Folia Oeconomica
Acta Universitatis Lodziensis
ISSN 0208-6018 e-ISSN 2353-7663

https://www.czasopisma.uni.lodz.pl/foe/
1(370) 2025
DOI: https://doi.org/10.18778/0208-6018.370.04

Magdalena Kawecka https://orcid.org/0000-0001-7671-8340
University of Economics in Katowice, Department of Demography and Economic Statistics
magdalena.kawecka@edu.uekat.pl

Qualifications and Skills Versus Future Competencies of Young Adult Poles

Abstract: In this article, we analyse the skills and competencies of young Polish adults in the context of a rapidly changing labour market. The research focuses on understanding how young adults, faced with technological progress, globalisation, and demographic changes, can adapt their skills to meet future job requirements. The study is based on data from the 2021 Human Capital Balance project. Using Principal Component Analysis (PCA) and the k-means clustering method, the aim is to identify key skills and competencies that will impact the future job prospects of young people. The importance of both technical professional skills and personal traits, such as communication, teamwork ability, creativity, and problem-solving approach, is highlighted in the work. In the context of the labour market, these skills are crucial for employees to perform tasks effectively and achieve professional success.

Keywords: qualifications, skills, young people, k-means, PCA

JEL: C38, E24, J13, J24

1. Introduction

The labour market is a dynamic environment, evolving in response to changing economic, social, and technological factors. These shifts affect employees, employers, and educational institutions that must continuously adjust to new demands. Key areas impacted by the changing environment include technology, technological advancement, encompassing automation, artificial intelligence, and robotics. These innovations introduce greater flexibility (e.g., remote work), influence business models, and elevate employee expectations (Worek, Turek, 2015:80–97; Worek, 2019a; Worek, 2019b). In the context of globalisation, the labour market becomes more competitive. Globalisation also drives cross-border recruitment (Stachowicz-Stanusch, Aleksander, 2018:485–497). Additionally, ageing populations and shifts in age structure impact the labour market. Demographic aspects, such as gender and age, are also crucial (Ratajczyk, 2007). In light of these changes, investing in skills and education becomes imperative. Simultaneously, employers and educational institutions must tailor their training programmes to provide relevant skills (Fong, Halfond, Schroeder, 2017; Osiecka-Chojnacka, 2007:2–3; Plewka, 2021). Equally important are shifts in organisational culture and management practices which alter workplace norms and expectations (Worek, Turek, 2015:80–97; Kwiatkowski, 2018:14–30).

The factors mentioned above are related to underlying competencies which constitute the subject of interest in various scientific fields. In labour market terms, ‘competencies’ refer to a blend of knowledge, abilities, and personal traits that help employees work effectively (Winterton, 2009:686–701). Competencies encompass not only technical professional skills but also personal characteristics such as communication, teamwork, creativity, and problem-solving (Stachowicz-Stanusch, Aleksander, 2018:485–497; Boyatzis, 2005:27). These can be grouped into ‘hard’ skills (technical, professional) and ‘soft’ skills (e.g., communication, teamwork, and problem-solving) (Bakhshi et al., 2017; Chałas, 2018:30–68). In the context of the labour market, they form a key element of professional success (Andruszkiewicz, Kulik-Grzybek, 2017:27–33).

The topic related to the skills and competencies of adults, including young individuals, is especially relevant in many research initiatives – such as the 2021 Human Capital Balance project – where youth skill gaps are closely examined. Moreover, numerous global and European initiatives – such as the Future of Jobs Report 2025 published by the World Economic Forum, the OECD Future of Work portal, and the European Skills Agenda – underscore the urgent need to align skills with the dynamic requirements of the labour market (ILO, 2021b; UNDP, MBRF, 2020; United Nations Development Programme and Mohammed bin Rashid Al Maktoum Knowledge Foundation, 2024). Additionally, national forecasts such as Syrowka (2025) emphasise the emergence of new green and digital occupations, underscoring the need for forward-looking skill development. National analyses further highlight digital adaptability, resilience and data literacy as critical future competencies (CERTES, 2025:12–18; Infuture Institue, 2021).

When considering competencies and skills, the aim of this study is to identify and analyse the competencies and skills of young adults in Poland which most strongly shape their future employability. The primary aim of the study is not merely to identify the key competencies among young adults in Poland, but to serve as a foundation for addressing broader challenges in the labour market. Specifically, the study aims to contribute theoretically by expanding the understanding of how competencies are formed and evolve in the context of rapid technological and economic changes, empirically by providing a detailed analysis of the current skill gaps, and practically by offering actionable recommendations for policymakers, educators, and employers to better align training and recruitment strategies with future market needs.

This topic was chosen due to its critical importance in the current socio-economic landscape. In an era of digital transformation, globalisation, and demographic shifts, ensuring that young people are equipped with the right skills is essential for economic growth and social inclusion. The theoretical framework of this study draws on seminal works in the human capital theory (e.g., Becker, 1964) and lifelong learning which underline the necessity of continuous skill development (Hyland, 2012:209–226; Kwiatkowska-Ciotucha et al., 2022). The theoretical underpinning of this research is grounded in the human capital theory and the concept of lifelong learning. These frameworks provide a robust basis for understanding the dynamics of skills development and the role of continuous education in enhancing employability (Maniak, 2015). Numerous international studies have demonstrated that a mismatch between the skills acquired and those demanded by the labour market can lead to significant socio-economic challenges. The study relies on data, encompassing both soft and hard skills, exploring how young adults can adapt to the changing work environment in Poland, while offering insights that may also apply elsewhere.

2. Literature review

The contemporary labour market is undergoing rapid transformations driven by technological advancements, globalisation, and demographic shifts. These changes have significantly altered the demand for skills, making continuous learning and adaptability more critical than ever. While traditional economic theories provide a foundation for understanding labour market dynamics, new challenges such as skills mismatches, automation, and labour market segmentation require a broader analytical perspective.

PARP's own analyses underscore the centrality of matching competencies to labour market needs. In the 2015 report, Kocór (2015a) identifies core skill clusters – digital literacy, problem solving, and interpersonal abilities – as critical for employability in Poland's evolving economy. A deeper examination of mismatch dynamics in the same year (Kocór, 2015b:12–25) reveals that misalignment between educational provision and employer expectations generates both overqualification and skill shortages, intensifying structural unemployment. Building on this, Kocór (2019) argues that persistent mismatches can lead alternately to surplus or deficit of specific competencies, amplifying labour market inefficiencies and calling for targeted policy interventions (Stec et al., 2018).

According to Jarmołowicz and Knapińska (2011), the human capital theory explains wage differentials based on education, qualifications, and professional experience. The pioneers of this theory – Gary Becker, Theodore W. Schultz, and Jacob Mincer – distinguished between general skills, which enhance job mobility, and specific skills, which are tied to positions, offering higher wages but limiting mobility. Human capital develops through individual investments in education and training. Becker (1964:635–638) emphasised that human capital investment became a central element of economic policy in the mid-20th century. He highlighted that knowledge, skills, and competencies constitute a fundamental form of capital that directly influences economic growth and competitiveness.

The dual labour market theory explains labour market segmentation, particularly its impact on young workers. The labour market consists of two sectors: the primary sector, which offers stable employment, competitive wages, and career development opportunities, and the secondary sector, which is characterised by job instability, lower wages, and underutilised qualifications (Bednarski, Kukulak-Dolata, 2021:85–96). Jarmołowicz and Knapińska (2011) note that entry into the primary sector requires highly specific qualifications, whereas the secondary sector primarily employs young workers, immigrants, and women, whose skills are often undervalued or become obsolete due to limited career progression opportunities.

The mismatch between young people's skills and labour market demands exacerbates issues arising from labour market segmentation (Bojanowska-Sosnowska, 2021:14). The matching theory and search and matching theory indicate that mismatches often stem from information asymmetry, misjudgement of skills by job seekers, or recruitment inefficiencies by employers (Jarmołowicz, Knapińska, 2011; Piróg, 2013). Young adults, being at the start of their careers, often lack a realistic understanding of labour market expectations, which hinders optimal employment and salary decisions (Kot, 2018:172–187). Additionally, job search costs – both financial and time-related – further limit labour market efficiency.

As Janicki (2007:100–105) points out, labour market segmentation has macroeconomic implications, creating a sustained demand for low-skilled labour, particularly in developed economies. Many low-prestige, low-wage jobs remain unfilled by native workers due to unfavourable working conditions and limited career prospects. Consequently, these positions are frequently occupied by young workers, immigrants, or low-skilled employees, leading to wider social inequalities and skill depreciation (Hamm, 2004:52–53).

Michael Spence's signalling theory (Spence, 2001; 2002) explains how formal education serves as a signal to employers, rather than directly reflecting an individual's actual competencies. Graduates use degrees and certificates to indicate their potential productivity, regardless of whether their acquired skills match employer expectations. This often leads to skills mismatches, where young professionals are employed in roles incompatible with their competencies, resulting in underemployment, frustration, and barriers to effective labour market entry.

Paprocka and Terlecki (2015) differentiate between hard and soft skills. Hard skills refer to technical knowledge, formal education, and professional experience, whereas soft skills encompass interpersonal abilities, critical thinking skills, and workplace adaptability (Armstrong, 2016; Fastnacht, 2006:109–114). Competencies extend beyond formal qualifications to include attributes that improve job performance and long-term employability (Nosal, 1997).

The Future of Jobs Report 2025 predicts that 22% of current jobs will undergo transformation by 2030, with 170 million new positions created and 92 million eliminated (World Economic Forum, 2025). Technology-driven sectors, such as AI, big data, and fintech, will experience the highest job growth, while roles in administration and customer service will decline. Analytical thinking, adaptability, and digital literacy will be key workforce competencies (OECD, 2023).

The OECD Skills Outlook 2023 finds a strong link between lifelong learning and lower unemployment, emphasising cognitive skills, problem-solving, and continuous learning (OECD, 2023). Similarly, the European Skills Agenda stresses the urgent need to address digital skills shortages (European Commission, 2020).

Lifelong learning, competency development, and flexible education models are essential for labour market resilience and economic competitiveness. Investing in education and training is crucial for minimising skills mismatches, preventing structural unemployment, and ensuring long-term labour market sustainability.

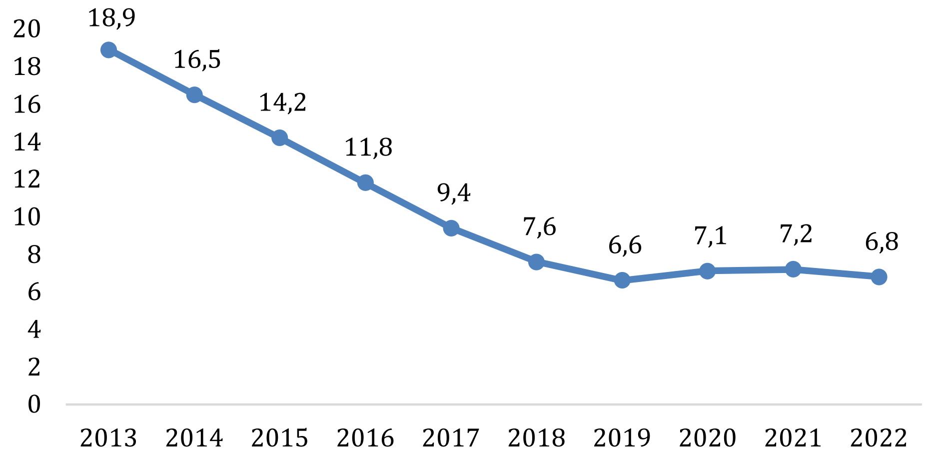

Figure 1. Unemployment rate of young people in Poland, age 15–29
Source: author's own analysis based on Eurostat data

This decline in the unemployment rate indicates an improvement in Poland's labour market conditions compared to the previous year (ILO, 2021c; 2022). However, persistent challenges remain, particularly for young individuals who face difficulties in securing stable employment and gaining relevant professional experience. Structural barriers, skill mismatches, and limited access to career development opportunities continue to hinder their labour market integration. Therefore, continuous monitoring of these indicators and a comprehensive analysis of labour market trends are essential for designing effective employment policies, enhancing youth employability, and fostering long-term economic stability.

3. Data characteristics

The data used in this study originate from the Human Capital Balance (HCB, Polish abbreviation BKL) research series, conducted by the Polish Agency for Enterprise Development (Polish abbreviation PARP) in collaboration with Jagiellonian University. These studies have been carried out since 2009 and include cross-sectional analyses that form the foundation of this research. The long-term nature of this study allows for the analysis of trends in young people's competencies and their evolution over time (PARP, 2022b: 7). Moreover, the growing demand for digital credentials in the Open Badges standard highlights the importance of integrating micro- and nano-certifications into lifelong-learning frameworks (Michalik, Felczak, 2022).

The data were collected using a structured Computer-Assisted Telephone Interview (CATI) survey (Jurek, Wiącek, Hubisz, 2012) conducted on a randomly stratified sample of 310 respondents aged 18–29. Stratification was based on gender, education level, and employment status to ensure representativeness. The data collection took place between X–Y 2021. Quality control measures included consistency checks, validation procedures, and statistical weighting to correct for sample imbalances. The methodology aligns with international standards used in labour market research, such as the OECD PIAAC framework, ensuring comparability with other large-scale workforce studies (Sitek et al., 2024; Rojek, 2023:157–160).

The HCB dataset was selected due to its reliability and comprehensive scope, covering various aspects of employment, qualifications, and socio-demographic characteristics. The data are based on a stratified sample, representative of young adults in Poland. To enhance representativeness, weighting adjustments were applied to reflect the structure of the target population (PARP, 2023:12; European Commission, 2022).

Missing values were systematically addressed: responses coded as –1, –3, –7, and –8 (missing responses, refusals) were replaced with the mean, following appropriate data cleaning procedures. Approximately 5.2% of responses contained missing values, primarily in the self-reported competency section. Although this percentage is low, it should be considered in the interpretation of results (Mazurkiewicz, 2021:27). The imputation of missing values may influence variance estimations, particularly in PCA analysis, as it reduces natural variability within the dataset and may introduce biases in factor loadings.

In the table below, all variables included in the study are listed, along with their codes.

Table 1. Variables used for the study

Variables	Code	Definition
Age* (*wiek_4k*)	1	18–29
	2	30–39
	3	40–49
	4	50–69
Gender (*m2*)**	0	Male
	1	Female
Employed BAEL** (*BAEL_praca*)	0	Unemployed or economically inactive
	1	Employed
Education level** (*wykszt_4k*)	1	Primary school or below
	2	Vocational school
	3	Secondary school
	4	Higher education (college or university)
Qualifications (*k02-k25*) Scale: 1 low, 2 basic, 3 medium, 4 high, 5 very high.	k01	Information analysis and drawing conclusions
	k02	Learning new things
	k03	Computer skills
	k04	Handling specialised computer programs
	k05	Machine operation
	k06	Assembly and repair of machinery and technical equipment
	k07	Basic accounting
	k08	Advanced mathematical calculations
	k09	Artistic skills
	k10	Physical fitness
	k11	Handling stressful situations
	k12	Willingness to take responsibility for task completion
	k13	Creativity
	k14	Time management and punctuality
	k15	Independent work organisation
	k16	Teamwork
	k17	Easy interpersonal communication
	k18	Being communicative and clear in conveying thoughts
	k19	Collaboration with people of different nationalities
	k20	Administrative work and documentation management
	k21	Coordinating the work of others
	k22	Conflict resolution among individuals
	k23	Fluent in spoken and written Polish (language proficiency)

Variables	Code	Definition
	k24	Willingness to travel frequently
	k25	Willingness to work unconventional hours as required by the employer
Knows a foreign language (*w9*)	0	No
	1	Yes
Has a driver's licence (*d11*)	0	No
	1	Yes
Would be able to learn a new profession (*s11*)	1	Definitely not
	2	Probably not
	3	Probably yes
	4	Definitely yes

* Age category used in the study: 18–29 years old. Data aggregated based on the respondent's self-reported year of birth, collected at the beginning of the questionnaire.
** The data necessary for understanding the respondents have not been included in the PCA/k-means analysis. Data aggregated based on information collected during the study in accordance with the questionnaire available on the PARP website

Source: based on the Human Capital Balance database provided by PARP (2021a; 2021b)

The variables in the dataset are measured on different scales:
– Nominal (e.g., gender, employment status, knowledge of foreign languages),
– Ordinal (e.g., education level, ability to learn new skills),
– Interval (e.g., age group),
– Ratio (e.g., number of years of education).

These scale definitions follow basic statistical concepts (StatSoft, 2023).

The competency scale (1–5) is ordinal and represents increasing levels of skills. However, these increments may not be equal in terms of actual skill development, as subjective perceptions of skill levels can vary among respondents (PARP, 2022:18). This scale has been consistently applied in previous editions of the HCB, allowing for cross-year comparisons and trend analysis (PARP, 2018:4–23). Furthermore, the comparability of these scales with international frameworks, such as the Future of Jobs Report 2025 (WEF), suggests their applicability in broader global workforce analyses.

Certain variables, such as age and gender, were excluded from PCA/k-means analysis due to their categorical nature. Instead, variables with high predictive potential regarding differences in competencies and employment were included. Preliminary correlation analysis indicated that some competencies were strongly correlated, influencing the dimensionality reduction process in PCA (Górniak, Kasparek, Jelonek, 2021:34).

The decision to focus on individuals aged 18–29 rather than broader or narrower cohorts (e.g., 15–29 or 20–24) is based on the specific characteristics of this demographic. Individuals in this age range represent a crucial transition period from education to stable employment, often experiencing their first long-term labour market interactions (PARP, 2023:19). Research also indicates that this group is more exposed to labour market instability and economic fluctuations compared to older cohorts (European Youth Portal, 2022: 8). Furthermore, international labour market research suggests that the 18–29 range better reflects youth employment trends without including adolescents still in compulsory education or those already fully integrated into the labour force (OECD, 2022:21).

Table 2. Demographic data count statistics – data distribution

Variables	Definition	Count
Age	18–29	310
Gender	Male	142
	Female	168
BAEL employed	Unemployed or economically inactive	178
	Employed	132
Education Level	Primary school or below	54
	Vocational school	26
	Secondary school	155
	Higher education (college or university)	75

Source: based on the Human Capital Balance database provided by PARP (2021a; 2021b)

Based on the available data from 2021, the age group of 18–29 years constituted 17.64% of the total sample. Women accounted for 54.19%, while men accounted for 45.81%, with 57.42% being employed and 42.58% being unemployed or economically inactive. The most common education levels were secondary (50.00%) and tertiary (24.19%). Additional statistical summaries:

– Mean age: 24.5 years (SD = 3.1 years),
– Employment rate: 57.42%,
– Proportion with a driver's licence: 68.25%.

It is essential to emphasise that the data from 2021 may reflect the short-term effects of the COVID-19 pandemic. Changes in employment, education, and remote work may have influenced the perception of competencies and young people's readiness to enter the labour market (Eurostat, 2022:11).

The COVID-19 pandemic had a particularly significant impact on young people in the EU labour market, especially in the NEET group (young people not in employment, education, or training). In 2020, youth unemployment rates increased across most EU countries, although there was a partial recovery in 2021 (ILO, 2021a; CEJSH, 2022:43–54; Subocz, 2022:29). In Poland, the increase in youth unemployment was lower than the EU average, indicating a better situation compared to Southern European countries (WUP Poznań, 2021:17). Additional data from Eurostat indicate that young workers were particularly affected in service industries, with employment rates dropping by an average of 4.5 percentage points in hospitality and retail sectors across the EU.

The HCB data provide a valuable foundation for analysing young people's competencies and employment challenges. However, further research is required on regional disparities, sectoral trends, and the long-term impact of the pandemic on career trajectories. Future analyses could also benefit from integrating additional datasets, such as OECD PIAAC or Eurostat's Labour Force Survey, allowing for a more comprehensive understanding of young adults' labour market conditions (OECD, 2022:35).

4. Research methodology

In the study, Principal Component Analysis (PCA) and k-means clustering were used to identify hidden structures within the skillset of young adults. PCA allowed for the transformation of correlated variables into orthogonal components, effectively reducing dimensionality and eliminating data redundancy before clustering. The k-means algorithm was then applied to classify respondents based on their competency profiles, enabling the identification of homogeneous groups with similar skillsets (Han, Pei, Kamber, 2017).

The utilisation of these advanced data analysis techniques, such as (specific methods used, e.g., machine learning clustering, PCA), enabled a more refined analysis of data patterns, facilitating deeper insights. While traditional methods (e.g., basic descriptive statistics or simple regression models) provide an initial understanding, these techniques allow for more complex pattern recognition and robust data segmentation.

The PCA method was utilised for dimensionality reduction, aiding in the understanding of data structure and relationships among variables. Principal Component Analysis identifies the most influential variables that contribute to the structure of the dataset, grouping correlated features into principal components where variance is maximised.

The fundamental theorem of PCA states that for any set of k variables $X_1, X_2, \ldots, X_k$, with their covariance matrix denoted as Σ, being an invertible matrix, it is always possible to transform these variables into a set of k low-correlated variables $Y_1, Y_2, \ldots, Y_k$ by performing appropriate rotations. It is important to note that the assumption of normality of the distribution is not required here. Equation 1 describes the first principal component as a linear combination of k original variables $X_1, X_2, \ldots, X_k$:

$$Y_1 = a_{11}X_1 + a_{12}X_2 + \ldots + a_{1k}X_k. \tag{1}$$

The second and subsequent principal components will be written similarly (2):

$$Y_2 = a_{21}X_1 + a_{22}X_2 + \ldots + a_{2k}X_k. \tag{2}$$

The coefficients a_{ij} appearing in the equations are constants (like regression equations). Linear combinations are obtained through axis rotations.

By utilising the k new independent variables $Y_1, Y_2, \ldots, Y_k$, the transformation ensures that the new variables remain orthogonal while capturing the full variance of the dataset. Each subsequent component captures a progressively smaller portion of variance (Pedregosa et al., 2011).

The k-means method is a technique belonging to the group of clustering algorithms which involves searching for and extracting groups of similar objects (clusters). It represents a group of non-hierarchical algorithms. Unlike hierarchical clustering, it requires the number of clusters to be defined beforehand.

The k-means algorithm groups data by attempting to divide the samples into *n* groups with equal variance while minimising the criterion known as inertia or within-cluster sum-of-squares (see equation 3 below). This iterative process reallocates data points to clusters until the intra-cluster variance is minimised. It seeks to maximise intra-cluster similarity while ensuring that clusters remain distinct from each other (Hartigan, Wong, 1979:100–105).

The k-means algorithm divides a set of samples into disjoint clusters, each of which is described by the mean value of the samples in that cluster. These means are often called 'cluster centroids,' and they serve as the representative points for each cluster in multidimensional space $NXKC_{\mu_j}X$ (Arthur, Vassilvitskii, 2007:1027–1035; Pedregosa et al., 2011):

$$\sum_{i=0}^{n} \min_{\mu \in C}(\| x_i - \mu_j \|^2). \tag{3}$$

Inertia can be seen as a measure of how internally consistent the clusters are. However, there are drawbacks to this approach, such as (Pedregosa et al., 2011):

1. 'Inertia assumes that clusters are convex and isotropic, which is not always the case. It responds poorly to elongated clusters, or manifolds with irregular shapes.
2. Inertia is not a normalised metric: lower values indicate better performance, with zero being optimal. However, in high-dimensional spaces, Euclidean distances tend to become inflated, which is an instance of the "curse of dimensionality." Applying a dimensionality reduction algorithm such as PCA before k-means clustering helps mitigate this issue and accelerates computations' (Pedregosa et al., 2011).

The study used selected Python libraries, including scikit-learn, pandas, matplotlib, scikit-image, and NumPy and the Statistica tool.

5. Research results

The first step in the analysis was to prepare the dataset and select the variables representing young individuals in the 18–29 age category. As detailed in the 'Data Characteristics' chapter, the study utilised data from the 2021 Human Capital Balance (HCB) project, conducted by the Polish Agency for Enterprise Development (PARP) in collaboration with Jagiellonian University. The dataset consists of 310 respondents, selected through random stratified sampling based on gender, education level, and employment status to ensure representativeness. The data collection took place between X and Y 2021, providing a snapshot of the skills and competencies of young adults at that time. These findings can be particularly valuable for policymakers, educators, and labour market analysts in designing tailored interventions for young workers. Identifying competency-based clusters enables the development of targeted training programmes and career guidance initiatives.

The applied methodology is rooted in economic and workforce theories, particularly the human capital theory (Becker, 1964), which emphasises the role of skills in economic productivity and employability. PCA and k-means clustering were applied to examine workforce segmentation in alignment with competency-based labour market research (Mishra, Tripathi, 2020; Dias, Silva, 2021). PCA was used to reduce dimensionality, transforming correlated variables into orthogonal components, effectively minimising redundancy before applying clustering techniques. The k-means algorithm was then utilised to classify respondents based on their competency profiles, enabling the identification of homogeneous skill-based groups among young adults.

Before performing PCA, all numerical variables were standardised to ensure comparability across different measurement scales. Since PCA is sensitive to variable magnitude, standardisation prevented certain variables from disproportionately influencing the results. Additionally, PCA requires the dataset to exhibit strong correlations to be effective. To confirm its suitability, the dataset was evaluated using statistical criteria, which confirmed that the method was appropriate for dimensionality reduction. The dataset includes a mix of measurement scales: nominal variables (e.g., gender, employment status), ordinal variables (e.g., education level, self-assessed competencies on a 5-point Likert scale), and interval variables (e.g., age groups). These distinctions were accounted for in the preprocessing stage to ensure analytical consistency.

The selection of the number of principal components was based on Kaiser's criterion (eigenvalues > 1) and the scree plot method (Jolliffe, Cadima, 2016). An 80% variance threshold was chosen, following best practices in factor analysis (Jolliffe, 2002), ensuring that key competency patterns were preserved while minimising noise and redundancy in the data. The transformed dataset was then used for clustering to enhance segmentation accuracy.

To determine the optimal number of clusters, the Elbow method and the Silhouette score were applied. Both approaches confirmed that a three-cluster solution provided the best separation of competency groups among young adults.

A comparison of clustering results before and after PCA demonstrated that PCA significantly improved cluster separation and reduced noise within the dataset. These steps allowed for a more structured and reliable identification of distinct skill clusters which are further analysed in the following sections. Similar clustering approaches have been used in workforce competency segmentation. For instance, Dias and Silva (2021) applied PCA to analyses of occupational skill structures, while Mishra and Tripathi (2020) used k-means clustering to classify job seekers based on their competency profiles. The findings in this study align with these methodologies, reinforcing the applicability of PCA and k-means in labour market research.

Such preparations allowed for grouping the data based on soft and hard skills, where:
1. Soft skills: k11, k12, k13, k14, k15, k16, k17, k18, k19, k21, k22, k24, k25, s11.
2. Hard skills: k01, k02, k03, k04, k05, k06, k07, k08, k09, k10, k23, k20, w9, d11.

Based on the pre-processed data, a correlation matrix was computed to identify relationships between variables (see the full correlation table in Appendix 1 at the end of this document). The most strongly correlated ones are:

Table 3. Pairs of variables that are most highly correlated in both data sets

Soft skills		
Variable 1	Variable 2	Correlation
k16	k17	0.696
k17	k18	0.682
k16	k18	0.620
k14	k15	0.593
k13	k12	0.569
Hard skills		
k06	k05	0.676
k03	k04	0.564
k02	k01	0.525
k02	k07	0.492
k08	k07	0.489

Source: author's own analysis based on PARP data (2021a; 2021b)

Results indicate significant relationships[1] between individual competencies, which can be useful in formulating training, recruitment, or development strategies (Table 3).

Table 4 presents the variance explained by each principal component as well as the cumulative explained variance for the components of soft and hard skills data.

Table 4. Explained variance and cumulative explained variance for PC soft and hard skills in 2021

Soft skills PC*	**Variance explained (Soft skills) [%]**	**Cumulative variance explained (Soft skills) [%]**	**Hard skills PC**	**Variance explained (Hard skills) [%]**	**Cumulative variance explained (Hard skills) [%]**
PC1	42.02	42.02	PC1	30.83	30.83
PC2	7.97	49.99	PC2	11.65	42.48
PC3	7.49	57.48	PC3	8.79	51.28
PC4	6.81	64.29	PC4	7.80	59.08
PC5	5.43	69.72	PC5	6.81	65.88

1 Results indicate statistically significant relationships ($p < 0.05$) between individual competencies, meaning these associations are unlikely due to chance. Statistical significance allows for verifying hypotheses about variable dependencies, which is crucial for shaping training, recruitment, and development strategies (Aczel, Sounderpandian, 2020).

Soft skills PC*	Variance explained (Soft skills) [%]	Cumulative variance explained (Soft skills) [%]	Hard skills PC	Variance explained (Hard skills) [%]	Cumulative variance explained (Hard skills) [%]
PC6	5.01	74.73	PC6	5.54	71.43
PC7	4.34	79.07	PC7	5.42	76.85
PC8	**4.06**	**83.13**	**PC8**	**4.59**	**81.44**
PC9	3.93	87.05	PC9	4.29	85.73
PC10	3.36	90.41	PC10	3.72	89.45
PC11	3.21	93.62	PC11	3.59	93.03
PC12	2.50	96.12	PC12	2.65	95.68
PC13	2.21	98.33	PC13	2.43	98.11
PC14	1.67	100	PC14	1.89	100

* PC – Principal Component.

Source: author's own analysis based on PARP data (2021a; 2021b)

These percentage values indicate the proportion of the total variance in the dataset that is accounted for by each principal component. In both datasets, several of the first components explain a significant portion of the variance, with the first component being the most significant.

The interpretation of principal components involves analysing the loadings (weights) of each original variable on the principal components (see Appendix 2). These loadings indicate the extent to which each variable contributes to the principal component. Accordingly, it is interpreted that:

6. Soft skills

– The first principal component (PC1) explains approximately 42.02% of the variance.

This component represents a general factor of soft skills because it has significant loadings from multiple soft skills variables. Negative signs indicate that as PC1 increases, these soft skills tend to decrease.

– The second principal component (PC2) explains approximately 7.97% of the variance.

This component, dominated by variables such as s11, k25 and k24, represents specific aspects of soft skills, such as social or interpersonal abilities.

– Third principal component (PC3) explains approximately 7.49% of the variance.

With significant loadings from variables such as s11 and k25, this can highlight a different dimension of interpersonal skills, distinct from what PC2 captures.

7. Hard skills

– The first principal component (PC1) explains approximately 30.83% of the variance.

This component has large negative loadings on all of the hard skills variables, suggesting that it may represent overall technical proficiency.

– The second principal component (PC2) explains approximately 11.65% of the variance.

With different loadings, such as positive on k05 and negative on k01 and k02, it may represent a specific subset of technical skills or knowledge.

– The third principal component (PC3) explains approximately 8.79% of the variance.

PCA is a statistical technique used to reduce the dimensionality of data by transforming original variables into new variables called principal components which are linear combinations of the original variables. These components are ordered by the amount of variance they explain, with the first component accounting for the most variance. The goal of PCA is to capture as much information as possible using fewer components, as indicated by the cumulative explained variance (see Figure 2). In interpreting loadings within PCA, the magnitude of a loading indicates the strength of the relationship between an original variable and a principal component, while the sign (positive or negative) indicates the direction of this relationship. A positive loading suggests that the variable contributes positively to the component, whereas a negative loading implies an inverse relationship. Understanding these loadings is crucial for interpreting the principal components and gaining deeper insights into the data structure (Jolliffe, 2002:164–167).

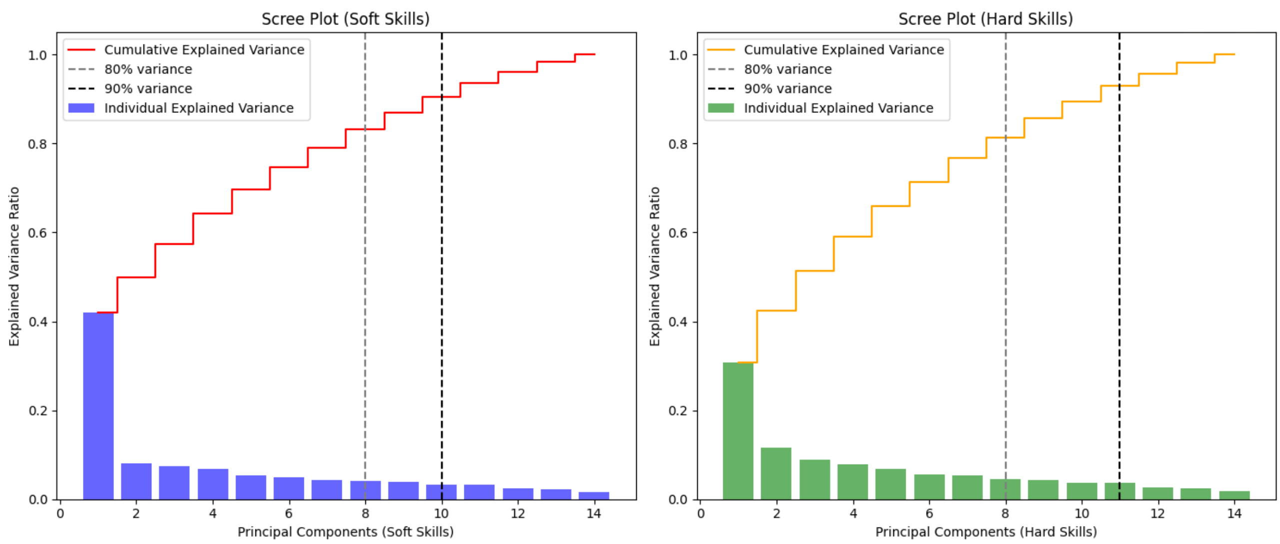

Figure 2. Scree plot PCA for soft and hard skills in 2021
Source: author's own analysis based on PARP data (2021a; 2021b)

On the above-presented Scree plots for soft skills (on the left) and hard skills (on the right), additional lines indicating how many components explain 80% and 90% of the variance have been added for soft skills and hard skills:

– The first black dashed line indicates the number of components needed to explain at least 80% of the variance.
– The second black dashed line shows the number of components needed to explain at least 90% of the variance.

The additional information on the charts provides a better understanding of how many components are needed to achieve a specific level of variance explanation, which is crucial when deciding on data dimensionality reduction.

According to the results in Figure 3 and based on the Elbow Method, 8 components were chosen for both soft and hard skills explaining over 80% of the variance.

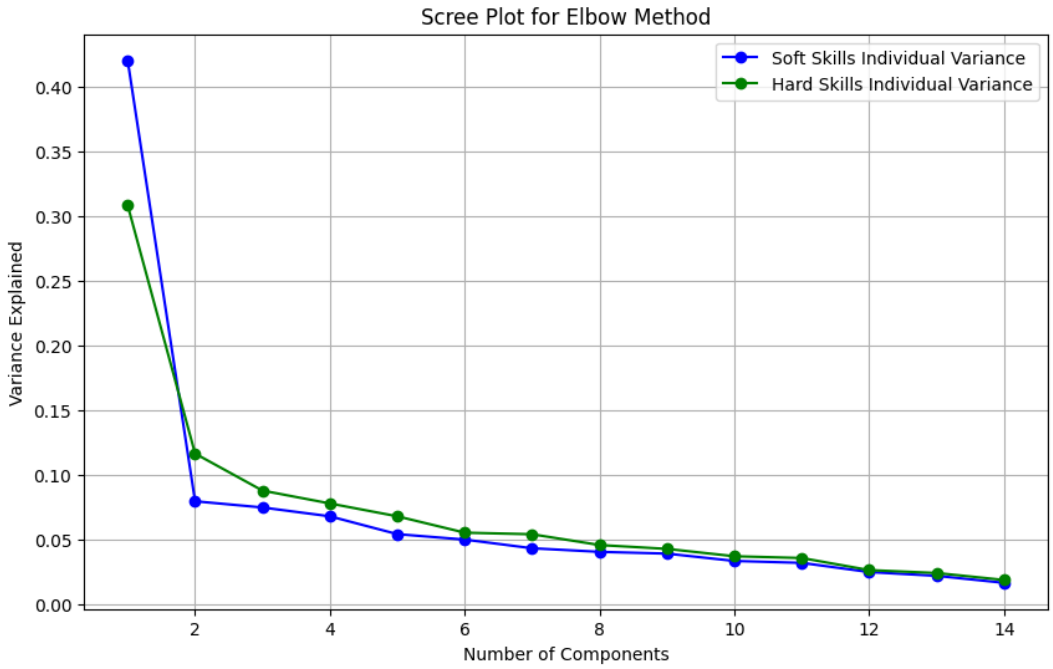

Figure 3. Scree Plot for the Elbow Method for soft and hard skills in 2021
Source: author's original analysis based on PARP data (2021a; 2021b)

Summarising the PCA analysis, the transformed data was used for further clustering analysis using the k-means algorithm. Since the data has been reduced to 8 dimensions, clustering should be more efficient and potentially more effective.

To choose the appropriate number of clusters, the Elbow Method (Figure 4) was applied. It involves analysing the plot of the sum of squared distances between points and their nearest centroids (the sum of squared errors, SSE) against the number of clusters. The point at which the plot stops decreasing sharply and starts showing a linear trend can be considered the 'elbow' and a good choice for the number of clusters. Silhouette analysis, which measures how well a point fits into its cluster compared to other clusters, was used as confirmation of the choice of the number of clusters. A high silhouette score indicates a well-fitted point to its cluster and well-separated clusters from each other (Figure 5).

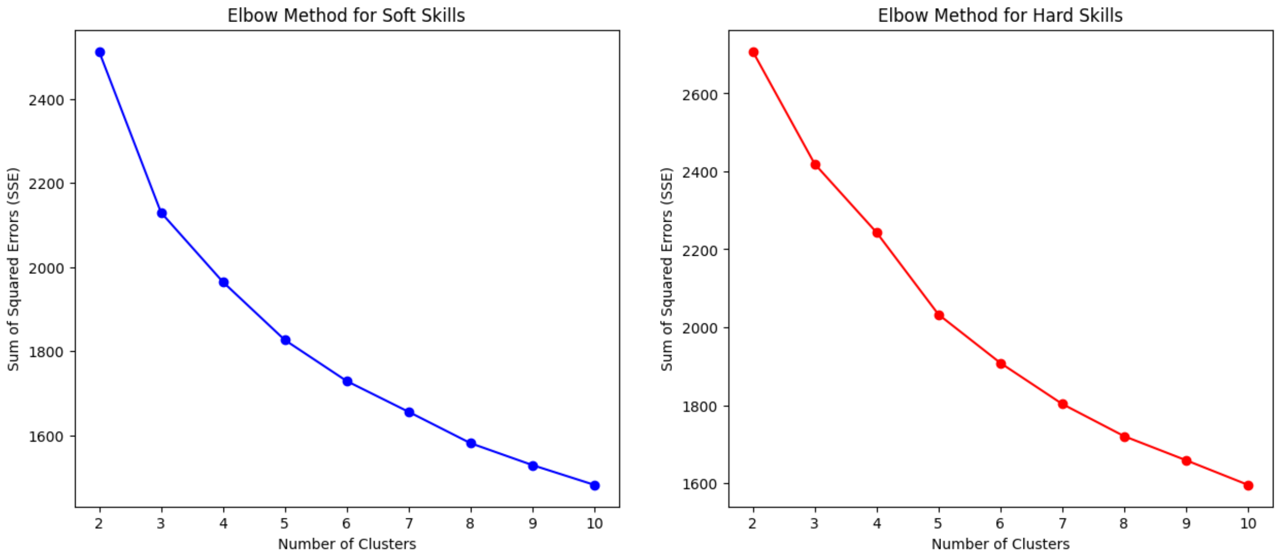

Figure 4. Elbow Method for soft and hard skills in 2021
Source: author's own analysis based on PARP data (2021a; 2021b)

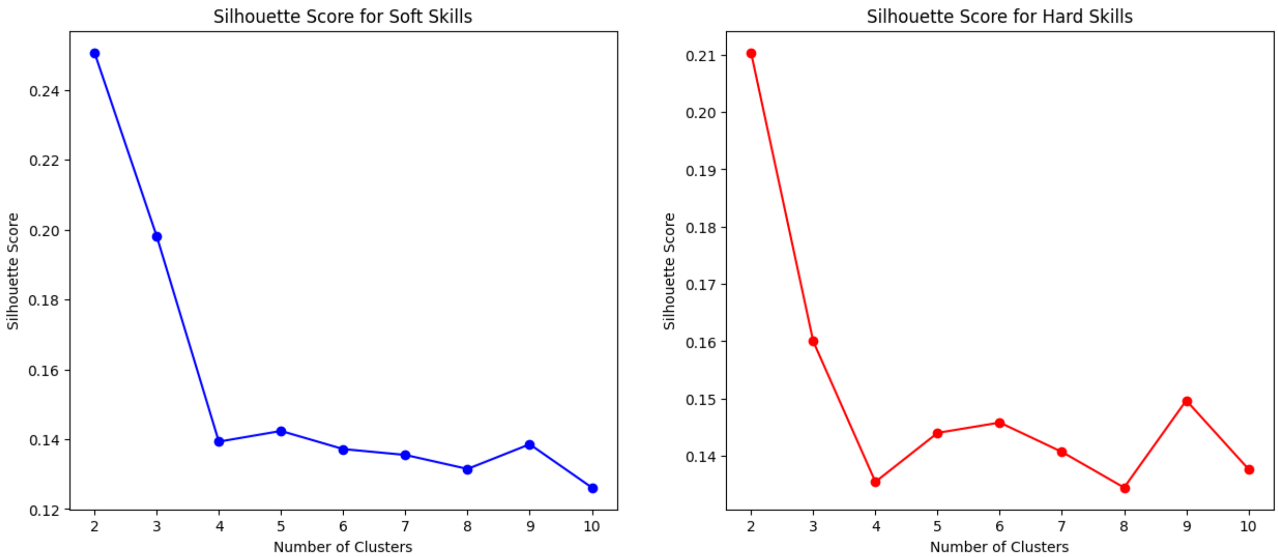

Figure 5. Silhouette for soft and hard skills in 2021
Source: author's own analysis based on PARP data (2021a; 2021b)

The Elbow Method indicates the presence of three clusters. This method is more subjective and relies on the visual interpretation of the plot. Choosing three clusters can be good for a simpler model with fewer clusters, which often facilitates interpretation and visualisation. However, the result of silhouette analysis suggests four clusters as the optimal number. Silhouette measures how well each point fits into its cluster compared to other clusters, so a higher score indicates better cluster separation and the distinctiveness of individual clusters.

The results of the average silhouette values for both clustering methods (with three and four clusters) for soft and hard competencies allowed for the selection of the recommended number of clusters (see Figure 5).

Table 5. Results of average silhouette values

	Three clusters	Four clusters
Soft skills	0.198	0.139
Hard skills	0.160	0.135

Source: author's own analysis based on PARP data (2021a; 2021b)

Based on these results, clustering with three clusters appears to be a better choice for both groups of competencies. It provides better separation and coherence of clusters. Therefore, three clusters will be used for further analysis.

Table 6. Results of average silhouette values

Soft skills cluster	0	1	2	Hard skills cluster	0	1	2
k11	– 0.156	0.654	– 1.317	k01	– 0.237	– 1.144	0.700
k12	– 0.193	0.678	– 1.256	k02	– 0.167	– 1.191	0.637
k13	– 0.123	0.578	– 1.217	k03	– 0.064	– 1.297	0.558
k14	– 0.088	0.509	– 1.145	k04	– 0.261	– 0.906	0.638
k15	– 0.136	0.631	– 1.322	k05	– 0.360	– 0.504	0.601
k16	– 0.178	0.746	– 1.504	k06	– 0.325	– 0.358	0.506
k17	– 0.184	0.706	– 1.370	k07	– 0.235	– 0.939	0.620
k18	– 0.184	0.687	– 1.316	k08	– 0.341	– 0.877	0.718
k19	– 0.126	0.574	– 1.197	k09	0.019	– 0.649	0.220
k21	– 0.179	0.659	– 1.252	k10	– 0.182	– 0.630	0.444
k22	– 0.255	0.700	– 1.100	k23	0.006	– 1.349	0.498
k24	– 0.135	0.459	– 0.837	k20	– 0.116	– 0.976	0.499
k25	– 0.228	0.505	– 0.638	w9	0.249	– 0.994	0.086
s11	– 0.054	0.175	– 0.309	d11	– 0.038	– 0.029	0.054

Source: author's own analysis based on PARP data (2021a; 2021b)

The Cluster Heatmap presented in Figure 6 allows us to observe how each feature contributes to the formation of clusters for competencies.

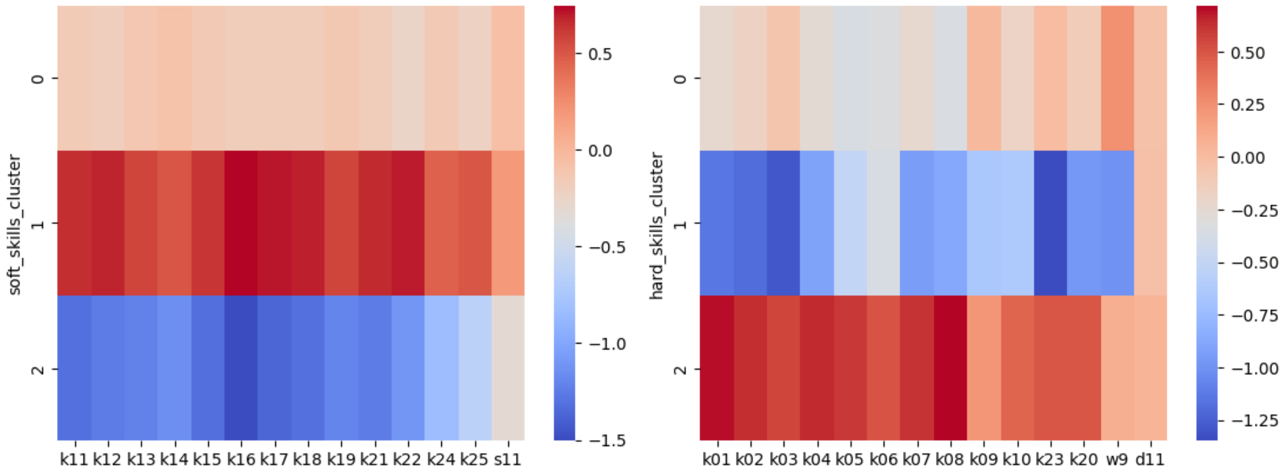

Figure 6. Heatmap of soft and hard skills clusters
Source: author's own analysis based on PARP data (2021a; 2021b)

The results of the average feature values for each cluster (see Table 6) in the soft and hard competency groups allow for the interpretation of the results, where:

8. Soft skills

- Cluster 0 (The average values of soft skill scores in Cluster 0 are slightly below zero, indicating marginally lower levels of these competencies compared to the overall average):
  - This cluster is characterised by young individuals with slightly below-average values in all soft skills.
  - It indicates individuals with moderate skills in stress management, creativity, teamwork, and communication.

Lower values in all soft skills may suggest a group of young people who may need development in interpersonal skills, creativity, time management, and adapting to unusual professional situations.

- Cluster 1 (average feature values are positive):
  - Individuals in this cluster demonstrate higher skills in all mentioned competencies.
  - High values suggest well-developed soft skills, such as coping with stress, creativity, effective communication, and teamwork.

Higher values in all competencies suggest a group of young people who excel in stressful situations, are effective in time management, creative, and adaptable in their work. They can be considered well-suited for teamwork and cross-cultural work environments.

- Cluster 2 (average feature values are significantly negative):
  - It is characterised by significantly below-average values in all soft skills.
  - Young individuals in this cluster may struggle in various areas, including stress management, creativity, teamwork, and communication.

Very low values in all features may indicate a group that can benefit from training in teamwork, communication, time management, and adapting to change.

These clusters provide insights into the strengths and weaknesses of young individuals in various soft skills areas, helping identify areas where they may need further development or training.

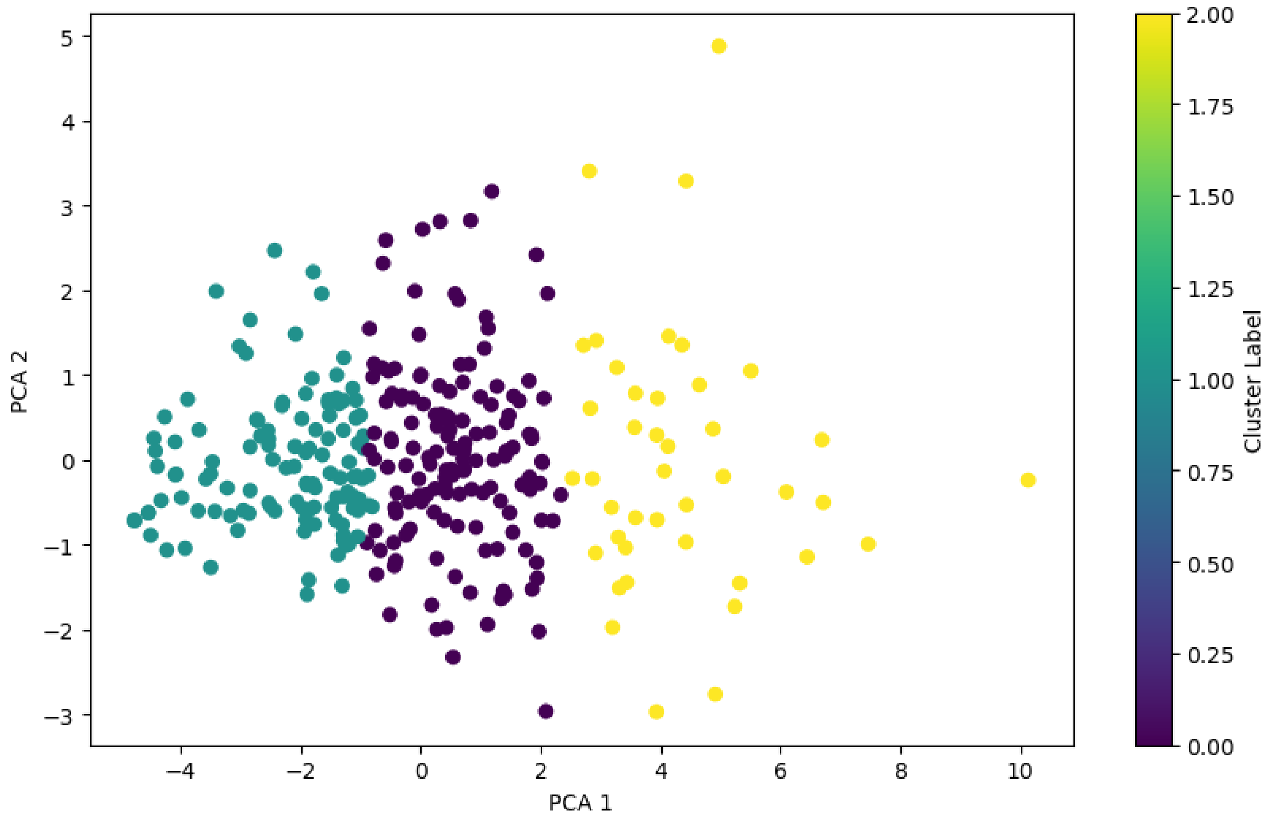

Figure 7. Visualisation of soft skills clusters
Source: author's own analysis based on PARP data (2021a; 2021b)

In summary, the distinction between clusters reveals varying levels of soft skills among young individuals. Cluster 1 represents a cohort with elevated proficiencies in soft skills, whereas Cluster 2 necessitates substantial enhancement in these domains. Cluster 0 epitomises the overall workforce population with middling soft skills. The clustering phenomena are aptly depicted in Figure 7, offering a visual encapsulation of these differentiations.

9. Hard skills

- Cluster 0 (average feature values are negative or close to zero):
  - This cluster is characterised by moderate or slightly below-average values in most of the hard skills.

- Young individuals in this cluster may possess average technical proficiency, basic computer literacy, and familiarity with specialised software to varying degrees.

Moderate values may indicate young individuals with average technical and analytical skills who may require development in specific areas.

- Cluster 1 (average feature values are significantly negative):
  - Individuals in this cluster exhibit low values across all hard skills.
  - Low values may indicate deficiencies in skills such as data analysis, computer, and technical equipment operation.

The lowest average values may suggest a group with the highest potential for development in technical, analytical, and administrative skills.

- Cluster 2 (average feature values are positive):
  - This cluster is characterised by higher values in all hard skills.
  - Individuals in this cluster appear to possess well-developed technical, analytical, and administrative skills.

The highest average values suggest a group with highly developed skills in these areas, indicating individuals with strong competencies in technical, administrative, and language-related skills.

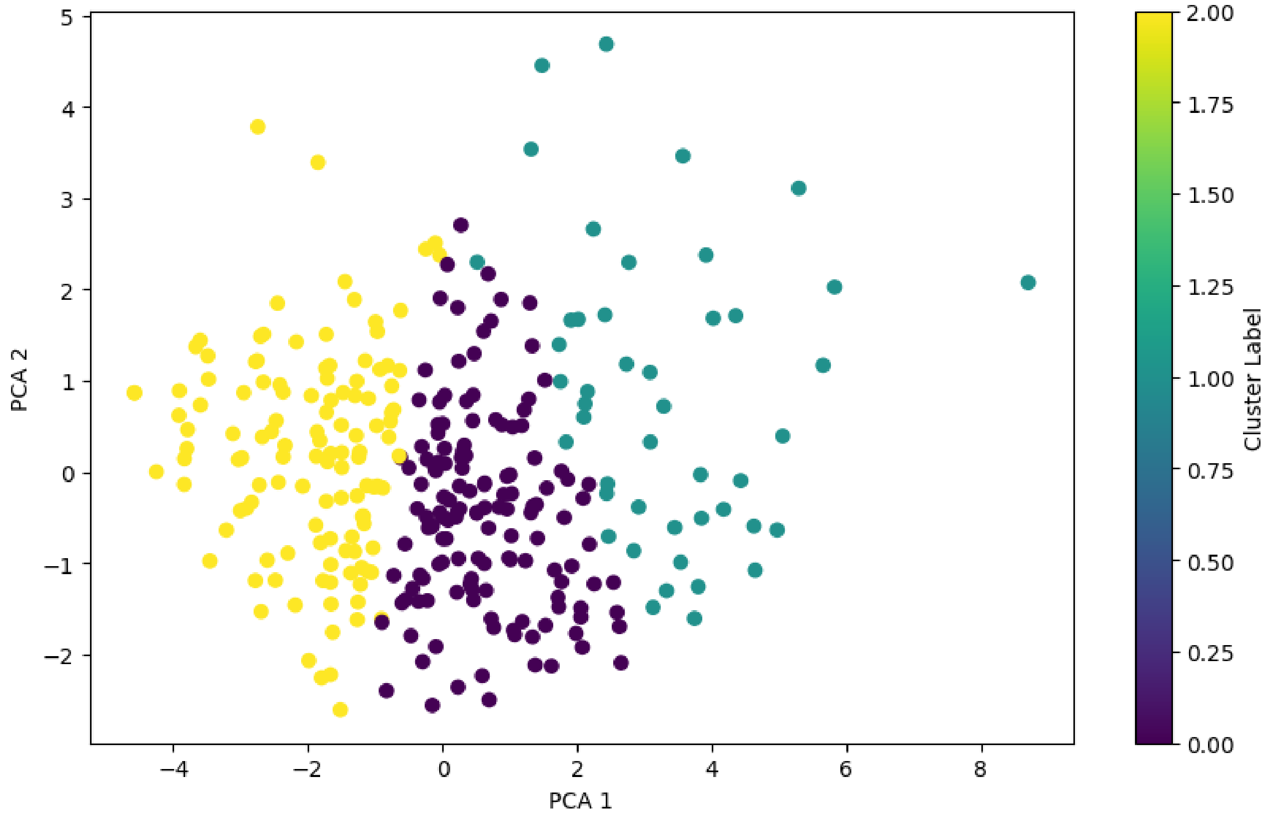

Figure 8. Visualization of hard skills clusters
Source: author's own analysis based on PARP data (2021a; 2021b)

In the case of hard skills, Cluster 1 indicates a group with low hard skills, while Cluster 2 represents individuals with high competencies in these areas. On the other hand, Cluster 0 may represent the general population of workers with average technical competencies (see Figure 8).

Although these findings are derived from the Polish labour market, similar competency patterns have been observed in EU-wide labour studies (OECD, 2022). This suggests that the identified skill clusters may be relevant in other labour markets with similar workforce structures and educational systems, providing broader applicability for competency-based workforce planning.

10. Conclusions

The main aim of this study was to examine the relationship between key skills and the employability of young people by identifying the most significant competency clusters using PCA and k-means clustering. The study focused on analysing the balance between technical vocational skills and personal characteristics such as communication, teamwork, creativity, and problem-solving abilities to determine their role in career success.

PCA was used to reduce the dimensionality of the data while retaining most of the data's variability. By transforming correlated variables into orthogonal components, PCA enabled a clearer identification of skill groupings relevant for workforce segmentation. The k-means clustering method was employed to group the data, enabling the identification and analysis of distinct competency profiles among young workers. As a result, these statistical methods facilitated the efficient analysis of large datasets and provided a structured framework for skill categorisation, supporting labour market planning.

The clustering revealed three main competency groups, each representing different levels of soft and hard skills. Cluster 1 was characterised by a high level of soft skills, Cluster 2 by an intermediate level, while Cluster 3 exhibited deficits in key interpersonal competencies. The differentiation in hard skills followed a similar pattern, with varying degrees of technical proficiency defining cluster boundaries. Such significant associations between individual competencies can be useful in designing targeted training programmes, optimising recruitment processes, and structuring skill development policies at institutional and national levels.

The study identified the strongest correlations between soft and hard skill sets, reinforcing the argument that both domains are interdependent in modern employment contexts.

The results confirm that while technical competencies remain crucial, the ability to communicate effectively, collaborate, and adapt to dynamic work environments is equally significant in securing stable employment. Furthermore, this research reinforces the need for an interdisciplinary approach in competency-building, where a combination of technical expertise and interpersonal effectiveness enhances employability and career growth.

The study also reaffirms the need for continuous learning and professional development as the labour market undergoes profound transformations due to automation, digitalisation, and globalisation. The ability to learn throughout one's life has emerged as a fundamental requirement, particularly in light of increasing job automation and shifting employment patterns. This aligns with research from the World Economic Forum (2021), which predicts that over 50% of employees will require significant upskilling by 2030 due to the impact of automation and AI.

As mentioned earlier, the findings underscore the necessity for revising educational and training curricula to align with market needs. There is a call for the implementation of competency-based learning models that integrate technical proficiency with cognitive and social abilities, ensuring that graduates are adequately prepared for workforce demands.

To address the challenges identified in this study, targeted actions ought to be implemented across different sectors of the labour market. Policymakers should focus on developing competency-based workforce strategies that emphasise continuous learning and cross-disciplinary skill development. One approach is to create subsidised training programmes aimed at young individuals with lower competency levels, particularly in soft skills such as communication, teamwork, and problem-solving. Educational institutions play a crucial role in shaping the future workforce. It is essential to redesign curricula to integrate both technical and soft skills, ensuring that graduates possess a balanced set of competencies. Expanding partnerships with industries can further enhance training programmes, aligning education with real-world labour market demands and increasing students' employability. For employers, investing in employee development through tailored upskilling programmes can help bridge competency gaps. Special emphasis should be placed on enhancing skills in communication, adaptability, and digital literacy, which are increasingly valued in modern workplaces. Providing structured training initiatives can not only improve workforce productivity but also facilitate smoother transitions for young employees entering the labour market.

The research confirms that skill-based segmentation in the labour market is a critical factor influencing employability, as demonstrated through the application of Principal Component Analysis (PCA) and k-means clustering. These methods enabled the identification of distinct competency clusters, highlighting the interplay between soft and hard skills in shaping workforce adaptability. Unlike previous studies, which often focused on either technical or interpersonal skills in isolation, this approach provides an integrated perspective, offering a more nuanced understanding of competency-based labour market segmentation. Employers increasingly prioritise a balance between technical acumen and interpersonal capabilities, reinforcing the necessity for holistic professional development strategies.

This study provides valuable guidance for educators and trainers regarding the most desirable skills in the labour market, which is essential for shaping effective training and educational programmes. Future studies should explore longitudinal data to track competency evolution over time and assess how emerging industries reshape skill demands. It is necessary to investigate how technological evolution, changes in the labour market structure, and global trends impact required skills. Additionally, studying the influence of factors such as globalisation and demographic changes is valuable.

While this study provides valuable insights into the competencies of young adults in the labour market, several limitations should be acknowledged. The time frame restriction is an important factor, as the analysis is based on data collected in 2021, reflecting a specific economic and post-pandemic context. Future research should examine labour market shifts in the aftermath of the pandemic to determine how skill demands continue to evolve in response to technological advancements and changing employment structures. Tracking competency evolution over time through longitudinal studies could provide deeper insights into how workers adapt to emerging industry needs. Although the study offers a comprehensive perspective on young adults' competencies, certain aspects of skill evolution and industry-specific requirements may require further exploration in future research.

Another limitation concerns the regional scope of the study, which focuses exclusively on Poland. Although some findings align with broader EU labour trends, conducting cross-country comparisons would offer a more comprehensive perspective on how competency dynamics vary across different economic and cultural environments. Expanding the research beyond national borders could enhance the generalizability of the results and inform international workforce policies. Studies by Eurostat and the European Centre for the Development of Vocational Training (CEDEFOP) indicate that national labour market conditions significantly influence skill mismatches, highlighting the importance of international comparisons.

Additionally, industry-specific analysis is needed to better understand how skill requirements differ across various economic sectors. Future studies could provide a sectoral breakdown, identifying distinct competency needs in industries such as high-tech, services, and manufacturing. As previous research suggests, the digital transformation is accelerating in industries like finance and healthcare, while manufacturing continues to demand specialized technical skills (OECD, 2022). Such an approach would offer more tailored recommendations for training programmes and workforce development strategies, ensuring that skill-building efforts align with sector-specific demands.

References

Aczel A.D., Sounderpandian J. (2020), *Statystyka w zarządzaniu*, Wydawnictwo Naukowe PWN, Warszawa.

Andruszkiewicz F., Kulik-Grzybek D. (2017), *Kompetencje warunkujące funkcjonowanie na współczesnym rynku pracy*, "Szkoła – Zawód – Praca", no. 13, pp. 26–34, https://cejsh.icm.edu.pl/cejsh/element/bwmeta1.element.desklight-8ccc8f19-4c15-4719-8817-9efb1ef071c8 [accessed: 15.03.2025].

Armstrong M., Taylor S. (2016), *Zarządzanie zasobami ludzkimi*, Wolters Kluwer, Warszawa.

Arthur D., Vassilvitskii S. (2007), *k-means++: The Advantages of Careful Seeding*, "Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms", pp. 1027–1035.

Bakhshi H., Downing J., Osborne M., Schneider P. (2017), *The Future of Skills: Employment in 2030*, Pearson and Nesta, London.

Becker G.S. (1964), *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*, University of Chicago – Department of Economics, University of Chicago – Booth School of Business, Chicago.

Bednarski M., Kukulak-Dolata I. (2021), *Segmentacja rynku pracy w Polsce*, "Rynek Pracy", vol. 176–177, no. 1–2, pp. 48–62.

Bojanowska-Sosnowska A. (2021), *Wykluczenie kompetencyjne i jego wpływ na rynek pracy w Polsce*, Instytut Pracy i Spraw Socjalnych, Warszawa.

Boyatzis R. (2005), *The competent manager: A model for effective performance*, John Wiley & Sons, New York.

CEJSH (2022), *Wpływ pandemii COVID-19 na sytuację osób młodych na europejskim rynku pracy – wybrane aspekty*, "Edukacja Ustawiczna Dorosłych", no. 2, pp. 43–54.

CERTES (2025), *Kompetencje Przyszłości 2025*, CERTES Consulting, Warszawa.

Chałas K. (2018), *Przyszłość należy do świadomych i aktywnych obywateli. Jakie kompetencje są im niezbędne? Jak skutecznie je kształtować? (Perspektywa aksjologiczna)*, [in:] S.M. Kwiatkowski (ed.), *Kompetencje przyszłości*, "Seria Naukowa", vol. 3, Fundacja Rozwoju Systemu Edukacji, Warszawa, pp. 30–68.

Dias J., Silva C. (2021), *Clustering analysis of youth employment trends: An application of k-means and PCA techniques*, "Journal of Labour Market Research", vol. 55(3), pp. 215–231.

European Commission (2020), *European Skills Agenda*, https://employment-social-affairs.ec.europa.eu/policies-and-activities/skills-and-qualifications/european-skills-agenda_en [accessed: 15.03.2023].

European Youth Portal (2022), *Wpływ COVID-19 na zatrudnienie młodzieży*, https://youth.europa.eu/news/effects-of-covid-19-youth-employment_pl [accessed: 15.03.2025].

Eurostat (2022), *COVID–19 strongly impacted young people's employment*, https://ec.europa.eu/eurostat/web/products-eurostat-news/-/ddn-20220331-1 [accessed: 15.03.2023].

Fastnacht D. (2006), *Miękkie kompetencje w zarządzaniu*, "Zeszyty Naukowe Wyższej Szkoły Zarządzania Ochroną Pracy w Katowicach", no. 1(2), pp. 109–114.

Fong J., Halfond J., Schroeder R. (2017), *The Changing Landscape for Professional and Continuing Education in the U.S.*, https://tiny.pl/c8n4d [accessed: 15.09.2023].

Górniak J., Kasparek K., Jelonek M. (2021), *Bilans Kapitału Ludzkiego 2021/2022*, PARP, Warszawa.

Hamm B. (2004), *Kapitał społeczny z punktu widzenia socjologicznego*, [in]: L. Frąckiewicz, A. Rączaszka (eds.), *Kapitał społeczny*, Wydawnictwo Akademii Ekonomicznej, Katowice, pp. 49–59.

Han J., Pei J., Kamber M. (2017), *Data Mining: Concepts and Techniques*, Morgan Kaufmann, Amsterdam–Boston–Heidelberg–London–New York–Oxford–Paris–Sandiego–San Francisco–Singapore–Sydney–Tokyo.

Hartigan J.A., Wong M.A. (1979), *A K-means Clustering Algorithm*, "Applied Statistics", vol. 28, pp. 100–105.

Hyland T. (2012), *Lifelong Learning, Mindfulness, and the Affective Domain of Education*, [in:] D.N. Aspin, J. Chapman, K. Evans, R. Bagnall (eds.), *Second International Handbook of Education Lifelong Learning*, vol. 26, Springer, London, pp. 209–226.

ILO (2021a), *ILO Monitor: COVID-19 and the World of Work. Eighth Edition, ILO Briefing Note*, International Labour Office, Geneva, https://www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/documents/briefingnote/wcms_824092.pdf [accessed: 14.09.2023].

ILO (2021b), *World Employment and Social Outlook: Trends 2021*, International Labour Office, Geneva, https://www.ilo.org/publications/world-employment-and-social-outlook-trends-2021 [accessed: 14.09.2023].

ILO (2021c), *Statistical ILO Brief: An update on the youth labour market impact of the COVID-19 crisis*, International Labour Office, Geneva, https://www.ilo.org/wcmsp5/groups/public/@ed_emp/documents/briefingnote/wcms_795479.pdf [accessed: 14.09.2023].

ILO (2022), *World Employment and Social Outlook: Trends 2022*, International Labour Office, Geneva, https://www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/---publ/documents/publication/wcms_834081.pdf [accessed: 14.09.2023].

Infuture Institue (2021), *Przyszłość edukacji. Scenariusze 2046*, Collegium Da Vinci and Infuture Institue, https://infuture.institute/raporty/przyszlosc-edukacji/ [accessed: 15.09.2024].

Janicki W. (2007), *Przegląd teorii migracji ludności*, "Annales Universitatis Mariae Curie-Skłodowska, Sectio B, Geographia, Geologia, Mineralogia et Petrographia", vol. 62, pp. 285–304.

Jarmołowicz W., Knapińska M. (2011), *Współczesne teorie rynku pracy a mobilność i przepływy pracowników w dobie globalizacji*, "Zeszyty Naukowe Polskiego Towarzystwa Ekonomicznego w Krakowie", no. 9, pp. 123–136.

Jolliffe I.T. (2002), *Principal Component Analysis (Springer Series in Statistics)*, Springer, New York–Berlin–Heidelberg–Hong Kong–London–Milan–Paris–Tokyo.

Jolliffe I.T., Cadima J. (2016), *Principal component analysis: a review and recent developments*, "Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences", vol. 374(2065), 20150202.

Jurek P., (2012), *Metody pomiaru kompetencji zawodowych*, "Zeszyt Informacyjno-Metodyczny – Doradcy Zawodowego", no. 54.

Kocór M. (2015a), *Kompetencje na rynku pracy. Raport PARP*, PARP, Warszawa.

Kocór M. (2015b), *Niedopasowanie kompetencyjne*, [in:] J. Górniak (ed.), *Polski rynek pracy – wyzwania i kierunki działań na podstawie badań Bilans Kapitału Ludzkiego 2010–2015*, PARP, Warszawa, pp. 12–25.

Kocór M. (2019), *Nadwyżka czy deficyt kompetencji? Konsekwencje niedopasowania na rynku pracy*, Wydawnictwo Uniwersytetu Jagiellońskiego, Kraków.

Kot P. (2018), *Związek samooceny ze skutecznością tranzycji z systemu edukacji na rynek pracy*, "Polskie Forum Psychologiczne", no. 23(1), pp. 172–187.

Kwiatkowska-Ciotucha D., Załuska U., Zięba M., Pietrus-Rajman M., Prokopowicz M. (2022), *Kształcenie ustawiczne a potrzeby rynku pracy i gospodarki*, Dobre Kadry Centrum Badawczo-Szkoleniowe, Wrocław.

Kwiatkowski S.M. (2018), *Kompetencje przyszłości*, [in:] S.M. Kwiatkowski (ed.), *Kompetencje przyszłości*, "Seria Naukowa", vol. 3, Fundacja Rozwoju Systemu Edukacji, Warszawa, pp. 14–30.

Maniak G. (2015), *Kształcenie przez całe życie – idea i realizacja. Polska na tle Unii Europejskiej*, "Studia Ekonomiczne", no. 214, pp. 128–139.

Mazurkiewicz A. (2021), *Analiza rynku pracy młodych Polaków*, Uniwersytet Ekonomiczny w Poznaniu, Poznań.

Michalik S., Felczak J. (2022), *Zapotrzebowanie na cyfrowe poświadczenia w standardzie Open Badges*, Instytut Badań Edukacyjnych, Warszawa.

Mishra S., Tripathi R. (2020), *Predicting employment barriers for youth using principal component analysis: Evidence from developing economies*, "International Journal of Social Economics", vol. 47(8), pp. 1125–1142.

Nosal C.S. (1997), *Psychologia decyzji kadrowych. Strategie, kryteria, procedury*, Wydawnictwo Profesjonalnej Szkoły Biznesu, Kraków.

OECD (2022), *Employment Outlook 2022. Building Back More Inclusive Labour Markets*, https://www.oecd.org/en/publications/oecd-employment-outlook-2022_1bb305a6-en.html [accessed: 15.03.2023].

OECD (2023), *OECD Skills Outlook 2023. Skills for a Resilient Green and Digital Transition*, https://www.oecd.org/en/publications/2023/11/oecd-skills-outlook-2023_df859811.html [accessed: 15.03.2023].

Osiecka-Chojnacka J. (2007), *Szkolnictwo zawodowe wobec problemów rynku pracy*, "Infos. Zagadnienia Społeczno-Gospodarcze", no. 16, pp. 2–3.

Paprocka I., Terlecki M. (2015), *Kompetencje twarde czy miękkie? Analiza ofert pracy pod kątem pożądanych przez pracodawców kompetencji zawodowych. Raporty z badań*, Uniwersytet Kazimierza Wielkiego w Bydgoszczy, Bydgoszcz.

PARP (2018), *Metodologia badania Bilans Kapitału Ludzkiego 2016–2023*, https://www.parp.gov.pl/component/publications/publication/metodologia-badania-bilans-kapitalu-ludzkiego-2016-2023 [accessed: 15.09.2023].

PARP (2021a), *Badanie osób dorosłych: Baza danych z badania ludności BKL edycja 2021*, https://www.parp.gov.pl/component/site/site/bilans-kapitalu-ludzkiego#metodologiabadaniabkl [accessed: 15.09.2023].

PARP (2021b), *Kwestionariusz z badania ludności BKL edycja 2021: Badanie osób dorosłych*, https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Fwww.parp.gov.pl%2Fimages%2Fpublications%2FBKL%2FKwestionariusz_z_badania_ludnoci_BKL_edycja_2021_1.docx&wdOrigin=BROWSELINK [accessed: 15.09.2023].

PARP (2022a), *Raport z badania pracodawców. Powrót do rzeczywistości? Drugi rok pandemii oczami polskich firm*, https://www.parp.gov.pl/publikacje/publication/raport-z-badania-pracodawcow-powrot-do-rzeczywistosci-drugi-rok-pandemii-oczami-polskich-firm [accessed: 15.09.2023].

PARP (2022b), *Raport z badania ludności w wieku 18–69 lat*, https://www.parp.gov.pl/publikacje/publication/raport-z-badania-ludnosci-w-wieku-18-69-lat-1 [accessed: 15.09.2023].

PARP (2023), *Indywidualne konta szkoleniowe i mikropoświadczenia – rozwiązania wspierające realizację idei kształcenia przez całe życie*, https://www.parp.gov.pl/publikacje/publication/indywidualne-konta-szkoleniowe-i-mikroposwiadczenia-rozwiazania-wspierajace-realizacje-idei-ksztalcenia-przez-cale-zycie [accessed: 15.03.2025].

Pedregosa F., Grisel O., Blondel M., Prettenhofer P., Weiss R., Dubourg V., Vanderplas J., Passos A., Cournapeau D. (2011), *Scikit-learn: Machine Learning in Python*, "Journal of Machine Learning Research", vol. 12, pp. 2825–2830, https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf [accessed: 14.09.2023].

Piróg D. (2013), *Wybrane teorie przechodzenia absolwentów szkół wyższych na rynek pracy w warunkach gospodarki opartej na wiedzy*, "Prace Komisji Geografii Przemysłu Polskiego Towarzystwa Geograficznego", no. 23, pp. 147–163.

Plewka C. (2021), *Kształcenie dualne jako sposób naprawiania systemu edukacji zawodowej*, "Około Pedagogii", no. 2, pp. 24–35.

Ratajczyk W. (2007), *Edukacja a współczesny rynek pracy*, "Acta Universitatis Lodziensis. Folia Oeconomica", vol. 210, pp. 65–83.

Rojek M. (2023), *Międzynarodowe programy kształcenia – nowy kontekst uczenia się i rozwoju nauczycieli*, "Kultura – Społeczeństwo – Edukacja", vol. 2(24), pp. 147–161.

Sitek M., Pokropek A., Penszko P., Chyl K., Haman J. (2024), *Główne wyniki międzynarodowego badania umiejętności dorosłych PIAAC 2023*, Instytut Badań Edukacyjnych, Warszawa.

Spence A.M. (2002), *Signaling in Retrospect and the Informational Structure of Markets*, "The American Economic Review", vol. 92(3), pp. 434–459.

Stachowicz-Stanusch A., Aleksander A., (2018), *Kompetencje przyszłości*, "Zeszyty Naukowe Politechniki Śląskiej, Seria: Organizacja i Zarządzanie", z. 121, pp. 485–497.
StatSoft (2023), *Podstawowe pojęcia statystyki*, http://www.statsoft.pl/textbook/stathome.html [accessed: 14.09.2023].
Stec K., Strzebońska A., Worek B., Antosz P. (2018), *Rozwój kompetencji – uczenie się osób dorosłych i podmioty oferujące usługi rozwojowe*, PARP, Warszawa.
Subocz E. (2022), *Wpływ pandemii COVID-19 na sytuację młodych na rynku pracy w Unii Europejskiej*, "Edukacja Ustawiczna Dorosłych", no. 2, pp. 43–54.
Syrowka K. (2025), *Future® 2025. Rynek pracy i zawody przyszłości*, Syrowka.com, Warszawa.
UNDP, MBRF (2020), *Global Knowledge Index 2020*, https://www.undp.org/publications/global-knowledge-index-2020 [accessed: 15.09.2023].
United Nations Development Programme and Mohammed bin Rashid Al Maktoum Knowledge Foundation (2024), *Global Knowledge Index 2024*, https://www.knowledge4all.com/ [accessed: 15.03.2025].
Winterton J. (2009), *Competence across Europe: Highest common factor or lowest common denominator?*, "Journal of European Industrial Training", vol. 33(8/9), pp. 686–701.
Worek B., Turek K. (2015), *Uczenie się przez całe życie – „akcelerator" rozwoju*, [in:] J. Górniak (ed.), *Polski rynek pracy – wyzwania i kierunki działań na podstawie badań Bilans Kapitału Ludzkiego 2010–2015*, PARP, Warszawa, pp. 80–97.
Worek B. (2019a), *Lifelong learning and its role in labor market*, Scholar, Warszawa.
Worek B. (2019b), *Uczące się społeczeństwo. O aktywności edukacyjnej dorosłych Polaków*, Wydawnictwo Uniwersytetu Jagiellońskiego, Kraków.
World Economic Forum (2021), *Upskilling for Shared Prosperity*, https://www.weforum.org/publications/upskilling-for-shared-prosperity/ [accessed: 15.03.2023].
World Economic Forum (2025), *The Future of Jobs Report 2025*, https://www.weforum.org/publications/the-future-of-jobs-report-2025/ [accessed: 15.03.2023].
WUP Poznań (2021), *Wpływ pandemii COVID-19 na postawy i motywacje młodych Wielkopolan*, https://wrot.umww.pl/wp-content/uploads/2022/12/Wplyw-pandemii-COVID-19-na-postawy-i-motywacje-mlodych-Wielkopolan-2021.pdf [accessed: 15.03.2024].

Appendix 1. Correlation matrix

Appendix 1.1. Soft skills correlation matrix

	k11	k12	k13	k14	k15	k16	k17	k18	k19	k21	k22	k24	k25	s11
k11	1	0.54	0.49	0.44	0.49	0.44	0.44	0.48	0.37	0.45	0.46	0.32	0.35	0.16
k12	0.54	1	0.57	0.55	0.61	0.50	0.46	0.53	0.39	0.50	0.47	0.31	0.36	0.19
k13	0.49	0.57	1	0.51	0.56	0.48	0.47	0.55	0.41	0.52	0.49	0.33	0.33	0.21
k14	0.44	0.55	0.51	1	0.64	0.46	0.41	0.50	0.34	0.46	0.44	0.25	0.26	0.15
k15	0.49	0.61	0.56	0.64	1	0.51	0.47	0.54	0.39	0.51	0.45	0.28	0.32	0.19
k16	0.44	0.50	0.48	0.46	0.51	1	0.55	0.54	0.39	0.46	0.47	0.26	0.31	0.17
k17	0.44	0.46	0.47	0.41	0.47	0.55	1	0.67	0.35	0.40	0.47	0.25	0.25	0.11
k18	0.48	0.53	0.55	0.50	0.54	0.54	0.67	1	0.42	0.47	0.50	0.28	0.28	0.15

	k11	k12	k13	k14	k15	k16	k17	k18	k19	k21	k22	k24	k25	s11
k19	0.37	0.39	0.41	0.34	0.39	0.39	0.35	0.42	1	0.42	0.40	0.33	0.29	0.23
k21	0.45	0.50	0.52	0.46	0.51	0.46	0.40	0.47	0.42	1	0.58	0.34	0.33	0.15
k22	0.46	0.47	0.49	0.44	0.45	0.47	0.47	0.50	0.40	0.58	1	0.31	0.29	0.15
k24	0.32	0.31	0.33	0.25	0.28	0.26	0.25	0.28	0.33	0.34	0.31	1	0.53	0.17
k25	0.35	0.36	0.33	0.26	0.32	0.31	0.25	0.28	0.29	0.33	0.29	0.53	1	0.19
s11	0.16	0.19	0.21	0.15	0.19	0.17	0.11	0.15	0.23	0.15	0.15	0.17	0.19	1

Source: author's own analysis based on PARP data (2021a; 2021b)

Appendix 1.2. Hard skills correlation matrix

	k01	k02	k03	k04	k05	k06	k07	k08	k09	k10	k23	k20	w9	d11
k01	1	0.60	0.53	0.49	0.33	0.20	0.52	0.46	0.27	0.33	0.54	0.50	0.32	0.25
k02	0.60	1	0.57	0.48	0.36	0.22	0.50	0.41	0.29	0.41	0.51	0.43	0.31	0.24
k03	0.53	0.57	1	0.76	0.31	0.16	0.48	0.50	0.27	0.39	0.50	0.53	0.42	0.30
k04	0.49	0.48	0.76	1	0.35	0.23	0.44	0.56	0.24	0.29	0.43	0.53	0.37	0.27
k05	0.33	0.36	0.31	0.35	1	0.69	0.32	0.32	0.13	0.30	0.23	0.21	0.12	0.28
k06	0.20	0.22	0.16	0.23	0.69	1	0.20	0.24	0.06	0.23	0.11	0.08	0.05	0.25
k07	0.52	0.50	0.48	0.44	0.32	0.20	1	0.55	0.23	0.30	0.53	0.49	0.29	0.21
k08	0.46	0.41	0.50	0.56	0.32	0.24	0.55	1	0.23	0.25	0.39	0.48	0.30	0.21
k09	0.27	0.29	0.27	0.24	0.13	0.06	0.23	0.23	1	0.22	0.30	0.28	0.22	0.07
k10	0.33	0.41	0.39	0.29	0.30	0.23	0.30	0.25	0.22	1	0.33	0.22	0.19	0.21
k23	0.54	0.51	0.50	0.43	0.23	0.11	0.53	0.39	0.30	0.33	1	0.48	0.35	0.19
k20	0.50	0.43	0.53	0.53	0.21	0.08	0.49	0.48	0.28	0.22	0.48	1	0.33	0.23
w9	0.32	0.31	0.42	0.37	0.12	0.05	0.29	0.30	0.22	0.19	0.35	0.33	1	0.21
d11	0.25	0.24	0.30	0.27	0.28	0.25	0.21	0.21	0.07	0.21	0.19	0.23	0.21	1

Source: author's own analysis based on PARP data (2021a; 2021b)

Appendix 2. Loadings (weights)

Appendix 2.1. Soft skills loadings

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14
k11	– 0.302	– 0.092	– 0.008	0.124	– 0.217	– 0.014	0.227	0.337	– 0.202	0.720	– 0.199	– 0.153	0.156	0.160
k12	– 0.293	– 0.095	0.027	0.171	– 0.120	0.337	0.166	0.168	– 0.545	– 0.534	– 0.115	0.015	0.273	– 0.145
k13	– 0.27	0.112	0.144	– 0.058	– 0.470	0.036	0.585	– 0.072	0.449	– 0.175	0.104	– 0.112	– 0.239	– 0.095
k14	– 0.262	– 0.018	0.094	0.534	0.171	– 0.460	0.043	– 0.011	0.098	– 0.042	– 0.288	0.491	– 0.145	– 0.193
k15	– 0.284	– 0.109	0.007	0.467	0.290	– 0.081	– 0.152	0.078	0.170	– 0.131	0.349	– 0.615	– 0.021	0.149
k16	– 0.329	0.141	– 0.129	– 0.103	0.128	0.300	– 0.116	0.226	– 0.102	– 0.030	0.109	0.294	– 0.608	0.441
k17	– 0.309	0.173	– 0.270	– 0.253	0.240	0.172	– 0.175	0.231	0.188	0.106	– 0.148	– 0.139	– 0.068	– 0.688
k18	– 0.309	0.244	– 0.271	– 0.137	0.206	0.001	0.108	– 0.057	0.316	– 0.059	0.101	0.268	0.638	0.321
k19	– 0.252	– 0.099	– 0.125	0.232	– 0.481	0.293	– 0.518	– 0.477	0.111	0.136	– 0.006	0.096	0.053	– 0.044
k21	– 0.284	0.184	0.169	– 0.220	– 0.096	– 0.392	– 0.075	– 0.125	– 0.398	0.117	0.627	0.110	0.021	– 0.216
k22	– 0.287	0.238	0.098	– 0.276	0.103	– 0.249	– 0.042	– 0.441	– 0.200	– 0.099	– 0.524	– 0.359	– 0.096	0.213
k24	– 0.199	– 0.469	0.290	– 0.374	– 0.202	– 0.260	– 0.361	0.383	0.227	– 0.212	– 0.109	0.042	0.097	0.115
k25	– 0.180	– 0.466	0.429	– 0.140	0.442	0.344	0.208	– 0.349	0.043	0.203	0.069	0.120	0.010	– 0.080
s11	– 0.058	– 0.558	– 0.696	– 0.134	– 0.006	– 0.244	0.223	– 0.184	– 0.112	– 0.057	0.058	– 0.015	– 0.141	– 0.007

Source: author's own analysis based on PARP data (2021a; 2021b)

Appendix 2.2. Hard skills loadings

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14
k01	– 0.361	– 0.134	0.020	0.185	0.096	0.116	0.046	0.060	– 0.193	0.160	– 0.567	0.616	0.133	– 0.082
k02	– 0.336	– 0.065	0.218	0.237	– 0.226	– 0.130	0.036	0.409	0.080	0.409	– 0.067	– 0.379	– 0.355	– 0.307
k03	– 0.320	– 0.118	– 0.129	– 0.309	0.241	0.007	– 0.388	– 0.029	0.351	0.003	0.302	0.352	– 0.465	– 0.089
k04	– 0.322	0.106	– 0.168	– 0.376	0.276	0.061	– 0.049	– 0.073	0.290	0.333	– 0.169	– 0.358	0.525	– 0.021
k05	– 0.278	0.512	0.076	– 0.132	– 0.045	– 0.104	0.114	0.141	– 0.191	0.166	0.064	0.077	– 0.178	0.697
k06	– 0.210	0.560	0.043	– 0.245	– 0.045	– 0.141	0.069	0.103	– 0.237	– 0.388	0.050	0.069	0.075	– 0.569
k07	– 0.307	– 0.090	– 0.190	0.443	– 0.032	– 0.227	0.122	0.172	0.110	– 0.047	0.571	0.178	0.441	0.039
k08	– 0.298	0.018	– 0.180	0.307	0.306	– 0.182	0.336	– 0.337	0.203	– 0.410	– 0.259	– 0.257	– 0.286	0.085
k09	– 0.123	– 0.170	0.607	– 0.270	– 0.187	0.159	0.491	– 0.062	0.386	– 0.149	0.085	0.146	0.086	0.042
k10	– 0.212	0.231	0.266	0.285	– 0.368	0.097	– 0.425	– 0.637	0.053	0.078	0.036	– 0.026	0.079	– 0.024
k23	– 0.303	– 0.254	0.108	– 0.035	– 0.101	0.264	– 0.409	0.343	– 0.130	– 0.546	– 0.109	– 0.257	0.125	0.242
k20	– 0.277	– 0.235	– 0.171	– 0.125	0.010	0.459	0.294	– 0.269	– 0.536	0.134	0.319	– 0.148	– 0.127	– 0.091
w9	– 0.128	– 0.404	– 0.009	– 0.346	– 0.26	– 0.710	– 0.037	– 0.218	– 0.261	– 0.002	– 0.078	– 0.016	0.049	0.059
d11	– 0.031	0.055	– 0.592	– 0.117	– 0.677	0.175	0.145	0.047	0.280	– 0.062	– 0.163	0.076	– 0.064	0.020

Source: author's own analysis based on PARP data (2021a; 2021b)

Kwalifikacje i zdolności w kontekście przyszłych kompetencji

Streszczenie: Artykuł skupia się na analizie kompetencji i zdolności młodych dorosłych Polaków w kontekście dynamicznie zmieniającego się rynku pracy. W obliczu postępu technologicznego, globalizacji i zmian demograficznych badanie koncentruje się na zrozumieniu, jak młodzi dorośli mogą dostosować swoje umiejętności do przyszłych wymagań zawodowych. Badanie opiera się na danych z Bilansu Kapitału Ludzkiego z 2021 roku. W badaniu zastosowano analizę głównych składowych (PCA) oraz metodę klasteryzacji k-średnich. Celem opracowania jest zidentyfikowanie kluczowych umiejętności i kompetencji, które będą miały wpływ na szanse młodych ludzi na rynku pracy w przyszłości. Artykuł podkreśla znaczenie zarówno technicznych umiejętności zawodowych, jak i cech osobistych, takich jak komunikacja, zdolność do pracy zespołowej, kreatywność oraz podejście do rozwiązywania problemów. W kontekście rynku pracy te umiejętności są kluczowe dla zdolności pracowników do efektywnego wykonywania zadań i osiągania sukcesów zawodowych.

Słowa kluczowe: kwalifikacje, umiejętności, młodzi ludzie, k-średnie, PCA

JEL: C38, E24, J13, J24

Received: 2024-04-09; revised: 2025-01-12. Accepted: 2025-05-21