Data Visualization : Quand les prénoms révèlent nos votes

Exploration de la sociologie française à travers l'analyse visuelle de données démographiques et électorales

o Introduction : L'art de faire parler les données

Votre prénom influence-t-il votre vote ? Cette question, a priori saugrenue, cache en réalité des enjeux sociologiques fascinants. À travers l'analyse visuelle de 2.5 millions d'enregistrements démographiques français, j'ai découvert des corrélations surprenantes entre les prénoms et les tendances politiques.

Cet article explore comment la data visualization peut révéler des insights inattendus sur notre société, en utilisant des techniques modernes de représentation de données pour rendre accessible une analyse statistique complexe.

📊 La donnée comme miroir de la société

Le prénom, marqueur sociologique

Les prénoms ne sont pas neutres. Ils reflètent :

• L'époque : modes et influences culturelles

• La géographie : traditions régionales

• Le milieu social : aspirations et références

• Les valeurs familiales : conservatisme vs modernité

Ces mêmes facteurs influencent les choix politiques. L'hypothèse devient alors : **peut-on prédire** les tendances de vote d'un territoire en analysant les prénoms de ses habitants ?

Méthodologie : de la théorie aux données

Pour tester cette hypothèse, j'ai développé une approche en trois étapes :

- 1. **Identification géographique** : Quel département est le plus représentatif d'un prénom donné
- 2. Analyse électorale : Quelles sont les tendances de vote de ce département ?
- 3. **Visualisation des corrélations** : Comment rendre ces liens visibles et compréhensibles ?

III Cartographie interactive : La France des prénoms

L'outil : Folium et la puissance des cartes

La première innovation a été de créer une carte interactive de France où chaque département peut être coloré selon sa "affinité" avec un prénom donné.

```
python
import folium
from folium import plugins
def create_prenoms_heatmap(prenom_data):
  """Créer une carte de chaleur des prénoms en France"""
  # Centre de la France
  m = folium.Map(location=[46.6, 1.9], zoom_start=6)
  # Données de chaleur basées sur les ratios
  heat data = [
    [dept['lat'], dept['lon'], dept['ratio']]
    for dept in prenom_data
  # Couche de chaleur
  plugins.HeatMap(heat_data, radius=25, blur=15).add_to(m)
  return m
```

Révélations cartographiques

Les premières visualisations ont révélé des patterns géographiques frappants :

Prénoms "parisiens": KEVIN, JESSICA → Concentration en Île-de-France et grandes métropoles **Prénoms "ruraux"**: MARIE, JEAN → Forte représentation dans la France périphérique **Prénoms** "**frontaliers"**: HANS (Alsace), JOSÉ (Sud-Ouest) → Influences culturelles transfrontalières

Cette géographie des prénoms épouse étonnamment celle du vote politique français.

Graphiques dynamiques : L'évolution temporelle

Analyse générationnelle avec Plotly

La dimension temporelle s'est révélée cruciale. Chaque prénom a sa "courbe de vie" qui raconte l'histoire sociale française.

```
python
import plotly.express as px
import plotly.graph_objects as go
def create_temporal_analysis(prenoms_data):
  """Visualisation de l'évolution des prénoms dans le temps"""
  fig = go.Figure()
  for prenom in prenoms_data:
    fig.add_trace(go.Scatter(
       x=prenom['annees'],
       y=prenom['frequence'],
       mode='lines+markers',
      name=prenom['nom'],
       line=dict(width=3),
       hovertemplate='<b>%{fullData.name}</b><br>' +
              'Année: %{x}<br>' +
              'Fréquence: %{y:.2f}%<extra></extra>'
    ))
  fig.update_layout(
    title="Évolution temporelle des prénoms (1946-2020)",
    xaxis_title="Année",
    yaxis_title="Fréquence (%)",
    hovermode='x unified',
    template='plotly_white'
  return fig
```

Les "générations de prénoms"

L'analyse temporelle révèle des cohortes distinctes :

Génération "Tradition" (1946-1965): Marie, Jean, Pierre, Anne

- Prénoms stables, référence chrétienne
- Corrélation avec vote conservateur/centriste

Génération "Libération" (1966-1985) : Christophe, Valérie, Laurent, Isabelle

- Diversification des choix, influence américaine
- Profil électoral varié, centre-gauche dominant

Génération "Mondialisation" (1986-2005): Kevin, Jessica, Dylan, Océane

- Influences internationales, nouveaux modèles
- Tendance vers les extrêmes politiques

Génération "Retour aux sources" (2006-2020) : Emma, Louis, Jade, Gabriel

- Synthèse tradition/modernité
- Profil électoral en cours de formation

Analyse des corrélations : Le "Prénom Score"

Création d'un indicateur composite

Pour quantifier la relation prénom-vote, j'ai développé le "Prénom Score" :

```
python
def calculate_prenom_score(prenom_data, election_data):
  """Calcul du score politique d'un prénom"""
  score = 0
  total_weight = 0
  for dept in prenom_data:
    dept_code = dept['code']
    prenom_ratio = dept['ratio']
    if dept_code in election_data:
       # Pondération par la représentativité du prénom
       weight = prenom_ratio * dept['population']
       # Score politique (-100 à +100)
       political_score = calculate_political_lean(election_data[dept_code])
       score += political_score * weight
       total_weight += weight
```

Résultats surprenants

Le classement des prénoms selon leur "couleur politique" révèle :

return score / total_weight if total_weight > 0 else 0

Top 5 "Prénoms de droite" :

- 1. HENRI (-87) Aristocratie française
- 2. CHARLES (-82) Tradition monarchiste

- 3. MARIE-CLAIRE (-78) Bourgeoisie catholique
- 4. BERNARD (-75) Classes moyennes conservatrices
- 5. FRANÇOIS (-71) Paradoxe de la tradition

Top 5 "Prénoms de gauche" :

- 1. MOHAMED (+89) Immigration, diversité
- 2. FATIMA (+85) Multiculturalisme
- 3. DYLAN (+73) Modernité populaire
- 4. JESSICA (+68) Nouvelles classes moyennes
- 5. KEVIN (+64) Phénomène générationnel

III Visualisations interactives : Dashboard complet

Interface utilisateur avec Streamlit

L'étape finale a consisté à créer un tableau de bord interactif permettant l'exploration libre des données :

```
import streamlit as st
import altair as alt
def create_dashboard():
  st.title(" Prénoms & Politique : Explorateur Interactif")
  # Sidebar avec contrôles
  st.sidebar.header("Paramètres d'analyse")
  selected_prenom = st.sidebar.selectbox(
    "Choisir un prénom",
    options=get_available_prenoms()
  year_range = st.sidebar.slider(
    "Période d'analyse",
    min_value=1946,
    max_value=2020,
    value=(1980, 2010)
  # Colonnes principales
  col1, col2 = st.columns(2)
  with col1:
    st.subheader("Carte géographique")
    create_folium_map(selected_prenom)
  with col2:
    st.subheader("Tendance politique")
    create_political_gauge(selected_prenom)
  # Graphique temporel
```

st.subheader("Évolution dans le temps")
create_time_series(selected_prenom, year_range)

Fonctionnalités avancées

Le dashboard intègre plusieurs outils d'analyse :

Comparaison multi-prénoms : Superposition de jusqu'à 5 prénoms pour analyser les différences générationnelles et géographiques.

Analyse de clustering : Regroupement automatique des prénoms selon leurs profils politiques similaires.

Prédiction territoriale : Estimation de la couleur politique d'un département basée sur la composition des prénoms.

o Insights et implications

Découvertes marquantes

L'analyse révèle plusieurs phénomènes sociologiques majeurs :

La "polarisation prénomnale" : Les prénoms se segmentent de plus en plus selon les clivages politiques, reflétant une société plus fragmentée.

L'effet "bobo" : Certains prénoms (Augustin, Margot, Théophile) concentrent les nouvelles classes urbaines éduquées, créant des îlots politiques spécifiques.

Le "vote communautaire" : Les prénoms d'origine étrangère montrent une cohérence géographique et politique forte, signe d'une intégration encore inégale.

La "revanche des Kevin": Les prénoms populaires des années 90 deviennent des marqueurs de protestation politique.

Limites et précautions

Cette analyse, bien qu'éclairante, comporte des limites importantes :

Corrélation ≠ **Causalité** : Les prénoms ne déterminent pas le vote, ils révèlent des appartenance sociales sous-jacentes.

Biais de généralisation : Les tendances observées reflètent des moyennes, pas des destins individuels.

Évolution temporelle: Les associations prénom-politique évoluent avec les générations.

Perspectives et extensions

Développements futurs

Cette recherche ouvre plusieurs pistes d'approfondissement :

Analyse intersectionnelle : Croiser prénoms, professions, et revenus pour affiner les profils sociologiques.

Modélisation prédictive : Développer des algorithmes de prédiction électorale basés sur les données démographiques.

Études comparatives : Étendre l'analyse à d'autres pays européens pour identifier les spécificités françaises.

Temps réel : Créer un système de veille des nouvelles tendances prénomnales comme indicateur d'évolution sociale.

Applications pratiques

Au-delà de l'aspect scientifique, cette méthode trouve des applications concrètes :

Marketing politique : Ciblage plus fin des campagnes électorales.

Études de marché : Segmentation consumer basée sur les marqueurs culturels.

Politiques publiques: Adaptation des services selon les profils territoriaux.

Recherche académique: Nouvel outil d'analyse sociologique quantitative.

9

Conclusion : La data au service de la compréhension sociale

Cette exploration des liens entre prénoms et politique démontre la puissance de la data visualization pour révéler les structures cachées de notre société. En transformant 2.5 millions de données administratives en visualisations interactives, nous avons pu mettre en lumière des phénomènes sociologiques complexes de manière accessible et engageante.

L'exercice rappelle que derrière chaque visualisation se cachent des choix méthodologiques, des biais potentiels, et surtout des histoires humaines. Les prénoms ne sont pas que des étiquettes : ils portent les espoirs, les valeurs et les appartenances de générations entières.

Dans un monde où la data science prend une place croissante, cette approche humaniste de l'analyse de données offre une voie prometteuse pour comprendre nos sociétés sans les réduire à des algorithmes. Car si les données peuvent prédire nos votes, elles révèlent surtout la richesse et la complexité de ce que nous sommes.

Cet article s'appuie sur une analyse de données publiques françaises (INSEE, ministère de l'Intérieur) et respecte l'anonymat des individus. Le code source et les données sont disponibles sur GitHub pour permettre la reproductibilité des résultats.

Mots-clés: data visualization, sociologie électorale, prénoms français, analyse politique, Python, cartographie interactive, Plotly, Folium, démographie