

Rapport Final – Data ANIP Challenge

David Akpovi Octobre 2025

Introduction

Ce projet s'inscrit dans le cadre du *Data Challenge* organisé pour l'Agence Nationale d'Identification des Personnes (ANIP) du Bénin.

L'objectif est de démontrer la capacité à transformer des données brutes provenant de multiples sources en un ensemble cohérent et analysable, permettant de produire des insights exploitables pour la prise de décision publique. Les analyses portent sur le Bénin, à travers trois grandes étapes :

1. **La collecte et la préparation des données.**
2. **L'exploration et l'analyse descriptive.**
3. **La visualisation et la narration des résultats à travers un tableau de bord interactif.**

Les données concernent le Bénin et couvrent les dimensions démographiques, économiques et sociales entre 1960 et 2024. Certaines séries, notamment démographiques, incluent des projections estimées jusqu'en 2025.

Tâche 1 – Collecte et Préparation des Données

Objectif

Collecter, nettoyer, harmoniser et consolider des données provenant de sources internationales variées afin de constituer un dataset complet et homogène sur le Bénin.

Méthodologie

Les sources principales utilisées incluent :

- **World Bank Open Data (WDI)** pour les indicateurs économiques et sociaux.
- **WHO Global Health Observatory** pour les données de santé.
- **UN Population Division** pour les données démographiques.

Les données ont été collectées via API et fichiers CSV, nettoyées à l'aide de pandas, et harmonisées à l'aide de fonctions personnalisées (`standardize_columns`, `harmonize_geo`, `harmonize_year`). Chaque table a été vérifiée sur la cohérence des clés iso3 et year avant consolidation.

Résultats

- Données finales : **dataset_final.csv** (complet, enrichi) et **dataset_final_core.csv** (version allégée à 7 colonnes principales).

- Un contrôle qualité a confirmé la cohérence temporelle (1960–2024) et l’absence d’anomalies majeures.

Livrables

- **Dataset harmonisé (CSV & Excel)** (ces fichiers sont placés dans le dossier de la Tâche 1 sur GitHub https://github.com/davidakpovi/Data-ANIP-Challenge/tree/main/Ta%CC%82che_1)
- **Glossaire des variables**
- **Notebook documenté**

Tâche 2 – Exploration et Analyse

Objectif

Explorer les tendances historiques, détecter les anomalies et construire des indicateurs dérivés pour une meilleure compréhension des dynamiques économiques, sociales et démographiques.

Principales Analyses Réalisées

- **Tendances temporelles** : visualisation de la croissance de la population, de l’espérance de vie et du PIB.
- **Détection d’anomalies** : vérification des ruptures de séries et valeurs aberrantes sur les données économiques.
- **Création de nouveaux indicateurs** :
 - Taux de croissance démographique annuel.
 - PIB par habitant (calculé à partir du PIB courant et de la population totale).
 - Ratio de mortalité maternelle normalisé par population.

Résultats

Les analyses ont permis de dégager les éléments suivants :

- **Une croissance démographique soutenue**, traduisant un fort dynamisme naturel.
- **Une amélioration sanitaire constante**, matérialisée par l’allongement de l’espérance de vie.
- **Une corrélation positive** entre l’évolution du PIB et les indicateurs de bien-être.
- **Des discontinuités économiques ponctuelles** (notamment au début des années 1990 et en 2020) suggérant des chocs externes.

Livrables

- **Dataset enrichi** incluant les indicateurs calculés.
- **Notebook analytique** avec détection d'anomalies et documentation méthodologique.
- **Export Excel multi-feuilles** pour intégration Power BI. (ces fichiers sont placés dans le dossier de la Tâche 2 sur GitHub https://github.com/davidakpovi/Data-ANIP-Challenge/tree/main/Ta%CC%82che_2)

Glossaire des variables principales

Nom de la variable	Définition	Unité	Source	Période	Niveau géographique
iso3	Code pays ISO à 3 lettres	—	Standard ISO	—	National
year	Année de référence	—	Toutes sources	1960–2024	National
population_total	Population totale estimée du Bénin	Nombre de personnes	World Bank (WDI)	1960–2024	National
gdp_current_usd	Produit Intérieur Brut à prix courants	USD	World Bank (WDI)	1960–2024	National
gdp_real_growth_pct	Taux de croissance du PIB réel	Pourcentage (%)	World Bank (WDI)	1960–2024	National
life_expectancy_years	Espérance de vie à la naissance	Années	World Bank / WHO	1960–2024	National
maternal_mortality_ratio	Taux de mortalité maternelle	Pour 100 000 naissances vivantes	WHO / World Bank	1960–2024	National

Ce glossaire présente les principales variables intégrées dans le dataset final et harmonisées selon un format commun. Les définitions respectent les standards des organisations internationales (World Bank, WHO, ONU).

Tâche 3 – Visualisation et Insights

Objectif

Transformer le dataset enrichi en un outil de visualisation interactif permettant une compréhension intuitive et stratégique des dynamiques du Bénin.

Approche

Power BI n'étant pas disponible sur macOS, j'ai réalisé les visualisations sous Jupyter Notebook à l'aide des bibliothèques Plotly et Matplotlib. Cette approche a permis de recréer un environnement interactif équivalent à Power BI, offrant une expérience fluide d'exploration des données.

Types de visualisations réalisées

- **Séries temporelles** : évolution du PIB, de la population et de l'espérance de vie.
- **Graphiques comparatifs** : corrélation PIB–espérance de vie.

- **Distribution des valeurs** : analyse de la mortalité maternelle dans le temps.
- **KPI visuels** : mise en avant des valeurs récentes et des taux de variation.

Principaux enseignements

- **Croissance démographique soutenue** : la population du Bénin augmente régulièrement sur la période observée.
- **Amélioration sanitaire** : l'espérance de vie progresse, signe d'un meilleur accès aux soins et à la qualité de vie.
- **Défis persistants** : la mortalité maternelle reste élevée malgré les progrès constatés.
- **Lien entre économie et santé** : la croissance économique s'accompagne d'une amélioration des indicateurs sociaux.
- **Vigilance économique** : certaines années révèlent des baisses du PIB réel, indiquant des périodes d'instabilité à investiguer.

Adaptation de Power BI à Jupyter Notebook

Comme je l'ai souligné plus haut, Power BI Desktop n'étant pas disponible sur macOS, les visualisations interactives de la Tâche 3, je les ai réalisées sous Jupyter Notebook à l'aide des bibliothèques Plotly et Matplotlib. Cette approche m'a permis de reproduire les fonctionnalités essentielles d'un dashboard Power BI :

- filtres dynamiques sur les années,
- comparaison d'indicateurs clés,
- narration visuelle intégrée,
- export HTML interactif.

Résultats

J'ai conçu deux tableaux de bord interactifs à l'aide de la bibliothèque Plotly afin de renforcer la dimension visuelle et analytique du projet. Le premier illustre l'évolution comparée des principaux indicateurs — population totale, PIB (en USD courants) et espérance de vie — sur la période observée. Le second met en avant les indicateurs clés de développement (KPI) sous forme de barres dynamiques, permettant une lecture claire et immédiate des valeurs les plus récentes. Ces deux visualisations ont été exportées au format HTML (`evolution_indicateurs.html` et `kpi_bar_interactif.html`), dans le dossier de la Tâche 3 sur GitHub https://github.com/davidakpovi/Data-ANIP-Challenge/tree/main/Ta%CC%82che_3, rendant le tableau de bord consultable directement depuis un navigateur web sans installation supplémentaire.

Recommandations et Conclusion

Recommandations principales :

1. Approfondir la granularité territoriale des données

Actuellement, les analyses portent sur des agrégats nationaux. Il est recommandé d'intégrer, à moyen terme, des données régionales ou communales issues de l'INSAE et des ministères sectoriels afin de révéler les disparités locales et de mieux cibler les politiques publiques.

2. Automatiser la mise à jour via des sources ouvertes

Les indicateurs utilisés (population, PIB, santé) peuvent être actualisés régulièrement à partir d'API officielles telles que celles de la Banque mondiale, du PNUD ou de l'OMS.

Mettre en place un script d'automatisation permettrait de réduire les délais de mise à jour et de fiabiliser le suivi des indicateurs stratégiques.

3. Intégrer un tableau de bord opérationnel pour le suivi du développement

Le tableau de bord construit dans le cadre de ce projet pourrait être déployé dans un environnement de suivi (Power BI par exemple) afin de fournir aux décideurs un accès rapide et interactif aux indicateurs socio-économiques du Bénin.

4. Renforcer l'interopérabilité avec les bases administratives sectorielles

Le croisement des données de population gérées par l'ANIP avec celles issues des bases administratives du Ministère de la Santé (indicateurs de mortalité, accès aux soins, couverture vaccinale) et du Ministère de l'Éducation (taux de scolarisation, alphabétisation, infrastructures éducatives) permettrait de construire des indicateurs intégrés de développement humain. Cette interopérabilité offrirait une vision plus fine des inégalités territoriales, faciliterait le ciblage des politiques publiques, et renforcerait la capacité de l'État à anticiper les besoins en matière de santé, d'éducation et de services sociaux.

5. Poursuivre les analyses longitudinales

Les tendances observées entre 1960 et 2024 méritent un suivi longitudinal pour évaluer l'impact des réformes économiques et sociales. Une mise à jour annuelle du tableau de bord permettrait de mesurer les progrès réels du Bénin vers les Objectifs de Développement Durable (ODD).

Conclusion

Ce projet m'a permis de démontrer une chaîne complète de traitement de données, depuis la collecte jusqu'à la production d'insights. Les jeux de données que j'ai obtenus sont prêts à être intégrés dans un environnement décisionnel ou à être utilisés pour des analyses approfondies sectorielles.

Le travail réalisé illustre ma capacité à :

- Collecter et harmoniser des sources internationales fiables.
- Produire des indicateurs interprétables et cohérents.
- Délivrer une narration claire et orientée vers la décision publique.

Ce rapport constitue ainsi une base solide pour la création d'un système de suivi dynamique du développement national au Bénin.