

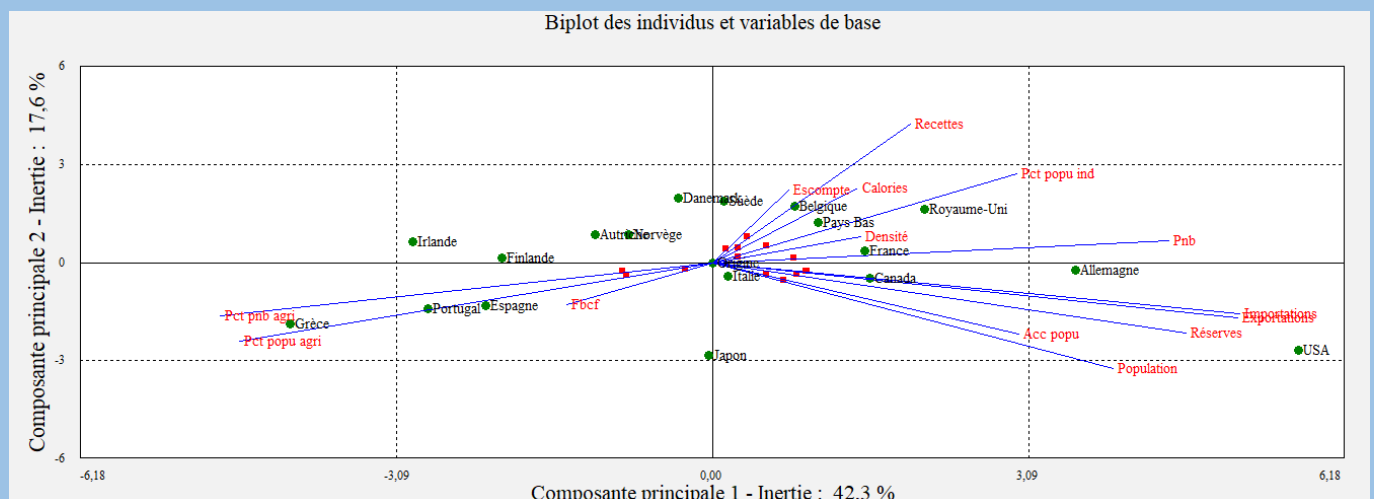
UNIWIN version 10.4.0

Analyses statistiques multivariées descriptives et explicatives

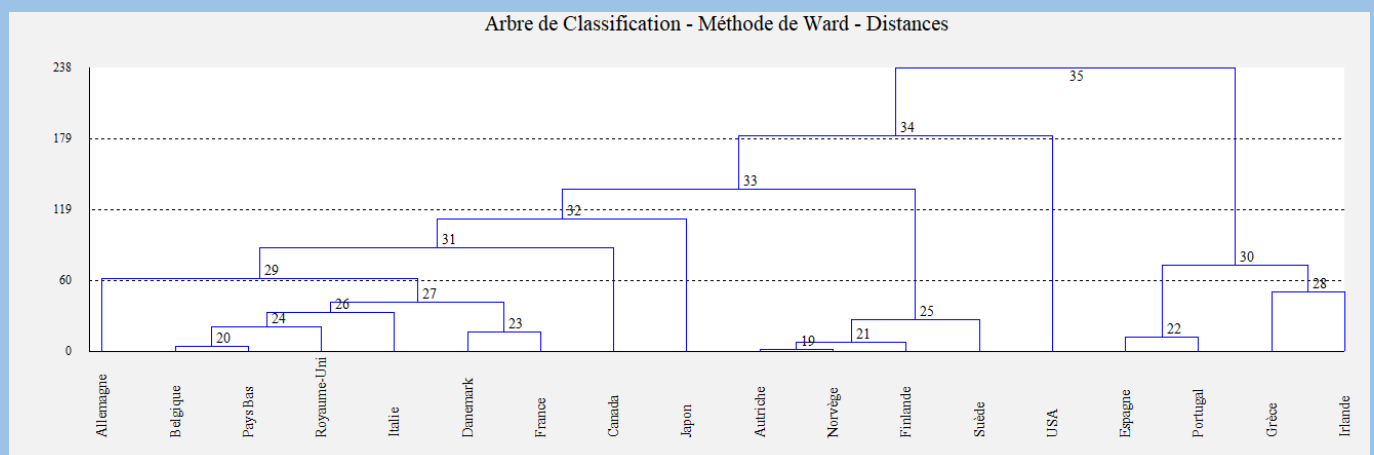
Les méthodes d'analyse de données multivariées sont aujourd'hui largement utilisées dans de nombreux secteurs, notamment en sciences sociales, en marketing, en économie, en finance, en contrôle de qualité et en analyse sensorielle.

UNIWIN apporte à la fois des méthodes usuelles (analyse en composantes principales, analyse des correspondances simples et multiples, analyse sur tableau de distances ou de dissimilarités, classification hiérarchique, ellipses de tolérance et de confiance, analyse factorielle discriminante globale et pas à pas, analyse discriminante bayésienne, analyse discriminante qualitative) et des méthodes spécialisées (analyse factorielle multiple, méthodes Statis et Statis duale, analyse procrustéenne généralisée, analyse factorielle de données mixtes, analyse en composantes principales symbolique, analyse en composantes principales fonctionnelle, scoring, analyse en composantes indépendantes, analyse NIPALS, classification par cartes auto-organisatrices, méthodes des K-moyennes et des K-médoïdes, classification par densité, méthode Apriori, régression sur composantes principales, régression typologique, régressions PLS1 et PLS2, réseaux de neurones artificiels, méthode KNN, méthode de classement bayésienne naïve, méthode SIMCA, arbres et forêts aléatoires de décision et de régression), séparateurs à vastes marges.

UNIWIN est un module complémentaire à STATGRAPHICS Centurion et il permet également d'analyser directement des fichiers MICROSOFT EXCEL (formats xlsx et xlsx).



UNIWIN fonctionne sur tout ordinateur sous Windows 7, 8, 10, 11, Server 2016, 2019 ou 2022 équipé d'au minimum 1 Go de Ram et disposant de 150 Mo de place disque. UNIWIN est livré avec une documentation au format électronique et est disponible en version monoposte et réseau flottant.



Des formations à l'utilisation du logiciel et aux méthodes d'analyse de données multivariées sont régulièrement organisées dans nos locaux et sur site, en intra et en inter-entreprises.

Gestion des données, impression, rapport

- Fichiers aux formats STATGRAPHICS et Excel
- Affichage des données et des rapports dans un tableur
- Export des rapports au format HTML

Graphiques et outils généraux

- Nuage à 2 dimensions, graphique en bulles
- Histogramme à classes d'amplitudes inégales
- Courbes de Gini et ROC
- Imputations de données manquantes
- Recodifications de variables alphabétiques
- Codage disjonctif complet
- Tableaux de distances
- Tris à plat et tris croisés
- Cartes géographiques thématiques (France, Europe)
- Composition graphique (wmf, emf)
- Graphiques aux formats emf, bmp, gif, jpeg, png, tiff
- Nombreuses options graphiques dont identification des points et zoom en pourcentage ou dynamique

Analyse en Composantes Principales

- Analyse sur corrélations de Pearson ou de Spearman, covariances ou données brutes
- Individus de base et supplémentaires
- Variables de base et supplémentaires
- Rapport détaillé pour les individus et les variables
- Cercles factoriels, plans factoriels, Biplot
- Graphiques avec symboles proportionnels
- Diagrammes des contributions, cosinus, distances

Analyse des Correspondances Simples

- Lignes et colonnes de base ou supplémentaires
- Rapport détaillé pour les lignes et les colonnes
- Plans factoriels des lignes, des colonnes, simultanés, barycentriques
- Graphiques avec symboles proportionnels
- Diagrammes des contributions, cosinus, distances

Analyse des Correspondances Multiples

- Individus et variables de base et supplémentaires
- Rapport détaillé pour les individus et les variables
- Plans factoriels des individus, des variables, simultanés, barycentriques
- Graphiques avec symboles proportionnels
- Diagrammes des contributions, cosinus, distances

Analyse sur Tableau de Distances / Dissimilarités

- Symétrisation des tableaux non symétriques
- Utilisation de la technique de la constante additive
- Rapport détaillé pour les individus et Plans factoriels

Classification Ascendante Hiérarchique

- Méthode des voisins réciproques (Ward)
- Données quantitatives standardisées ou non, qualitatives
- Tableau de contingence (lignes ou colonnes)
- Rapport général sur la classification
- Dendrogramme vertical ou horizontal avec zoom
- Troncature et rapport d'analyse des classes formées
- Nuages codifiés à deux dimensions

Ellipses de Tolérance ou de Confiance

- Ellipses de tolérance contenant un % de la population
- Ellipses de confiance autour de centres de gravité

Analyse Factorielle Multiple

- Tableaux d'individus observés sur des groupes de variables qui peuvent être différents
- Rapport détaillé pour le compromis, l'intrastructure et l'interstructure
- Graphique des corrélations des variables (compromis)
- Graphique des individus (compromis)

- Graphique des groupes de variables (interstructure)
- Graphique des trajectoires des individus (intrastructure)

Méthodes Statis et Statis duale

- Tableaux d'individus observés sur des groupes de variables qui peuvent être différents (Statis)
- Tableaux de variables observées sur des ensembles d'individus qui peuvent être différents (Statis duale)
- Rapport détaillé compromis, intrastructure et interstructure
- Graphique des individus ou des variables (compromis)
- Graphique des groupes de variables (interstructure)
- Graphique des trajectoires des individus ou des variables (intrastructure)

Analyse Factorielle de Données Mixtes

- Etude simultanée de variables quantitatives et qualitatives
- Combine les méthodes ACP et ACM
- Rapport détaillé pour les variables et les individus
- Cercles factoriels des variables quantitatives
- Représentation simultanée des variables quantitatives et qualitatives
- Plans factoriels des individus

Analyse Discriminante Pas à Pas

- Analyse de 2 groupes ou plus
- Sélection pas à pas ascendante et descendante des variables les plus explicatives des groupes
- Rapport détaillé sur les étapes de la sélection et résumé

Analyse Factorielle Discriminante

- Analyse de 2 groupes ou plus
- Populations d'apprentissage, de validation et de prévision
- Rapport détaillé sur l'analyse et les classements
- Cercle des corrélations des variables
- Plans factoriels des individus codifiés
- Courbe ROC

Analyse Discriminante Bayésienne

- Analyse de 2 groupes ou plus
- Populations d'apprentissage, de validation et de prévision
- Données centrés-réduites ou non
- Test de Box pour l'égalité des matrices des covariances
- Analyse linéaire et analyse quadratique
- Rapport détaillé sur l'analyse et les classements
- Cercle des corrélations des variables
- Plans factoriels des individus codifiés
- Courbe ROC

Analyse Discriminante Qualitative

- Analyse de 2 groupes ou plus
- ACM sur les variables explicatives qualitatives
- Populations d'apprentissage, de validation et de prévision
- Rapport détaillé sur l'analyse et les classements
- Représentation des variables qualitatives
- Plans factoriels des individus codifiés

Scoring

- Populations d'apprentissage et de prévision
- Rapport détaillé sur l'analyse et les classements
- Histogramme des coefficients de la fonction de score
- Calcul de la sensibilité et de la spécificité
- Courbes de densité et de répartition à 2 et 3 zones
- Courbes ROC et Lift
- Simulation interactive du score d'un individu

Régression sur Composantes Principales

- Méthodes d'ACP et de régression combinées
- Rapport général détaillé (individus, variables, Anova)
- Graphiques des R2 et des R2 ajustés
- Cercles et plans factoriels
- Graphiques de la régression (observé/ajusté, composant, résidus)

Interface avec le langage R

- Utilisation de R depuis UNIWIN
- Chargement des packages
- Liste des packages installés

Analyse NIPALS

- Méthode séquentielle d'ACP avec données manquantes
- Validation croisée (méthode de Krzanowski)
- Valeurs propres
- R-carrés
- Poids des variables
- Corrélations des variables
- Scores des observations
- Cosinus carrés, cosinus carrés cumulés
- Distances carrées à l'origine
- Distances normalisées au modèle
- Données reconstituées
- Ecart
- Diagramme des inerties
- Cercle des corrélations
- Graphique des scores des observations
- Graphique des contributions
- Graphique des cosinus carrés cumulés
- Graphique des distances carrées à l'origine
- Graphique des distances normalisées au modèle

Analyse en Composantes Indépendantes

- Utilisation des packages R 'ica' et 'JADE'
- Séparation aveugle de signaux
- Algorithmes : JADE, FastICA (estimation en parallèle ou par déflation), Infomax (méthode de Newton ou descente du gradient), SOBI (frjd)
- Choix du nombre de composantes
- Centrage-réduction ou non des données
- Matrice de rotation initiale
- Matrice de mélange
- Matrice de démixage
- Composantes principales
- Matrice de blanchiment
- Matrice de rotation
- Composantes indépendantes
- Variances des composantes
- Résidus
- Coefficients de Durbin-Watson
- Graphique des données initiales
- Graphique d'une composante principale
- Graphique d'une composante indépendante
- Histogramme des données initiales
- Histogrammes des données estimées

Analyse procrustéenne généralisée (GPA)

- Utilisation du package R 'shapes'
- Distances initiales entre les configurations
- Distances finales entre les configurations
- Configurations finales (observations)
- Configurations finales (configurations)
- Consensus et distances au consensus
- Inerties ACP (observations et configurations)
- Corrélations des variables
- Scores du consensus
- Scores des observations x configurations
- Scores des configurations
- Graphique des distances des configurations au consensus
- Graphique des corrélations des variables x configurations
- Plan factoriel du consensus des observations
- Plan factoriel des observations x configurations
- Plan factoriel des configurations

Classification par cartes auto-organisatrices

- Utilisation du package R 'kohonen'
- Cartes SOM de Kohonen
- Choix des nombres de nœuds en X et en Y
- Choix de la topologie (hexagonale ou rectangulaire)
- Choix de la fonction de voisinage (gaussienne ou circulaire)
- Centrage-réduction ou non des données
- Distance euclidienne au carré, euclidienne ou valeur absolue
- Nombre maximum de présentations
- Plage pour le taux d'apprentissage alpha
- Choix de la racine aléatoire
- Effectifs dans les nœuds
- Affectations des nœuds
- Affectations des observations
- Profil des nœuds
- Graphique de la progression de l'apprentissage
- Graphique des effectifs dans les nœuds
- Graphique des observations dans les nœuds
- Graphique des distances entre les nœuds
- Graphique de la qualité des nœuds
- Graphique des profils des nœuds
- Carte thermique individuelle ou globale
- Classification ascendante hiérarchique

Réseaux de neurones artificiels

- Utilisation du package R 'neuralnet'
- Réponse quantitative ou qualitative
- Nombre de neurones par couche cachée
- Seuil d'arrêt
- Nombre maximal d'étapes
- Nombre de répétitions pour l'apprentissage du réseau
- Méthode : classement ou régression
- Pré-traitement de données
- Algorithmes rprop+, rprop-, backprop, sag, slr
- Fonction d'erreur quadratique ou entropie croisée
- Fonction d'activation logistique ou tangente hyperbolique
- Sortie linéaire ou non
- Choix de la racine aléatoire
- Jeux d'apprentissage, de validation et de prévision
- Poids ajustés
- Poids généralisés
- Résultats pour le jeu d'apprentissage
- Matrice de confusion pour le jeu d'apprentissage
- Résultats pour les jeux de validation et de prévision
- Matrice de confusion pour le jeu de validation
- Sensibilités et spécificités
- Graphique du réseau
- Graphique des poids généralisés
- Graphique valeurs prévues versus valeurs observées
- Graphique résidus versus valeurs observées
- Graphique résidus versus valeurs prévues
- Courbes ROC pour les jeux d'apprentissage et de validation

Analyse en Composantes Principales Symbolique

- Analyse sur tableau de données individus x variables ou sur tableau de données intervalles
- Composantes principales (classique et symbolique)
- Coordonnées des variables (classique et symbolique)
- Cosinus carrés
- Contributions
- Distances carrées à l'origine
- Diagramme des inerties
- Cercle des corrélations des variables (classique)
- Plan factoriel des variables (symbolique)
- Plan factoriel des concepts (classique et symbolique)

Régressions PLS1 et PLS2

- Utilisation du package R 'plsdepot'
- Régression PLS1 (un seul Y)
- Régression PLS2 (plusieurs Y)
- PRESS, R2
- Q2, Q2 cumulés
- R2X, R2Y
- Corrélations X vs t, Y vs t (PLS2)
- Corrélations X vs u, Y vs u (PLS2)
- Coefficients de régression standardisés
- Coefficients de régression non standardisés
- Poids w
- Poids w*
- Scores X (t)
- Scores Y (u)
- Poids des variables X
- Poids des variables Y
- Corrélations variables – composantes
- Valeurs observées et estimées pour chaque Y
- Résidus pour chaque Y
- Distances au modèle en X
- Distances au modèle en Y
- T2 de Hotelling (PLS1)
- VIP (PLS2)
- Diagramme des R2X cumulés
- Diagramme des R2Y cumulés (PLS2)
- Diagramme des Q2 cumulés
- Diagramme des coefficients standardisés
- Cercle des corrélations
- Plan factoriel des observations (t vs t)
- Plan factoriel des observations (u vs t)
- Plan factoriel des observations (u vs u)
- Diagramme des poids des variables X et Y
- Graphique des T2 de Hotelling (PLS1)
- Graphique des distances au modèle en X
- Graphique des distances au modèle en Y
- Graphique des VIP (PLS2)
- Graphique observés vs estimés
- Graphique des résidus

Méthode DBSCAN

- Utilisation du package R 'dbscan'
- Classification par la méthode 'Density-Based Spatial Clustering of Applications with Noise'
- Classification des observations
- Synthèse de la classification
- Graphique des distances K-NN
- Graphique des distances d'accessibilité par densité
- Graphique des coefficients de silhouette
- Nuages de points des classes
- Nuages de points des classes avec enveloppes

Méthode KNN

- Utilisation du package R 'caret'
- Détermination du K optimal
- Classement pour le jeu d'apprentissage (classement)
- Matrice de confusion pour le jeu d'apprentissage (classement)
- Sensibilités, spécificités pour le jeu d'apprentissage (classement)
- Classement pour le jeu de prévision (classement)
- Valeurs observées, estimées, résidus pour le jeu d'apprentissage (régression)
- Valeurs estimées pour le jeu de prévision (régression)
- Graphique de l'évolution de la précision (classement)
- Graphique des données du jeu d'apprentissage (classement)
- Graphique des données du jeu de prévision (classement)
- Graphique de la matrice de confusion (classement)
- Courbe ROC (classement)
- Graphique de l'évolution du R-carré (régression)
- Graphique des valeurs estimées vs observées pour le jeu d'apprentissage (régression)
- Graphique des résidus vs valeurs estimées pour le jeu d'apprentissage (régression)

Analyse APRIORI

- Utilisation du package R 'arules'
- Libellés des modalités
- Motifs fréquents et motifs fréquents fermés
- Détails des règles
- Règles pour la colonne cible
- Occurrences et fréquences des modalités de la colonne cible
- Graphique du lift vs conséquent x antécédent
- Graphique de la confiance vs conséquent x antécédent
- Graphique du support vs conséquent x antécédent
- Graphique du lift par rapport au support
- Graphique du lift par rapport à la confiance
- Graphique du lift par rapport au support pour la colonne cible
- Graphique du lift par rapport à la confiance pour la colonne cible

Méthode bayésienne naïve

- Utilisation du package R 'naivebayes'
- Facteur de classement
- Informations issues de R
- Classement et synthèse apprentissage
- Classement et synthèse validation
- Classement prévision
- Graphiques des lois marginales
- Graphiques des lois conditionnelles
- Graphiques des données (apprentissage, validation, prévision)
- Graphiques des données de prévision
- Courbe ROC

Méthode SIMCA

- Utilisation du package R 'mdatools'
- Facteur de classement
- Tableaux des inerties
- Poids des variables
- Scores des observations
- Résumé du classement (apprentissage)
- Détails du classement (apprentissage)
- Matrice de confusion (apprentissage)
- Distances entre les modèles
- Pouvoirs discriminants des variables
- Détails du classement (prévision)
- Diagrammes des inerties
- Graphiques des poids des variables
- Graphiques des scores des observations
- Graphiques des distances entre les modèles
- Graphiques des pouvoirs discriminants des variables
- Graphiques de Cooman

Arbres de décision et de régression

- Utilisation des packages R 'rpart' et 'rpart.plot'
- Graphique des coefficients de complexité
- Résultats et règles pour l'arbre complet
- Graphique des importances des variables
- Résultats et règles pour l'arbre élagué
- Classement apprentissage, validation, prévision
- Valeurs observées, estimées, résidus (régression)
- Graphique de l'arbre avec possibilité de zoomer
- Graphique des valeurs estimées et des résidus (régression)
- Courbe ROC (classement)

Forêts aléatoires de décision et de régression

- Utilisation du package R randomForest
- Importances des variables explicatives
- Classement apprentissage, validation, prévision
- Matrice de confusion (classement)
- Courbe ROC (classement)
- Valeurs observées, estimées, résidus (régression)
- Graphique des R-carrés (régression)

Classification par la méthode des K-moyennes

- Utilisation des packages R 'stats' et 'cluster'
- Nombre optimal de classes (KM)
- Centroides standardisés des classes (KM)
- Centroides non standardisés des classes (KM)
- Classification du jeu d'apprentissage (KM)
- Résumé de la classification (KM)
- Statistiques pour les classes (KM)
- Contributions des variables aux classes (KM)
- Contributions des classes aux variables (KM)
- Classification du jeu de prévision (KM)
- Indices de la classification (CAH)
- Centroides non standardisés des classes (CAH)
- Classification (CAH)
- Résumé de la classification (CAH)
- Statistiques pour les classes (CAH)
- Contributions des variables aux classes (CAH)
- Contributions des classes aux variables (CAH)
- Diagramme des inerties intra-classes (KM)
- Diagramme des inerties expliquées (KM)
- Diagramme des indices de Calinski-Harabasz (KM)
- Diagramme des coefficients moyens de silhouette (KM)
- Graphique des coefficients individuels de silhouette (KM)
- Nuage de points des classes (KM)
- Diagramme des indices de la classification (CAH)
- Diagramme de l'arbre de classification (CAH)
- Nuage de points des classes (CAH)

Classification par la méthode des K-médoïdes

- Utilisation des packages R 'stats' et 'cluster'
- Standardisation des données (PAM)
- Nombre optimal de classes (PAM)
- Médoïdes des classes (PAM)
- Classification du jeu d'apprentissage (PAM)
- Résumé de la classification (PAM)
- Statistiques pour les classes (PAM)
- Contributions des variables aux classes (PAM)
- Contributions des classes aux variables (PAM)
- Classification du jeu de prévision (PAM)
- Indices de la classification (CAH)
- Centroides des classes (CAH)
- Classification (CAH)
- Résumé de la classification (CAH)
- Statistiques pour les classes (CAH)
- Contributions des variables aux classes (CAH)
- Contributions des classes aux variables (CAH)
- Diagramme des coefficients moyens de silhouette (PAM)
- Graphique des coefficients individuels de silhouette (PAM)
- Nuage de points des classes (PAM)
- Diagramme des indices de la classification (CAH)
- Diagramme de l'arbre de classification (CAH)
- Nuage de points des classes (CAH)

Régression typologique

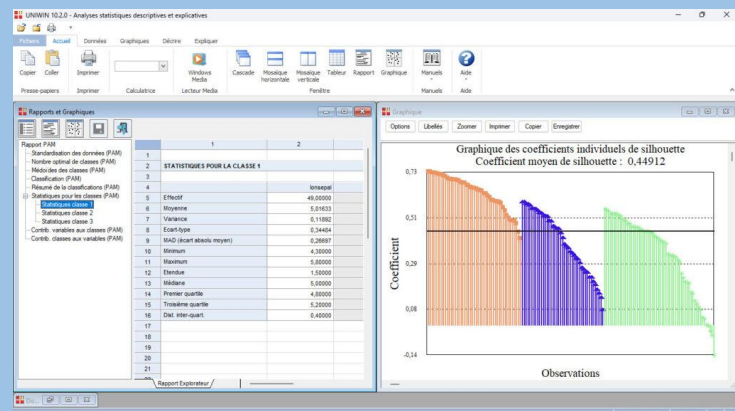
- Somme globale des carrés des résidus
- Somme des carrés des résidus et R2
- Coefficients des régressions
- Statistiques pour les classes
- Résultats apprentissage
- Résultats prévision
- Somme globale des carrés des résidus vs nombre de classes
- Somme globale des carrés des résidus vs itérations
- R-carré global vs itérations
- Graphique Y observé vs X (apprentissage)
- Graphique Y prévu vs Y observé (apprentissage)
- Graphique Y observé vs X (prévision)

Analyse en composantes principales fonctionnelle

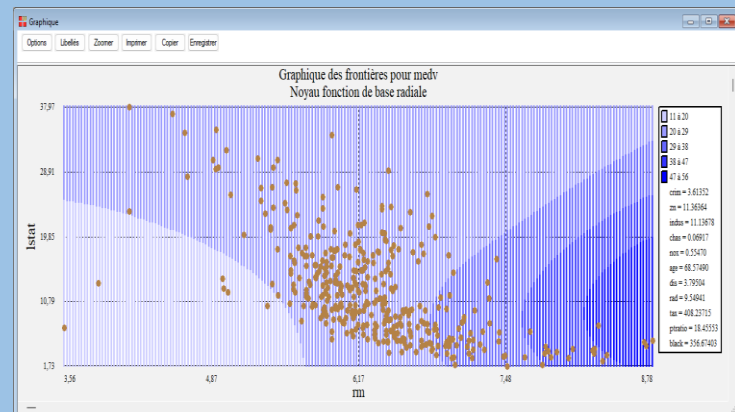
- Indices généralisés de validation croisée
- Erreurs quadratiques moyennes
- Valeurs propres et inerties
- Fonctions propres
- Effets des fonctions propres sur la fonction propre moyenne
- Scores des observations
- Données observées
- Données observées centrées (si centrage)
- Base des splines cubiques ou des séries de Fourier
- Indices généralisés de validation croisée
- Erreurs quadratiques moyennes
- Données estimées
- Résidus
- Moyennes des données observées et estimées
- Moyenne globale, observé et estimé pour une observation
- Diagramme des inerties
- Fonctions propres
- Effets des fonctions propres
- Scores des observations

Séparateurs à vastes marges (SVM)

- Paramètres optimaux des noyaux
- Exactitude (classement)
- Erreur quadratique (régression)
- Importance des variables
- Vecteurs supports
- Valeurs de décision (classement)
- Prévisions
- Matrice de confusion (classement)
- Sensibilité, spécificité (classement)
- Courbe ROC
- Graphique des frontières
- Graphique observés vs ajustés (régression)
- Graphique des résidus (régression)



Graphique K-médoïdes



Graphique Séparateurs à vastes marges