

Les éléments à télécharger sont disponibles à l'adresse suivante :

**<http://www.editions-eni.fr>**

Saisissez la référence ENI de l'ouvrage **EPPEN** dans la zone de recherche et validez. Cliquez sur le titre du livre puis sur le lien de téléchargement.

## Introduction : présentation de Pentaho

1. Qu'est-ce que l'Informatique Décisionnelle ? . . . . . 25
2. Environnement logiciel requis pour utiliser au mieux ce livre . . . . . 28
3. Version Pentaho utilisée . . . . . 29
4. Description générale de l'offre Pentaho . . . . . 30
5. Objectifs de ce livre . . . . . 32

## Chapitre 1

### Prise en main rapide de Pentaho

1. Objectifs du chapitre . . . . . 35
2. Installation du package Pentaho Business Intelligence Server . . . . . 36
  - 2.1 Pré-requis Java . . . . . 37
  - 2.2 Installation de la suite Pentaho . . . . . 37
  - 2.3 Lancement de Pentaho Business Intelligence Server . . . . . 39
  - 2.4 Connexion à l'interface Pentaho User Console (PUC) . . . . . 40
  - 2.5 Connexion à l'interface Pentaho Administration Console (PAC) . . . . . 41
    - 2.5.1 Administration des utilisateurs . . . . . 44
    - 2.5.2 Administration des sources de données . . . . . 46
    - 2.5.3 Administration des services . . . . . 48
    - 2.5.4 Administration de la planification des tâches . . . . . 49

**Chapitre 2****Exemples Pentaho**

|                                                                          |    |
|--------------------------------------------------------------------------|----|
| 1. Objectifs du chapitre . . . . .                                       | 51 |
| 2. Exemples Pentaho via l'interface Pentaho User Console (PUC) . . . . . | 51 |
| 2.1 Interface Pentaho User Console (PUC) . . . . .                       | 51 |
| 2.2 Quelques exemples significatifs. . . . .                             | 52 |
| 2.2.1 Market Analysis by Year . . . . .                                  | 52 |
| 2.2.2 Product Line Analysis . . . . .                                    | 54 |
| 2.2.3 Chart examples . . . . .                                           | 56 |
| 2.2.4 Chart Pick List . . . . .                                          | 57 |
| 2.2.5 Flash Chart List . . . . .                                         | 58 |
| 2.2.6 Inventory . . . . .                                                | 59 |

**Chapitre 3****Modélisation en Informatique Décisionnelle**

|                                                                                                |    |
|------------------------------------------------------------------------------------------------|----|
| 1. Principes de modélisation . . . . .                                                         | 61 |
| 1.1 Principes de base du modèle relationnel . . . . .                                          | 61 |
| 1.2 Exemple de modèle relationnel . . . . .                                                    | 63 |
| 1.2.1 Les tables de la base de données . . . . .                                               | 63 |
| 1.2.2 Opérateur de projection . . . . .                                                        | 64 |
| 1.2.3 Opérateur de sélection . . . . .                                                         | 65 |
| 1.2.4 Opérateur de jointure . . . . .                                                          | 66 |
| 1.2.5 Opérateur d'union . . . . .                                                              | 67 |
| 1.2.6 Opérateur d'intersection . . . . .                                                       | 68 |
| 1.2.7 Opérateur de différence . . . . .                                                        | 69 |
| 1.2.8 Opérateur de produit cartésien . . . . .                                                 | 70 |
| 1.3 Pourquoi le modèle relationnel n'est pas adapté<br>à l'Informatique Décisionnelle. . . . . | 71 |
| 1.3.1 Approche opérationnelle . . . . .                                                        | 72 |
| 1.3.2 Approche décisionnelle . . . . .                                                         | 74 |

- 2. Architecture type d'un système décisionnel . . . . . 75
- 3. Modélisation en univers décisionnel . . . . . 76
  - 3.1 Notion de cube . . . . . 76
  - 3.2 Modélisation conceptuelle. . . . . 76
    - 3.2.1 Concept de fait. . . . . 77
    - 3.2.2 Concept de dimension . . . . . 77
    - 3.2.3 Concept de hiérarchie de dimension . . . . . 78
    - 3.2.4 Modèles en étoile, en flocon et en constellation . . . . . 78
    - 3.2.5 Cas particulier des dimensions à évolution lente. . . . . 80
  - 3.3 Méthodes de modélisation . . . . . 81

## Chapitre 4

### Présentation de l'ETL PDI

- 1. Principes de fonctionnement des ETL (Extract-Transform-Load) . . 83
  - 1.1 Définition générale . . . . . 83
  - 1.2 Les phases Extract-Transform-Load. . . . . 85
- 2. Installation de Pentaho Data Integration (PDI) . . . . . 86
  - 2.1 Téléchargement . . . . . 86
  - 2.2 Installation . . . . . 87
  - 2.3 Lancement de Pentaho Data Integration . . . . . 87
- 3. Présentation générale de l'ETL PDI . . . . . 88
  - 3.1 Fonctionnalités principales de PDI . . . . . 88
  - 3.2 Accès aux exemples. . . . . 89

## Chapitre 5

### Mise en œuvre de PDI

- 1. Transformation PDI n°1 . . . . . 91
  - 1.1 Présentation générale des transformations . . . . . 91
  - 1.2 Présentation des données . . . . . 91
    - 1.2.1 Structure de table *COUREUR* (198 enregistrements) . 92

|        |                                                                          |     |
|--------|--------------------------------------------------------------------------|-----|
| 1.2.2  | Structure de table <i>EQUIPE</i> (22 enregistrements) . . . . .          | 93  |
| 1.2.3  | Structure de table <i>EQUIPE_BUDGET</i><br>(22 enregistrements). . . . . | 93  |
| 1.2.4  | Structure de table <i>ETAPE</i> (21 enregistrements). . . . .            | 93  |
| 1.2.5  | Structure de table <i>PAYS</i> (26 enregistrements). . . . .             | 93  |
| 1.3    | Objectif de la transformation . . . . .                                  | 93  |
| 1.4    | Présentation détaillée de la transformation PDI . . . . .                | 94  |
| 1.4.1  | Fenêtre de lancement de PDI . . . . .                                    | 94  |
| 1.4.2  | Écran de bienvenue. . . . .                                              | 95  |
| 1.4.3  | Création d'une nouvelle transformation . . . . .                         | 95  |
| 1.4.4  | Placement des étapes dans la transformation . . . . .                    | 97  |
| 1.4.5  | Paramétrage de l'étape Extraction depuis table . . . . .                 | 100 |
| 1.4.6  | Paramétrage de l'étape Insertion dans table. . . . .                     | 103 |
| 1.4.7  | Notes explicatives. . . . .                                              | 105 |
| 1.4.8  | Sauvegarde de la transformation. . . . .                                 | 105 |
| 1.4.9  | Onglet de navigation . . . . .                                           | 106 |
| 1.4.10 | Exécution de la transformation. . . . .                                  | 106 |
| 2.     | Transformation PDI n°2 . . . . .                                         | 108 |
| 2.1    | Objectif de la transformation . . . . .                                  | 108 |
| 2.2    | Paramétrage de la source de données Microsoft SQL Server. . . . .        | 109 |
| 2.3    | Étapes et liens de la transformation . . . . .                           | 110 |
| 3.     | Transformation PDI n°3 . . . . .                                         | 111 |
| 3.1    | Objectif de la transformation . . . . .                                  | 111 |
| 3.2    | Tri . . . . .                                                            | 112 |
| 3.3    | Les étapes et liens de la transformation. . . . .                        | 113 |
| 4.     | Transformation PDI n°4 . . . . .                                         | 113 |
| 4.1    | Objectif de la transformation . . . . .                                  | 113 |
| 4.2    | Sélection . . . . .                                                      | 114 |
| 4.3    | Étapes et liens de la transformation . . . . .                           | 115 |
| 5.     | Transformation PDI n°5 . . . . .                                         | 116 |
| 5.1    | Objectif de la transformation . . . . .                                  | 116 |
| 5.2    | Étapes et liens de la transformation . . . . .                           | 117 |

|      |                                                                |     |
|------|----------------------------------------------------------------|-----|
| 6.   | Transformation PDI n°6 . . . . .                               | 118 |
| 6.1  | Objectif de la transformation . . . . .                        | 118 |
| 6.2  | Lecture du fichier CSV . . . . .                               | 119 |
| 6.3  | Insertion des données dans la table MySQL . . . . .            | 123 |
| 6.4  | Étapes et liens de la transformation . . . . .                 | 124 |
| 7.   | Transformation PDI n°7 . . . . .                               | 124 |
| 7.1  | Objectif de la transformation . . . . .                        | 124 |
| 7.2  | Script SQL . . . . .                                           | 126 |
| 7.3  | Étapes et liens de la transformation . . . . .                 | 127 |
| 8.   | Transformation PDI n°8 . . . . .                               | 127 |
| 8.1  | Objectif de la transformation . . . . .                        | 127 |
| 8.2  | Script SQL . . . . .                                           | 128 |
| 8.3  | Étapes et liens de la transformation . . . . .                 | 129 |
| 9.   | Transformation PDI n°9 . . . . .                               | 129 |
| 9.1  | Objectif de la transformation . . . . .                        | 129 |
| 9.2  | Étape de lecture d'un fichier texte de longueur fixe . . . . . | 130 |
| 9.3  | Étape d'affichage d'une trace . . . . .                        | 133 |
| 9.4  | Étapes et liens de la transformation . . . . .                 | 134 |
| 10.  | Transformation PDI n°10 . . . . .                              | 134 |
| 10.1 | Objectif de la transformation . . . . .                        | 134 |
| 10.2 | Étape d'accès à une base de données Microsoft Access. . . . .  | 135 |
| 10.3 | Étape d'agrégation de valeurs. . . . .                         | 136 |
| 10.4 | Étape d'écriture dans un fichier texte . . . . .               | 137 |
| 10.5 | Étapes et liens de la transformation . . . . .                 | 137 |
| 11.  | Transformation PDI n°11 . . . . .                              | 138 |
| 11.1 | Objectif de la transformation . . . . .                        | 138 |
| 11.2 | Étape de lecture . . . . .                                     | 138 |
| 11.3 | Étapes et liens de la transformation . . . . .                 | 139 |
| 12.  | Transformation PDI n°12 . . . . .                              | 139 |
| 12.1 | Objectif de la transformation . . . . .                        | 139 |
| 12.2 | Étape de lecture . . . . .                                     | 140 |
| 12.3 | Étape d'écriture dans un classeur Microsoft Excel . . . . .    | 140 |

|      |                                                 |     |
|------|-------------------------------------------------|-----|
| 12.4 | Étapes et liens de la transformation            | 141 |
| 13.  | Transformation PDI n°13                         | 141 |
| 13.1 | Objectif de la transformation                   | 141 |
| 13.2 | Étape de lecture d'un classeur Microsoft Excel. | 142 |
| 13.3 | Étapes et liens de la transformation            | 143 |
| 14.  | Transformation PDI n°14                         | 143 |
| 14.1 | Objectif de la transformation                   | 143 |
| 14.2 | Fichier XML en lecture                          | 144 |
| 14.3 | Étapes et liens de la transformation            | 147 |
| 15.  | Transformation PDI n°15                         | 147 |
| 15.1 | Objectif de la transformation                   | 147 |
| 15.2 | Recherche                                       | 148 |
| 15.3 | Étapes et liens de la transformation            | 150 |
| 16.  | Transformation PDI n°16                         | 150 |
| 16.1 | Objectif de la transformation                   | 150 |
| 16.2 | Jointure.                                       | 151 |
| 16.3 | Étapes et liens de la transformation            | 152 |
| 17.  | Transformation PDI n°17                         | 152 |
| 17.1 | Objectif de la transformation                   | 152 |
| 17.2 | Listage des noms des sous-répertoires           | 153 |
| 17.3 | Étapes et liens de la transformation            | 154 |
| 18.  | Transformation PDI n°18                         | 154 |
| 18.1 | Objectif de la transformation                   | 154 |
| 18.2 | Listage des noms des fichiers d'un répertoire   | 155 |
| 18.3 | Étapes et liens de la transformation            | 155 |
| 19.  | Tâche PDI n°19                                  | 156 |
| 19.1 | Objectif de la tâche.                           | 156 |
| 19.2 | Récupération depuis un serveur FTP (download).  | 157 |
| 19.3 | Étapes et liens de la tâche                     | 159 |
| 20.  | Tâche PDI n°20                                  | 160 |
| 20.1 | Objectif de la tâche.                           | 160 |

- 20.2 Transfert vers un serveur FTP (upload) . . . . . 160
- 20.3 Étapes et liens de la tâche . . . . . 162
- 21. Transformation PDI n°21 . . . . . 162
  - 21.1 Objectif de la transformation . . . . . 162
  - 21.2 Clonage. . . . . 163
  - 21.3 Étapes et liens de la tâche . . . . . 164
- 22. Transformation PDI n°22 . . . . . 164
  - 22.1 Objectif de la transformation . . . . . 164
  - 22.2 Dédoublonnage. . . . . 165
  - 22.3 Étapes et liens de la tâche . . . . . 166
- 23. Transformation PDI n°23 . . . . . 166
  - 23.1 Objectif de la transformation . . . . . 166
  - 23.2 Script SQL de création de la table EQUIPE\_BUDGET23 . . . 167
  - 23.3 Champ calculé . . . . . 167
  - 23.4 Étapes et liens de la tâche . . . . . 168
  - 23.5 Contenu de la table EQUIPE\_BUDGET23 après exécution . 169
- 24. Transformation PDI n°24 . . . . . 169
  - 24.1 Objectif de la transformation . . . . . 169
  - 24.2 Éclatement . . . . . 170
  - 24.3 Étapes et liens de la tâche . . . . . 171

**Chapitre 6**

**Présentation de WAQR**

- 1. Solutions de génération de rapports sous Pentaho. . . . . 173
- 2. Présentation du générateur de rapports intégrés WAQR. . . . . 174

## Chapitre 7

### Mise en œuvre de WAQR

|                                                                 |     |
|-----------------------------------------------------------------|-----|
| 1. Liste des points de vente . . . . .                          | 175 |
| 1.1 Objectif. . . . .                                           | 175 |
| 1.2 Présentation de la base de données SampleData . . . . .     | 176 |
| 1.3 Accès à l'interface de Pentaho User Console (PUC) . . . . . | 176 |
| 1.4 Création d'une liste des points de vente . . . . .          | 179 |
| 2. Liste des employés . . . . .                                 | 183 |
| 2.1 Objectif. . . . .                                           | 183 |
| 2.2 Conception du rapport via l'assistant WAQR . . . . .        | 184 |
| 2.3 Compte-rendu d'exécution . . . . .                          | 184 |
| 3. Liste des règlements de clients . . . . .                    | 185 |
| 3.1 Objectif. . . . .                                           | 185 |
| 3.2 Conception du rapport via l'assistant WAQR . . . . .        | 185 |
| 3.3 Compte-rendu d'exécution . . . . .                          | 186 |
| 4. Liste des règlements de clients (France) . . . . .           | 186 |
| 4.1 Objectif. . . . .                                           | 186 |
| 4.2 Conception du rapport via l'assistant WAQR . . . . .        | 186 |
| 4.3 Compte-rendu d'exécution . . . . .                          | 187 |
| 5. Liste des produits . . . . .                                 | 187 |
| 5.1 Objectif. . . . .                                           | 187 |
| 5.2 Conception du rapport via l'assistant WAQR . . . . .        | 188 |
| 5.3 Compte-rendu d'exécution . . . . .                          | 188 |

## Chapitre 8

### Présentation de PRD

|                                                                     |     |
|---------------------------------------------------------------------|-----|
| 1. Présentation du générateur de rapports Pentaho Report Designer . | 189 |
| 2. Installation du package Pentaho Report Designer . . . . .        | 191 |



**Chapitre 9**

**Mise en œuvre de PRD**

- 1. Un premier rapport basé sur des données MySQL . . . . . 193
  - 1.1 Objectif. . . . . 193
  - 1.2 Présentation de la base de données classic\_models. . . . . 194
  - 1.3 Construction du rapport via l'assistant . . . . . 195
  - 1.4 Exécution du rapport généré par l'assistant . . . . . 204
  - 1.5 Amélioration de la présentation  
du rapport généré par l'assistant . . . . . 205
  - 1.6 Compte-rendu d'exécution . . . . . 207
- 2. Tri des données . . . . . 207
  - 2.1 Mise en place du tri . . . . . 207
  - 2.2 Compte-rendu d'exécution . . . . . 208
- 3. Regroupement des données. . . . . 209
  - 3.1 Mise en place du regroupement. . . . . 209
  - 3.2 Compte-rendu d'exécution . . . . . 211
- 4. Sélection statique des données . . . . . 212
  - 4.1 Mise en place de la sélection statique . . . . . 212
  - 4.2 Compte-rendu d'exécution . . . . . 213
- 5. Sélection dynamique des données  
(paramètre zone de texte). . . . . 213
  - 5.1 Mise en place de la sélection dynamique. . . . . 213
  - 5.2 Compte-rendu d'exécution . . . . . 216
- 6. Sélection dynamique des données  
(paramètre liste déroulante) . . . . . 216
  - 6.1 Mise en place de la sélection dynamique. . . . . 216
  - 6.2 Compte-rendu d'exécution . . . . . 219
- 7. Sélection dynamique des données  
(paramètre case à cocher) . . . . . 220
- 8. Un rapport basé sur des données SQL Server . . . . . 222
  - 8.1 Objectif. . . . . 222
  - 8.2 Définition de la Data Source et du Query . . . . . 222

|      |                                                             |     |
|------|-------------------------------------------------------------|-----|
| 8.3  | Compte-rendu d'exécution . . . . .                          | 225 |
| 9.   | Formatage des dates et des montants . . . . .               | 225 |
| 9.1  | Objectif. . . . .                                           | 225 |
| 9.2  | Définition du Query . . . . .                               | 226 |
| 9.3  | Alignement des éléments . . . . .                           | 226 |
| 9.4  | Formatage des montants. . . . .                             | 227 |
| 9.5  | Formatage des dates . . . . .                               | 229 |
| 9.6  | Compte-rendu d'exécution . . . . .                          | 230 |
| 10.  | Totalisation en pied de groupe . . . . .                    | 230 |
| 10.1 | Objectif. . . . .                                           | 230 |
| 10.2 | Définition de la Data Source et du Query . . . . .          | 231 |
| 10.3 | Choix des champs et du regroupement . . . . .               | 231 |
| 10.4 | Choix des champs et du regroupement . . . . .               | 231 |
| 10.5 | Affichage du rapport à la fin de la phase assistée. . . . . | 232 |
| 10.6 | Mise en place d'un cumul de montant par client. . . . .     | 233 |
| 10.7 | Compte-rendu d'exécution . . . . .                          | 234 |
| 11.  | Champs calculés . . . . .                                   | 234 |
| 11.1 | Objectif. . . . .                                           | 234 |
| 11.2 | Mise en place d'un champ calculé . . . . .                  | 235 |
| 11.3 | Compte-rendu d'exécution . . . . .                          | 238 |
| 12.  | Connexion à une source de données de type table . . . . .   | 238 |
| 12.1 | Objectif. . . . .                                           | 238 |
| 12.2 | Définition de la Data Source et du Query . . . . .          | 240 |
| 12.3 | Compte-rendu d'exécution . . . . .                          | 243 |
| 13.  | Utilisation d'une jointure SQL dans le Query . . . . .      | 243 |
| 13.1 | Objectif. . . . .                                           | 243 |
| 13.2 | Définition du Query . . . . .                               | 245 |
| 13.3 | Regroupement sur le champ officeCode . . . . .              | 245 |
| 13.4 | Compte-rendu d'exécution . . . . .                          | 245 |
| 14.  | Sous-rapports . . . . .                                     | 246 |
| 14.1 | Objectif. . . . .                                           | 246 |
| 14.2 | Définition du Query du rapport principal . . . . .          | 246 |

- 14.3 Sélection des champs et regroupement pour le rapport principal . . . . . 246
- 14.4 Compte-rendu d'exécution intermédiaire du rapport principal . . . . . 247
- 14.5 Intégration du sous-rapport dans le rapport principal . . . . . 247
- 14.6 Définition du Query du sous-rapport . . . . . 248
- 14.7 Placement des champs dans le sous-rapport . . . . . 248
- 14.8 Compte-rendu d'exécution intermédiaire du rapport principal . . . . . 250
- 14.9 Lien entre le rapport principal et le sous-rapport . . . . . 250
- 14.10 Intégration du paramètre point\_vente dans la requête SQL . . . . . 252
- 14.11 Compte-rendu d'exécution intermédiaire du rapport principal . . . . . 252
- 14.12 Gestion des couleurs . . . . . 253
- 14.13 Compte-rendu d'exécution final du rapport . . . . . 254
- 15. Publication des rapports dans le portail Web Pentaho . . . . . 254
  - 15.1 Objectif . . . . . 254
  - 15.2 Mot de passe de publication et authentification de l'utilisateur . . . . . 255
  - 15.3 Lancement du serveur Web Pentaho BI Server . . . . . 256
  - 15.4 Publication d'un rapport depuis Pentaho Report Designer . . . . . 257
  - 15.5 Contrôle de la publication sur le serveur Web Pentaho BI Server . . . . . 259
- 16. Liens hypertextes dans les rapports (Drill Down) . . . . . 260
  - 16.1 Objectif . . . . . 260
  - 16.2 Structure du rapport principal . . . . . 261
  - 16.3 Définition du Query du rapport principal . . . . . 261
  - 16.4 Compte-rendu d'exécution intermédiaire du rapport principal . . . . . 262
  - 16.5 Structure du rapport secondaire . . . . . 262
  - 16.6 Définition du Query du rapport secondaire . . . . . 263
  - 16.7 Compte-rendu d'exécution intermédiaire du rapport secondaire . . . . . 263

|       |                                                                                     |     |
|-------|-------------------------------------------------------------------------------------|-----|
| 16.8  | Mise en place du paramètre dans le rapport secondaire . . . .                       | 264 |
| 16.9  | Intégration du paramètre dans le Query<br>du rapport secondaire . . . . .           | 265 |
| 16.10 | Mise en place du lien hypertexte dans le rapport principal .                        | 265 |
| 16.11 | Publication des deux rapports<br>sur le serveur Web Pentaho BI Server . . . . .     | 266 |
| 16.12 | Contrôle des publications sur le serveur<br>Web Pentaho BI Server . . . . .         | 267 |
| 17.   | Conception d'un graphique de type Secteurs. . . . .                                 | 268 |
| 17.1  | Objectif. . . . .                                                                   | 268 |
| 17.2  | Structure du rapport n°1 . . . . .                                                  | 269 |
| 17.3  | Définition du Query du rapport n°1 . . . . .                                        | 269 |
| 17.4  | Sélection des champs, regroupement et<br>totalisation pour le rapport n°1 . . . . . | 270 |
| 17.5  | Compte-rendu d'exécution du rapport n°1. . . . .                                    | 271 |
| 17.6  | Structure du rapport n°2 . . . . .                                                  | 271 |
| 17.7  | Compte-rendu d'exécution du rapport n°2. . . . .                                    | 272 |
| 17.8  | Structure du rapport n°3 . . . . .                                                  | 272 |
| 17.9  | Définition du Query du rapport n°3 . . . . .                                        | 272 |
| 17.10 | Élaboration d'un graphique en Secteurs. . . . .                                     | 273 |
| 17.11 | Compte-rendu d'exécution du rapport n°3. . . . .                                    | 274 |

## Chapitre 10

### Présentation de Mondrian/JPivot

|     |                                                                          |     |
|-----|--------------------------------------------------------------------------|-----|
| 1.  | Présentation générale . . . . .                                          | 275 |
| 1.1 | Architecture simplifiée de Mondrian . . . . .                            | 275 |
| 1.2 | Serveur Mondrian . . . . .                                               | 276 |
| 1.3 | Interface de restitution JPivot . . . . .                                | 277 |
| 2.  | Installation de Mondrian/JPivot . . . . .                                | 278 |
| 2.1 | Mise en place d'un serveur J2EE Tomcat . . . . .                         | 278 |
| 2.2 | Déploiement de l'application Mondrian<br>sur le serveur Tomcat . . . . . | 281 |

**Chapitre 11**

**Mise en œuvre de Mondrian/JPivot**

- 1. Présentation de la base de données . . . . . 285
  - 1.1 La base de données AdventureWorks. . . . . 285
  - 1.2 Structure des tables utilisées . . . . . 286
    - 1.2.1 Accès aux structures détaillées des tables . . . . . 286
    - 1.2.2 Structure de table factresellersales (faits) . . . . . 287
    - 1.2.3 Structure de table dimreseller . . . . . 287
    - 1.2.4 Structure de table dimproduct . . . . . 287
    - 1.2.5 Structure de table dimproductsubcategory . . . . . 288
    - 1.2.6 Structure de table dimproductcategory. . . . . 288
    - 1.2.7 Structure de table dimtime. . . . . 288
  - 1.3 Les requêtes MDX et les schémas associés. . . . . 289
  - 1.4 Paramétrages complémentaires et recommandations. . . . . 289
    - 1.4.1 Paramétrages complémentaires . . . . . 289
    - 1.4.2 Recommandations . . . . . 291
- 2. Les requêtes MDX . . . . . 292
  - 2.1 Requête n°1. . . . . 292
    - 2.1.1 Objectif . . . . . 292
    - 2.1.2 Schéma Mondrian  
(script mondrian\_adventureworks\_schema\_01.xml) . . . 292
    - 2.1.3 Commentaires . . . . . 293
    - 2.1.4 Contraintes de nommage . . . . . 293
    - 2.1.5 Nom du schéma Mondrian. . . . . 293
    - 2.1.6 Nom des cubes dans le schéma Mondrian. . . . . 294
    - 2.1.7 Table des faits dans le schéma Mondrian . . . . . 294
    - 2.1.8 Déclaration des mesures non calculées . . . . . 294
    - 2.1.9 Requête MDX  
(script mondrian\_adventureworks\_mdx\_01.jsp) . . . . 294
    - 2.1.10 Commentaires . . . . . 296
    - 2.1.11 Blocs d'instructions obligatoires . . . . . 296
    - 2.1.12 Code MDX. . . . . 297
    - 2.1.13 Résultat de l'exécution sous JPivot . . . . . 298

|       |                                                             |     |
|-------|-------------------------------------------------------------|-----|
| 2.2   | Requête n°2. . . . .                                        | 298 |
| 2.2.1 | Objectif . . . . .                                          | 298 |
| 2.2.2 | Schéma Mondrian . . . . .                                   | 298 |
| 2.2.3 | Définition des dimensions dans le schéma . . . . .          | 299 |
| 2.2.4 | Déclaration des dimensions dans le cube . . . . .           | 300 |
| 2.2.5 | Requête MDX . . . . .                                       | 301 |
| 2.2.6 | Résultat de l'exécution sous JPivot . . . . .               | 301 |
| 2.3   | Requête n°3. . . . .                                        | 302 |
| 2.3.1 | Objectif . . . . .                                          | 302 |
| 2.3.2 | Schéma Mondrian . . . . .                                   | 302 |
| 2.3.3 | Définition des dimensions dans le schéma . . . . .          | 303 |
| 2.3.4 | Requête MDX . . . . .                                       | 303 |
| 2.3.5 | Résultat de l'exécution sous JPivot . . . . .               | 304 |
| 2.4   | Requête n°4. . . . .                                        | 304 |
| 2.4.1 | Objectif . . . . .                                          | 304 |
| 2.4.2 | Schéma Mondrian . . . . .                                   | 304 |
| 2.4.3 | Requête MDX . . . . .                                       | 305 |
| 2.4.4 | Résultat de l'exécution sous JPivot . . . . .               | 305 |
| 2.5   | Requête n°5. . . . .                                        | 306 |
| 2.5.1 | Objectif . . . . .                                          | 306 |
| 2.5.2 | Schéma Mondrian . . . . .                                   | 306 |
| 2.5.3 | Requête MDX . . . . .                                       | 306 |
| 2.5.4 | Résultat de l'exécution sous JPivot . . . . .               | 307 |
| 2.6   | Requête n°6. . . . .                                        | 307 |
| 2.6.1 | Objectif . . . . .                                          | 307 |
| 2.6.2 | Schéma Mondrian . . . . .                                   | 307 |
| 2.6.3 | Requête MDX . . . . .                                       | 308 |
| 2.6.4 | Résultat de l'exécution sous JPivot . . . . .               | 309 |
| 2.7   | Requête n°7. . . . .                                        | 309 |
| 2.7.1 | Objectif . . . . .                                          | 309 |
| 2.7.2 | Schéma Mondrian . . . . .                                   | 310 |
| 2.7.3 | Éléments vus dans le cadre des précédents schémas . . . . . | 310 |
| 2.7.4 | Dimension Produits . . . . .                                | 311 |

|        |                                                       |     |
|--------|-------------------------------------------------------|-----|
| 2.7.5  | Requête MDX                                           | 311 |
| 2.7.6  | Résultat de l'exécution sous JPivot                   | 312 |
| 2.8    | Requête n°8.                                          | 312 |
| 2.8.1  | Objectif                                              | 312 |
| 2.8.2  | Schéma Mondrian                                       | 313 |
| 2.8.3  | Requête MDX                                           | 313 |
| 2.8.4  | Résultat de l'exécution sous JPivot                   | 314 |
| 2.9    | Requête n°9.                                          | 314 |
| 2.9.1  | Objectif                                              | 314 |
| 2.9.2  | Schéma Mondrian                                       | 315 |
| 2.9.3  | Spécificités de la gestion de la dimension temporelle | 315 |
| 2.9.4  | Requête MDX                                           | 316 |
| 2.9.5  | Résultat de l'exécution sous JPivot                   | 317 |
| 2.10   | Requête n°10.                                         | 317 |
| 2.10.1 | Objectif                                              | 317 |
| 2.10.2 | Schéma Mondrian                                       | 317 |
| 2.10.3 | Requête MDX                                           | 318 |
| 2.10.4 | Résultat de l'exécution sous JPivot                   | 318 |
| 2.11   | Requête n°11.                                         | 319 |
| 2.11.1 | Objectif                                              | 319 |
| 2.11.2 | Schéma Mondrian                                       | 319 |
| 2.11.3 | Requête MDX                                           | 320 |
| 2.11.4 | Résultat de l'exécution sous JPivot                   | 320 |
| 2.12   | Requête n°12.                                         | 321 |
| 2.12.1 | Objectif                                              | 321 |
| 2.12.2 | Schéma Mondrian                                       | 321 |
| 2.12.3 | Requête MDX                                           | 321 |
| 2.12.4 | Résultat de l'exécution sous JPivot                   | 322 |
| 2.13   | Requête n°13.                                         | 322 |
| 2.13.1 | Objectif                                              | 322 |
| 2.13.2 | Schéma Mondrian                                       | 323 |
| 2.13.3 | Requête MDX                                           | 323 |
| 2.13.4 | Résultat de l'exécution sous JPivot                   | 323 |

|        |                                               |     |
|--------|-----------------------------------------------|-----|
| 2.14   | Requête n°14. . . . .                         | 324 |
| 2.14.1 | Objectif . . . . .                            | 324 |
| 2.14.2 | Schéma Mondrian . . . . .                     | 324 |
| 2.14.3 | Requête MDX . . . . .                         | 324 |
| 2.14.4 | Résultat de l'exécution sous JPivot . . . . . | 325 |
| 2.15   | Requête n°15. . . . .                         | 325 |
| 2.15.1 | Objectif . . . . .                            | 325 |
| 2.15.2 | Schéma Mondrian . . . . .                     | 325 |
| 2.15.3 | Requête MDX . . . . .                         | 326 |
| 2.15.4 | Résultat de l'exécution sous JPivot . . . . . | 326 |
| 2.16   | Requête n°16. . . . .                         | 327 |
| 2.16.1 | Objectif . . . . .                            | 327 |
| 2.16.2 | Schéma Mondrian . . . . .                     | 327 |
| 2.16.3 | Requête MDX . . . . .                         | 328 |
| 2.16.4 | Résultat de l'exécution sous JPivot . . . . . | 329 |
| 2.17   | Requête n°17. . . . .                         | 329 |
| 2.17.1 | Objectif . . . . .                            | 329 |
| 2.17.2 | Schéma Mondrian . . . . .                     | 330 |
| 2.17.3 | Requête MDX . . . . .                         | 331 |
| 2.17.4 | Résultat de l'exécution sous JPivot . . . . . | 331 |
| 2.18   | Requête n°18. . . . .                         | 332 |
| 2.18.1 | Objectif . . . . .                            | 332 |
| 2.18.2 | Schéma Mondrian . . . . .                     | 332 |
| 2.18.3 | Requête MDX . . . . .                         | 333 |
| 2.18.4 | Résultat de l'exécution sous JPivot . . . . . | 333 |
| 2.19   | Requête n°19. . . . .                         | 334 |
| 2.19.1 | Objectif . . . . .                            | 334 |
| 2.19.2 | Schéma Mondrian . . . . .                     | 334 |
| 2.19.3 | Requête MDX . . . . .                         | 335 |
| 2.19.4 | Résultat de l'exécution sous JPivot . . . . . | 335 |
| 2.20   | Requête n°20. . . . .                         | 336 |
| 2.20.1 | Objectif . . . . .                            | 336 |
| 2.20.2 | Schéma Mondrian . . . . .                     | 336 |



- 2.20.3 Requête MDX . . . . . 337
- 2.20.4 Résultat de l'exécution sous JPivot . . . . . 338
- 3. Les principales fonctions MDX . . . . . 338
  - 3.1 Fonctions MDX de gestion des hiérarchies . . . . . 339
  - 3.2 Fonctions MDX de gestion des niveaux . . . . . 339
  - 3.3 Fonctions MDX de gestion des membres. . . . . 339
  - 3.4 Fonctions numériques MDX . . . . . 340
  - 3.5 Fonctions MDX de gestion de sets . . . . . 340

**Chapitre 12**

**Modélisation Mondrian via PSW**

- 1. Aide la modélisation par Pentaho Schema Workbench (PSW) . . . . 341
  - 1.1 Installation . . . . . 342
  - 1.2 Lancement de Pentaho Schema Workbench . . . . . 342
- 2. Premier exemple de modélisation d'un schéma Mondrian via PSW . . . . . 342
  - 2.1 Fenêtre initiale de PSW . . . . . 342
  - 2.2 Chargement du schéma élaboré au chapitre  
Mise en œuvre de Mondrian/JPivot. . . . . 343
  - 2.3 Contrôle de l'accès à la base de données . . . . . 345
  - 2.4 Test local du schéma sous PSW. . . . . 346
- 3. Deuxième exemple de modélisation d'un schéma Mondrian via PSW . . . . . 347
  - 3.1 Présentation du schéma Schema\_AdventureWorks\_21 . . . . 347
  - 3.2 Requête MDX utilisable sur le schéma  
Schema\_AdventureWorks\_21 . . . . . 348
  - 3.3 Résultat de l'exécution sous JPivot . . . . . 349
  - 3.4 Création du schéma Schema\_AdventureWorks\_21 . . . . . 349
  - 3.5 Création du cube Ventes indirectes . . . . . 350
  - 3.6 Définition de la table des faits associée au cube . . . . . 351
  - 3.7 Définition des mesures [CA] et [Frais de port] . . . . . 352

|      |                                                                                          |     |
|------|------------------------------------------------------------------------------------------|-----|
| 3.8  | Définition de la mesure calculée<br>[Pourcentage frais de port] . . . . .                | 353 |
| 3.9  | Définition de la dimension [Types de grossiste] . . . . .                                | 354 |
| 3.10 | Définition de la hiérarchie de la dimension<br>[Types de grossiste]. . . . .             | 356 |
| 3.11 | Définition des niveaux de la hiérarchie de la dimension<br>[Types de grossiste]. . . . . | 357 |
| 3.12 | Déclaration de la table associée<br>à la dimension [Types de grossiste] . . . . .        | 358 |
| 3.13 | Définition de la dimension [Temps] . . . . .                                             | 359 |
| 3.14 | Définition de la hiérarchie [Hiérarchie temps]. . . . .                                  | 360 |
| 3.15 | Définition des niveaux de la hiérarchie<br>de la dimension [Temps]. . . . .              | 360 |
| 3.16 | Déclaration de la table associée à la dimension [Temps] . . . . .                        | 362 |
| 3.17 | Utilisation des dimensions dans le cube . . . . .                                        | 363 |
| 3.18 | Contrôle final du schéma . . . . .                                                       | 366 |
| 3.19 | Accès au code source du schéma . . . . .                                                 | 367 |
| 4.   | Intégration de Mondrian dans le portail Pentaho . . . . .                                | 368 |
| 4.1  | Publication du schéma Mondrian<br>dans le portail depuis PSW . . . . .                   | 368 |
| 4.2  | Mise en place d'une analyse Mondrian<br>dans le portail Pentaho . . . . .                | 373 |
| 5.   | Présentation de l'interface de navigation JPivot . . . . .                               | 376 |
| 5.1  | Exemple support de la présentation de JPivot . . . . .                                   | 376 |
| 5.2  | Barre d'icônes . . . . .                                                                 | 377 |
| 5.3  | Éditeur MDX . . . . .                                                                    | 377 |
| 5.4  | Accès au cube . . . . .                                                                  | 379 |

## Chapitre 13

### Agrégation de données via PAD

|     |                                                       |     |
|-----|-------------------------------------------------------|-----|
| 1.  | Présentation de Pentaho Aggregator Designer . . . . . | 387 |
| 1.1 | Principe général. . . . .                             | 387 |

- 1.2 Installation du package Pentaho Aggregator Designer . . . . . 388
- 2. Optimisation d'un schéma Mondrian . . . . . 388
  - 2.1 Présentation du jeu de données . . . . . 388
  - 2.2 Présentation du schéma Mondrian à optimiser . . . . . 389
  - 2.3 Optimisation sous Pentaho Aggregator Designer . . . . . 390

**Chapitre 14**

**Intégration des composants via PDS**

- 1. Présentation de Pentaho Design Studio . . . . . 399
- 2. Installation de Pentaho Design Studio . . . . . 400
  - 2.1 Téléchargement . . . . . 400
  - 2.2 Installation . . . . . 401
  - 2.3 Lancement de Pentaho Design Studio . . . . . 401
- 3. Présentation rapide de l'environnement initial . . . . . 402
  - 3.1 Le répertoire workspace . . . . . 402
  - 3.2 Les fenêtres Eclipse . . . . . 403
  - 3.3 La création du projet Pentaho Solutions . . . . . 403
- 4. Étude d'une première action pré-enregistrée . . . . . 405
  - 4.1 Le dossier de l'action HelloWorld . . . . . 406
  - 4.2 Le code de l'action HelloWorld . . . . . 406
  - 4.3 L'action HelloWorld vue au travers de Pentaho Design Studio . . . . . 407
- 5. Création d'une première action . . . . . 411
- 6. Structure d'une action . . . . . 414
  - 6.1 Inputs . . . . . 415
  - 6.2 Outputs . . . . . 416
  - 6.3 Process Actions . . . . . 416
- 7. Une action listant des données extraites d'une table MySQL . . . . . 416

## Chapitre 15

### Intégration de rapports BIRT

1. Générateur de rapports BIRT ..... 421
  - 1.1 Présentation générale ..... 421
  - 1.2 Architecture de BIRT ..... 422
  - 1.3 Installation de BIRT ..... 424
  - 1.4 Intégration du BIRT Viewer dans Pentaho ..... 424
    - 1.4.1 Étape 1 : Déploiement de BIRT Viewer 2.6.1 dans Pentaho 3.6 ..... 424
    - 1.4.2 Étape 2 : Déploiement d'un plug-in sur le serveur Pentaho. .... 427
    - 1.4.3 Étape 3 : Gestion des drivers JDBC sur le serveur Pentaho. .... 427
    - 1.4.4 Étape 4 : Déploiement des rapports BIRT ..... 428
2. Élaboration de rapports BIRT ..... 428
  - 2.1 Liste des points de vente (données MySQL) ..... 428
    - 2.1.1 Base de données utilisée ..... 428
    - 2.1.2 Construction du rapport en mode assisté ou non ... 429
    - 2.1.3 Objectif du rapport. .... 429
    - 2.1.4 Le paramétrage du lien à la base de données MySQL classic\_models. .... 429
    - 2.1.5 Création d'un nouveau projet sous BIRT ..... 432
    - 2.1.6 Création d'un nouveau rapport sous BIRT ..... 433
    - 2.1.7 Création de la Data Source (connexion aux données) 433
    - 2.1.8 Création de la Data Set (jeu de données) ..... 434
    - 2.1.9 Mise en page du rapport. .... 436
    - 2.1.10 Prévisualisation du rapport ..... 439
  - 2.2 Liste des employés (données MySQL) ..... 439
    - 2.2.1 Base de données utilisée ..... 439
    - 2.2.2 Objectif du rapport. .... 440
    - 2.2.3 Création de la Data Source (connexion aux données) 440
    - 2.2.4 Création de la Data Set (jeu de données) ..... 441
    - 2.2.5 Mise en page du rapport. .... 442

- 2.3 Liste des employés version 2 (données MySQL) . . . . . 444
  - 2.3.1 Objectif du rapport. . . . . 444
  - 2.3.2 Création de la Data Set (jeu de données) . . . . . 444
  - 2.3.3 Mise en page du rapport. . . . . 445
- 2.4 Liste des paiements (données MySQL) . . . . . 446
  - 2.4.1 Objectif du rapport. . . . . 446
  - 2.4.2 Création de la Data Set (jeu de données) . . . . . 446
  - 2.4.3 Mise en page du rapport. . . . . 447
- 2.5 Liste des employés version 3 (données MySQL) . . . . . 449
  - 2.5.1 Objectif du rapport. . . . . 449
  - 2.5.2 Création de la Data Set (jeu de données) . . . . . 449
  - 2.5.3 Mise en page du rapport. . . . . 449
  - 2.5.4 Prévisualisation du rapport (saisie du paramètre) . . . 451
- 3. Déploiement de rapports BIRT sous Pentaho . . . . . 452
  - 3.1 Récupération des rapports sous Eclipse BIRT . . . . . 452
  - 3.2 Stockage des rapports dans le répertoire de publication Pentaho . . . . . 452
  - 3.3 Lancement des rapports . . . . . 454

**Chapitre 16**

**Analyse de données via Weka**

- 1. Présentation générale de Weka . . . . . 455
  - 1.1 Fonctionnalités de Weka . . . . . 455
  - 1.2 Installation de Weka . . . . . 456
  - 1.3 Espace de travail de Weka . . . . . 457
  - 1.4 Le format de fichier ARFF . . . . . 460
- 2. Traitements Weka . . . . . 462
  - 2.1 Arbre de décision. . . . . 462
    - 2.1.1 Approche statistique. . . . . 462
    - 2.1.2 Les données à traiter. . . . . 462
    - 2.1.3 Accès aux données sous Weka . . . . . 463
    - 2.1.4 Sélection des attributs . . . . . 466

|       |                                                                         |     |
|-------|-------------------------------------------------------------------------|-----|
| 2.1.5 | Visualisation des nuages de points . . . . .                            | 468 |
| 2.1.6 | Filtrage des données . . . . .                                          | 469 |
| 2.1.7 | Définition d'un arbre de décision . . . . .                             | 472 |
| 2.2   | Régression. . . . .                                                     | 476 |
| 2.2.1 | Approche statistique. . . . .                                           | 476 |
| 2.2.2 | Les données à traiter. . . . .                                          | 477 |
| 2.2.3 | Accès aux données sous Weka . . . . .                                   | 478 |
| 2.2.4 | Définition d'une régression linéaire. . . . .                           | 478 |
| 3.    | Utilisation du workflow de Weka. . . . .                                | 479 |
| 4.    | Connexion de Weka à des sources de données JDBC . . . . .               | 486 |
| 4.1   | Installation du pilote d'accès . . . . .                                | 486 |
| 4.2   | Correspondances des types de données . . . . .                          | 487 |
| 4.3   | Constitution de la base de données de test . . . . .                    | 488 |
| 4.4   | Utilisation de Weka à partir d'une base de données<br>via JDBC. . . . . | 489 |

## Chapitre 17

### Gestion des métadonnées via PME

|     |                                                           |     |
|-----|-----------------------------------------------------------|-----|
| 1.  | Présentation de Pentaho Metadata Editor . . . . .         | 493 |
| 1.1 | Principe général. . . . .                                 | 493 |
| 1.2 | Installation du package Pentaho Metadata Editor . . . . . | 494 |
| 2.  | Construction d'une vue métier sous PME . . . . .          | 494 |
| 2.1 | Ouverture du modèle metadata.xmi . . . . .                | 494 |
| 2.2 | Connexion à une source de données . . . . .               | 497 |
| 2.3 | La gestion des droits d'accès . . . . .                   | 498 |
| 2.4 | Définition d'un Business Model (vue métier). . . . .      | 500 |
| 2.5 | Référencement des tables dans la vue métier. . . . .      | 501 |
| 2.6 | Mise en place des relations entre les tables . . . . .    | 502 |
| 2.7 | Mise en place des catégories . . . . .                    | 503 |
| 2.8 | Publication sur le serveur Pentaho. . . . .               | 505 |

- 3. Utilisation de la vue métier sous WAQR ..... 506
  - 3.1 Ouverture de la console PAC. .... 506
  - 3.2 Création rapide d'un rapport ..... 508
- 4. Utilisation de la vue métier sous Pentaho Report Designer (PRD). 509
  - 4.1 Lancement de PRD ..... 509
  - 4.2 Conception d'un rapport PRD basé sur la vue métier AdventureWorks ..... 510

**Annexes**

- Glossaire ..... 513
  
- Index ..... 521

