



## Quantifying Historical Migrations Using a Multi-Step Probabilistic Algorithm and Surname Distributions over the Centuries: A Case Study of Malopolska

Michal Kordys, Independent scholar, Poland, [michal.kordys01@gmail.com](mailto:michal.kordys01@gmail.com)

---

Contemporary studies on historical migrations gradually demystified the idea that the population in the past was stable and homogenous, and unveiled dynamic and complex connections throughout the whole of Europe. However, even if we are in the era of digitalization and big data, there is no standardized methodology to investigate historical migrations, and there are major challenges in quantifying and providing consistent estimates of migration groups. This research aims to tackle these challenges by proposing a replicable approach, consisting of a multi-step algorithm that uses probabilistic analysis to quantify migration and estimate the most likely migrants' origin based on surname distributions over the centuries. Particularly, the focus is on the Malopolska case study, on both international and Polish migrations to the region from the 1500s to the Great War. The proposed approach to quantification of historical migrations can be a helpful tool to use in combination with existing methodologies to validate and increase the accuracy of the estimates of other case studies in Europe.

---

Les études contemporaines sur les migrations historiques ont peu à peu démystifié l'idée selon laquelle la population passée était stable et homogène, et ont mis en lumière des connexions dynamiques et complexes à l'échelle de l'Europe entière. Cependant, même à l'ère de la numérisation et des mégadonnées, il n'existe pas de méthodologie standardisée pour étudier les migrations historiques, et il reste de grandes difficultés à quantifier et fournir des estimations cohérentes des groupes de migrants. Cette recherche vise à relever ces défis en proposant une approche reproductible, fondée sur un algorithme en plusieurs étapes qui utilise une analyse probabiliste pour quantifier la migration et estimer l'origine la plus probable des migrants à partir des distributions de noms de famille au fil des siècles. L'étude se concentre en particulier sur le cas de la Malopolska, en analysant les migrations internationales et polonaises vers cette région, du XVI<sup>e</sup> siècle jusqu'à la Grande Guerre. L'approche proposée pour la quantification des migrations historiques peut constituer un outil utile, à combiner avec les méthodologies existantes, afin de valider et d'affiner les estimations d'autres études de cas en Europe.

---



## Introduction

Studying historical migrations is critical to understanding history and the roots of modern societies (Guo et al. 2015). In the past, historical migrations were investigated by qualitative analysis and archives with a limited amount of records (Guo et al. 2015). Nowadays, with the gradual digitalization of historical data, there are more and more possibilities to tackle historical migrations more quantitatively and longitudinally (Guo et al. 2015). However, even modern migration analysis is still challenged by the lack of data “availability,” “reliability,” and “comparability” across countries (Barni and Extra 2009). Factors such as limited access to archives, unstructured records, different paces of digitalization across countries, various definitions of ethnicity, and political and geographical changes over the centuries can lead to scattered and inconsistent results. In fact, until now, we have not had a “systematic” methodology to quantify and make estimates on historical migrations (Hogan and Kertzer 1985; Guo et al. 2015), especially when we want to analyze migration patterns over an extended period of time.

An interesting study that addressed these challenges and proposed a way to overcome the lack of a systematic approach to the quantification of historical migration was conducted by (Guo et al. 2015). The authors utilized big data available on a genealogy website (family trees created by users) to analyze migrations, especially within the United States. The authors compared collected birth records with the US 1880 census, confirming that family trees help analyze historical migrations. A few years later, the same authors expanded the study on family trees with records of people born in North America and Europe from 1630 to 1930 and created a framework to collect and clean the data and fuzzy link trees to create the largest longitudinal geo-social network at the time (Koylu et al. 2020). Again, they validated the results with the US 1880 census. The use of genealogy websites is steadily becoming a popular source for investigating population patterns longitudinally over a long period of time (Koylu and Kasakoff 2022; Koylu and Kasakoff 2024); however, we still observe a main focus on the United States.

In Europe, retrieving data from official sources is far more complex: countries adopt different record methodologies (Baffour and Valente 2012), making it difficult to assess cross-border mobility. Moreover, the first unified European census was only available in 2011 (Eurostat 2025), and therefore, it cannot be used to validate historical migrations. Moreover, census data shows only a snapshot in time and does not provide enough information to analyze and validate historical migration patterns over centuries. For these reasons, in Europe, it is still difficult to introduce a systematic approach to studying historical migrants across countries, and more commonly, we observe studies focusing on internal migrations or specialized research topics (Drouhot et al. 2022). A few relevant examples consist of historical studies using surname distributions: one

concerning the whole of Spain to investigate overall internal migrations over the past five centuries (Rodríguez-Díaz, José Blanco-Villegas, and Manni 2017), one concerning the Barcelona Area between 1451 and 1900 to cluster and identify migrations of Catalan origin (Jordà, Pujadas-Mora, and Cabré 2016), and one concerning migration in the Dijonnais region from 1376 to 1610 with focus on anthropotonyms (Darlu et al. 2012).

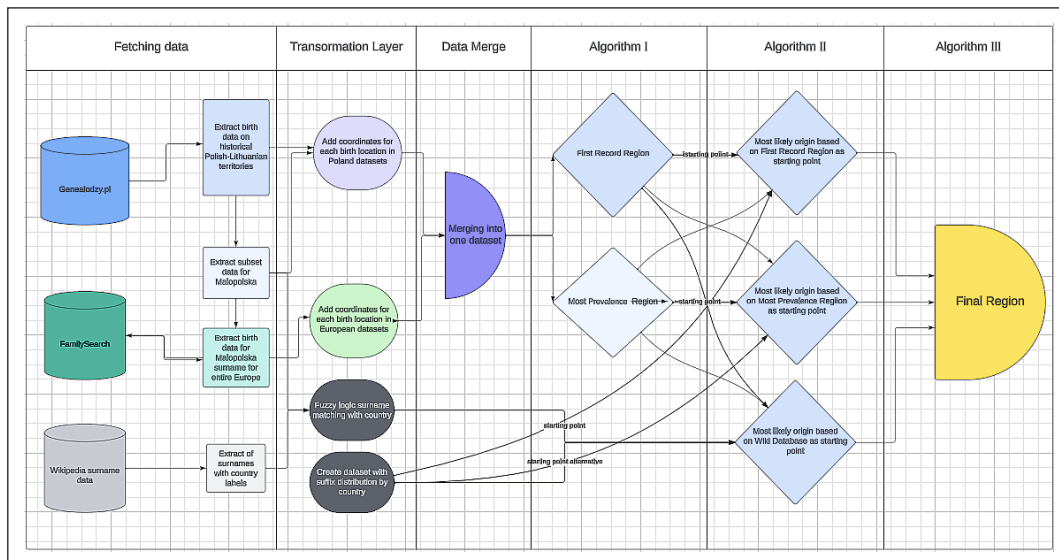
Therefore, we are still in need, especially in Europe, for a more replicable and systematic approach to quantify international migrations over the centuries. This study aims to address this research gap and propose a new methodology combining several successful approaches from previous studies. Particularly, the new methodology takes advantage of digitalized historical registers and big data from online and genealogy sources to analyze surnames' distributions over the centuries. Birth records are used to focus on long-term migrants rather than temporary migrants, who usually are unmarried and without children (Hayhoe 2016). Considering the biases that come with every source, multiple data sources are used and cross-validated, and a multi-step probabilistic algorithm is designed to minimize human biases when establishing surnames' origin and provide more relevant estimates on historical migrations. This automated search for surnames' origin across sources is helpful in reducing biases due to, for example, the lack of a generally accepted concept of ethnicity in the past (Conrad, Graf, and Wille 2023), etymological inconsistencies, or vague assignments by contemporary records (for example, Polish contemporary records mentioning "Olendrzy," which means Dutch, might refer to Dutch migrants but also Germans, Belgians, or any other migrant who settled following settling rights and laws granted initially to Dutch settlers) (Wójcik-Żołądek 2014). As each source has its intrinsic bias, the proposed approach can be used in combination with other existing methodologies as an additional dimension to validate and increase the accuracy of migration estimates.

In order to test this new methodology, we consider the case study of Malopolska. Malopolska from the 1500s to 1914 is an interesting case study, showing dynamic migrations from multiple groups such as Ukrainians, Belarusians, Germans (Conrad, Graf, and Wille 2023), Dutch (Conrad, Graf, and Wille 2023), Italians, French (Kowalski 2019), Czech (Hanczakowski 2019), Scots (Devine and Hesse 2012), and more. During the 1500s and 1600s, Malopolska was a capital region of the Kingdom of Poland and was an attractive destination for Western Europeans: it had wealthy trading centres, it was free from religious wars and counter-reformations (Wasiak 2022), and political power was also strong, especially under King Sigismund (Nowakowska 2018). The following century marked the slow decline of the Polish-Lithuanian Commonwealth and the rise of colonial empires in the Northwest as alternative migration destinations (Kowalski 2016). During the partitions of the Polish-Lithuanian Commonwealth in the late 1700s,

Southern Poland fell under the Habsburgs, and the new ruler, Maria Theresa (and later Joseph the Second), started a colonization campaign in the eastern provinces of the Empire, bringing around 38,000 migrants to the East of the Empire (Wójcik-Żołądek 2014). The following century, up to the Great War, presented a strong emigration trend, and between 1870 and 1914, approximately 3.5 million people left Poland (IPN 2024). These migration patterns described by historical studies are compared with our results in order to evaluate overlaps and validate our proposed methodology. In the future, we aim to extend our approach to other European case studies.

## Research process and methods

Firstly, the research commences with the acquisition of historical birth records in Malopolska from the 1500s to the Great War. The use of historical birth records allows the exclusion of temporary migrants and focuses on locals and long-term migrants who settled in the region and had children. For each unique surname found in the Malopolska archive, we collect all the available birth locations associated with that surname across Europe over the centuries. Secondly, coordinates are added to standardize how locations are encoded and better customize regions of origin. Finally, a multi-step probabilistic algorithm is designed to analyze surnames' distributions and establish in an automated fashion the most likely region of surnames' origin, quantifying the presence of different migration groups in Malopolska over time. It is important to note that the algorithms use multiple online sources to cross-validate surnames' origin and reduce intrinsic biases of each data source. Python 3 programming language is used to perform all these steps (**Figure 1**).



**Figure 1:** Process diagram.



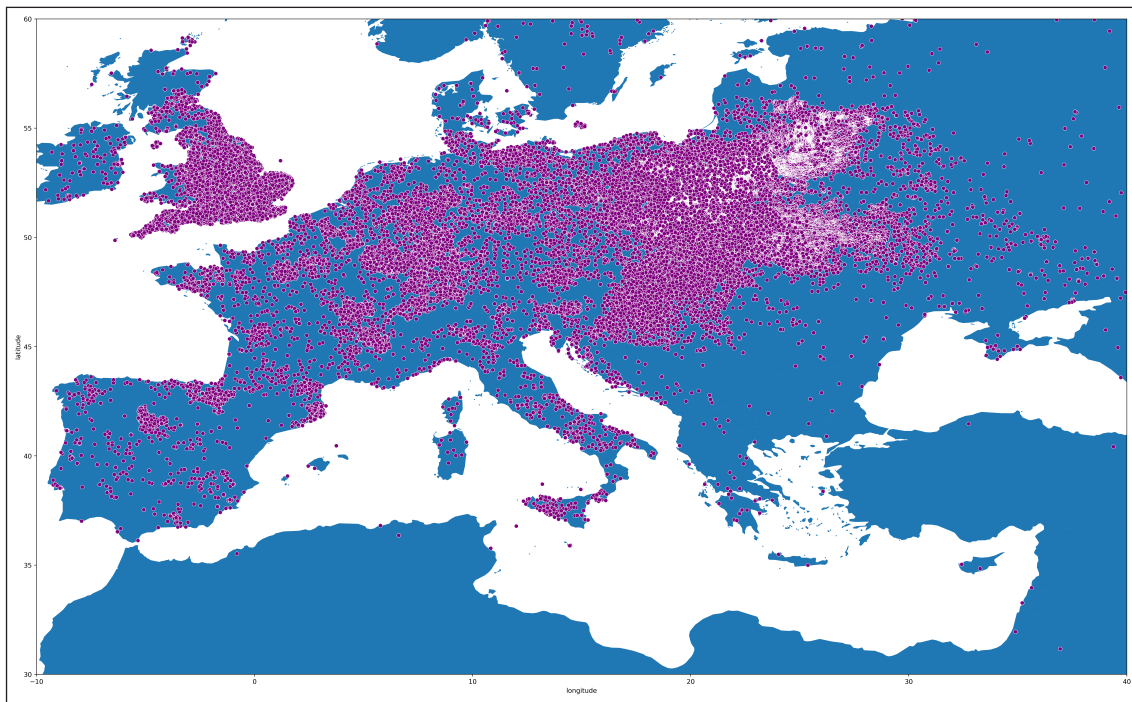
### **Data sources**

First of all, data is collected from all these sources:

- **Genealodzy.pl** (Geneteka 2025): Genealodzy is an open-source database with historical records from Poland and Polish–Lithuanian Commonwealth territories. This digital archive is used to find birth records in Malopolska between the 1500s and the Great War. The most common surnames found in this archive (50,227 surnames, comprising more than 80% of all the records) are used for further analysis.
- **FamilySearch.org** (FamilySearch 2025): FamilySearch is an open-source database with various worldwide historical records on population. This source is used to find all the available records on each surname found in Malopolska birth records. This information is relevant to finding the geographical distribution of a given surname in the whole of Europe over time.
- **Wikipedia** (Wikipedia 2025): Wikipedia is an open-source platform containing a well-structured dataset of the most typical surnames for each European country. This data is used as an additional source to assess a surname's origin by both analyzing suffixes and adopting Fuzzy Logic (Python library: thefuzz) (TheFuzz 2025) to find a probabilistic match with surnames from Malopolska birth records. However, as Wikipedia records are not historical but are based on current surname distribution by country, the origin might be biased. This is why the study assigns higher weights and importance to older historical records.
- **Genezanazwisk.pl** (Naruszewicz–Duchlińska 2012): Geneza Nazwisk is an open-source database containing detailed surname etymology and origin. It is used as an additional source to validate surname origin. This website has been chosen over similar sources created by other scholars as the database structure is suitable for automated web scraping.
- **Automated Google search** (Google LLC 2025): I designed a script to collect information when searching for a given surname's origin on the Google search engine. This source is used as an additional method to assess surnames' origin, but as the results can come from any website and not necessarily verified sources, less weight and importance are given compared to older historical records.

It is important to mention that both Genealodzy and FamilySearch are run by the Church of Jesus Christ of Latter-Day Saints. Particularly, in the Genealodzy archive, which is our starting point for collecting surnames of people inhabiting Malopolska between the 1500s and 1914, we would expect to see more records from communities

other than Christian. For example, only 0.14% of records of the entirety of Poland are labelled as Jewish. This number goes down to 0.07% for Malopolska. It is far below the estimates of 10% (Conrad, Graf, and Wille 2023). In addition to that, we did not find records labelled as Muslim. Therefore, the reader who wants to obtain a percentage distribution for all migrant groups inhabiting Malopolska over the centuries should adjust the results by the estimated size of these communities (**Figure 2**).



**Figure 2:** Occurrences of all surnames found in Malopolska across Europe.

### **Coordinates**

After collecting the input data, coordinates are added to all the locations found for a given surname in the whole of Europe over time. This step is essential to standardize locations and solve inevitable inconsistencies in the naming conventions. In fact, borders and regions' names have mutated drastically over the centuries, and on the other hand, each digital archive follows different guidelines and structures. The Python module used for finding and adding coordinates is `geopy.geocoders` (GeoPy Contributors 2018). Coordinates are then used to establish the diocese where the record is from and map it. Shapefile maps are open-source maps of fourteenth-century dioceses created by Stanford University (Stanford University Libraries 2025), and they are used in this study to provide more granularity (for example, to divide Italy into Northern Italy, Eastern Italy, and other parts) and customization of the regions, which would be otherwise unavailable in other maps.

### Algorithm I

After regions are defined and all available birth locations are assigned to their regions, the study proceeds with probabilistic estimates of surnames' origin through a multi-step algorithm. For simplicity, it has been divided into three parts. The initial estimates are produced in this first part of the algorithm, and next, they will be sequentially validated by Algorithm II and finally by Algorithm III. First of all, for each surname, the region of the oldest birth record is identified. Secondly, distributions across all the regions are calculated based on all the historical records available on that given surname. We acknowledge that a surname may have originated from one region and then thrived in another, and therefore, the distributions have been adjusted with logarithmic weights to ensure the oldest record receives greater importance in estimating surnames' origin. Weights are assigned to the distributions following the formula below:

$$\exp\left(-\frac{(Rw - Ro)}{(Rn - Ro)}\right),$$

where w belongs to (0,n)

Ro- Is the year of the first birth record found

Rn- Is the year of the last birth record found

Rw- Is the year of a given birth record

Once weights are assigned, the algorithm proceeds with a so-called Forward Walk to identify the most likely region of origin. The benefit of the Forward Walk is that it can capture long-term population movements replicating, to a certain extent, the steps taken by people while spreading throughout Europe over the centuries. In order to achieve that, a 2D matrix of the regions is designed so that regions that are geographically bordering each other are also bordering in the matrix (**Figure 3**). For each surname, two Forward Walks are performed. One starts from the region of the first birth record available, and one starts from the region with the highest prevalence of that given surname. In both cases, the algorithm starts analyzing the neighbouring regions in the matrix. Suppose the total percentage distribution of the starting region together with its neighbouring regions sums up to at least 90%. In that case, the algorithm stops as the high distribution around the starting point provides enough confidence in the region of origin.

On the contrary, if the sum of the distributions does not reach at least 90%, the algorithm keeps walking forward to the neighbouring region with the highest prevalence of that given surname and keeps summing up all the distributions of the neighbouring regions. This forward walk keeps moving forward as described until the

[ 0 , 1 , 2 , 3 , 4 , 5 , 6 , 7 , 8 , 9 ]
[ 10 , 11 , 12 , 13 , 14 , 15 , 16 , 17 , 18 , 19 ]
[ 20 , 21 , 22 , 23 , 24 , 25 , 26 , 27 , 28 , 29 ]
[ 30 , 31 , 32 , 33 , 34 , 35 , 36 , 37 , 38 , 39 ]
[ 40 , 41 , 42 , 43 , 44 , 45 , 46 , 47 , 48 , 49 ]
[ 50 , 51 , 52 , 53 , 54 , 55 , 56 , 57 , 58 , 59 ]
[ 60 , 61 , 62 , 63 , 64 , 65 , 66 , 67 , 68 , 69 ]
[ 70 , 71 , 72 , 73 , 74 , 75 , 76 , 77 , 78 , 79 ]
[ 80 , 81 , 82 , 83 , 84 , 85 , 86 , 87 , 88 , 89 ]

**Figure 3:** Matrix representing defined regions in Europe. Each number represents one region.

sum of the distributions reaches at least 90%, or until there are no further regions to investigate. The total distribution reached is the new confidence score assigned to the starting point region (region of the oldest record or region of highest prevalence). This output will then be used by the second algorithm together with other sources to validate surname origin estimates.

### **Algorithm II**

The second part of the multi-step algorithm is validating the estimates of surnames' origin generated by Algorithm I. In order to do that, for each surname, the algorithm performs three validations called Three Passes, which, based on comparisons with other sources, reinforce more or less the estimates on surnames' origin. To minimize biases, each Pass begins the validation across the same sources from a different starting point.

The Highest Prevalence Pass uses as a starting point the region of the highest prevalence of a given surname and its respective confidence score assigned by Algorithm I. As a first step, the algorithm checks whether this region matches the one where the surnames' suffix is most likely to belong. If there is a match, a multiplier is used to recalculate the score and adjust the confidence in the region. If there is a match with another high-prevalence region where the surname has been recorded, the percentage

distributions of other regions are adjusted accordingly, and all the estimates are then normalized. Based on these results, the starting point region can be either confirmed or reassigned. Then, the algorithm compares the region with the one of the first birth record, and if there is a match, the score is adjusted as in the previous step. Afterward, the algorithm compares the region with the one identified by the automated Google search and again adjusts the score if there is a match. Next, the algorithm compares the region with the one identified using Fuzzy Logic to find a probabilistic match with the most common surnames by country stored in Wikipedia, and the score is adjusted accordingly. Finally, the last score adjustment is given by comparing it with the region assigned by Geneza Nazwisk.

The First Record Pass uses the region of the first birth record and its respective confidence score assigned by Algorithm I as a starting point. As an initial step, the algorithm checks whether this region matches the one where the surnames' suffix is most likely to belong. If there is a match, a multiplier is used to recalculate the score and adjust the confidence in the region. Secondly, the algorithm compares the region with the one from the highest prevalence pass, and if there is a match, the score is adjusted as in the previous step. Finally, the algorithm follows the same steps as above. It compares the region with the one from the automated Google search, the one from the Wikipedia dataset, and the one assigned by Geneza Nazwisk, again adjusting the final score accordingly.

The Suffix Pass uses as starting point the region where the suffix of a given surname is the most common. Firstly, the algorithm compares the region with the one with the highest prevalence, and if there is a match, the score is adjusted. Secondly, the algorithm compares the region with the one of the first birth record, and if there is a match, the score is adjusted. Finally, the algorithm follows the same steps as above. It compares the region with the one from the automated Google search, the one from the Wikipedia dataset, and the one assigned by Geneza Nazwisk, adjusting the final score in the same way as above. When all Three Passes are completed, the regions and their adjusted scores are used to feed Algorithm III, which will make a final decision on surnames' origin.

### ***Algorithm III***

The last part of the multi-step algorithm consists of a set of rules that make decisions on the most likely region of origin of a given surname based on the regions and their respective scores produced by the Three Passes in Algorithm II. The prediction accuracy can be derived from the rule numbers: the first rules hold greater accuracy than the last ones. The following rules are applied, depending on the results previously obtained for each surname:

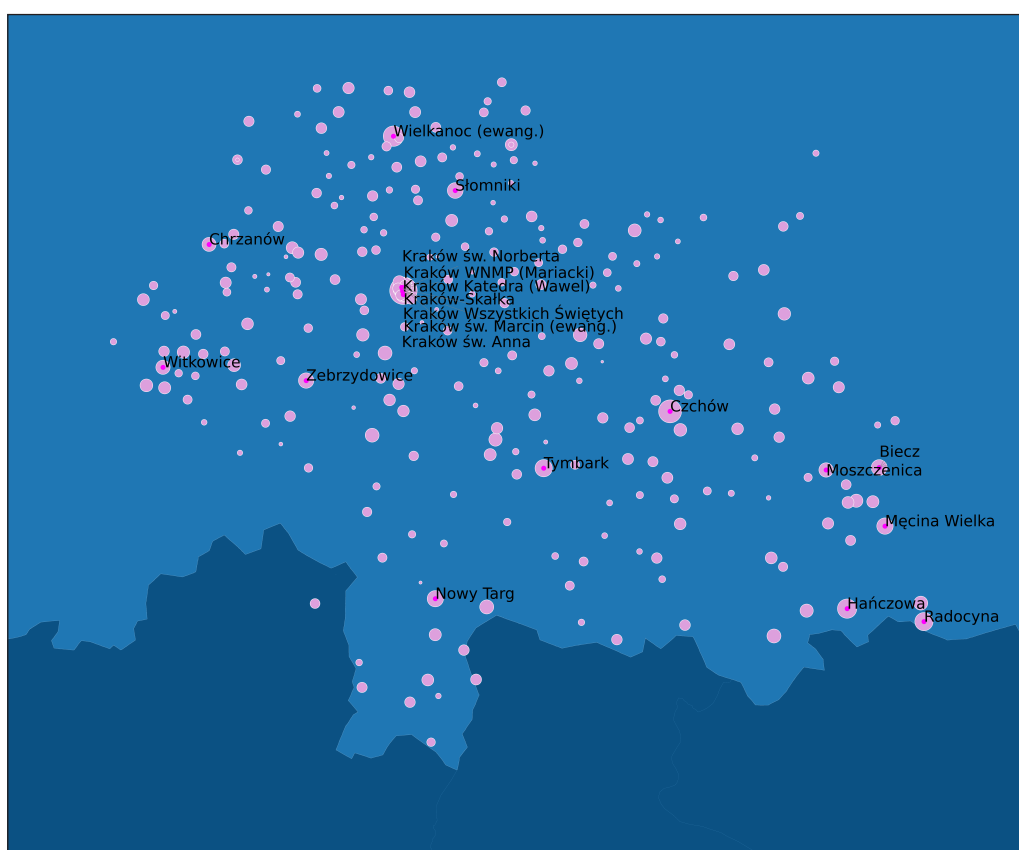
- **Rule 1:** If the region from the First Record Pass matches the region from the Highest Prevalence Pass, the final region of origin is the same.
- **Rule 2:** If the region from the First Record Pass does not match the region from the Highest Prevalence Pass, and both have the same confidence score, the algorithm takes the region of the first birth record as the final region of origin.
- **Rule 3:** If the region from the First Record Pass does not match the region from the Highest Prevalence Pass, and they do not have the same confidence score, the algorithm picks the region with the highest score.
- **Rule 4:** If the region from the First Record Pass and the region from the Highest Prevalence Pass are not available due to a lack of historical records, the algorithm takes the region from the Suffix Pass as the final region. However, this rule is applied only when a given suffix has high and medium confidence and therefore has more than 100 records available in Wikipedia database.
- **Rule 5:** Similarly to Rule 4, if the region from the First Record Pass and the region from the Highest Prevalence Pass are not available due to lack of historical records, the algorithm takes the region from the Suffix Pass as the final region. However, this time the rule applies only to suffixes which have between 30 and 100 records available in Wikipedia database.
- **Rule 6:** If surnames are not clearly identified, suffixes are primarily used to identify the region of origin of a given surname. However, the set of suffixes used in this rule are based on researches conducted by scholars. The most important suffixes are the following:
  - Rus suffixes: -WICZ (Gliński 2016; Baranivska 2021)
  - Ukrainian suffixes: -KO and -UK (Baranivska 2021)
  - German suffixes: -ACH, -HN, -HL, -EIT, -ERT, -DEL, -INK, -KEL, -MAN, -MANN, -SCH, -ICH, -HEL, -ATZ (Baker 2005) or containing characters such as “Ä, Ö, Ü, ä, ö, ü, ß” (only if a surname has been previously identified as belonging to Eastern or Central European regions)
  - Polish suffixes: -SKI, -SKA, -CKI, -CKA, -ZKI, -ZKA (Gliński 2016)
  - Neofici group: Neofici consist of people who converted to the Catholic faith and after baptism took a new surname. Frequently, the new surnames were created from weekday names or months (Gawryszczak 2014).

Once the most likely region of origin is found for all the surnames available in Malopolska birth records from the 1500s to the Great War, we can quantify the historical migrations over the centuries calculating the percentage of migrants coming from different regions in Europe.



## Research results

The final outcome of the multi-step algorithm provides estimates on the percentage of migrants settled in Malopolska over the centuries and coming from different regions in Europe. The results show a mosaic of people with different heritages throughout Malopolska. **Figure 4** shows places with the highest percentage of foreign surnames (the size of a dot represents the percentage distribution of non-Polish surnames). We summarized the results into three historical periods: from the 1500s to 1650, from 1650 to 1750, and finally from 1750 to 1914. For each period, we include a historical overview to provide context, the most relevant migration groups studied in previous researches as a benchmark, and the output of our multi-step algorithm with percentage estimates of the major migration groups.



**Figure 4:** Map of surname distribution (locations where the foreign surname percentage distribution is greater than 1%).

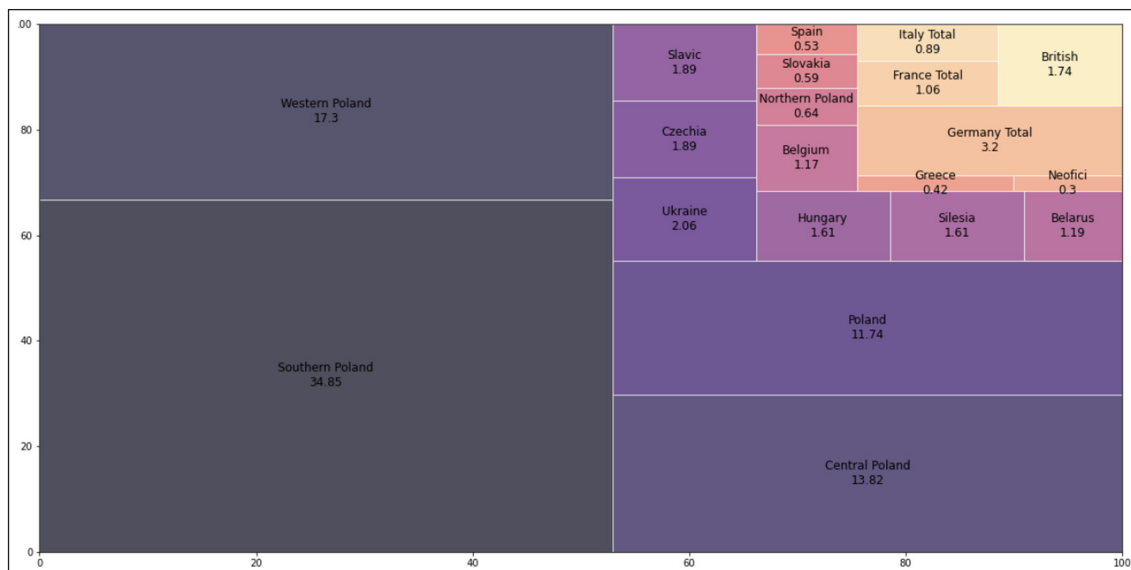
### *Migrations until 1650*

Between the 1500s and 1600s, Malopolska was the capital region of the Kingdom of Poland and was an attractive destination for Western Europeans for multiple reasons. On one side, Metropolitan Krakow was a wealthy trading centre that offered great

opportunities for traders, artisans, artists, and individuals seeking a better life. On the other side, the Commonwealth was free from religious wars and counter-reformations (Wasiak 2022), making it a refuge for migrants escaping violence in the West (Davies 1996, 504). Political power was also strong, proven by the Prussian Homage to the Kingdom of Poland in 1525 in Krakow and King Sigismund's success in secularizing the Teutonic Order, which stunned even the Pope (Nowakowska 2018).

According to Conrad, Graf, and Wille (Conrad, Graf, and Wille 2023), the entire Kingdom of Poland before the union with Lithuania in 1569 was estimated to have 70% of Poles, 15% of Ruthenians, and approximately 10% of Germans, the rest of the population being Armenians, Jews, Tatars, Vlachs (Romanians, Moldavians), and other (Conrad, Graf, and Wille 2023). Prior to the Golden Age in the 1400s, a group finding refuge from religious intolerance consisted of Czech Hussite migrants (Hanczakowski 2019). Later historical sources point to other migrant groups, such as the Dutch, predominantly protestants escaping religious prosecution and wars (Kowalski 2019). Kowalski describes Germans, Italians, French, and various ethnicities being welcomed by the protestant community of Krakow (Kowalski 2019). Finally, the mass migration of Scots is another key event in Polish history (Devine and Hesse 2012), with an estimated presence of 30,000 Scottish families in Poland (Kowalski 2016). In the 1500s, goods from Scotland constituted around 10% of total Gdansk imports, and in the first half of the 1600s, Scottish and English ships accounted for 12% of those entering Gdansk harbour (Kowalski 2016).

We can review the overall migration patterns described in previous studies and loosely compare our results (**Figure 5**) with the study conducted by Conrad, Graf, and Wille (Conrad, Graf, and Wille 2023) on the entirety of Poland before the union with Lithuania. The estimates provided by our algorithm show that up to 1650 Polish surnames constituted 77.7% of all available records (comparable with the estimate of 70% provided by the aforementioned authors). Northwestern Europeans (particularly ethnic Germans) were found to be most prevalent in this period, with a presence of 7.81% (relatively close to the 10% of Germans estimated by Conrad, Graf, and Wille). Eastern European surnames were 5.82%, and 2.68% were from the neighbouring Kingdom of Hungary (labels: Hungary, Slovakia, Romania, Croatia). The results also show a presence of 3.25% of Ukrainian and Belarus surnames: Conrad, Graf, and Wille estimated 15% of Ruthenians in the whole of Poland, and as expected, the number turned out to be far lower in Malopolska as the group was native to Red Rus and mostly present in the East. Surnames from the British Isles were nearly 1.75%, overlapping with the documented Scott migration to Poland (Devine and Hesse 2012). To put numbers into context, this is similar to the number of Poles in the United Kingdom in 2017 at the peak of migration (1.51%) (Statista 2020; Statista 2021). Other relevant migration groups found were Czech (1.85%), French (1.06%), and Italian (0.89%).



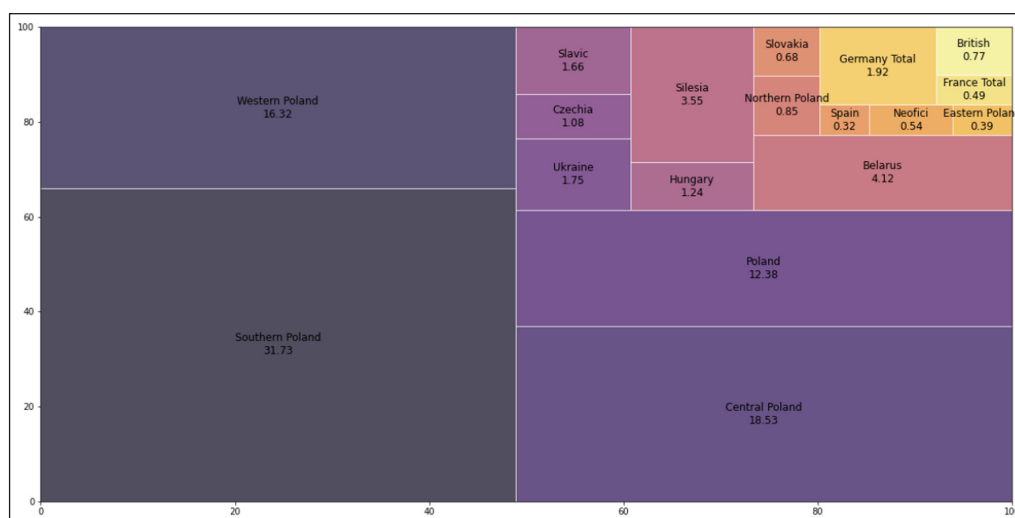
**Figure 5:** Distribution of surnames by region for the period until 1650 (values greater than 0.3%).

### ***Migrations between 1650 and 1750***

The following century marked the slow decline of the Polish–Lithuanian Commonwealth. According to a study conducted by Malinowski and Luiten van Zanden, the economic growth between 1662 and 1776 was half of that between 1500 and 1578 (Malinowski and Luiten van Zanden 2017). This can be attributed to climate change (Izdebski and Guzowski 2025) and the growing instability. Moreover, the rise of colonial empires in the Northwest and alternative migration destinations (Kowalski 2016) affected the migration landscape in Malopolska. In addition, Socinians were expelled from Poland during the wars with Sweden (Conrad, Graf, and Wille 2023), together with some other protestants (including some Scots) who supported the Swedish king (Devine and Hesse 2012). Dutch settlers (Olender) continued to migrate mainly to Northern and Central Poland (Marszał 2022). In this period, an important geopolitical change occurred behind the southern border, where Habsburgs steadily increased their possessions, and people faced the counter-reformation laws of the new Catholic rulers (Michels 2024).

In the results of this study (**Figure 6**), we can observe indeed an initial decline in surnames from regions such as France, Belgium, and the British Isles. However, German surnames rebounded between 1700 and 1750, which might be associated with a constant influx of Dutch settlers (as previously mentioned, this label was used for multiple ethnic groups) (Wójcik-Żołądek 2014). Similarly, French surnames rebounded, especially those originating from the area around the Rhine River Valley neighbouring Germany. Slovak (0.68%) and Silesia (3.55%) surnames slowly increased as both regions became Habsburg domains. Another group that steadily rose was the Rus (labelled “Belarus”), which reached 4.12% by the end of 1750. This increase is likely associated with the loss of

territories by the Polish–Lithuanian Commonwealth in the East. During this period, we also observe a growth in the number of “Neofici” surnames (Jewish converts), who could benefit from Chapter XII and Article 7 of the III Lithuanian Statute from 1588, stating that Jews who converted could join the ranks of nobility (Lipińska and Lipiński 2021).



**Figure 6:** Distribution of surnames by region for the period 1650–1750 (values greater than 0.27%).

### ***Migrations between 1750 and 1914***

Between 1772 and 1795, during the partitions of the Polish–Lithuanian Commonwealth, Southern Poland fell under the Habsburgs, and the new ruler, Maria Theresa, started a colonization campaign in the eastern provinces of the Empire, including Southern Poland. It is estimated that 38,000 migrants moved to the East of the Empire (Wójcik–Żołądek 2014). Her successor, Joseph the Second, continued these policies with new campaigns in the 1780s, bringing around 14,400 people to Southern Poland and 170 new settlements, out of which 120 were purely German speaking (Lepucki 1938). The following century, up to the Great War, presented a strong emigration trend from Poland. It is estimated that between 1870 and 1914, approximately 3.5 million people left Poland (Suleja 2022). In the Galicja region especially, the emigration started in 1880, and this wave notably affected the birth rate (Jura 2002). In the same period, especially between 1850 and 1914, an intense Jewish migration from Southern Poland and from the entire Galicja region to Hungary can be observed, as reported by the study of Csíki (Csíki 2022). The author also states that between 1900–1910, both Poles and Jews (approximately 135,000 and 120,000 people, respectively) left Galicja for Hungary, Silesia, and the Americas.

In general, this emigration trend can also be observed in this study’s results (Figure 7, Figure 8). Ukrainian and Belarusian (Rus) surnames declined compared

to previous periods, with Ukrainians levelling around 1.75–1.77%, while Belarusian presence kept decreasing until 1914, reaching 3.08%. French, Belgian, Dutch, and British groups declined further, as did other Western surnames. Malopolska was no longer an attractive migration destination. The only steady increase was among Slovakian and Silesian surnames, as those territories and Malopolska were both part of the Habsburg Empire. In addition to that, the number of German surnames grew compared to the previous period, together with Swiss surnames, which reached 0.28% at its peak in 1800–1850. Interestingly, in the records, we notice a rise in the occurrence of local Southern Polish surnames at the expense of generic Polish surnames ending with suffixes such as -SKI and -CKI.

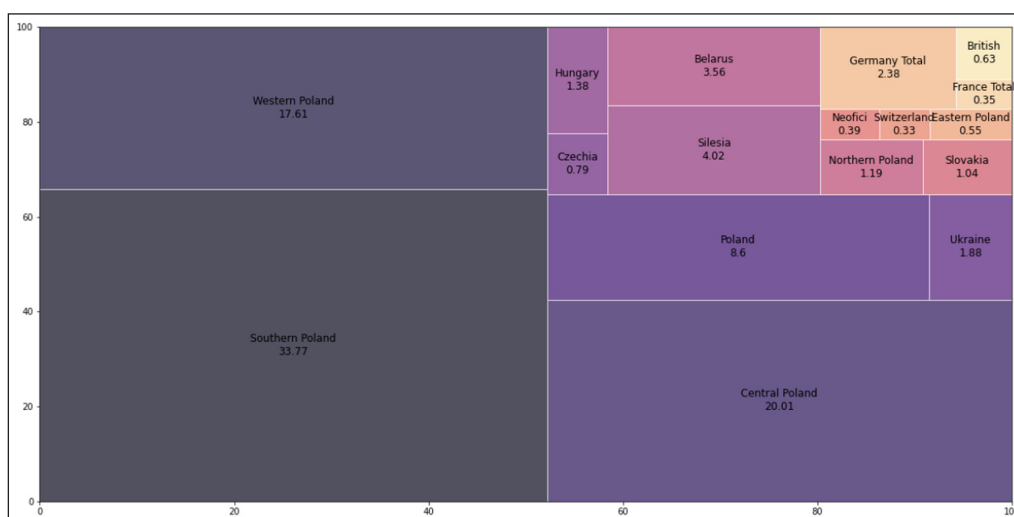


Figure 7: Distribution of surnames by region for the period 1750–1850 (values greater than 0.27%).

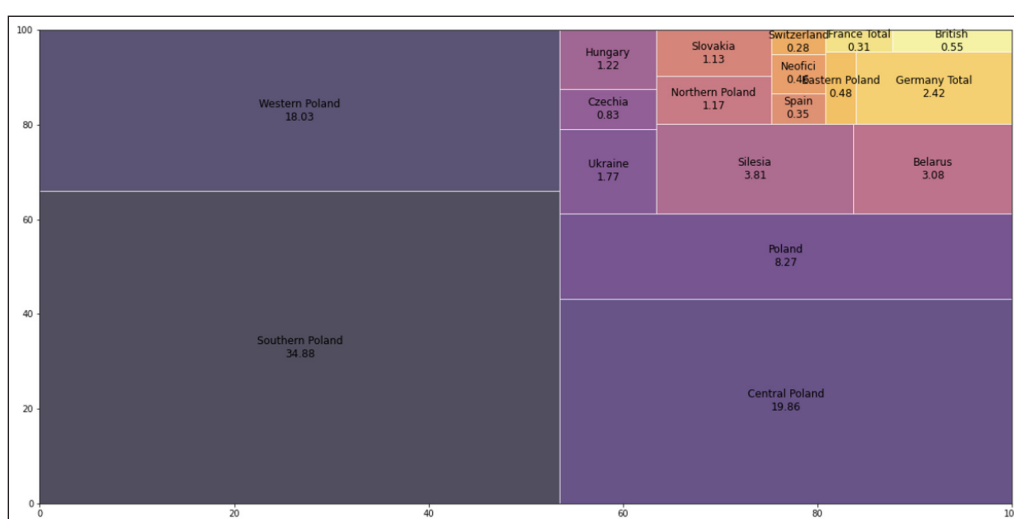
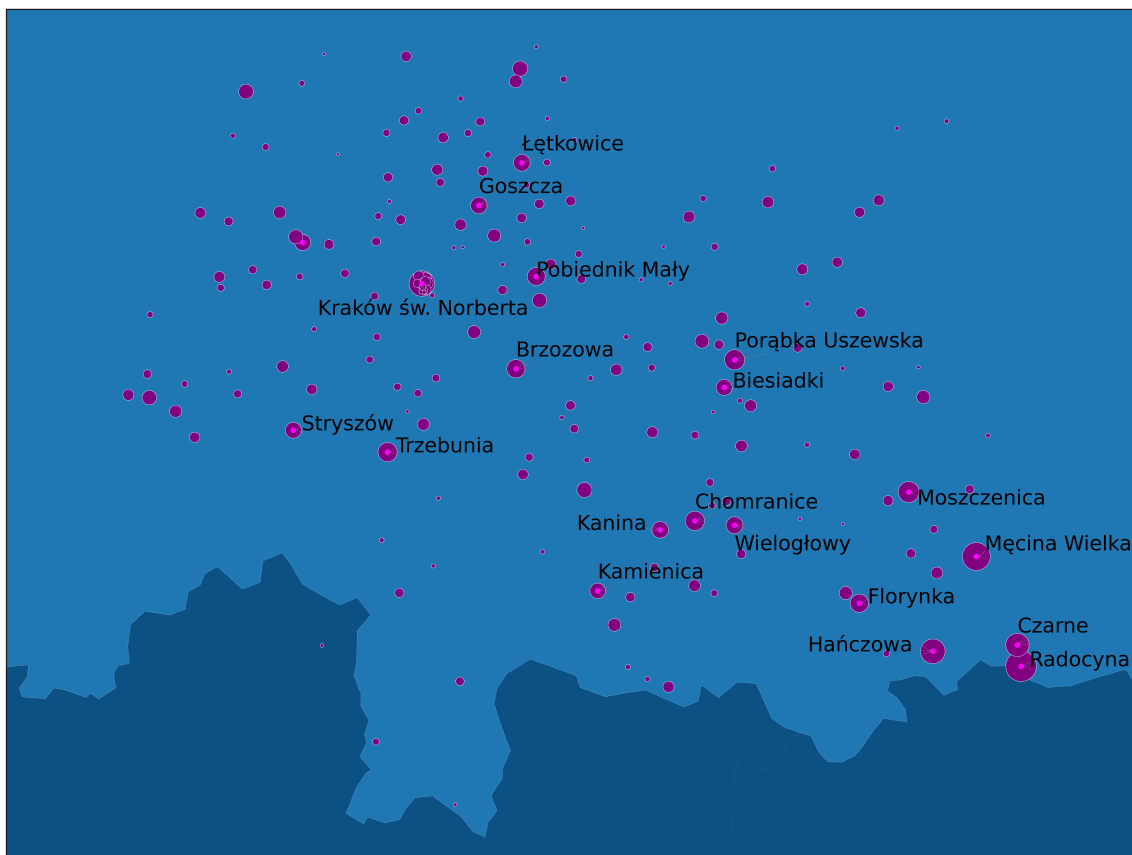


Figure 8: Distribution of surnames by region for the period 1850–1914 (values greater than 0.3%).

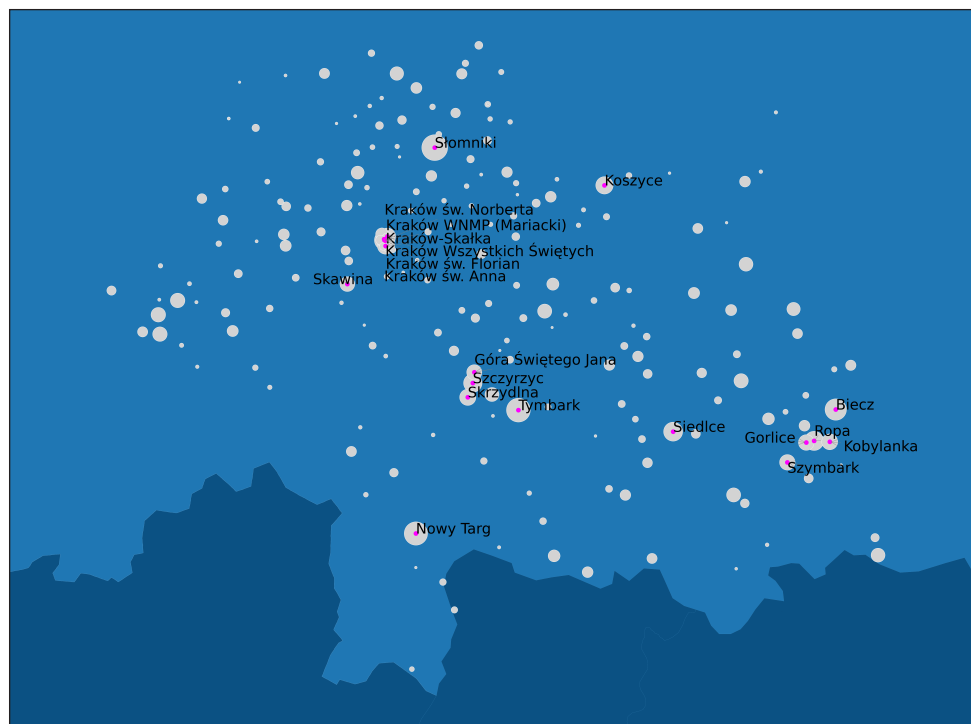
### Major migration destinations

After reviewing the distributions of the surnames throughout Malopolska over the centuries, we observed that certain migration groups tended to settle in specific areas in Malopolska. Ukrainian surnames tended to spread around the Northwest–Southeast axis and penetrate deeper into the Carpathian Mountains (**Figure 9**). Rus surnames appeared with the highest percentage in Krakow, Oświęcim, Tymbark, and Gorlice (**Figure 10**). German surnames aligned with the Northwest–Southeast line, which correlates with the trading routes cutting through the region (Krakow, Bochnia) (**Figure 11**). They also appeared in towns set up on Magdeburgian law (for example, Czchów), but they are quite uncommon in southern mountainous towns. Swiss surnames, similarly to German surnames, exhibited a unique alignment with the Northwest–Southeast line, especially in the location of Tuchów (**Figure 12**). Surnames from the British Isles were concentrated in the Northwest and some in the Gorlice region, including Roznowice and Ropa (**Figure 13**). In fact, Gorlice was a region of activity for Canadian–Scottish investors in the Oil industry (Miasto Gorlice 2025). Finally, surnames from the Kingdom of Hungary spread throughout Malopolska, with some visible prevalence in the West, Tatra Mountains, and Podhale (**Figure 14, Figure 15**).

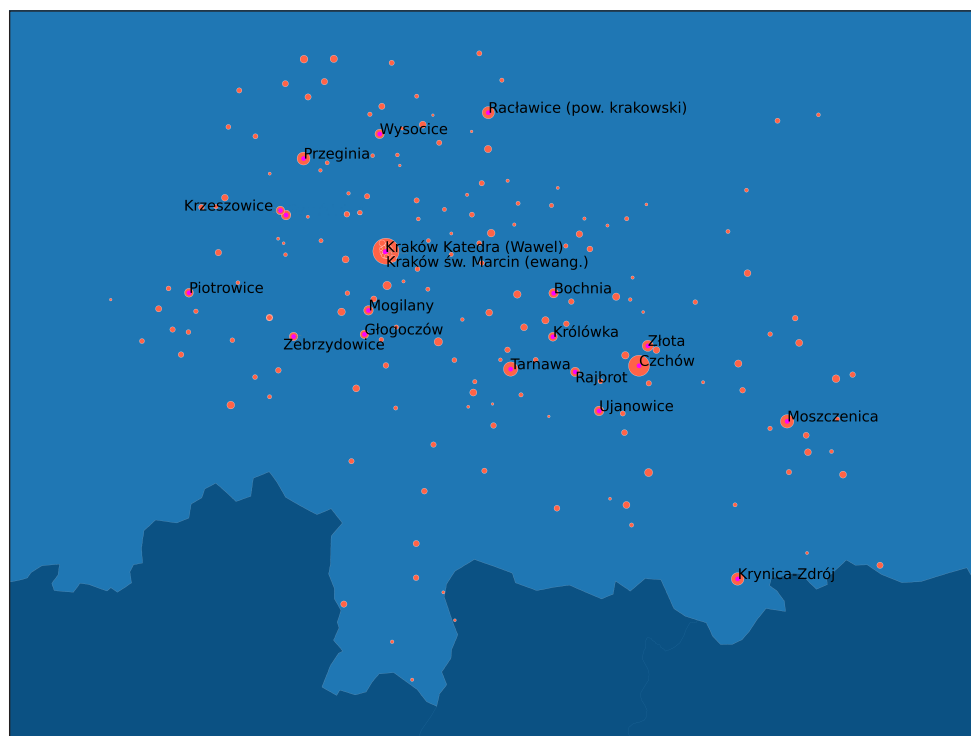


**Figure 9:** Distribution of Ukrainian surnames for parishes where their account is greater than 1%.





**Figure 10:** Distribution of Rus (label: Belarus) surnames for parishes where their account is greater than 1%.



**Figure 11:** Distribution of surnames for all combined four German regions for parishes where their account is greater than 1%.



Figure 12: Distribution of Swiss surnames for parishes where their account is greater than 1%.



Figure 13: Distribution of British surnames for parishes where their account is greater than 1%.

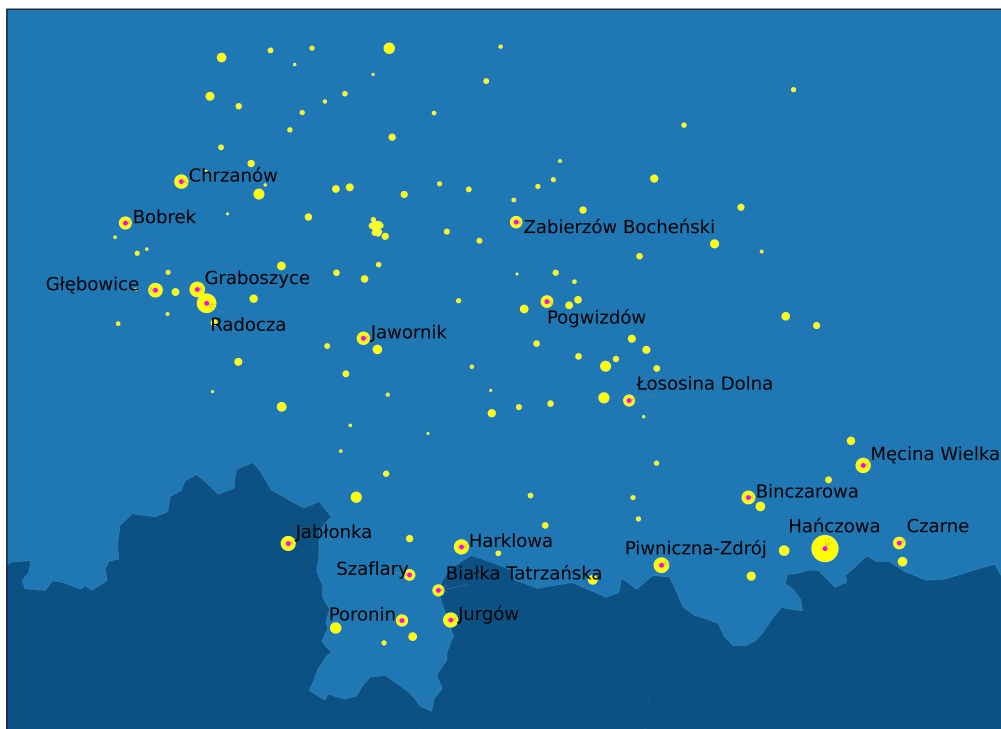


Figure 14: Distribution of Hungarian surnames for parishes where their account is greater than 1%.

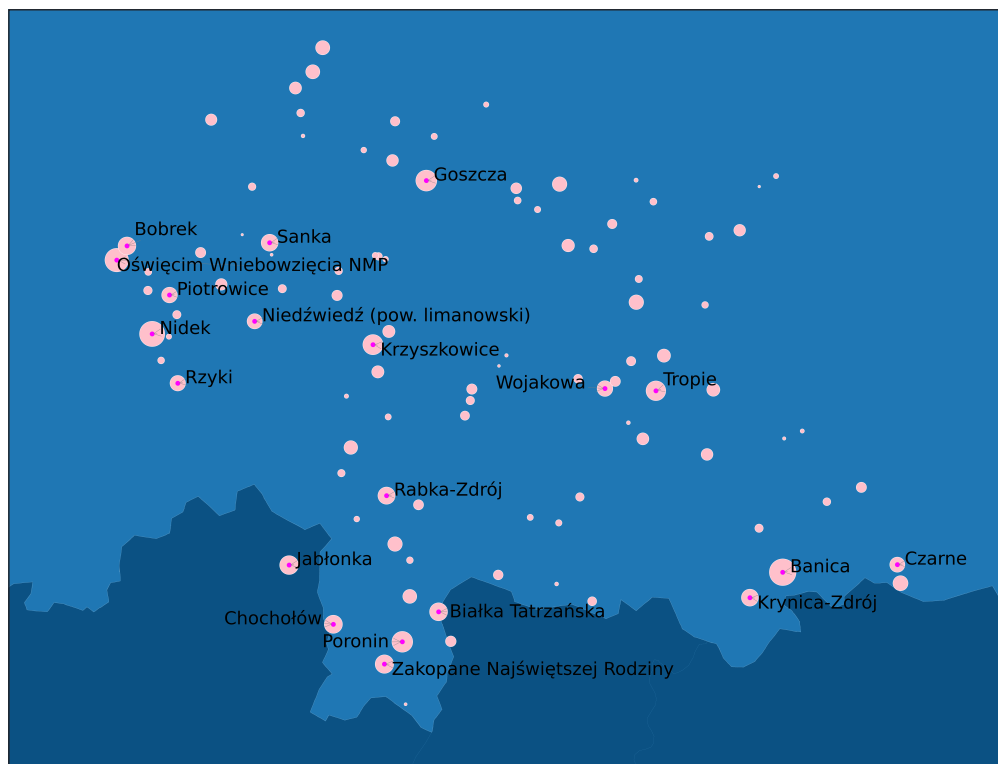


Figure 15: Distribution of Slovak surnames for parishes where their account is greater than 1%.

## Conclusion

In summary, the study of historical migrations has gradually mutated over the years, embracing new tools that come with the era of digitalization. However, to these days, we still have no “systematic” methodology to quantify historical migrations over the years (Guo et al. 2015) and to provide consistent estimates on migration groups (Hogan and Kertzer 1985), especially in Europe, where we still struggle with the lack of data “availability,” “reliability,” and “comparability” across countries (Barni and Extra 2009). To address these challenges, our study proposes a replicable and systematic approach to quantify historical migrations across Europe, using a multi-step algorithm that uses probabilistic analysis to estimate the most likely migrants’ origin based on surname distributions over the centuries. The algorithm uses multiple online sources to cross-validate surnames’ origin and reduce intrinsic biases of each data source.

The methodology consists of several sequential steps. First, historical birth records are extracted from a historical register. For each unique surname found in the archive, all the available birth locations associated with that surname across Europe and over the centuries are collected from an open-source genealogy website. Secondly, coordinates are added to standardize all the locations and customize the regions. Finally, a multi-step probabilistic algorithm is designed to analyze surnames’ distributions and establish in an automated fashion the most likely region of surnames’ origin, quantifying the presence of different migration groups over time. After identifying the region of the oldest birth record, calculating distributions across regions, and assigning logarithmic weights to give higher importance to older records, Algorithm I performs a so-called Forward Walk on a 2D matrix to capture long-term population movements and identify the most likely region of origin. These initial estimates are provided to Algorithm II, which performs three validations called Three Passes. Each Pass compares initial estimates with other sources to reinforce more or less the confidence in surnames’ origin. Each Pass uses the same sources to minimize biases but starts the validation from a different starting point. Finally, based on the estimates produced by the Three Passes, Algorithm III makes a decision on the most likely region of origin of a given surname, following a set of Rules.

The case study analyzed to test this methodology was focused on international and Polish migrations in Malopolska from the 1500s to the Great War. We could observe an overlap between some historical estimates available for the entirety of Poland and our results. Moreover, the overall migration patterns and ethnic groups identified by our algorithm are corroborated by the trends and events described in history. In more detail, our estimates confirmed that the period up to 1650 was indeed the Golden Age for migrations in Malopolska, as it was an attractive destination from economic, social,

and political perspectives. We found that up to 1650, Polish surnames constituted 77.7% of all available records, documented at 70% in the entirety of Poland in a similar period (Conrad, Graf, and Wille 2023). Northwestern European surnames (particularly ethnic German) were 7.81%, close to 10% of Germans in Poland reported in previous research (Conrad, Graf, and Wille 2023). Eastern European surnames were 5.82%, and 2.68% were from the neighbouring Kingdom of Hungary. The results also found 3.25% of Ukrainian and Belarus surnames, which is expected to be far less than in the whole of Poland, where previous studies accounted for 15% of Ruthenians (Conrad, Graf, and Wille 2023), mostly present in the East. Surnames from the British Isles were nearly 1.75%, overlapping with the documented Scott migration to Poland (Devine and Hesse 2012). Other relevant migration groups recorded were Czech (1.85%), French (1.06%) and Italian (0.89%).

In the following century (1650–1750), with the slow decline of the Polish–Lithuanian Commonwealth and the rise of colonial empires in the Northwest as alternative migration destinations (Kowalski 2016), our results also underlined a decline of surnames from regions such as France, Belgium, and the British Isles. However, at the end of the 1700s, Southern Poland fell under the Habsburgs, and the colonization campaign of the East of the Habsburgs Empire increased again the number of migrants (Wójcik–Żołądek 2014). Similarly, our results showed a rebound of German and French surnames. On the contrary, the period from 1850 to the Great War presented a strong emigration trend, with approximately 3.5 million people leaving Poland between 1870 and 1914 (Suleja 2022). This trend is also visible in our results, revealing a decline in Ukrainian, Belarusian, French, Belgian, Dutch, British surnames and more Western groups. In summary, all these dynamic migration patterns in Malopolska, which started with a period of splendor, followed by an initial decline, a rebound, and a final decline, are all captured by our algorithm.

Although the results are promising, some enhancements could increase the accuracy and confidence in our proposed methodology. For instance, clustering of similar surnames to find a common origin (for example, surnames such as Müller, Muller, Miller, and Miler, currently treated as separate surnames by the algorithm). Additionally, including records from Jewish and Muslim communities would provide a more comprehensive picture of the historical migrations, as, unfortunately, no open-source records for Malopolska could be found at this time. Additionally, deeper comparisons with existing case studies conducted by previous researchers on migration and surname distribution could be useful to compare methodologies and potentially combine them for better estimates. Another interesting future development could be investigating whether results from the algorithm can lead to identifying DNA origins, considering that surnames tracking has already proved to be a valid alternative, as

demonstrated by Guglielmino and De Silvestri (Guglielmino and De Silvestri 1995) as early as 1995. Finally, we hope this methodology will be used further to investigate new case studies in Europe and start building a greater collection of historical migrations in Europe over the years, which would provide a better understanding of the magnitude of migrations flows, which are otherwise difficult to quantify considering only one region.

In conclusion, European history is intrinsically complex and “polyvocal” (Hansen et al. 2023), and in the past years, there have been several debates and reflections on how it should be studied and researched (Hansen et al. 2023). Particularly, a new “multi-perspective” approach has emerged, which focuses on “mutual interaction, exchange, and transnational contact [...] without overlooking local specificities” (Hansen et al. 2023). In line with this approach, our results highlight how “macro” and local historical events are not isolated but have a cross-regional impact and strongly influence individuals’ choices and movements over the years. In fact, in its specific way, the case of Malopolska shows how past events happening throughout Europe brought people from different places to meet, assimilate, and build a rich and multifaceted heritage. We hope this research and future developments can contribute to further building a “multi-perspective” European history.

---



## Disclosure statement

Python scripts used in the study are available on Gitlab: <https://gitlab.com/michalk1888/surname-migration>. Dataset (in .pkl format) with final results is available upon request, after stating purpose of the request. Due to large volume, it could not be uploaded on Gitlab.

## Competing interests

The author has no competing interests to declare.

## Contributions

### Editorial

#### Section Editor

Frank Onuh, The Journal Incubator, University of Lethbridge, Canada

#### Copy Editor

Christa Avram, The Journal Incubator, University of Lethbridge, Canada

#### Layout Editor

A K M Iftekhar Khalid, The Journal Incubator, University of Lethbridge, Canada

---

## References

- Baffour, Bernard, and Paolo Valente. 2012. "An Evaluation of Census Quality." *Statistical Journal of the IAOS* 28 (3–4): 121–135. Accessed June 15, 2025. <https://doi.org/10.3233/sji-2012-0752>.
- Baker, Theola Walder. 2005. "Internal Dialectal Clues in German Surnames." *Missouri State Genealogical Association Journal* 25 (3): 172–177. Accessed June 15, 2025. [https://www.mosga.org/upload/journal/Volume\\_25,\\_2005reduced.pdf](https://www.mosga.org/upload/journal/Volume_25,_2005reduced.pdf).
- Baranivska, Oksana. 2021. "Wpływy wschodniosłowiańskie (ukraińskie) w nazwiskach mieszkańców województwa kujawsko-pomorskiego." *Postscriptum Polonistyczne* 3 (1): 121–130. Accessed June 15, 2025. <https://journals.us.edu.pl/index.php/PPol/article/view/11116>.
- Barni, Monica, and Guus Extra, eds. 2009. *Mapping Linguistic Diversity in Multicultural Contexts*. De Gruyter Mouton.
- Conrad, Benjamin, Tobias P. Graf, and Arndt Wille. 2023. "Interethnic Relations in Early Modern History (ca. 1500–1800)." In *The European Experience: A Multi-Perspective History of Modern Europe, 1500–2000*, edited by Jan Hansen, Jochen Hung, Jaroslav Ira, Judit Klement, Sylvain Lesage, Juan Luis Simal, and Andrew Tompkins, 167–176. Open Book Publishers. Accessed June 15, 2025. <https://doi.org/10.11647/OBP.0323.16>.
- Csíki, Tamás. 2022. "The Immigration of Galician Jews to Hungary in the Age of the Austro-Hungarian Monarchy, 1867–1914." *Studia Historyczne* 62 (4[248]): 43–61. Accessed June 15, 2025. <https://doi.org/10.12797/sh.62.2019.04.03>.

- Darlu, Pierre, Gerrit Bloothoof, Alessio Boattini, Leendert Brouwer, Matthijs Brouwer, Guy Brunet, et al. 2012. "The Family Name as Socio-Cultural Feature and Genetic Metaphor: From Concepts to Methods." *Human Biology* 84 (2): 169–214. Accessed June 15, 2025. <https://doi.org/10.1353/hub.2012.a479284>.
- Davies, Norman. 1996. *Europe: A History*. Oxford University Press.
- Devine, Tom M., and David Hesse, eds. 2012. *Scotland and Poland: Historical Encounters, 1500–2010*. John Donald.
- Drouhot, Lucas G., Emanuel Deutschmann, Carolina V. Zuccotti, and Emilio Zagheni. 2022. "Computational Approaches to Migration and Integration Research: Promises and Challenges." *Journal of Ethnic and Migration Studies* 49 (2): 389–407. Accessed June 15, 2025. <https://doi.org/10.1080/1369183x.2022.2100542>.
- Eurostat. 2025. "Population and Housing Censuses." European Union. Accessed June 15. <https://ec.europa.eu/eurostat/web/population-demography/population-housing-censuses>.
- FamilySearch. 2025. "What Will You Discover about Your Ancestors?" Accessed June 15. <https://www.familysearch.org/>.
- Gawryszczak, Anna. 2014. "Neofici żydowscy Włodzi w XIX wieku." *Acta Universitatis Lodziensis, Folia Historica* 93: 29–42. Accessed June 15, 2025. <https://doi.org/10.18778/0208-6050.93.04>.
- Geneteka. 2025. *Genealogiczna kartoteka – baza urodzeń, małżeństw i zgonów*. Polskie Towarzystwo Genealogiczne (Polish Genealogical Society). Accessed June 15. <https://geneteka.genealodzy.pl/index.php?op=gt&lang=pol&w=06mp>.
- GeoPy Contributors. 2018. "Welcome to GeoPy's Documentation." Accessed June 15, 2025. <https://geopy.readthedocs.io/en/stable/>.
- Gliński, Mikołaj. 2016. "A Foreigner's Guide to Polish Surnames." *#language & literature* (blog), January 8. Culture.pl. Accessed June 15, 2025. <https://culture.pl/en/article/a-foreigners-guide-to-polish-surnames>.
- Google LLC. 2025. Google.com. Accessed June 15. <https://www.google.com/>.
- Guglielmino, C. Rosalba, and Annalisa De Silvestri. 1995. "Surname Sampling for the Study of the Genetic Structure of an Italian Province." *Human Biology* 67 (4): 613–628. Accessed June 15, 2025. <https://www.jstor.org/stable/41465411>.
- Guo, Diansheng, Alice Kasakoff, Caglar Koylu, Yuan Huang, and Jack Grieve. 2015. "Historical Population Informatics: Comparing Big Data of Family Trees and the U.S. 1880 Census for Migration Analysis." Paper presented at the First International Workshop on Population Informatics for Big Data (PopInfo'15), Sydney, Australia, August 10. Accessed June 15, 2025. <https://dmm.anu.edu.au/popinfo2015/papers/1-guo2015popinfo.pdf>.
- Hanczakowski, Michał. 2019. "Czescy emigranci i ich wpływ na polską kulturę przełomu XVI i XVII wieku na przykładzie rodziny Rybińskich i Jana Łasickiego." *Postscriptum Polonistyczne* 24 (2): 215–232. Accessed June 15, 2025. <https://journals.us.edu.pl/index.php/PPol/article/view/9431>.
- Hansen, Jan, Jochen Hung, Jaroslav Ira, Judit Klement, Sylvain Lesage, Juan Luis Simal, et al. 2023. "Introduction." In *The European Experience: A Multi-Perspective History of Modern Europe, 1500–2000*, edited by Jan Hansen, Jochen Hung, Jaroslav Ira, Judit Klement, Sylvain Lesage, Juan Luis

Simal, and Andrew Tompkins, xv–xx. Open Book Publishers. Accessed June 15, 2025. <https://doi.org/10.11647/OBP.0323.88>.

Hayhoe, Jeremy. 2016. *Strangers and Neighbours: Rural Migration in Eighteenth-Century Northern Burgundy*. University of Toronto Press.

Hogan, Dennis P., and David I. Kertzer. 1985. "Longitudinal Approaches to Migration in Social History." *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 18 (1): 20–30. Accessed June 15, 2025. <https://doi.org/10.1080/01615440.1985.10594145>.

IPN (Instytut Pamięci Narodowej). 2024. "Włodzimierz Suleja: Za chlebem." Accessed June 15, 2025. <https://ipn.gov.pl/pl/historia-z-ipn/archiwum/156883,Wlodzimierz-Suleja-Za-chlebem.html>.

Izdebski, Adam, and Piotr Guzowski. 2025. "Zmiany klimatu a upadek cywilizacji. Wyniki badań polskich naukowców publikuje PNAS." Uniwersytet Jagielloński w Krakowie. Accessed June 15. [https://www.uj.edu.pl/wiadomosci/-/journal\\_content/56\\_INSTANCE\\_d82IKZvhit4m/10172/139254869](https://www.uj.edu.pl/wiadomosci/-/journal_content/56_INSTANCE_d82IKZvhit4m/10172/139254869).

Jordà, Joan Pau, Joana Maria Pujadas-Mora, and Anna Cabré. 2016. "Surnames and Migrations: The Barcelona Area (1451–1900)." In *Names and Their Environment: Proceedings of the 25th International Congress of Onomastic Sciences*, edited by Carole Hough and Daria Izdebska, 131–143. Accessed June 15, 2025. [https://ddd.uab.cat/pub/capli/2016/174338/InternationalCongressOnomasticSciences\\_2016\\_Jorda\\_Pujadas\\_Cabre.pdf](https://ddd.uab.cat/pub/capli/2016/174338/InternationalCongressOnomasticSciences_2016_Jorda_Pujadas_Cabre.pdf).

Jura, Jerzy. 2002. "Emigracja z Galicji w drugiej połowie XIX i na początku XX wieku na przykładzie wybranych powiatów." *Zeszyty Naukowe Ostrołęckiego Towarzystwa Naukowego* 16, 227–240. Accessed June 15, 2025. [https://bazhum.muzhp.pl/media/texts/zeszyty-naukowe-ostroleckiego-towarzystwa-naukowego/2002-tom-16/zeszyty\\_naukowe\\_ostroleckiego\\_towarzystwa\\_naukowego-r2002-t16-s227-240.pdf](https://bazhum.muzhp.pl/media/texts/zeszyty-naukowe-ostroleckiego-towarzystwa-naukowego/2002-tom-16/zeszyty_naukowe_ostroleckiego_towarzystwa_naukowego-r2002-t16-s227-240.pdf).

Kowalski, Waldemar. 2016. *The Great Immigration: Scots in Cracow and Little Poland, circa 1500–1660*. Studies in Central European Histories 63. Brill.

———. 2019. "The Reformation and Krakow Society, c. 1517–1637: Social Structures and Ethnicities." In *Krakau – Nürnberg – Prag: Stadt und Reformation. Krakau, Nürnberg und Prag (1500–1618)*, edited by Michael Diefenbacher, Olga Fejtová, and Zdzisław Noga, 129–146. Pavel Mervart.

Koylu, Caglar, and Alice Kasakoff. 2022. "Measuring and Mapping Long-Term Changes in Migration Flows Using Population-Scale Family Tree Data." *Cartography and Geographic Information Science* 49 (2): 154–170. Accessed June 15, 2025. <https://doi.org/10.1080/15230406.2021.2011419>.

———. 2024. "Population-Scale Kinship Networks." *International Encyclopedia of Geography*, 1–12. Accessed June 15, 2025. <https://doi.org/10.1002/9781118786352.wbieg2193>.

Koylu, Caglar, Diansheng Guo, Yuan Huang, Alice Kasakoff, and Jack Grieve. 2020. "Connecting Family Trees to Construct a Population-Scale and Longitudinal Geo-Social Network for the U.S." *International Journal of Geographical Information Science* 35 (12): 2380–2423. Accessed June 15, 2025. <https://doi.org/10.1080/13658816.2020.1821885>.

Lepucki, Henryk. 1938. *Działalność kolonizacyjna Marii Teresy i Józefa II w Galicji 1772–1790*. Instytut Popierania Polskiej Twórczości Naukowej.

Lipińska, Izabela, and Marian Lipiński. 2021. "Nobilitacje neofitów w Polsce w latach 1764–1765." *Wielkopolskie Towarzystwo Genealogiczne Gniazdo*. Accessed June 15, 2025. <http://www.wtg->

gniazdo.org/upload/opracowania/Nobilitacje\_neofitow\_w\_Polsce\_w\_latach\_1764-1765\_artykul\_Lipinski.pdf.

Malinowski, Mikołaj, and Jan Luiten van Zanden. 2017. "Income and Its Distribution in Preindustrial Poland." *Cliometrica* 11: 375–404. Accessed June 15, 2025. <https://doi.org/10.1007/s11698-016-0154-5>.

Marszał, Tadeusz. 2022. "German Immigrants in Central Poland in the Late 18th and Early 19th Centuries." *European Spatial Research and Policy* 29 (1): 25–51. Accessed June 15, 2025. <https://doi.org/10.18778/1231-1952.29.1.02>.

Miasto Gorlice. 2025. "Oil Industry in Gorlice." Accessed June 15. <https://www.gorlice.pl/pl/420/Oil-industry-in-gorlice.html>.

Michels, Georg B. 2024. "Rebels and Turcophiles? The Hungarian Protestant Clergy's Resistance against the Habsburg Counter Reformation." *Austrian History Yearbook* 55: 36–59. Accessed June 15, 2025. <https://doi.org/10.1017/S0067237824000067>.

Naruszewicz-Duchlińska, Alina. 2012. "Strona główna." Accessed June 15, 2025. <http://www.genezanazwisk.pl/>.

Nowakowska, Natalia. 2018. "Martin Luther's Polish Revolution." *OUPblog* (blog), June 25. Accessed June 15, 2025. <https://blog.oup.com/2018/06/martin-luthers-polish-revolution/>.

Rodríguez-Díaz, Roberto, María José Blanco-Villegas, and Franz Manni. 2017. "From Surnames to Linguistic and Genetic Diversity: Five Centuries of Internal Migrations in Spain." *JASs Reports: Journal of Anthropological Sciences* 95: 249–267. Accessed June 15, 2025. <https://doi.org/10.4436/JASS.95020>.

Stanford University Libraries. 2025. "Dioceses, Medieval Europe, 1200–1500." Accessed June 15. <https://searchworks.stanford.edu/view/pt668dz7698>.

Statista. 2020. "Total Population of the United Kingdom (UK) from 2015 to 2025." Accessed June 15, 2025. <https://www.statista.com/statistics/263754/total-population-of-the-united-kingdom/>.

———. 2021. "Number of Polish Nationals Resident in the United Kingdom from 2008 to 2021." Accessed June 15, 2025. <https://www.statista.com/statistics/1061639/polish-population-in-united-kingdom/>.

TheFuzz. 2025. "thefuzz 0.22.1." Accessed June 15. <https://pypi.org/project/thefuzz/>.

Wasiak, Kornelia. 2022. "Fight for Religious Tolerance during the First Polish Interregnum (1572–1573)." *Codrul Cosminului* 28 (2): 253–268. Accessed June 15, 2025. <https://doi.org/10.4316/CC.2022.02.01>.

Wikipedia. 2025. Wikipedia.com. Accessed June 15. <https://www.wikipedia.org/>.

Wójcik-Żołądek, Monika. 2014. "Współczesne procesy migracyjne: definicje, tendencje, teorie." *Studia BAS* 4 (40): 9–35. Accessed June 15, 2025. [https://orka.sejm.gov.pl/wydbas.nsf/0/59A988BD7B830629C1257DDA0044E261/%2524File/Strony%2520odStudia\\_BAS\\_40-2.pdf](https://orka.sejm.gov.pl/wydbas.nsf/0/59A988BD7B830629C1257DDA0044E261/%2524File/Strony%2520odStudia_BAS_40-2.pdf).

