

RAPPORT ANALYTIQUE — ATELIER DÉVELOPPEMENT

ANALYSE EXPLORATOIRE (EDA) DES RÉSULTATS DE LA PRÉSIDENTIELLE 2022 (TOUR 1 & TOUR 2)

1) Contexte et objectif

Dans ce travail, je réalise une analyse exploratoire (EDA) sur les résultats agrégés des élections présidentielles 2022, en comparant le Tour 1 et le Tour 2. L'objectif est de comprendre les tendances de participation / abstention, ainsi que la qualité du vote via les votes blancs et votes nuls, et d'identifier des écarts territoriaux. Cette étude reste volontairement descriptive : aucune prédition n'est réalisée.

2) Choix raisonné du dataset

Pertinence

J'ai choisi un dataset officiel de résultats électoraux agrégés par bureau de vote (Présidentielle 2022 T1/T2). Ce dataset est pertinent car il permet :

- de mesurer précisément l'abstention et la participation (via inscrits, votants, abstentions),
- d'analyser les votes blancs et nuls, utiles pour comprendre certaines formes de rejet,
- de comparer ces indicateurs selon les territoires (départements, communes, bureaux).

Limites

- Les données sont agrégées : il n'y a pas d'information socio-démographique (âge, niveau de vie, etc.). On peut décrire les écarts, mais pas expliquer les causes profondes.
- Une partie des territoires a des contextes très spécifiques (ex : Français de l'étranger, certains territoires ultramarins) : il faut les interpréter avec précaution.

Biais potentiels

- Les comparaisons "moyenne par bureau de vote" ne sont pas strictement équivalentes à des taux pondérés par population (un gros bureau "pèse" autant qu'un petit dans une moyenne simple).
- Certains résultats peuvent être influencés par des facteurs locaux : logistique, accessibilité au vote, organisation du bureau.

Ce que ce dataset permet de comprendre

Ce dataset permet surtout de repérer des patterns : évolution entre tours, zones à forte abstention, hausse des votes blancs/nuls, dispersion entre bureaux, et segmentation simple par taille.

3) PROBLÉMATIQUE ANALYTIQUE (SANS PRÉDICTION)

Comment évoluent la participation, l'abstention, les votes blancs et les votes nuls entre le Tour 1 et le Tour 2 de la Présidentielle 2022, et observe-t-on des disparités territoriales et des profils de bureaux de vote spécifiques ?

DURÉE ET PÉRIODE

4) Méthodologie (nettoyage + préparation)

- Chargement des données en filtrant uniquement 2022_pres_t1 et 2022_pres_t2.
 - Contrôles de qualité : valeurs manquantes, doublons, cohérence métier (ex : liens entre inscrits/votants/abstentions), types de variables.
 - Création de variables dérivées pour faciliter l'analyse :
 - taux de participation / abstention (en %),
 - votes blancs et nuls en % des votants,
 - catégorisation de l'abstention (classes),
 - segmentation par taille de bureau (selon le nombre d'inscrits).
 - Production d'un fichier final nettoyé exporté en CSV.
- Outils utilisés : Python (Pandas) + visualisations (Matplotlib/Seaborn).

5) RÉSULTATS — INSIGHTS CLÉS (ARGUMENTÉS PAR LES DONNÉES)

Insight 1 — Participation en baisse au 2^e tour

En moyenne par bureau de vote, la participation passe de 75,88 % (T1) à 74,50 % (T2), soit -1,37 point.

L'abstention augmente mécaniquement de 24,12 % à 25,50 %, soit +1,37 point.

Cela montre une légère démobilisation au second tour, même si l'écart reste modéré.

Insight 2 — Hausse très nette des votes blancs et nuls au 2^e tour

Les votes blancs (en % des votants) passent de 1,58 % (T1) à 6,68 % (T2), soit +5,10 points.

Les votes nuls passent de 0,78 % à 2,52 %, soit +1,74 point.

Cette évolution est très marquante : au second tour, les électeurs votent davantage “blanc”, et on observe aussi plus de bulletins invalidés. On peut l'interpréter comme un signal de rejet ou de tension autour du choix final.

NSIGHT 3 — ABSTENTION TRÈS VARIABLE SELON LES TERRITOIRES (TOUR 2)

4) Méthodologie (nettoyage + préparation)

- Chargement des données en filtrant uniquement 2022_pres_t1 et 2022_pres_t2.
- Contrôles de qualité : valeurs manquantes, doublons, cohérence métier (ex : liens entre inscrits/votants/abstentions), types de variables.
- Création de variables dérivées pour faciliter l'analyse :
 - taux de participation / abstention (en %),
 - votes blancs et nuls en % des votants,
 - catégorisation de l'abstention (classes),
 - segmentation par taille de bureau (selon le nombre d'inscrits).
- Production d'un fichier final nettoyé exporté en CSV.
Outils utilisés : Python (Pandas) + visualisations (Matplotlib/Seaborn).

NSIGHT 4 — “POCHES LOCALES” VISIBLES DANS LA DISTRIBUTION (TOUR 2)

La distribution de l'abstention au Tour 2 est assez dispersée :

- 1er quartile : ~19,62 %
- médiane : ~23,30 %
- 3e quartile : ~28,67 %

On observe aussi une proportion non négligeable de bureaux très hauts :

- environ 7,42 % des bureaux dépassent 40 % d'abstention,
- environ 2,94 % dépassent 50 %.
- Ces chiffres montrent l'existence de zones “à risque” en termes de mobilisation.

Insight 5 — Vote blanc très hétérogène selon les territoires (Tour 2)

Le vote blanc (en % des votants) est très variable selon les territoires :

- élevé dans des départements comme Ariège : ~11,47 % et Hautes-Pyrénées : ~10,46 %,
- très bas dans Wallis-et-Futuna : ~1,71 % ou Polynésie française : ~2,36 % (et Mayotte : ~3,93 %).
- Cela suggère que le vote blanc n'a pas partout la même signification politique ou sociale : il dépend probablement du contexte local.

NSIGHT 6 — EFFET DE LA TAILLE DU BUREAU (TOUR 2)

En segmentant les bureaux par taille (selon le nombre d'inscrits), l'abstention moyenne augmente avec la taille :

- petits bureaux (≤ 500 inscrits) : ~21,35 %
- 501–1000 inscrits : ~27,49 %
- 1001–2000 inscrits : ~27,89 %
- très grands bureaux (> 2000 inscrits) : ~57,55 % (mais catégorie rare : seulement 128 bureaux)

On peut retenir surtout les trois premières catégories, car elles représentent la grande majorité des bureaux. L'idée importante est que les bureaux plus "grands" semblent plus exposés à l'abstention (hypothèse possible : zones denses/urbaines, à confirmer).

INSIGHT 7 — ABSTENTION ET VOTE BLANC : RELATION TRÈS FAIBLE

La corrélation entre abstention (%) et blancs (% des votants) est très faible sur l'ensemble : $r \approx 0,03$.

Elle reste faible par tour : $T1 \approx 0,05$ et $T2 \approx -0,06$.

Cela indique que l'abstention et le vote blanc sont deux comportements différents. Les actions pour les réduire ne sont pas forcément les mêmes.

6) Signaux faibles / patterns à surveiller

- Certains profils de territoires cumulent abstention élevée et hausse des blancs/nuls : cela peut indiquer une défiance ou un rejet plus durable.
- Les territoires spécifiques (ex : Français de l'étranger, certains territoires ultramarins) ont des niveaux très élevés : ils méritent une analyse séparée (logistique, distance, accès au vote).

7) Contraintes méthodologiques (limites d'interprétation)

- Données agrégées : pas d'explication socio-démographique possible.
- Analyse descriptive : corrélation \neq causalité.
- Les comparaisons par département peuvent dépendre du choix d'agrégation (moyenne par bureau vs pondération par inscrits).

8) INTERPRÉTATION STRATÉGIQUE ET RECOMMANDATIONS

Implications (politiques publiques)

- La hausse des votes blancs et nuls au second tour peut signaler une difficulté d'adhésion à l'offre politique finale. Cela pose un enjeu de confiance et de pédagogie électorale.
- Les écarts territoriaux montrent que la mobilisation ne peut pas être traitée uniquement à l'échelle nationale : il faut une stratégie locale.

Recommandations opérationnelles

- Cibler prioritairement les zones à forte abstention (actions locales : communication, accès au vote, relais terrain).
- Mettre en place une communication simple sur la validité du bulletin pour limiter les votes nuls.
- Suivre séparément abstention et vote blanc (car ce sont deux phénomènes différents).

Axes d'amélioration / besoins en données complémentaires

- Croiser avec des données INSEE (densité, âge, niveau de vie, mobilité) pour mieux expliquer les écarts observés.
- Comparer avec un scrutin précédent (ex : 2017) pour distinguer ce qui est structurel de ce qui est conjoncturel.