

CARNET TECHNIQUE: DASHBOARD SMARTMARKET

Projet : Développement d'une application de Business Intelligence (BI)

Technologie : Python (Pandas, Streamlit, Plotly)

Par : Soraya Codo

1. Problème : Hétérogénéité des formats de données

- **Description :** Les données proviennent de trois services différents avec des formats distincts : un fichier CSV (Navigation), un fichier JSON (Campagnes régies) et un fichier Excel (CRM).
- **Solution apportée :** Utilisation de la bibliothèque pandas avec ses moteurs spécialisés (`read_csv`, `read_json`, `read_excel`).
- **Justification :** Centraliser l'ingestion dans un seul script Python permet de créer un pipeline de données reproductible sans avoir à convertir manuellement les fichiers avant le traitement.

2. Problème : Incohérence des nomenclatures (Data Mapping)

- **Description :** La colonne identifiant le secteur d'activité s'appelle `industry` dans le CRM mais doit être affichée sous le nom `sector` dans le dashboard. De même, les budgets sont nommés `budget` dans le JSON mais manipulés comme des coûts (`cost`) par le marketing.
- **Solution apportée :** Utilisation de la méthode `.rename(columns={...})` immédiatement après le chargement pour harmoniser les DataFrames.
- **Justification :** La transformation des variables dès l'entrée du pipeline de traitement évite les erreurs de type `KeyError` lors des calculs ultérieurs en prenant en compte le vocabulaire métier clair accessible à la cible du rapport (marketing)

3. Problème : Gestion du périmètre temporel (Data Filtering)

- **Description :** Les sources de données contiennent des historiques s'étalant sur plusieurs mois, alors que l'analyse est strictement limitée au mois de septembre 2025.
- **Solution apportée :** Conversion des colonnes de dates au format `datetime` via `pd.to_datetime()`, puis application d'un masque booléen pour isoler l'intervalle du 2025-09-01 au 2025-09-30.
- **Justification :** Le filtrage par masque booléen est la méthode la plus performante en Pandas pour les grands volumes de données et assure l'intégrité des KPIs présentés.

4. Problème : Expérience utilisateur et clarté des KPIs

- **Description :** Les sigles marketing (CTR, CPL, MQL...) peuvent être mal interprétés par des utilisateurs non spécialistes, rendant le dashboard difficile à utiliser sans assistance.
- **Solution apportée :** Intégration d'un glossaire dynamique dans la barre latérale via `st.sidebar.info`.
- **Justification :** En intégrant la documentation directement dans l'outil, on améliore l'autonomie du décideur et on réduit les risques d'erreurs d'interprétation des graphiques.