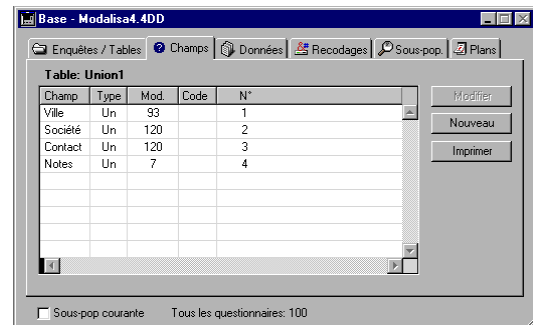
FROM Clients

UNION SELECT Fournisseurs.Ville,
Fournisseurs.Société, Fournisseurs.Contact,
Fournisseurs.Notes

FROM Fournisseurs

ORDER BY Ville, Société;

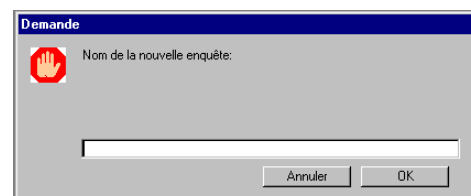
Le résultat de cette union est la table nommée Union1 constitué de l'extrapolation d'enregistrements des deux tables (clients & fournisseurs). Le nombre de tuples de la table Union1 est égale à 100 comme le mentionne l'écran base ci-dessous.



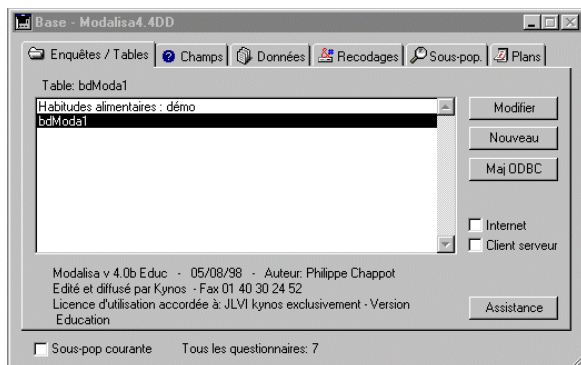
LES SEPARATEURS

INTEGRATION DANS MODALISA

Une fois le typage des données fini, cliquer sur le bouton "OK", puis saisir un nom d'enquête avant de valider par "OK".



MISE A JOUR ODBC

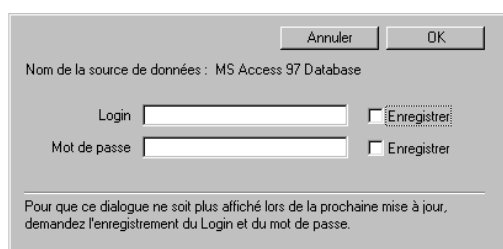


Le bouton Maj ODBC a pour fonction la mise à jour des données intégrées pour la réalisation d'analyses dans Modalisa.

La Maj ODBC, ne sous entend pas une mise à jour conventionnelle à savoir une modification, une suppression ou une création de tuples dans une base de données.

La mise à jour pour Modalisa correspond à une nouvelle interrogation de la base de données permettant une réactualisation des données.

Il est donc possible de mettre à jour les fiches intégrées dans Modalisa.



Par exemple : sur une base de données Access, la création, la suppression ou la modification d'une ou de plusieurs fiches dans le fichier clients, est ainsi prise en compte dans Modalisa.

La mise à jour ODBC concerne uniquement les données, dans le cas d'un changement de la structure des tables de la base de données, l'utilisateur doit créer une nouvelle structure d'enquête correspondante dans Modalisa.

SELECTION ALEATOIRE

On sait qu'il n'est pas nécessaire de traiter les 3 millions d'individus présents sur une base pour réaliser des analyses statistiques probantes.

Ainsi le principe du tirage aléatoire d'un nombre significatif d'individus permet de travailler sur une portion des données de la base.

La fonction de Sélection aléatoire permet de choisir parmi tous les "individus" présent dans la base un nombre compris entre 1 et le double du nombre réel.

On n'expliquera pas dans cette documentation purement procédurale les principes de la sélection

aléatoire. Ceci d'autant qu'il existe des ouvrages qui font foi sur le chapitre.

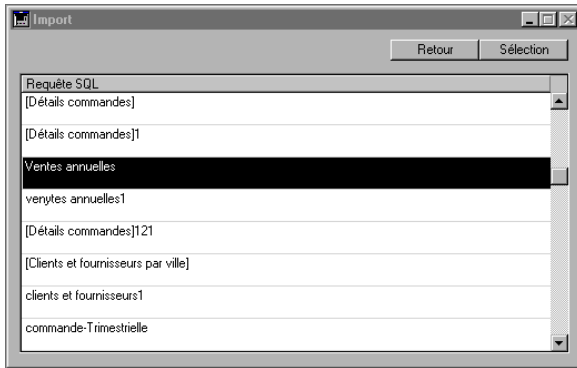
Cette fonctionnalité permet de réaliser des sélections faibles dans une base importante puis de procéder à des analyses fondées.

REUTILISATION D'UNE REQUETE

Modalisa enregistre les requêtes SQL qui ont servi à réaliser des enquêtes. Dès lors, il est possible de réutiliser ces requêtes en mode SQL direct.

Pour accéder à ces requêtes, Cliquer sur le bouton Charger du Mode SQL direct, sélectionner un nom de requête dans la liste active puis cliquer sur le bouton Selection.

La requête apparaîtra dans la fenêtre Requête SQL du mode direct.



Il est alors possible d'effectuer des modifications afin de rectifier d'éventuels erreurs ou d'utiliser la structure de la requêtes pour d'autres objectifs.

En effet la commande SELECT peut prendre un aspect extrêmement compliqué, comportant des attributs appartenant à plusieurs tables, des fonctions, des requêtes imbriquées...

Le langage SQL permet la liaison de plusieurs tables en utilisant 3 techniques différentes : la jointure, les requêtes imbriquées et les opérateurs ensemblistes.

Un élément d'une expression logique peut prendre plusieurs formes:

=, <>, <, >, <=, >=

BETWEEN

IN, MATCH

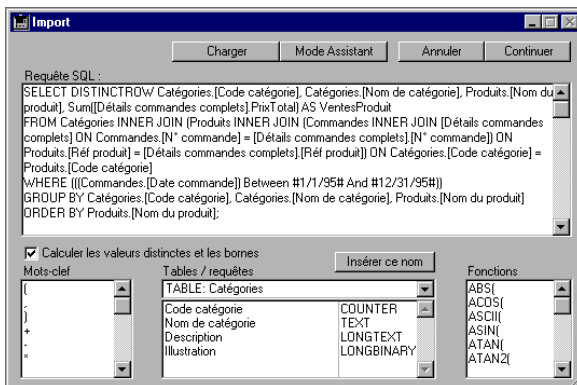
LIKE

IS NULL

ALL, ANY

EXIST

UNIQUE



L'opération la plus courante dans une base de données est la recherche des informations par une sélection selon différents critères. Cette opération est assurée par la commande SELECT.

A la commande SELECT (sélectionner les variables) vient se greffer la clause FROM pour donner l'information : dans quelle(s) table(s) les données sont stockées ?

Ces deux clauses sont généralement suivies d'une clause WHERE qui définit la condition à remplir par les valeurs d'un enregistrement (tuple).

La clause WHERE permet de spécifier sur quels variables se fait la jointure, et de définir un critère de sélection combiné avec la condition de jointure par un opérateur AND.

La clause WHERE permet de spécifier un critère de sélection, appelé prédicat. Si le prédicat est vérifié par l'enregistrement (tuple), celui-ci fera partie du résultat. On réalise ainsi une opération relationnelle de sélection. Le prédicat est une expression logique composée d'une suite de conditions (prenant chacune la valeur "vrai" ou "faux"), combinées entre elles par les opérateurs logiques AND, OR ou NOT.

NOTIONS SUR SQL

SQL : System Query Language (traduction littérale: Langage de requête de système). C'est un langage d'interrogation structuré de base de données développé par IBM et devenu un standard de fait.

Aperçu historique

La norme SQL-92 ou SQL-2 - produit de l'effort des comités de normalisation ANSI et ISO - devenait une véritable référence pour le monde informatique.

Malheureusement, le respect de cette norme impose aux constructeurs d'énormes modifications au niveau de leurs SGBD. La conformité de ceux-ci au nouveau texte nécessite de nombreuses années de travail. Ce constat a poussé les constructeurs à rapprocher rapidement leurs divers dialectes sans attendre la conformité à la nouvelle norme (SQL-2).

Un groupe de quarante constructeurs s'est réunis sous le nom de SAG (SQL Access Group).

Le SAG travaille sur la normalisation de l'accès aux bases de données, en particulier via le langage SQL. Plutôt que de modifier chaque dialecte pour le rendre conforme à un standard, il s'agit de développer un langage SQL commun, et pour chaque dialecte, un programme de conversion du dialecte en SQL commun et inversement.

En 1993, Microsoft lança ODBC, ensemble de primitives basées sur le travail du SAG, rapidement suivi par Borland (en 1994) avec l'IDAPI (Integrated Database Application Programming Interface). Ces spécifications consistent en une bibliothèque de fonctions SQL intégrables dans un programme hôte. Très rapidement le succès de ces bibliothèques a été fulgurant, surtout dans le contexte client/serveur où il s'agit d'accéder aux données de plusieurs serveurs présentant des dialectes SQL différents [MAR. 94].

SQL

SQL est utilisé pour extraire, mettre à jour et gérer des données informatiques. Les requêtes écrites avec SQL sont équivalentes, dans un langage classique comme Xbase, C, COBOL,... à plusieurs dizaine de lignes d'instructions.

Une requête SQL est une requête créée via l'utilisation d'une instruction SQL. Parmi les requêtes SQL spécifiques, on trouve la requête Union, la requête SQL direct, la requête Définition des données et la sous-requête.

Une instruction/chaîne SQL est une expression qui définit une commande SQL, telle que SELECT, UPDATE ou DELETE, et contient des clause telles que WHERE, ORDERBY... Les instructions SQL sont généralement utilisées dans les requêtes et les fonctions de regroupement.

Avec la commande d'interrogation du langage SQL SELECT, l'utilisateur peut obtenir l'ensemble des tuples qui satisfont un critère de qualification. Ainsi on peut manipuler divers tuples provenant de tables différentes.

Modalisa 4 n'autorise pas les commandes de modification à savoir l'ajout (INSERT), la mise à jour (UPDATE) et la suppression (DELETE).

La commande SELECT

Une requête SQL doit toujours commencer par le mot-clef **SELECT**. C'est la fonction d'interrogation qui permet de voir tout ou une partie d'un fichier.

SELECT, est la véritable clé de voûte du LMD (langage de manipulation de données) de SQL.

La forme simplifiée de la commande SELECT appelé aussi SELECT de base permet d'effectuer les opérations de projection, de sélection et la combinaison des deux.

La syntaxe de ce SELECT de base appelé aussi SFW (SELECT FROM WHERE) est la suivante :

SELECT [ALL | DISTINCT] clause_de_sélection

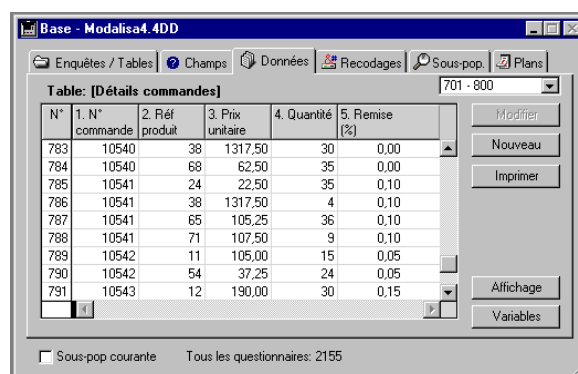
FROM nom_table

[**WHERE** condition]

Exemple : **SELECT * FROM** nom_table

L'* est la fonction SQL qui, employée après SELECT, provoque la visualisation du fichier tout entier.

Exemple : dans la base de données Comptoir de Microsoft Access, la Table Détails commandes contient les champs suivants : N° Commande, Réf produit, Prix Unitaire, Quantité et Remise.



The screenshot shows a window titled 'Base - Modalisa4.4DD'. The main area displays a table with the following data:

N°	1. N° commande	2. Réf produit	3. Prix unitaire	4. Quantité	5. Remise (%)
783	10540	38	1317,50	30	0,00
784	10540	68	62,50	35	0,00
785	10541	24	22,50	35	0,10
786	10541	38	1317,50	4	0,10
787	10541	65	105,25	36	0,10
788	10541	71	107,50	9	0,10
789	10542	11	105,00	15	0,05
790	10542	54	37,25	24	0,05
791	10543	12	190,00	30	0,15

At the bottom of the window, it indicates 'Sous-pop courante' and 'Tous les questionnaires: 2155'.

La table Détails commandes est constituée de 2155 tuples (enregistrements). Pour visualiser les quatre premières colonnes uniquement, il suffit d'écrire en mode SQL direct de Modalisa :

SELECT N° commande, Réf produit, Prix unitaire, Quantité **FROM** Détails Commandes

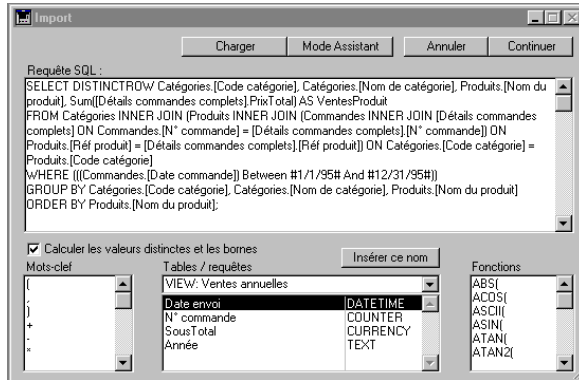
Pour connaître le produit le plus coûteux, utiliser la fonction **MAX**.

SELECT MAX (Prix unitaire) **FROM** Détails commandes

Pour visualiser tout les champs du fichier détails commandes :

SELECT * FROM Détails commandes

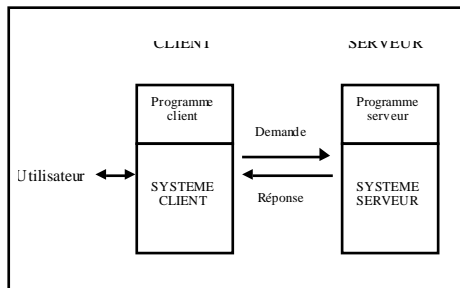
Remarque : la création ou la suppression de fichiers (tables) de bases de données ainsi que la création de nouvelles variables ne peut se faire sous Modalisa.



Dans la fenêtre ci-dessus, la requête SQL, écrite sous le mode SQL direct de Modalisa, permet à l'utilisateur de visualiser les ventes annuelles des différents produits de la base de données Comptoir d'Access.

L'ARCHITECTURE CLIENT-SERVEUR

Le mot client désigne le système sur lequel s'exécute le programme client. Le mot serveur indique le système sur lequel fonctionne le programme serveur.



L'architecture client-serveur

Dans ce type d'architecture, l'utilisateur dialogue avec l'application à partir du client. Il transmet des requêtes au serveur qui déclenche le programme serveur, et renvoie les résultats au client.

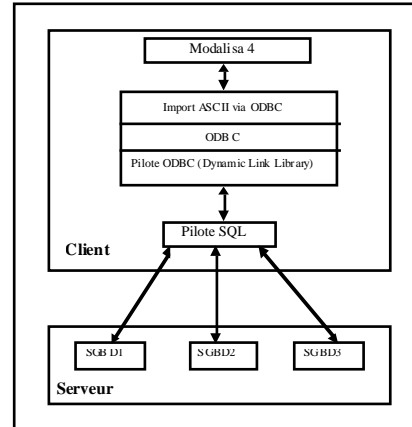
Les serveurs de base de données offrent des services classiques d'accès à une ou des bases de données : recherches, mise à jour, confidentialité. Plusieurs clients peuvent se partager le même serveur, celui-ci est alors capable de gérer l'atomicité des transactions, les accès concurrents et la reprise après panne.

En général, chaque éditeur de SGBD fournit un moyen de construire des applications client-serveur en utilisant une bibliothèque de fonctions appelée API. Les programmes clients adressent leurs requêtes au serveur via des fonctions contenues dans l'API.

OCI d'Oracle, DB-Lib de Sybase et SQL/Services de Rdb sont des API propriétaires parce que chaque API

permet uniquement d'accéder à la base pour laquelle elle a été développée par l'éditeur du SGBD.

Les systèmes ouverts, de nos jours, nécessitent de développer des applications capables d'accéder à des données réparties entre plusieurs SGBD différents. C'était le but que s'est fixé le SAG, à savoir, proposer une interface unique d'accès à des SGBD différents. Les spécifications du SAG étaient à la base d'ODBC de Microsoft et d'IDAPI de BORLAND, Novell et IBM.



L'architecture ODBC de Microsoft

ODBC est actuellement universellement admis, sa disponibilité ainsi que son accessibilité à différents SGBD ont incité l'auteur de Modalisa à l'intégrer dans la version 4 du logiciel.

NOTIONS SUR L'ODBC

ODBC (**OPEN DATABASE CONNECTIVITY**) est une bibliothèque d'une cinquantaine de fonctions qui peuvent être appelées à partir d'applications s'exécutant sous Windows ou Macintosh. Ce protocole standard permet aux applications d'accéder à divers fichiers ou serveurs de base de données externes.

ODBC est une API normalisée, proposée par Microsoft depuis 1993, permettant l'accès à tous les SGBD. Il utilise le langage SQL comme standard d'accès.

API : (Application Programming Interface) Interface de programmation d'application. Ensemble de routines disponibles dans une application pouvant être utilisés par les programmeurs de logiciels lors de la conception d'une application.

La norme ODBC est maintenant universellement admise. La maîtrise d'une seule bibliothèque permet l'utilisation de n'importe quel SGBD, même non relationnel. En effet, il suffit qu'il existe un pilote (driver) ODBC qui transforme les requêtes SQL en instructions compréhensibles par les SGBD comme Oracle, Ingres, Sybase... Il existe également des pilotes pour des pseudo-SGBD comme Microsoft Access ou des tableurs tel que Excel ou encore des gestionnaires de fichiers évolués comme dBase.

Un pilote ODBC est une bibliothèque de liaison dynamique (DLL: **D**ynamic-**l**ink **l**ibrary) qui permet de connecter une source de données ODBC spécifique à une autre application (client). Il est possible d'utiliser le pilote ODBC fourni avec l'application, opter pour un autre pilote, ou encore écrire son propre pilote.

Cette approche est particulièrement utile pour les environnements client-serveur car le client peut envoyer plusieurs requêtes simultanément et gérer les résultats d'une façon asynchrone.

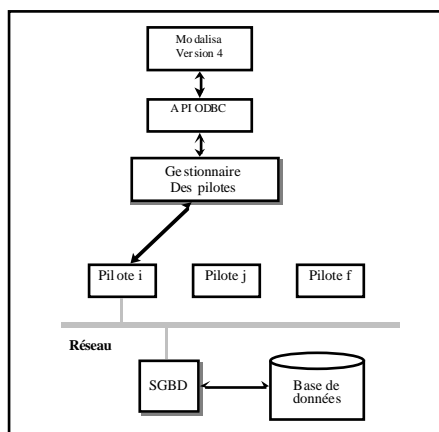
L'ODBC permet aux développeurs d'écrire des applications indépendantes du gestionnaire de données.

les mots de passe utilisateur et serveur, et éventuellement des informations relatives au verrouillage. Toutefois, il existe une fonction intégrée de gestion des connexions qui place de tels informations en antémémoire et les réutilise lorsque cela s'avère nécessaire.

Principes de fonctionnement de l'ODBC

Le développeur dispose d'un kit de développement de logiciels (SDK pour Software Development Kit) grâce auquel il écrit ses applications utilisant les fonctions de la bibliothèque ODBC.

Chaque instruction SQL possède un environnement qui spécifie la source (DATA SOURCE) des données, c'est-à-dire le gestionnaire cible de l'instruction. Celle-ci effectue un appel à un service de la couche DRIVER MANAGER (gestionnaire des pilotes ODBC) qui détermine quel pilote spécifique est nécessaire pour exécuter la requête. Le pilote approprié traduit la requête ODBC en une requête compatible avec le SGBD cible. La requête traduite est envoyée via le réseau au gestionnaire qui l'exécute. Le résultat de la requête fait le chemin en sens inverse et est traduit par le pilote approprié pour être compréhensible par l'application locale [MAR.94].



Fonctionnement de l'API ODBC

Connexion aux données ODBC

L'établissement d'une connexion à un serveur exige du temps et de la mémoire tant au niveau du client qu'à celui de serveur.

Pour établir une connexion à une source de données ODBC, il faut fournir les emplacements des fichiers,

EXEMPLE D'UTILISATION D' ODBC

Dans sa version 4, Modalisa importe via son module ODBC aussi bien les fichiers provenant d'autres applications Macintosh que ceux d'applications non Macintosh comme dBase, Access, Oracle...

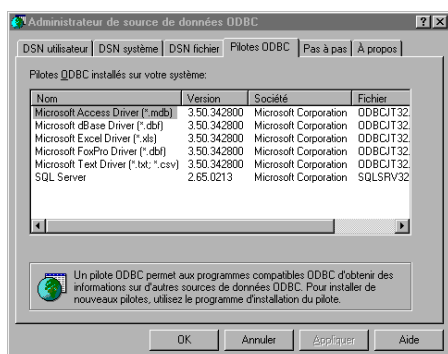
Il est important de rappeler que l'importation de données, via ODBC, est possible sur des différents type d'environnement (ordinateur personnel, réseau local et réseau distant).

Avec la version 4 de Modalisa, les conditions préalables à l'importation des fichiers se sont vue amincies, en effet, il n'est plus nécessaire de savoir si le fichier a été enregistré en format ASCII (texte). Le module ODBC permet l'importation des fichiers avec des extensions telles que .xls, .dbf, .mdb, .txt...

Recueil de données dans ACCESS

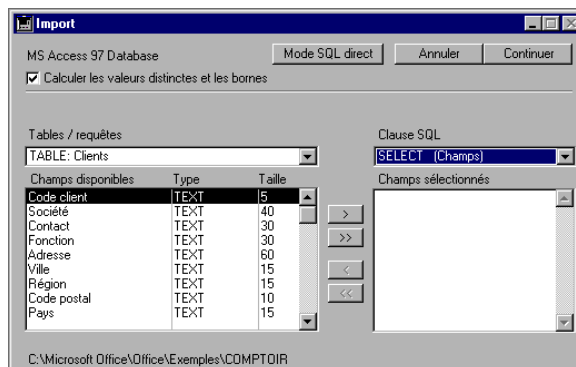
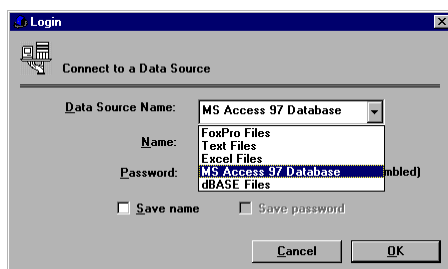
Cet exemple montre comment importer dans Modalisa les fichiers de la base de données Comptoir fournie comme exemple avec l'application Microsoft Access.

Pour commencer, vérifier que le pilote Microsoft Access (*.mdb) existe dans la l'onglet Pilote ODBC de la fenêtre Administrateur de source de données ODBC.



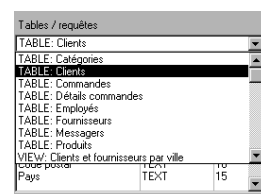
Si le pilote est absent, il suffit de l'installer (voir § Installation d'ODBC sur Windows).

Dans Modalisa, sélectionner l'article Import ASCII du menu Fichier puis pointer sur Importer des données via ODBC de la fenêtre Import et enfin mentionner le nom de la source de données correspondant au pilote MS Access 97 Database dans la fenêtre Login (pour plus de détails voir § Importation des données via ODBC).



Modalisa charge tous les fichiers ayant l'extension *.mdb (base de données de Microsoft Access).

Un simple clique sur l'option Tables/Requêtes fait apparaître la liste de tous les fichiers ayant l'extension *.mdb de la base de données Comptoir de MS Access

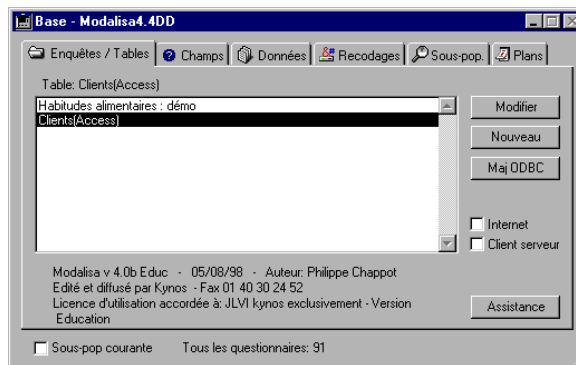


Lorsqu'un fichier est sélectionné, Modalisa visualise dans le cadre spécifique le type, la taille et l'intitulé des champs disponibles correspondant à ce fichier.

A ce stade, on peut utiliser le mode assistant ou le mode SQL direct (§ Sélection des données) pour sélectionner des champs provenant d'un ou de plusieurs tables.

Dans cet exemple, sélectionner la table CLIENTS

Une fois le typage des données fini, nommer la nouvelle enquête CLIENTS(Access) et valider par "OK".



Modalisa importe les données et fait apparaître la nouvelle enquête sur l'écran Base-Modalisa4.4DD.

Dans la fenêtre base ci-dessus on distingue l'enquête CLIENTS(Access).

Cliquer sur l'onglet Champs pour visualiser les différents champs de la table CLIENTS(Access). On distingue les intitulés des champs, le Type, le nombre de modalités, le code et un n° d'ordre. L'utilisateur peut imprimer le contenu de cette fenêtre pare contre il ne lui est pas possible d'activer le bouton Nouveau pour la simple raison que la table est liée à une requête SQL(ODBC).

Base - Modalisa4.4DD

Enquêtes / Tables Champs Données Recodages Sous-pop. Plans

Table: Clients(Access)

Champ	Type	Mod.	Code	N°
Code client	Text			1
Société	Un	91		2
Contact	Un	91		3
Fonction	Un	9		4
Adresse	Text			5
Ville	Un	69		6
Région	Un	18		7
Code postal	Text			8
Pays	Un	21		9
Téléphone	Text			10
Fax	Text			11

Sous-pop courante Tous les questionnaires: 91

L'onglet Données, permet à l'utilisateur de visualiser les données de la table importée.

Base - Modalisa4.4DD

Enquêtes / Tables Champs Données Recodages Sous-pop. Plans

Table: Clients(Access)

N°	1. Code client	2. Société	3. Contact	4. Fonction	5. Adresse
1	ALFKI	Alfredi Futterkiste	Maria Anders	Représentante(e)	Obere Str. 57
2	ANATR	Ana Trujillo Emparedados y	Ana Trujillo	Propriétaire	Avda. de la Co
3	ANTON	Antonio Moreno Taquería	Antonio Moreno	Propriétaire	Maladerso 23
4	AROUT	Around the Horn	Thomas Hardy	Représentante(e)	120 Hanover S
5	BERGS	Berglunds snabbköp	Christina Berglund	Acheteur	Berguvsvägen
6	BLAUS	Blauer See Delikatessen	Hanna Moos	Représentante(e)	Forststr. 57
7	BLENP	Blondel père et fils	Fidélrique Citeaux	Directeur du marketing	24, place Kléber
8	BOLID	Bólido Comidas preparadas	Martin Sommer	Propriétaire	C/Araquil, 67
9	BONAP	Bon app'	Laurence Leblhan	Propriétaire	12, rue des Bo
10	BOTTM	Bottom-Dollar Markets	Elizabeth Lincoln	Chef comptable	23 T sawassen
11	BSBEV	B's Beverages	Victoria Ashworth	Représentante(e)	Fairview Cir
12	CACTU	Cactus Comidas para llevar	Paticio Simpson	Assistent(e) des ventes	Cerrito 333
13	CENTC	Centro comercial Modurreu	Francisco Chang	Directeur du marketing	Sieras de Gran
14	CHOPS	Chop-suey Chinese	Yang Wang	Propriétaire	Hauptstr. 29
15	COMMI	Comércio Mineiro	Pedro Alfonso	Assistent(e) des ventes	Av. dos Lusíac
16	CONSH	Consolidated Holdings	Elizabeth Brown	Représentante(e)	Berkeley Gard
17	DRACD	Drachenblut Delikatessen	Sven Ottlieb	Acheteur	Walsenweg 21
18	DUMON	Du monde entier	Janine Labrun	Propriétaire	67, rue des Cir
19	EASTC	Eastern Connection	Ann Devon	Assistent(e) des ventes	35 King Georg

Sous-pop courante Tous les questionnaires: 91

Tri à plat

Le tri à plat est un dénombrement des effectifs effectué pour les modalités d'une seule variable. Le tri à plat sur la variable Fonction permet d'observer la répartition de 91 clients sur les différentes modalités : 20 clients sont des représentants, 17 sont des propriétaires, 10 chef comptable...

Tri à plat / Clients(Access)

Popul. totale Données brutes Tri Modalités

Ss rép. exclues % sur réponses % cumulés

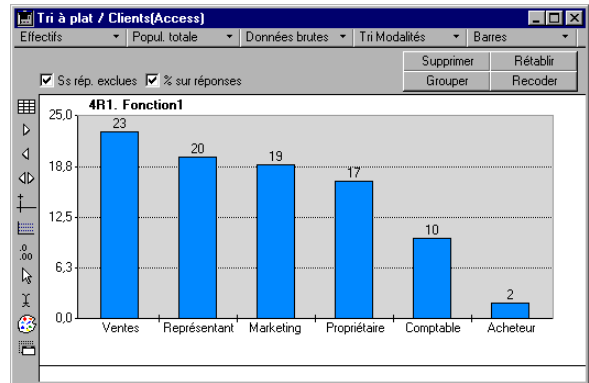
Supprimer Rétablir
Grouper Recoder

4. Fonction

	Effectifs	%	Cumul
Directeur du marketing	12	13,2	13,2
Chef des ventes	11	12,1	25,3
Représentant(e)	20	22,0	47,3
Propriétaire	17	18,7	66,0
Chef comptable	10	11,0	77,0
Assistent du marketing	1	1,1	78,1
Acheteur	2	2,2	80,3
Assistent(e) marketing	6	6,6	86,9
Assistent(e) des ventes	12	13,2	100,1
TOTAL	91	100,0	

Le tri à plat donne une idée immédiate sur la répartition des modalités de la variable étudiée.

Ainsi, on peut faire de recodages si cela s'avère nécessaire afin de donner plus de sens à l'analyse. On peut aussi faire des regroupements de modalités. Dans cet exemple les modalités Représentant et Représentant(e) peuvent être grouper sous la modalité Représentant.



Les modalités peuvent être représentées par un graphique de type diagramme en barres (exemple ci-dessus) ou de type secteurs.

Tri croisé

Le tri croisé est un tableau de contingences à deux entrées représentant pour chaque case les effectifs correspondant aux modalités croisées des deux variables. Dans cet exemple on croise la variable Pays (en lignes) avec la variable Fonction (en colonnes) comme l'indique la fenêtre suivante

Tris croisés / Clients(Access)

Effectifs Popul. totale Données brutes Tous Khi2

PEM global Base Répondants

Ss rép. exclues

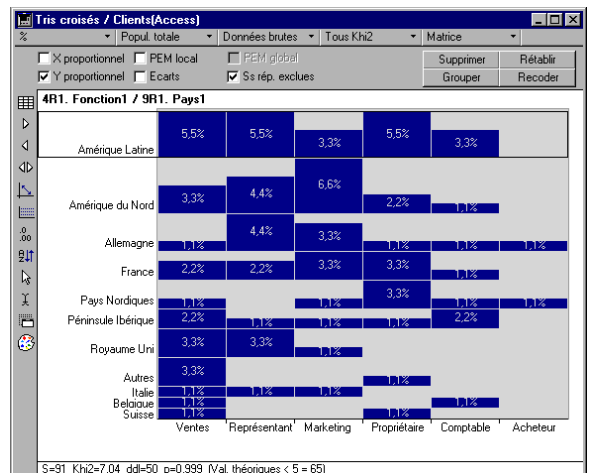
Supprimer Rétablir
Grouper Recoder

SR1. Pays1 / 4R1. Fonction1

	Ventes	Représentant	Marketing	Propriétaire	Comptable	Acheteur	Total
Amérique Latine	5	5	3	5	3		21
Amérique du Nord	3	4	6	2	1		16
Allemagne	1	4	3	1	1	1	11
France	2	2	3	3	1		11
Pays Nordiques	1		1	3	1	1	7
Péninsule Ibérique	2	1	1	1	2		7
Royaume Uni	3	3	1				7
Autres	3			1			4
Italie	1	1	1				3
Belgique	1				1		2
Suisse	1			1			2
Total	23	20	19	17	10	2	91

Khi2=7,04 ddl=50 p=0,999 (Val. théoriques < 5 = 65)

Le poids des réponses obtenues par le tri croisé peut être représenté par des graphiques en barres, barres empilées proportionnelles, lignes ou matrice.



Importation de données depuis un macintosh

Jusqu'à la date de publication de ce document, Microsoft (éditeur du Microsoft Access) n'a pas écrit un pilote (driver) ODBC MS Access Data Base pour les applications Macintosh.

Pour palier à l'absence de ce pilote ODBC, les utilisateurs de Macintosh dans un environnement hétérogène peuvent procéder comme suit : enregistrer les tables/requêtes (fichiers) d'Access sous les formats suivants:

- texte (*.txt, *.csv, *.tab, *.asc).
- dBase et FoxPro (*.dbf)

Si on prend l'exemple de la base de données Comptoir de MS Access. L'utilisateur de Modalisa sous MasOs dans un environnement hétérogène devrait :

- enregistrer les différents fichiers (tables et requêtes) de la base de données sous format dBase dans un sous répertoire spécifique.
- Configurer, dans les Tableaux de bord, l'**ODBC Setup PPC**. pointer Microsoft 3.01 dBase PPC comme New Data Source et à l'aide du bouton **Configure** Sélectionner le répertoire contenant les fichiers *.dbf de la base de données Microsoft Access.

Pour le reste des étapes voir le paragraphe Exemple d'utilisation d'ODBC.

**Kynos 3, rue des Montiboefus
75020 Paris
téléphone 33 (01) 40 30 23 23
télécopie 33 (01) 40 30 24 52
<http://www.modalisa.com>
email : jlvi@modalisa.com**