

# Guide Utilisateur de l'Application Data Extractor



التجاري وفا بنك  
Attijariwafa bank

Application développée et Guide rédigé par Larabi Ilyas, Stagiaire élève ingénieur chez Attijariwafa Bank.

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Prérequis</b>	<b>3</b>
2.1	Exigences logicielles .....	3
2.2	Exigences système .....	3
<b>3</b>	<b>Installation et configuration</b>	<b>5</b>
3.1	Téléchargement de l'application .....	5
3.2	Configuration de pbi-tools.....	6
<b>4</b>	<b>Utilisation de Data Extractor</b>	<b>6</b>
4.1	Lancement de l'application .....	6
4.2	Sélection d'un fichier .pbix.....	6
4.3	Traitement du fichier .....	6
4.4	Visualisation des sorties .....	7
<b>5</b>	<b>Dépannage</b>	<b>7</b>
5.1	Erreurs courantes .....	7
5.2	Messages d'erreur.....	7
5.3	Consultez la documentation.....	8
<b>6</b>	<b>Notes supplémentaires</b>	<b>8</b>

# 1 Introduction

Ce guide présente l'utilisation d'une application développée pour automatiser l'extraction des données granulaires et des indicateurs clés de performance (KPIs) à partir d'un fichier PBIX (Power BI Desktop). L'objectif est de faciliter l'analyse et l'exploitation des éléments internes du modèle de données sans passer par Power BI.

L'application génère deux fichiers Excel distincts :

`Extracted_Data.xlsx` :

Contient deux feuilles :

Données granulaires : extrait les tables et colonnes depuis le fichier `DataModelSchema.json`.

KPIs : liste les indicateurs extraits depuis le fichier `Layout.json`.

`Data_Structure.xlsx` :

Regroupe l'intégralité des informations disponibles dans le `DataModelSchema.json`, offrant une vision complète de la structure du modèle.

Ce processus permet un accès rapide, structuré et exploitable aux données contenues dans un fichier PBIX, en réponse à des besoins d'analyse, d'audit ou de documentation du modèle.

## 2 Prérequis

Avant d'utiliser Data Extractor, assurez-vous que les exigences suivantes sont remplies :

### 2.1 Exigences logicielles

- **Power BI Desktop** : Assurez-vous que la dernière version est installée, car L'application peut échouer avec des fichiers `.pbix` obsolètes.
- **Microsoft Excel** : Nécessaire pour visualiser les fichiers de sortie générés. Assurez-vous qu'Excel est fermé avant de lancer l'application pour éviter les erreurs.
- **pbi-tools et pbi-tools.core** : Téléchargez les deux exécutables (`pbi-tools.exe` et `pbi-tools.core.exe`) depuis la source officielle et ajoutez les dans `PATH` puis placez-les dans le dossier Bureau ou Téléchargements pour la première exécution.  
L'application les déplacera dans le sous-dossier `pbi-tools` du répertoire Data Extractor par la suite.

### 2.2 Exigences système

- **Système d'exploitation** : Windows 10 ou ultérieur (recommandé).
- **Python 3.13** : Environnement d'exécution utilisé pour développer l'application.
- **Librairies** :  
**import json** : Permet de lire et écrire des données au format JSON.

**import pandas as pd** : Fournit des structures de données (DataFrames) pour manipuler et analyser des tableaux de données.

**import os** : Offre des fonctions pour interagir avec le système de fichiers (chemins, répertoires).

**from openpyxl import Workbook** : Crée et gère des fichiers Excel (.xlsx) dans des classeurs.

**from openpyxl.styles import PatternFill, Border, Side, Font** : Ajoute des styles (remplissage, bordures, polices) aux cellules Excel.

**from openpyxl.utils import get\_column\_letter** : Convertit les indices de colonnes en lettres (ex. : 1 → "A") pour Excel.

**import hashlib** : Génère des hachages pour des identifiants uniques, utilisé ici pour les couleurs.

**import collections** : Fournit des conteneurs avancés pour organiser les données.

**import re** : Permet l'utilisation d'expressions régulières pour analyser et extraire des motifs dans les textes.

**import zipfile** : Gère la compression et décompression de fichiers ZIP.

**import shutil** : Offre des outils pour copier, déplacer ou supprimer des fichiers et dossiers.

**import tempfile** : Crée des fichiers et dossiers temporaires pour le traitement intermédiaire.

**import tkinter as tk** : Crée une interface graphique de base pour l'application.

**from tkinter import filedialog, messagebox** : Ajoute des boîtes de dialogue pour sélectionner des fichiers et afficher des messages.

**import subprocess** : Exécute des commandes externes (ex. : pbi-tools) dans le système.

**from datetime import datetime** : Gère les dates et heures pour les logs ou les timestamps.

- Les bibliothèques Python de base comme json, os, collections, re, zipfile, shutil, tempfile, tkinter, subprocess, et datetime sont incluses par défaut avec une installation standard de Python 3.13 et ne nécessitent pas d'installation supplémentaire. En revanche, les bibliothèques externes pandas, openpyxl, et leurs dépendances (comme openpyxl.styles et openpyxl.utils) doivent être installées via la commande **pip install pandas openpyxl** dans un terminal avant d'exécuter l'application, afin de garantir leur disponibilité.

## **3 Installation et configuration**

Suivez ces étapes pour configurer l'application Data Extractor :

### **3.1 Téléchargement de l'application**

1. Obtenez l'exécutable Data Extractor (`Data_Extractor.exe`) depuis la source fournie (par exemple, package de distribution officiel).
2. Enregistrez l'exécutable dans un dossier dédié, par exemple `C:\Data_Extractor`.

### 3.2 Configuration de pbi-tools

1. Téléchargez les deux exécutables `pbi-tools.exe` et `pbi-tools.core.exe` depuis le site officiel <https://pbi.tools/>
2. Placez les deux exécutables dans le dossier Bureau ou Téléchargements.
3. L'application les déplacera automatiquement dans le sous-dossier `pbi-tools` du répertoire Data Extractor lors de la première exécution.

## 4 Utilisation de Data Extractor

Cette section explique comment exécuter l'application et traiter un fichier .pbix.

### 4.1 Lancement de l'application

1. Double-cliquez sur le fichier `Data_Extractor.exe` dans le répertoire Data Extractor.
2. Une interface graphique Tkinter apparaîtra.

### 4.2 Sélection d'un fichier .pbix

1. Cliquez sur le bouton « Parcourir » dans l'interface pour ouvrir une boîte de dialogue de sélection de fichier.
2. Naviguez jusqu'à l'emplacement de votre fichier .pbix et sélectionnez-le.
3. Assurez-vous que le fichier .pbix a été créé ou mis à jour avec la dernière version de Power BI Desktop pour éviter les problèmes de compatibilité.

### 4.3 Traitement du fichier

1. Dès la sélection d'un fichier .pbix, l'application commence automatiquement le processus d'extraction.
2. L'application effectuera les actions suivantes :
  - Extraire `Layout.json` du fichier .pbix et l'enregistrer dans le sous-dossier `JSON files` du répertoire Data Extractor.
  - Filtrer et extraire les KPIs (mesures et agrégations) de `Layout.json` dans un `DataFrame pandas`.
  - Utiliser `pbi-tools.exe` pour extraire les fichiers du .pbix dans un dossier temporaire.
  - Compiler les fichiers extraits en un fichier .pbit à l'aide de `pbi-tools.core.exe`.
  - Extraire `DataModelSchema.json` du fichier .pbit et l'enregistrer dans le sous-dossier `JSON files`.
  - Extraire les tables et les colonnes de `DataModelSchema.json` dans un second `DataFrame pandas`.
  - Exporter les `DataFrames` dans deux fichiers Excel :

- `Extracted_Data.xlsx` :
    - Feuille « Données Granulaires » : Contient les tables et les colonnes.
    - Feuille « KPIs » : Contient les KPIs extraits.
  - `Data_Structure.xlsx` : Contient toutes les données de `DataModelSchema.json` y compris les relations, partitions, tables, colonnes, hiérarchies, etc.
3. Le processus peut prendre quelques secondes en raison des opérations de `pbi-tools` et `pbi-tools.core`.

## 4.4 Visualisation des sorties

1. En cas de succès, un message de confirmation apparaîtra.
2. Naviguez jusqu'au répertoire Data Extractor pour trouver :
  - Le sous-dossier `JSON files` contenant `Layout.json` et `DataModelSchema.json`.
  - Les fichiers `Extracted_Data.xlsx` et `Data_Structure.xlsx` dans le répertoire principal.
3. Ouvrez les fichiers Excel pour examiner les données extraites.

## 5 Dépannage

Les problèmes courants et leurs solutions sont listés ci-dessous.

### 5.1 Erreurs courantes

- **Excel est ouvert** : L'application échouera si Excel est en cours d'exécution. Fermez Excel et réessayez.
- **Fichier .pbix obsolète** : Assurez-vous que le fichier .pbix est compatible avec la dernière version de Power BI Desktop.
- **pbi-tools ou pbi-tools.core manquant** : Vérifiez que `pbi-tools.exe` et `pbi-tools.core.exe` sont dans le dossier Bureau ou Téléchargements pour la première exécution.
- **Échec de l'extraction** : Vérifiez si le fichier .pbix est corrompu ou restreint en accès.

### 5.2 Messages d'erreur

- L'application affiche des messages de succès ou d'erreur via des pop-ups Tkinter.
- En cas d'erreur (par exemple, échec de l'extraction JSON ou problème de génération Excel), un message détaillé vous guidera vers le problème.

### 5.3 Consultez la documentation

- Pour les problèmes non résolus, reportez-vous à ce guide utilisateur ou à toute documentation supplémentaire fournie avec l'application Data Extractor.

## 6 Notes supplémentaires

- L'application crée un sous-dossier `pbi-tools` et un sous-dossier `JSON files` dans le répertoire Data Extractor lors de l'exécution.
- Les fichiers temporaires sont automatiquement supprimés après le traitement.
- Assurez-vous d'avoir suffisamment d'espace disque pour les fichiers `.pbix` volumineux, car l'extraction temporaire peut nécessiter un espace de stockage important.