



Le Sphinx

Liste des fonctions

Le Sphinx Développement
Parc Altaïs
74650 Chavanod

Tel. : 04 50 69 82 98
contact@lesphinx.eu

Sphinx Déclic (version Windows), Plus² et Lexica sont des versions plus ou moins élaborées du logiciel. Plus² inclut toutes les fonctions de Déclic, et Lexica les fonctions de Plus² et Déclic. Toutes les versions permettent d'élaborer un questionnaire, de saisir les réponses, ou de l'administrer par Internet et de le dépouiller.

Plus² ajoute des fonctions avancées d'analyse de données multivariées et permet de traiter tout type de données existantes. Il offre en outre des possibilités avancées pour les enquêtes Internet (contrôle d'accès, scénarisation...)

Lexica ajoute des fonctions d'analyse de données textuelles élaborées applicables à tout type de corpus.

Certaines fonctions spécifiques aux modes d'administration font l'objet de modules complémentaires.

Les fonctions disponibles uniquement dans Plus² ou Lexica ou dans un module complémentaire sont indiquées par une mention en italique.

Sommaire

Conception d'enquêtes et élaboration de questionnaires	2
1. Définition des questions	2
2. Structuration du questionnaire	2
3. Mise en forme automatique du questionnaire.....	2
4. Mise en page personnalisée de formulaires papier ou écran	2
Diffusion et collecte	3
1. Modes de diffusion Internet, Opérateurs, Téléphone, PDA, Scanner	3
2. Saisie, collecte de données et consultation des données	3
3. Mise en ligne du questionnaire	3
Traitements statistiques.....	4
1. Dépouillements	4
2. Recodification.....	4
3. Analyses univariées	4
4. Analyses bivariées	5
5. Analyses multivariées (Plus ²).....	5
6. Présentations graphiques	6
7. Datamining et scoring	6
Analyse des textes	6
1. Dépouillement tableau	6
2. Analyse de contenu.....	6
3. Production de verbatim	6
4. Analyse des données textuelles	7
Gestion de données.....	7
Expertises.....	8

Conception d'enquêtes et élaboration de questionnaires

1. Définition des questions

Contenu : Libellé, nom de variable, modalités de réponses.

Type : Fermée unique, multiple, ordonnée, échelle, numérique, ouverte, date, code, clé.

Contrôle : réponse obligatoire, plage de valeurs admissibles, longueur d'un code, nombre de réponses possibles.

Questions tableau : libellé du groupe, liste des questions et des modalités.

Questions « autre précisez » : ajout automatique d'une question texte.

Barème : association de valeurs aux modalités des échelles ou des questions fermées.

Bibliothèque de questions : liste de questions de référence, insertion automatique dans le questionnaire, ajout d'éléments à la bibliothèque depuis le questionnaire.

2. Structuration du questionnaire

Parties : titre, consignes

Déplacement des questions une à une, par partie ou par série de questions.

Duplication de questions.

Renvois (Aller à ... après ...) : renvois sur un ou plusieurs critères.

Restrictions (Ne présenter ... que si ...) : conditions sur un ou plusieurs critères. Les restrictions s'appliquent également aux légendes présentes dans les formulaires multimédias.

Boutons de navigation pour la réalisation de formulaires écran ou de scripts (Internet, Cati, Opérateur...).

3. Mise en forme automatique du questionnaire

Mise en page paramétrable (nombre de colonnes, disposition des questions, numérotation...) .

Inclusion dans un document Word ou Open Office éditable

Impression papier.

Préparation des formulaires destinés à être mis en ligne.

4. Mise en page personnalisée de formulaires papier ou écran

Formulaires multiples : Plusieurs formulaires peuvent être définis selon les modes de diffusion utilisés.

Format et style graphique : les caractéristiques du formulaire (dimensions papier/écran, couleurs de fond, insertion de logo ...) sont directement paramétrables.

Personnalisation des formulaires par la mise en forme des questions, titres et légendes et leur disposition en mode wyswyg :

- Insertion d'illustrations (smileys, images à cliquer, graphismes...)
- Présentation aléatoire des modalités de réponses
- Insertion de consignes ou graphismes interactifs (Internet, script cati)
- Contrôle de la navigation (Internet, script cati).

- Boutons d'action (accès direct à une page du questionnaire, à un site web, à un autre formulaire, ...)

Diffusion et collecte

1. Modes de diffusion Internet, Opérateurs, Téléphone, PDA, Scanner

Internet : Publication de l'enquête et obtention en retour des liens vers le formulaire et/ou les résultats (*Plus²*).

Saisie opérateur : Préparation d'écrans de saisie pour opérateurs de saisie (*Plus², Module opérateurs*).

Cati : Production de script pour postes de télé-acteurs (*Plus², Gestionnaire Cati*).

PDA : Production de scripts pour Pocket PC (*Plus²*).

Scanner : Production de formulaires destinés à être lus par scanner avec ajout automatique du cadre de repérage, et du système d'identification (*Plus², Module scanner*).

2. Saisie, collecte de données et consultation des données

Saisie rapide en mode manuel : frappe au kilomètre, gestion automatique des renvois et restrictions, contrôle automatique de la saisie (plages admissibles, contraintes d'intégrité).

Saisie dans un environnement personnalisé avec possibilité de rappel des réponses antérieures, consignes conditionnelles, alertes.... (Internet, script cati).

Consultation en mode tableur ou fiche ou liste, recherche par individu, strate ou profil multicritère.

Contrôle d'accès Internet par code ou mot de passe individuel. (*Plus²*)

Saisie scanner : reconnaissance automatique des cases à cocher et en option des nombres et des caractères, journal de saisie et gestion des erreurs (*Module scanner*).

Saisie par répondeur vocal. (*Plus²*)

Saisie par Pocket PC. (*Plus²*)

Suppression d'observations : plage de numéros, profil multicritères, liste d'identificateurs

Redressement d'échantillon par extraction, substitution ou pondération.

Données manquantes : Repérage et gestion des remplacements par la valeur moyenne, la valeur la plus proche ou toute autre règle définie par l'utilisateur. (*Plus²*)

Doublons : Repérage automatique et suppression. (*Plus²*)

Gestion de panel : extraction de panélistes, enrichissement du panel. (*Plus²*)

Double-saisie: contrôle de qualité de la saisie par comparaison des fichiers de réponses. (*Plus²*)

3. Mise en ligne du questionnaire

Publication automatique : hébergement sur SphinxOnline ou sur une machine équipée du serveur d'enquêtes.

Insertion du questionnaire dans une page html. (*Plus²*)

Contrôle d'accès par code collectif ou individuel, publication de la liste des codes. (*Plus²*)

Définition des requêtes communiquées sous forme de liens utilisés pour la diffusion (saisir, consulter, modifier, voir les tables, voir les tableaux de bord). (*Plus²*)

Choix du mode d'affichage : html, java, xml. (*Plus²*)

Définition des actions après enregistrement : restitution des réponses, redirection vers une page web, exécution d'un programme... (*Plus²*)

Installation des fichiers de l'enquête sur le serveur avec ou sans les réponses locales.

Traitements statistiques

1. Dépouillements

Question par question : tableaux, graphiques, légende.

Plan de dépouillement : définition des traitements à effectuer et production dans Word du rapport correspondant.

Listes structurées : restitution des données brutes pour les variables sélectionnées sous forme de listes structurées selon les catégories d'une variable de l'étude.

Traitement par strate ou profils multicritères.

Tableaux de bord : Mise en page de tableaux, graphiques, légendes, commentaires et tout traitement statistique (univarié, bivarié, multivarié) ou tableaux personnalisés. Datamining (graphes de relation et arbres de décision). Résultats calculés automatiquement sur toutes les données collectées ou sur la strate considérée.

Actualisation automatique du contenu des tableaux de bord selon le profil analysé ou l'enrichissement des données.

Production des résultats : inclusion de listes de données brutes, rapports, ou tableaux de bord dans Word, Power Point, Open Office ou dans des pages html.

2. Recodification

Transformation des variables nominales : regroupement de modalités, dichotomisation, conversion des questions à réponses fermées ordonnées en échelles. (*Plus²*)

Création de nouvelles variables nominales construites par combinaisons d'autre variables et ou définitions de profils. (*Plus²*)

Calcul de nouvelles variables automatiques définies par programme (*Plus²*)

Mise en classe des variables numériques

Interprétation des variables codes par usage de dictionnaires de codes : code postal, département, région.

Recodification des dates en mois, jour de la semaine, jour du mois...

Recodification des variables textes :

- recherche des mots les plus fréquents
- recherche des mots d'un dictionnaire
- affectation à des catégories thématiques

3. Analyses univariées

Tri à plat : calcul de fréquences, pourcentages (par rapport aux observations ou citations), intervalles de confiance.

Calcul de moyenne, écart-type et indicateurs de dispersion.

Interprétation des codes et des dates.

Comparaison à des valeurs de référence.

Test d'écart à l'équi-répartition et à la normalité.

Présentation des tableaux : affichage ou non des non réponses, pourcentages cumulés, tri des tableaux selon l'ordre croissant ou décroissant.

Simplification des tableaux : suppression, regroupement de modalités, limitation du nombre de lignes, affichage des lignes présentant un effectif minimum...

Edition des résultats sous forme de légende : affichage des valeurs statistiques à l'intérieur d'un texte personnalisé

4. Analyses bivariées

Test de corrélation (Pearson), représentation en nuage de points et droite de régression.

Test de correspondance (chi2) tableau croisé, identification des cellules significatives (contribution au chi2), pourcentages en ligne ou en colonnes, représentation par carte factorielle des correspondances,

Test de comparaison de moyennes (Fisher) tableau de moyennes, représentation par profils, identification des cellules significatives (test t).

Paramétrage des tests statistiques : Fixation des seuils de significativités

Tableau des tests bivariés : Valeurs des tests pour l'ensemble des couples de variables.

Construction et calcul de tableaux : tableaux de groupes de variables, tableaux composés (définition par l'utilisateur des lignes, colonnes et contenu des cellules).

Paramétrage de l'affichage du contenu des tableaux pour signaler les valeurs remarquables (affichage conditionnel de signes graphiques, coloriage des cellules ou libellés du tableau).

5. Analyses multivariées (Plus²)

Régression multiple : coefficient de corrélation multiple, équation de régression multiple et caractérisation des paramètres, procédure pas-à-pas et simulation par ajout / suppression de variables explicatives. (Plus²)

MANOVA : Analyse multiple de la variance (à 2 facteurs). Construction du tableau des moyennes. Mesure des effets directs et des interactions. (Plus²)

Classification automatique : recherche automatique des meilleures partitions par la méthode des centres mobiles (k mean). Caractérisation des partitions obtenues par le nombre de classes par l'homogénéité, la répartition, la dispersion et un indicateur synthétique. Création d'une nouvelle variable caractérisant la partition. (Plus²)

Analyse en composantes principales : Analyse de tableaux de moyennes ou de tables d'individus. Calcul des composantes, affichage de la carte obtenue par projection des variables et des individus sur les 2 premières composantes ou celles choisies par l'utilisateur.

Affichage de variables illustratives, tableaux des contributions. (Plus²)

Analyse factorielle :

Analyse factorielle des correspondances (tableau croisé) ou analyse factorielle des correspondances multiples (table d'occurrence par individu). Calcul des axes factoriels, affichage de la carte sur les 2 premiers axes ou ceux choisis par l'utilisateur. Tableau des contributions.

Analyse factorielle multiple : affichages de variables illustratives. (Plus²)

Typologie : Définition de typologie à partir d'une carte d'ACP ou AFC par agrégation des points proches de centres de classes positionnés par l'utilisateur. (Plus²)

6. Présentations graphiques

Expression graphique des tableaux à plat : Secteur, barre, histogramme, anneau, courbe, surface, baromètre, radar, boîte à moustaches.

Expression graphique des tableaux croisés : Carte factorielle, profil, graphe importance/performance, graphe de relations, arbre de décision.

Représentation cartographique des données géographiques : cartes de France, des régions et départements incluses.

Aides à la lecture : Coloriage selon les valeurs des tableaux/graphiques.

7. Datamining et scoring

Graphes de relations : Production automatique du graphe de toutes les relations significatives entre les variables de l'étude. Recherche des circuits et définition des groupes de variables proches. Mise en évidence des variables structurantes ou de l'ensemble des variables influençant une variable à expliquer. (*Plus²*)

Arbre de décision : Construction d'arbres de décision par décomposition automatique ou manuelle d'une population cible en vue de mettre en évidence les éléments les plus significatifs qui la composent. (*Plus²*)

Scoring : Calcul automatique de nouvelles variables définies par un barème, une expression ou un programme écrit en visual basic. Le contenu de ces variables est disponible pour la scénarisation et l'interactivité des enquêtes en ligne, les scripts cati et la restitution des résultats d'évaluation individuelle.

Analyse des textes

1. Dépouillement tableau

Tableau de réponses identiques : texte, fréquences, pourcentages

Tableau de mots clés : mots classés par fréquence d'apparition ou mots appartenant à un dictionnaire.

2. Analyse de contenu

Questions ouvertes : Recodification des réponses aux questions ouvertes dans les enquêtes

Interview non directives documents : recodification des textes importés.

Définition de la grille de codification comme liste de questions. Mise en forme du « code book » comme formulaire présentant le texte, les informations de contexte, les éléments de la grille, et des possibilités d'annotations marginales et d'enregistrement des passages remarquables.

Procédure de découpage en unités de signification plus fines.

3. Production de verbatim

Extractions : Sélection d'extraits et production de listes de citations structurées par contexte.

Sélection selon les variables de contexte (identité du locuteur, catégorie...), le contenu (mots ou ensemble de mots présents dans le texte)

Classement par indicateurs lexicaux (longueur, banalité, intensité d'un champ lexical). (*Plus²*)

4. Analyse des données textuelles

Dénombrement des mots du texte selon le paramétrage des caractères séparateurs. (*Lexica*)

Analyse syntaxique : production d'une version lemmatisée et/ou catégorisée du corpus. (*Lexica*)

Calcul de lexiques : réduit, lemmatisé, par catégories grammaticales (noms, verbes, adjectifs...).

Navigation lexicale : parcours du texte par entrée lexicale, (*Plus²*) production des concordances d'un mot et production des lexiques relatifs (lexique à gauche et à droite du mot pivot). (*Lexica*)

Recherche de mots dans le lexique et le texte : forme entière, par début, partie ou selon un modèle. (*Plus²*)

Recherche de formes grammaticales : verbes, substantifs, adjectifs, Noms propres. (*Lexica*)

Recherche des locutions et mots composés : recherche manuelle ou par dictionnaire. (*Lexica*)

Regroupement d'éléments du lexique : regroupement manuel des mots marqués, regroupement automatique par racine (stemmatisation), dictionnaires et thésaurus. (*Plus²*)

Mise en évidence des segments répétés : définition des segments selon la longueur et/ou les mots contenus. (*Lexica*)

Cartographie des associations lexicales et cartes cognitives : analyse factorielle des correspondances pour un ensemble d'éléments lexicaux bruts, lemmatisés ou par catégories grammaticales (*Lexica*)

Définition de classes thématiques : méthode interactive à partir d'une carte factorielle des associations lexicales ou automatique par la méthode des nuées dynamiques. (*Lexica*)

Caractérisation de l'usage d'un mot par son contexte. (*Lexica*)

Bilan lexical par contexte et calcul de lexiques par catégories de contexte. Calcul des indices de spécificité lexicale. (*Lexica*)

Production de listes de mots spécifiques et représentation sous forme de carte factorielle des spécificité lexicales et zones de langage. (*Lexica*)

Réponses caractéristiques et production des extraits les plus significatifs par contexte. (*Lexica*)

Calcul d'indicateur lexicaux : longueur, richesse, banalité. (*Lexica*)

Recodification selon la présence de mots marqués : création d'une variable nominale dont les modalités sont les mots sélectionnés ou de plusieurs variables d'intensité lexicale mesurant le poids de chaque mot marqué. (*Lexica*)

Création de corpus expurgés : réduits aux mots marqués ou à certaines catégories. (*Lexica*)

Eclatement du corpus en fragments de textes plus petits : phrases ou tout autre découpage défini par un séparateur. (*Lexica*)

Gestion de données

Importation de fichiers de données : importation automatique de fichiers Excel, Access, CSV, TXT et de toute source de données ODBC Reconnaissance automatique des champs et affectation d'un type de variable. (*Plus²*)

Liaison dynamique d'un questionnaire avec un fichier/une source de données. (*Plus²*)

Relation avec d'autres logiciels : Lecture /Ecriture des fichiers SPSS, TRIPLE-S, Alceste (*Plus²*)

Importation depuis le web : base de données Sphinx publiée sur SphinxOnline ou un serveur d'enquêtes propre, aspiration de sites web ou de listes de liens provenant d'un moteur de recherche. (*Plus²*)

Rassemblement de plusieurs enquêtes : observations récoltées avec un même questionnaire, observations correspondant à la partie commune de plusieurs questionnaires, observations relatives au même individu (repéré par un identifiant) recueilli à partir de 2 questionnaires différents. (*Plus²*)

Gestion de panel : extraction aléatoire de panélistes en application de quotas, couplage avec l'enquête, enrichissement du panel à partir des remontées de l'enquête. (*Plus²*)

Expertises

Evaluation du questionnaire : en fonction des objectifs, mode de collecte et cible indiqués par l'utilisateur, le logiciel propose un diagnostic sur

- le nombre questions, leurs types et leur formulation.
- la complexité du questionnaire
- le temps estimé de saisie

Evaluation des données : calcul des taux de retour et de remplissage. Repérage des questions insuffisamment documentées (valeurs manquantes) et des individus présentant des valeurs extrêmes ou aberrantes. Caractérisation des données en terme de concentration, dispersion et richesse.

Suggestions pour les analyses : en fonction de la nature des données (type de question qualité des réponses) et des variables signalées comme essentielles, les conseils, mises en garde et suggestions sont proposés sur les possibilités d'analyse par strate, calcul de pourcentage, tests bi variés et analyses multi variées.