

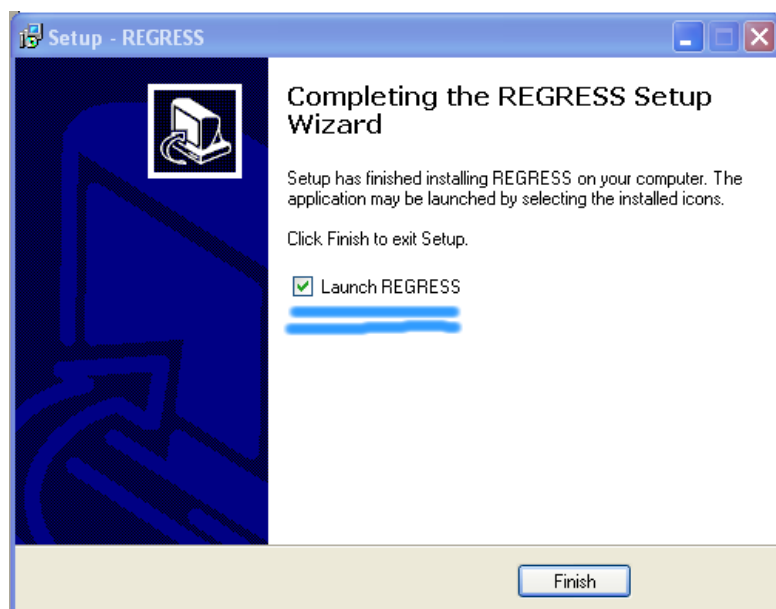
Installation du logiciel

Téléchargez le script d'installation (SETUP_REGRESS.EXE), exécutez-le sur votre machine. Le logiciel s'installe sur votre disque dur, normalement dans un sous-répertoire de « PROGRAM FILES », vous avez la possibilité de modifier cette option. Le programme d'installation respecte les standards de WINDOWS, il crée les éléments adéquats dans le menu DEMARRER.



Charger des données

A la fin de l'installation, vous avez la possibilité d'exécuter directement le logiciel.

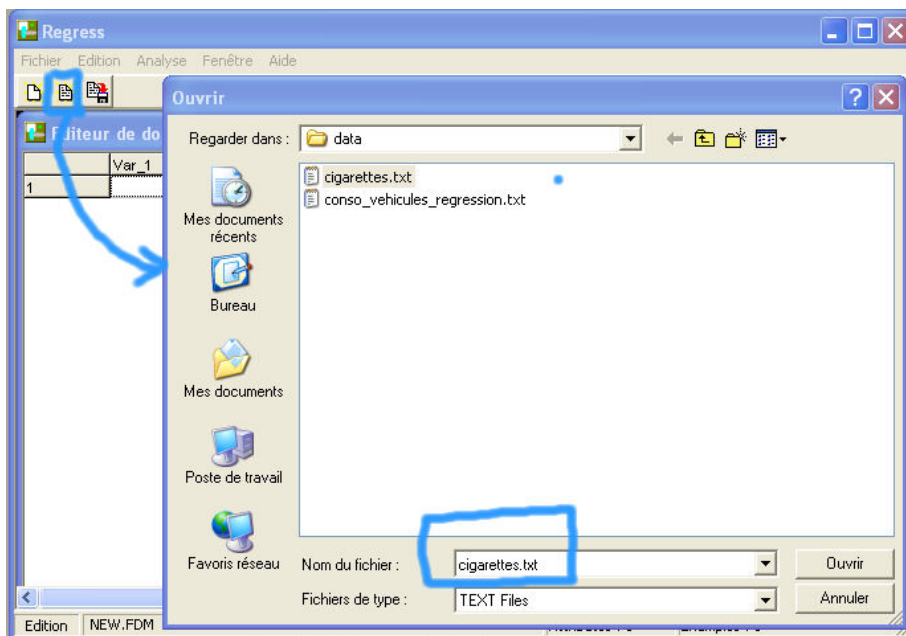


Vous aboutissez alors sur la fenêtre principale de REGRESS. Avant d'initier tout traitement, il est nécessaire de charger des données dans la grille qui s'apparente à une grille de tableur.

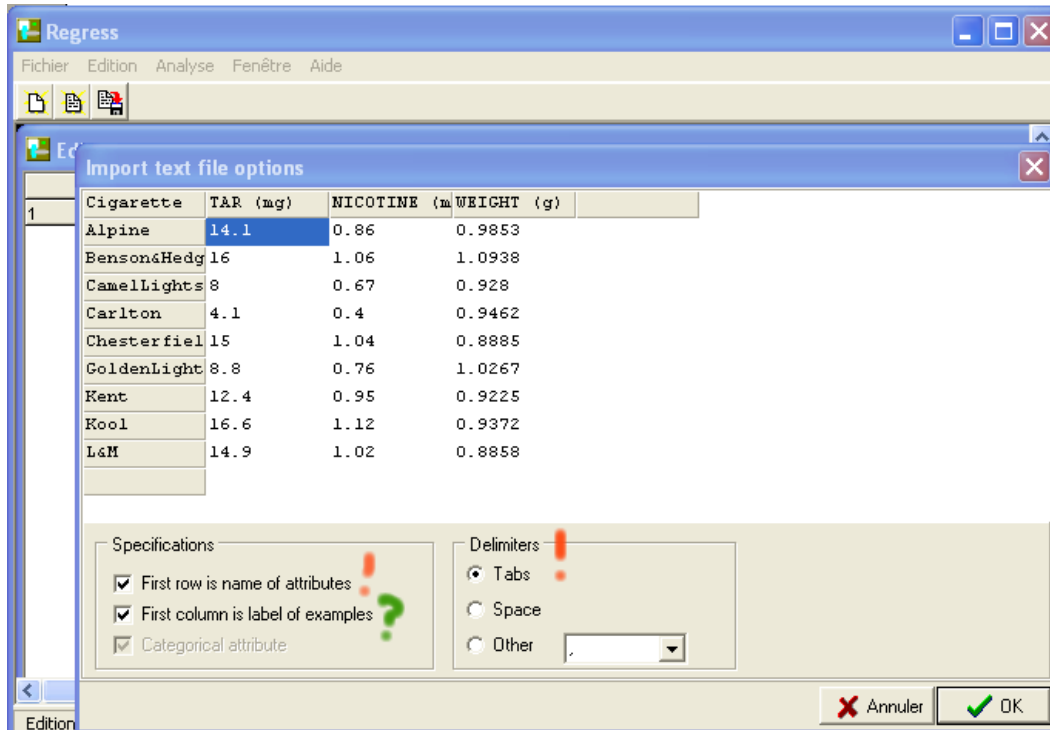
Parmi les formats de fichiers disponibles, nous recommandons le format texte avec « séparateur tabulation », c'est un format normalisé reconnu par une grande majorité des

logiciels d'économétrie, il est facile de le générer à partir d'un tableau comme EXCEL par exemple (cf. Enregistrer Sous / Texte Séparateur Tabulation dans EXCEL).

Pour charger des données, activez le menu FICHIER / OUVRIR, les deux fichiers exemples livrés avec le logiciel se trouvent dans le sous-répertoire DATA du répertoire principal du logiciel -- PROGRAM FILES\REGRESS si vous n'avez pas modifié les options par défaut de l'installation.



Un panneau vous guidera alors dans la spécification des caractéristiques de vos données.



La première option à définir est le caractère délimitant les colonnes de vos données -- DELIMITERS. Bien souvent, surtout si vos données proviennent d'un tableau, il s'agira du caractère TABULATION. Ensuite, dans un second temps, nous devons signaler que la première ligne contient le nom des variables (FIRST ROW IS NAME OF ATTRIBUTES). Si vous ne cochez pas

cette option, le logiciel attribue automatiquement le nom des variables de manière séquentielle VAR_1, VAR_2, etc.

Ces deux options se retrouvent presque systématiquement dans les fichiers de données. La troisième option – FIRST COLUMN IS LABEL OF EXAMPLES -- en revanche est un peu plus délicate : il s'agit de signaler que la première colonne représente les étiquettes désignant les observations. Cela peut être le nom des individus par exemple si chaque ligne représente des personnes, dans notre cas il s'agit du nom des marques de cigarettes. Cette caractéristique n'est pas toujours présente dans les fichiers de données, il vous appartient de bien le vérifier avant de cocher cette option. De plus, si vous avez effectivement des étiquettes désignant les observations dans votre fichier de données, il est obligatoire qu'elles soient dans la première colonne, cette contrainte semble assez raisonnable dans la pratique.

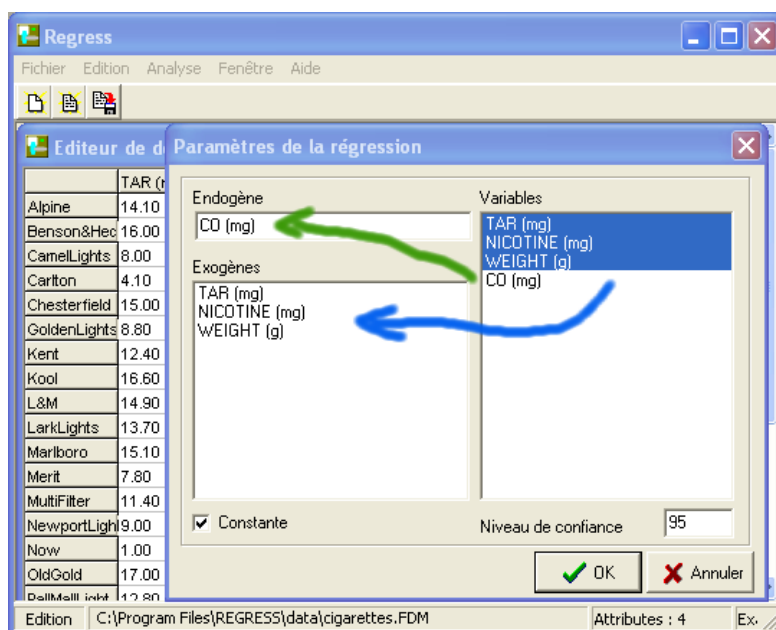
Remarque : Séparateur décimal

Quelle que soit la configuration de votre système d'exploitation, REGRESS considère toujours que le point décimal est le caractère « . » (Norme anglo-saxonne). Il vous revient donc d'effectuer les modifications adéquates si vous exportez les données à partir d'un tableur générant des valeurs avec un point décimal tel que la « , ». Utilisez pour cela un éditeur de texte quelconque, et lancez tout simplement un rechercher remplacer.

Régression linéaire multiple

Une fois les données chargées dans la grille, 24 observations et 4 variables dans notre cas, vous pouvez demander l'exécution d'une régression linéaire multiple.

Pour ce faire, activez le menu ANALYSE / REGRESSION MULTIPLE LINEAIRE, une boîte de dialogue vous permet de définir les paramètres du modèle à estimer.



A l'aide du glisser-déposer, sélectionnez les variables comme suit : CO est la variable endogène ; TAR, NICOTINE et WEIGHT sont les exogènes.

En cliquant sur le bouton OK, vous lancez l'exécution des calculs.

Variable endogène : CO (mg)

	Somme des carrés	ddl
Expliqués	386.8456	3.0000
Résiduels	26.9039	20.0000
Total	413.7496	23.0000

Coefficient de détermination : 0.9350
Coefficient de détermination corrigé : 0.9252

	Coefficients	B. Basse	B. Haute	STD	T de Student	R.Critique
TAR (mg)	0.8876	0.4798	1.2953	0.1955	4.5405	0.0002
NICOTINE (mg)	0.5185	-6.2658	7.3027	3.2523	0.1594	0.8749
WEIGHT (g)	2.0793	-4.5507	8.7094	3.1784	0.6542	0.5204
Constante	-0.5517	-6.7497	5.6463	2.9713	-0.1857	0.8546

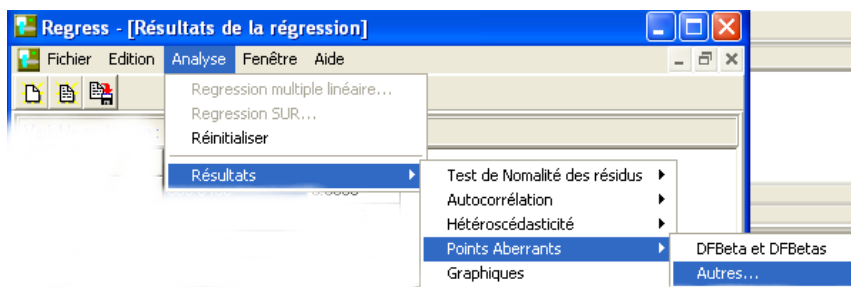
Plusieurs fenêtres sont automatiquement générées, la plus importante, du moins dans un premier temps, est la fenêtre RESULTATS DE LA REGRESSION, elle regroupe : le tableau d'analyse de variance, la valeur du R^2 et du R^2 corrigé, les coefficients et leur évaluation individuelle (test de significativité).

Les autres fenêtres présentent des résultats spécifiques. Il faut se méfier tout particulièrement de la fenêtre NORMALITE DES RESIDUS, les tests proposés (SKEWNESS, KURTOSIS et LAMBDA DE BERA-JARQUE) sont très imprécis, ils ne sont dignes de foi que lorsque l'effectif est élevé. Nous devons les lire donc avant tout à titre d'information, pour se donner une idée de la forme de la distribution des résidus. Les tests de SHAPIRO-WILK, LILLIEFORS, ANDERSON DARLING et D'AGOSTINO sont plus indiqués si l'on veut évaluer la normalité de manière plus rigoureuse.

Détecter les points atypiques

Elément fondamental d'évaluation de la régression, la détection des points atypiques est malheureusement peu présente dans les logiciels de régression. REGRESS propose une série d'indicateurs qui permettent d'attirer l'attention du praticien sur certaines observations qui peuvent jouer un rôle exagéré dans la régression.

Activez le menu indiqué dans la copie d'écran ci-dessous. Attention, il est important que la régression ait été lancée au préalable, ce qui est le cas dans notre exemple ; dans le cas contraire, ce menu n'est pas accessible.



Une nouvelle fenêtre apparaît alors, elle liste une série d'indicateurs pour l'ensemble des observations, les valeurs signalées de rouge sont celles qui excèdent les seuils de coupure.

Coupure	>0.3333	>0.8165	>2.0930	<=[0.50,1.50]	<0.6957	>10.0435	>2.8661
	Hi	DFFITS	RSTUDENT	COVRATIO	Wilks	Mahalanobis	Cook
Alpine	0.1707	0.3653	-0.8050	1.2946	0.8653	3.5729	0.0340
Benson&Hec	0.1849	0.0561	0.1177	1.5019	0.8505	4.0344	0.0008
CamelLights	0.0857	0.3848	1.2568	0.9756	0.9540	1.1059	0.0360
Carlton	0.1669	0.0568	0.1269	1.4686	0.8694	3.4495	0.0008
Chesterfield	0.1335	0.0527	-0.1343	1.4115	0.9042	2.4327	0.0007
GoldenLights	0.1572	0.3159	-0.7313	1.3037	0.8794	3.1484	0.0255
Kent	0.0979	0.1658	-0.5033	1.2908	0.9413	1.4309	0.0071
Kool	0.1196	0.1364	-0.3700	1.3551	0.9187	2.0324	0.0049
L&M	0.1336	0.1266	0.3223	1.3865	0.9041	2.4362	0.0042
LarkLights	0.0763	0.2935	-1.0210	1.0735	0.9639	0.8609	0.0215
Marlboro	0.2315	0.4577	-0.8340	1.3835	0.8020	5.6692	0.0532
Merit	0.0948	0.3901	1.2051	1.0102	0.9445	1.3482	0.0372
MultiFilter	0.2599	1.3848	-2.3368	0.6039	0.7723	6.7700	0.3920
NewportLight	0.1554	0.0358	-0.0834	1.4515	0.8813	3.0909	0.0003
Now	0.4603	0.5718	-0.6191	2.1004	0.5631	17.8105	0.0843
OldGold	0.2553	0.8400	1.4347	1.0925	0.7771	6.5848	0.1675
PallMallLight	0.3157	0.6578	-0.9686	1.4795	0.7141	9.1913	0.1085
Raleigh	0.1948	0.7517	1.5280	0.9592	0.8402	4.3675	0.1324
SalemUltra	0.1525	0.2555	-0.6023	1.3431	0.8844	3.0017	0.0169
Tareyton	0.0687	0.2325	0.8557	1.1333	0.9718	0.6673	0.0137
TrueLight	0.0961	0.0626	0.1921	1.3477	0.9432	1.3819	0.0010
ViceroyRichL	0.0695	0.2802	1.0250	1.0639	0.9709	0.6875	0.0196
VirginiaSlims	0.0835	0.4301	-1.4249	0.8925	0.9563	1.0479	0.0440
WinstonLight	0.2356	1.2262	2.2090	0.6437	0.7977	5.8227	0.3148

Il est possible de sauver ces données dans un fichier au format texte ou au format binaire natif de REGRESS (FDM). Pour ce faire, cliquez avec le bouton droit de la souris sur la grille de résultats, et sélectionnez l'option SAVE dans le menu contextuel qui apparaît.

0.0527	-0.1343	1.4115	0.9042	2.43
0.3159	-0.7313	1.3037	0.8794	3.148
0.1658	-0.5033	1.2908	0.9413	1.430
0.1364	-0.3700	1.3551	0.9187	2.032
0.1266	0.3223	1.3865	0.9041	2.436
0.2935	-1.0210	1.0735	0.9639	0.860
0.4577	-0.8340	1.3835	0.8020	5.669
0.3901	1.2051	1.0102	0.9445	1.348
1.3848	-2.3368	0.6039	0.7723	6.770
0.0358	-0.0834	1.4515	0.8813	3.090
0.5718	-0.6191	2.1004	0.5631	17.81
0.8400	1.4347	1.0925	0.7771	6.584

Tous ces calculs ont été reproduits dans un classeur EXCEL disponible sur le site web du cours (<http://chirouble.univ-lyon2.fr/~ricco/cours/> -- Section **Econométrie**). Il vous sera ainsi possible de retracer le détail des calculs avec les formules associées.