Clustering top1

- Votes :

\*The clustering algorithm led to three distinct clusters with different characteristics in terms of voting preferences, as well as social and economic background. For now, only the candidate who finished in the first position has been considered.\*

\*The first cluster, \*\*Cluster 0\*\*, shows a strong preference for candidate Macron in the first round (he finished first in 44.25% of the cities in Cluster 0). We may assume, based on the results, that this cluster is center-right oriented, given that Fillon finished first in 21% and Le Pen in 25.7% of the cities.\*

\*On the other hand, \*\*Cluster 1\*\* demonstrated an almost absolute preference for Le Pen, with her finishing in first position in approximately 81.11% of the cities within this cluster. Based on the results this cluster can be qualified as extrem-right - right\*

\*The \*\*Cluster 2\*\* appears to be relatively balanced in terms of political preferences. Although Le Pen secured 39.31% of the cities, Macron and Fillon obtained 22.5% and 21.56%, respectively, while Mélenchon garnered around 15%.\*

\*Let’s investigate further into the characteristics.\*

- Age :

\*The age distribution of the population across clusters provides significant insight into the demographics of the communes. Each cluster highlights a different focus on age categories, which may have an impact on social behaviors and voting patterns.\*

\*In \*\*Cluster 0\*\*, the most prominent age group is the population between 40 and 54 years old, making up 28.9% of the cluster. Additionally, the elderly population (aged 65 to 79) constitutes 22%, while the younger population, those aged between 55 and 64, accounts for 17.9%. The youngest age group (18 to 24) is notably smaller, contributing only 7.42%.\*

\*In \*\*Cluster 1\*\*, the distribution remains quite similar, with the largest segment still being the population aged 40 to 54, which makes up 28% of the cluster. However, the cluster also has a notable representation of people aged 25 to 39 (22.7%) and a relatively high share of individuals between 55 and 64 years old (17.8%). Again, the youngest age group (18 to 24) remains smaller, at just 7.34%.\*

\*For \*\*Cluster 2\*\*, there is a slight shift with more balanced age categories. Here, the population aged 40 to 54 continues to dominate (24.4%), while the 65 to 79 group is well represented (23.3%). The 25 to 39 age group holds a significant portion (19.7%), and the population aged 55 to 64 shows a slight decrease compared to the other clusters (16.1%).\*

- Education :

\*This figure shows the distribution of educational levels across three clusters, highlighting the variations in education background among the different clusters.\*

\*In \*\*Cluster 0\*\*, a quarter (25.9%) of the population has a qualification equivalent to CAP or BEP, while 18.5% have a Baccalaureate or professional diploma. Another 15.6% have no qualifications beyond the CEP level. The remaining population is distributed across various higher education levels, with 14.3% having a Bac +2 and 10.7% having a Bac +3 or +4.\*

\*In \*\*Cluster 1\*\*, the distribution leans more towards higher education, with 32.2% of individuals having a CAP or BEP equivalent qualification and 22.9% holding no qualifications beyond CEP. The proportion of people with higher education is relatively lower, with 11% having a Bac +2 and 6.45% with a Bac +5 or more.\*

\*In \*\*Cluster 2\*\*, the highest proportion of the population holds a CAP or BEP equivalent (29.7%), followed by 27.4% having no qualifications beyond CEP. Similar to \*\*Cluster 0\*\*, higher education qualifications such as Bac +2, Bac +3, or more are also prevalent but less common, with Bac +2 at 8.95% and Bac +3 or +4 at 6.49%.\*

\*This chart illustrates how education levels vary between clusters, with some showing a strong presence of higher education, while others have a higher concentration of individuals with basic or no qualifications.\*

* Socio-pro :

**This figure presents the distribution of socio-professional categories across three distinct clusters, highlighting the varying occupational backgrounds of the populations within these clusters.**

*In****Cluster 0****, retirees constitute the largest proportion at 30.2%, followed by intermediaries at 20.8%. There is also a significant presence of employees (17.2%) and workers (12.9%). The smallest categories are professionals, with just 5.3% being executives or higher intellectual professions, and a minimal proportion of farmers (1.64%).*

*In****Cluster 1****, retirees also dominate with 31% of the population. Employees form the second-largest group at 20.5%, while 19.6% are intermediate professionals. The proportion of workers is lower at 16.2%, and, as with Cluster 0, executives make up 4.45%, while farmers are the least represented group with just 2.59%.*

*In****Cluster 2****, retirees take up an even larger proportion at 45%, indicating an older population in this cluster. Employees are next at 15.1%, while intermediate professionals make up 14.7%. Workers and executives account for smaller shares at 10.7% and 5.23%, respectively. Like the other clusters, the proportion of farmers remains small, representing 4.21% of the population.*

*This chart underscores the variation in socio-professional categories across clusters, with retirees consistently being the largest group, but their proportion is particularly high in Cluster 2. Meanwhile, intermediate professionals and workers form varying but significant portions in each cluster.*

* Urban density
* *\*The* **\*\*Cluster 0\*\*** *is mainly characterized by urban areas, with a significant proportion of urban belts (18%) and large urban centers (7.71%), alongside notable presence in rural areas like bourgs ruraux (21%) and small towns (3.71%). This suggests that Cluster 0 represents a mix of larger cities, their urban peripheries, and nearby rural areas, indicating that it encompasses both highly urbanized zones and smaller surrounding towns.\**
* *\*The* **\*\*Cluster 1\*\*** *stands out due to its predominance of rural areas with dispersed housing (69.9%), which dominates the cluster, while intermediate urban centers (1.96%) are present but much less significant. This indicates that* **\*\*Cluster 1\*\*** *is primarily rural, with small concentrations of urban areas, suggesting this cluster represents areas where population density is lower, and urbanization is minimal.\**
* *\*The* **\*\*Cluster 2\*\*** *is defined by its strong concentration of highly dispersed rural areas (39.7%), coupled with rural areas with dispersed housing (41.3%). These figures highlight that* **\*\*Cluster 2\*\*** *is the most rural of all, with sparse populations spread across vast areas, indicating little to no urban influence in this cluster.\**
* *\*Overall,* **\*\*Cluster 0\*\*** *represents urban areas and their surrounding belts,* **\*\*Cluster 1\*\*** *captures predominantly rural zones with occasional urban centers, while Cluster 2 is almost entirely rural, with scattered and highly dispersed populations.\**
* Income :

**This bar chart illustrates the median disposable income across three clusters, showing distinct income levels among the groups.**

*In****Cluster 0****, the median disposable income stands at the highest level of €24,328.08, well above the overall average. This suggests that the communes in this cluster are relatively wealthier compared to the other clusters.*

*For****Cluster 1****, the median disposable income is €20,873.89, which is slightly below the mean median income of €21,519.77. This indicates a relatively average income level for the communes in this cluster.*

*In****Cluster 2****, the median disposable income drops to €19,357.32, significantly below the average. The communes in this cluster represent areas with lower disposable incomes.*

*The red dashed line highlights the overall mean median income of €21,519.77, providing a clear comparison of how each cluster fares against the average income level.*

Clustering top3

- Votes :

\*The clustering algorithm led to four distinct clusters with varying characteristics in terms of voting preferences. In this analysis, we are considering the total points each candidate accumulated based on their ranking within each cluster.\*

\*The first cluster, \*\*Cluster 0\*\*, displays a relatively balanced voting preference. Macron leads with 14k points, followed by Le Pen at 10k points, and Fillon with around 11k points. This cluster appears to be politically diverse, suggesting a center-right to right-leaning electorate. Other candidates such as Mélenchon and Dupont-Aignan also received significant support, further emphasizing the diverse political tendencies of \*\*Cluster 0\*\*.\*

\*In contrast, \*\*Cluster 1\*\* is dominated by Le Pen, who secured close to 20k points. Both Macron and Fillon follow with approximately 16k points each, reflecting a strong competition between far-right and centrist candidates. Mélenchon holds a solid base of support but is overshadowed by the dominant figures of Le Pen, Macron, and Fillon. The results indicate that \*\*Cluster 1\*\* is strongly right-oriented, with a clear preference for far-right ideologies.\*

\*The \*\*Cluster 2\*\* demonstrates a decisive preference for Le Pen, who accumulated over 30k points, significantly outpacing other candidates. Macron and Fillon each earned around 20k points, while Mélenchon also shows strong support. This cluster highlights a more polarized voting behavior, with a heavy inclination toward far-right but also notable centrist and left-leaning support, indicating diverse political dynamics.\*

\*Finally, \*\*Cluster 3\*\* stands out due to Lassalle’s unexpected lead, with 25 points, surpassing even mainstream candidates like Fillon, who earned 18 points. Unlike the previous clusters, Le Pen and Macron do not dominate, showing lower scores. The success of Lassalle suggests that \*\*Cluster 3\*\* may represent rural areas or regions where alternative political figures gained traction, deviating from the broader national patterns observed in the other clusters.\*

\*These distinct voting behaviors across clusters highlight the varying political landscapes and preferences within the electorate. Further investigation into the socio-economic factors of these clusters could provide more insights into these patterns.\*

* Age

\*The clusters identified in this analysis highlight significant differences in the age distribution across French communes. Each cluster represents a distinct composition in terms of age categories, suggesting that voting preferences may correlate with the demographics of each cluster.\*

\*In \*\*Cluster 0\*\*, the largest age group is between 40 and 54 years old, accounting for 28.9% of the population. The population aged 65 to 79 also makes up a considerable portion (22%), followed by those aged 55 to 64 (18%). The younger population, aged between 18 and 24, is the smallest group, representing only 7.53%.\*

\*In \*\*Cluster 1\*\*, the dominant age group is again those aged between 40 and 54, making up 24.3% of the population. The younger group, aged 25 to 39, is also significant, at 19.7%. The cluster shows a more balanced distribution across age groups compared to \*\*Cluster 0\*\*, with a notable representation of the population over 65, which accounts for 23.3%.\*

\*In \*\*Cluster 2\*\*, the 40 to 54 age group continues to dominate, comprising 28.1% of the population, while the next most represented group is the elderly population (65 to 79 years old), at 22.8%. The 25 to 39 age group also holds a significant share, while the youngest group remains underrepresented.\*

\*Finally, \*\*Cluster 3\*\* presents a similar pattern with the highest concentration in the 40 to 54-year-olds (27.8%), followed closely by the population aged 55 to 64 (17.9%) and the elderly (65 to 79 years old) at 21.4%. The younger population (18 to 24 years old) again represents the smallest share.\*

\*The analysis of age categories across these clusters indicates that the middle-aged and older populations are heavily represented across the clusters, while the younger population (18 to 24) remains relatively small in all cases.\*

* Education

\*The chart above illustrates the average percentage of different education levels across four clusters. The data provides a breakdown of the educational attainment within each cluster, giving insights into the socio-educational characteristics of these groups.\*

\*The \*\*Cluster 0\*\* shows a diverse educational background, with 25.5% of the population having achieved a vocational education (CAP/BEP) and 18.4% holding a high school diploma (Baccalauréat or equivalent). Interestingly, 14.4% of this cluster has no formal education, suggesting a mix of both skilled and unskilled labor backgrounds.\*

\*In \*\*Cluster 1\*\*, the distribution is slightly different, with 27.6% of the population holding a vocational qualification (CAP/BEP) and 29.7% having completed high school. Notably, 8.9% of the cluster has a postgraduate degree, highlighting a higher level of advanced education compared to \*\*Cluster 0\*\*. This suggests that \*\*Cluster 1\*\* may have a more educated population overall.\*

\*We can see that \*\*Cluster 2\*\* has the highest percentage of individuals with higher education, as 32.1% of the population has at least a high school diploma and 11.1% possess a postgraduate qualification (Bac +4 or higher). This indicates that Cluster 2 has a more highly educated population overall, likely contributing to a more skilled and professional workforce.\*

\*Lastly, \*\*Cluster 3\*\* shows a similar trend to Cluster 2, with 32% of the population holding a high school diploma and 22.8% having completed vocational training (CAP/BEP). In contrast to other clusters, this group has a higher percentage of individuals without a formal education (9.49%), which could point to economic or geographical factors affecting access to education in certain areas.\*

\*In conclusion, the educational composition across these clusters varies significantly, indicating different socio-economic contexts. \*\*Cluster 2\*\* and \*\*Cluster 3\*\* stand out for their higher levels of educational attainment, while \*\*Cluster 0\*\* and \*\*Cluster 1\*\* represent a more diverse mix of education levels. Further analysis of these educational trends could reveal more about the voting preferences and socio-economic dynamics within each cluster.\*

* Socio-pro

\*The chart provides a breakdown of the socio-professional categories for each cluster, revealing the diversity in professional backgrounds across the different clusters.\*

\*In \*\*Cluster 0\*\*, a significant portion of the population is made up of retirees (30.1%), followed by employees (21%) and workers (17%). The mix of retirees and workers suggests that this cluster may consist of older individuals with a working-class background. Additionally, the relatively low percentage of managers and professionals (13.4%) further supports this hypothesis.\*

\*For \*\*Cluster 1\*\*, retirees represent the largest group at 44.9%, which suggests that this cluster is predominantly composed of older individuals. The second largest group is employees (15.1%), with a smaller proportion of intermediate professionals (14.9%). Interestingly, this cluster has a very low percentage of workers and farmers, indicating that it is less focused on manual labor or agriculture.\*

\*In \*\*Cluster 2\*\*, the distribution is more evenly spread out. Retirees still form a large group (30.9%), but there is also a substantial representation of employees (19.6%) and intermediate professionals (20.2%). This cluster appears to be more diverse in terms of socio-professional categories, with a notable presence of workers (16.4%) as well.\*

\*Finally, \*\*Cluster 3\*\* mirrors Cluster 1 with a high percentage of retirees (45%) and a significant share of intermediate professionals (12.9%). However, this cluster also has a larger percentage of artisans and shopkeepers (10.7%) compared to the other clusters, suggesting a more entrepreneurial or self-employed population.\*

\*Overall, the chart illustrates that \*\*Cluster 0\*\* and \*\*Cluster 2\*\* have a more diverse range of socio-professional categories, while \*\*Cluster 1\*\* and \*\*Cluster 3\*\* are dominated by retirees and professionals. This distribution provides insight into the economic and social structure of each cluster and how it may influence other variables such as political preferences.\*

* Urban density

\*The \*\*Cluster 0\*\* is characterized by a mix of bourgs ruraux (21.2%), urban belts (19.9%), small towns (3.9%), and a notable share of large urban centers (8.33%). This cluster captures a blend of rural and semi-urban regions, with some influence from larger urban areas.\*

\*The \*\*Cluster 1\*\* is primarily rural, with a significant presence of rural areas with very dispersed housing (38.9%) and rural areas with dispersed housing (41.5%). This indicates that the population is spread across wide, low-density areas.\*

\*The \*\*Cluster 2\*\* is also predominantly rural, defined by rural areas with dispersed housing (70.2%), highlighting its low-density rural nature.\*

\*The \*\*Cluster 3\*\* stands out for its dominance of rural areas with very dispersed housing (66.7%) and rural areas with dispersed housing (33.3%), making it a highly rural cluster with sparse populations.\*

\*In summary, \*\*Cluster 0\*\* represents regions with a diverse mix of rural towns and urban influences, \*\*Cluster 1\*\* is primarily rural with very dispersed populations, \*\*Cluster 2\*\* is highly rural with dispersed housing, and \*\*Cluster 3\*\* features the sparsest population distribution.\*

* Income

\*The chart provides an overview of the median disposable income for the top 3 clusters, highlighting the income disparities across these clusters.\*

\*In \*\*Cluster 0\*\*, the median disposable income is the highest at €24,565.33, which indicates that this cluster consists of individuals with relatively higher purchasing power. This suggests that Cluster 0 may represent more affluent populations, where income levels are above the national average.\*

\*For \*\*Cluster 1\*\*, the median disposable income drops to €19,340.68, marking a notable decrease compared to \*\*Cluster 0\*\*. This cluster represents a group with lower disposable income, indicating that \*\*Cluster 1\*\* might consist of individuals or households with more limited financial means.\*

\*In \*\*Cluster 2\*\*, the median disposable income rises again to €20,914.45, placing this cluster slightly above the overall mean. This cluster seems to have a more balanced income distribution, representing middle-income groups.\*

\*Lastly, \*\*Cluster 3\*\* has the lowest median disposable income at €18,548.89. This indicates that \*\*Cluster 3\*\* is composed of individuals or households with lower economic power compared to the other clusters.\*

\*The red dashed line across the chart shows the overall mean median disposable income, which is €20,848.94. This benchmark allows us to see how each cluster compares to the average, with \*\*Cluster 0\*\* and \*\*Cluster 2\*\* being above the mean, while \*\*Cluster 1\*\* and \*\*Cluster 3\*\* fall below it.\*

\*In summary, this chart helps to highlight the economic diversity between the clusters, ranging from affluent populations in \*\*Cluster 0\*\* to those with lower financial resources in \*\*Cluster 3\*\*. The median disposable income provides valuable insights into the socioeconomic structure of each cluster.\*

**Clustering top1 - Conclusion:**

**Cluster 0** reflects a center-right leaning, with Macron leading, followed by strong support for Fillon and Le Pen. This cluster represents more affluent and urban areas, where middle-aged populations are dominant, and education levels show a blend of vocational and higher education qualifications. It also has a notable representation of retirees and workers, highlighting a mix of urban influence and rural characteristics.

**Cluster 1** is heavily dominated by Le Pen’s far-right vote, indicating a right-leaning cluster, with older populations primarily residing in rural areas. Retirees make up the largest proportion, alongside fewer workers, pointing to an economically weaker, rural demographic with lower levels of higher education attainment.

**Cluster 2** presents a balanced political landscape, with Le Pen leading but also strong centrist and left-wing representation. This cluster is more rural, with a significant percentage of retirees and lower education levels, indicating a demographic of older, less educated populations in dispersed rural areas.

**Clustering top3 - Conclusion:**

**Cluster 0** is politically diverse, with Macron leading, followed by Le Pen and Fillon. It represents middle-income regions with a blend of urban and rural characteristics, marked by balanced age categories and varied educational backgrounds, featuring both vocational and higher qualifications.

**Cluster 1** is right-leaning, dominated by Le Pen, with competitive support from Macron and Fillon. This cluster has a strong presence of retirees and a relatively rural population, showcasing lower incomes and a more traditional socio-economic profile.

**Cluster 2** exhibits a far-right inclination with a heavy preference for Le Pen. It is highly rural, with low-density populations, and is economically weaker compared to other clusters. Education levels are lower here, and the cluster has a prominent share of retirees.

**Cluster 3** is uniquely characterized by Lassalle’s surprising lead, deviating from national patterns, with rural demographics driving its voting behavior. It has the lowest income levels, high retiree percentages, and fewer urban influences.

**Global Conclusion:**

The two clustering approaches provide valuable insights into the socio-political landscape of the communes. **Clustering top1**, focusing on first-place candidates, highlights clear political preferences tied to economic conditions and urbanization levels, with **Cluster 0** representing affluent, urbanized areas, **Cluster 1** showcasing far-right rural regions, and **Cluster 2** featuring more balanced preferences.

In **Clustering top3**, we delve deeper into the cumulative voting preferences, which reveal even greater political diversity within regions. **Cluster 0** continues to show a diverse political landscape tied to middle-income, semi-urban regions, while **Cluster 1** and **Cluster 2** remain largely rural with far-right tendencies.

The analysis of education, socio-professional categories, urban density, and income levels further solidifies the characterization of clusters. Overall, this dual clustering provides a comprehensive view of the interplay between political preferences, socio-economic factors, and geographic distribution across the French communes.