

RAPPORT DE SYNTHÈSE — CHALLENGE 1 (WPP2024)

Ce rapport de synthèse (3–5 pages) présente les choix de visualisation, les insights principaux tirés des données démographiques WPP2024 traitées dans le répertoire `n1/`, et des recommandations actionnables destinées aux décideurs publics. Le texte est écrit en français et conçu pour être directement utilisable dans un dossier de présentation ou comme note décisionnelle.

1. CONTEXTE ET OBJECTIFS

Le Challenge 1 vise à nettoyer et enrichir le jeu de données démographiques WPP2024 afin de produire des indicateurs fiables pour l'analyse territoriale et la prise de décision. Les livrables principaux sont des jeux de données nettoyés et enrichis (`WPP2024_DEMOGRAPHIC_CLEAN.csv`, `WPP2024_DEMOGRAPHIC_ENRICHED.csv`) et une série de visuels (évolution de la population, espérance de vie, corrélations, cartes régionales).

Objectifs clés pour les décideurs : - Comprendre les dynamiques de population par pays et région. - Identifier les anomalies et tendances préoccupantes (chutes/hausse abruptes, espérance de vie en baisse). - Prioriser les interventions (santé publique, planification urbaine, allocation budgétaire).

2. CHOIX DE VISUALISATION — PRINCIPES ET JUSTIFICATION

Les choix de visualisation ont été guidés par trois principes : clarté, comparabilité, et actionnabilité.

1. Séries temporelles (ligne) — Population totale et espérance de vie : - Pourquoi : Montrer l'évolution dans le temps, détecter ruptures de tendance, et comparer trajectoires entre pays ou régions. - Usage : Graphiques multi-séries avec échelles partagées ou facettées pour conserver la comparabilité. - Détail technique : lignes lissées (par ex. rolling mean 3 ans) pour atténuer le bruit, avec surcouches pour anomalies identifiées.

2. Cartes choroplèthes — Distribution spatiale (population 2023, indice de développement régional) : - Pourquoi : Les cartes mettent en évidence les disparités géographiques et facilitent la priorisation territoriale. - Usage : Échelles de couleur perceptuellement uniformes (Viridis/Cividis) et classes définies par quantiles pour éviter la domination des outliers. - Détail technique : prévoir popups interactifs (PowerBI / HTML export) montrant séries historiques et indicateurs clés au survol.

3. Matrice de corrélation (heatmap) — Relations entre indicateurs : - Pourquoi : Identifier variables fortement liées (ex. espérance de vie ↔ PIB/habitant proxy) pour réduire la dimensionnalité ou prioriser mesures politiques. - Usage : Heatmap ordonnée par clustering hiérarchique pour révéler groupes d'indicateurs cohérents.

4. Graphiques de distribution (boxplots / violin) — Comparaisons régionales : - Pourquoi : Montrer la dispersion et les outliers dans les indicateurs entre régions. - Usage : Boxplots par région pour espérance de vie, taux de fertilité, etc., afin d'identifier inégalités internes.

5. Indicateurs agrégés et tableaux synthétiques (Small multiples / KPI tiles) : - Pourquoi : Fournir aux décideurs des « at-a-glance » métriques (croissance annuelle moyenne, part de population âgée, espérance de vie) pour la prise de décision rapide.

Chaque visuel est accompagné d'un commentaire synthétique (insight + action recommandée) dans le rapport et le dashboard PowerBI.

3. INSIGHTS PRINCIPAUX

Les insights suivants sont extraits du jeu de données nettoyé et enrichi. Ils synthétisent tendances globales, variations régionales et anomalies identifiées.

1. Croissance démographique mondiale : - Observation : La population mondiale augmente mais avec un rythme de croissance hétérogène — forte croissance en certaines régions d'Afrique subsaharienne, ralentissement dans plusieurs pays à revenus élevés. - Implication : Besoin de renforcer les capacités de services de base (santé, éducation, logement) dans les régions à forte croissance.

2. Espérance de vie — disparités régionales persistantes : - Observation : Une amélioration générale de l'espérance de vie sur les dernières décennies, mais avec des écarts significatifs entre régions (Afrique subsaharienne vs. Europe/Amériques). Certaines baisses locales ont été détectées et marquées comme anomalies. - Implication : Les baisses locales nécessitent des enquêtes ciblées (p. ex. crise sanitaire, conflits, crise économique).

3. Corrélations structurantes : - Observation : Corrélation positive notable entre indicateurs de développement socioéconomique (proxys) et espérance de vie. Certains indicateurs démographiques (taux de fertilité) montrent des corrélations négatives avec l'espérance de vie. - Implication : Les politiques de développement intégrées (santé, éducation, économie) restent prioritaires pour améliorer la longévité.

4. Concentration urbaine et pression sur services : - Observation : Les données montrent des augmentations rapides de population urbaine pour plusieurs pays, concentrant les enjeux de logement et d'infrastructures. - Implication : Planification urbaine proactive et investissements en infrastructures critiques sont nécessaires.

5. Anomalies détectées (exemples) : - Valeurs d'espérance de vie > 120 ou hausses de population ponctuelles non compatibles avec les tendances historiques — potentiellement erreurs de saisie ou ruptures réelles (migrations massives, catastrophes). - Recommandation : Validation métier des anomalies avant correction automatique.

4. RECOMMANDATIONS ACTIONNABLES POUR DÉCIDEURS PUBLICS

Les recommandations sont classées par horizon (court, moyen, long terme) et priorisées selon l'impact probable et la faisabilité.

A. Court terme (0–12 mois)

1. Valider les anomalies critiques : - Action : Former un petit comité technique (données + expertise métier) pour examiner les anomalies listées dans `WPP2024_DEMOGRAPHIC_ANOMALIES.csv` et décider : corriger, conserver, ou approfondir. - Raison : Éviter décisions basées sur données erronées.

2. Déployer tableaux de bord KPI pour suivi rapide : - Action : Publier un dashboard PowerBI avec KPI tiles (croissance annuelle, part population 65+, espérance de vie) mis à jour régulièrement. - Raison : Faciliter la prise de décision opérationnelle.

B. Moyen terme (1–3 ans)

3. Prioriser investissements dans les régions à forte croissance : - Action : Utiliser cartes choroplèthes pour orienter l'allocation budgétaire vers santé primaire, accès à l'eau, et infrastructures scolaires/urbaines. - Raison : Réduire les vulnérabilités liées à la croissance démographique rapide.

4. Lancer études approfondies sur zones à baisse d'espérance de vie : - Action : Enquête de terrain, analyses de cause (santé, environnement, conflits) pour chaque zone identifiée. - Raison : Adapter les politiques publiques aux causes réelles (p. ex. interventions sanitaires ciblées).

C. Long terme (3+ ans)

5. Intégrer modèles prospectifs dans la planification : - Action : Développer modèles démographiques projetant population, dépendance démographique, et besoins en services sur 10–30 ans. - Raison : Planifier investissements structurants (transport, énergie, santé) et anticiper pressions futures.

6. Renforcer capacités de gouvernance de la donnée : - Action : Mettre en place standards de qualité, pipeline ETL récurrent, tests automatisés et gouvernance des métadonnées (glossaire vivant). - Raison : Assurer fiabilité continue des indicateurs utilisés pour la décision publique.

5. MESURES D'IMPACT ET INDICATEURS DE SUIVI

Pour évaluer l'efficacité des actions recommandées, suivre : - Pourcentage d'anomalies

validées/traitées chaque trimestre. - Taux d'accès aux soins primaires dans les régions prioritaires (annuel). - Évolution de l'indice de dépendance démographique (part 65+ vs. 15–64) sur 5 ans. - Variation de la croissance urbaine et du taux de logements adéquats.

6. ANNEXES — NOTES TECHNIQUES

1. Données et scripts : les jeux sources et scripts de traitement se trouvent dans `n1/` (`Tache1_ANIP_WPP.py`, etc.). 2. Reproductibilité : suivre `n1/README.md` et `n1/requirements.txt` pour recréer l'environnement et exécuter le pipeline. 3. Visualisations disponibles : `n1/output/graphs/`, `n1/output/plots/`, `n1/output/powerbi_full/`.

Si vous le souhaitez, je peux : - Convertir ce rapport en PDF prêt à imprimer (3–5 pages) ; - Générer une version PowerPoint avec les visuels et messages clés ; - Créer le commit Git (message suggéré : "Add synthesis report for Challenge 1 — visualisations, insights, recommendations").

Quelle action voulez-vous que je fasse ensuite ?