


A DATA MINING APPROACH TO UNDERSTAND FEMALE ENTREPRENEURS ACROSS THE GLOBE

Abu Hanifah Ayob^A, Zulaiha Ali Othman^B, Sabrina Tiun^C, Zurina Muda^D



ARTICLE INFO	ABSTRACT
<p>Article history: Received: May, 15th 2024 Accepted: July, 15th 2024</p>	<p>Purpose: The aim of this study is to examine the pattern of women-owned businesses particularly whether country, age, education, citizenship, social class, income, and religion are associated with female entrepreneurship.</p>
<p>Keywords: Women; Self-Employment; Data Mining; Classification; World Values Survey.</p> 	<p>Theoretical Framework: The study develops a global demographic model of female entrepreneurship using data mining techniques.</p> <p>Design/Methodology/Approach: The study employs a data mining approach on the World Values Survey Wave 7 2017-2020 data that includes 19575 working women in 56 countries.</p> <p>Findings: The results show that the backgrounds of female entrepreneurs are largely heterogenous from one country to another. The most significant factor that influences entrepreneurship is education level while the least is citizenship status.</p> <p>Research, Practical & Social Implications: The study contributes to model a specific set of attributes of female entrepreneurs for each country. Governments could establish appropriate policies to encourage female entrepreneurship that fits with the local environment and is unique to each country.</p> <p>Originality/Value: The study contributes to the use of data mining techniques for producing more comprehensive findings than statistical approaches that are commonly employed in business economics research.</p> <p>Doi: https://doi.org/10.26668/businessreview/2024.v9i8.3469</p>

UMA ABORDAGEM DE MINERAÇÃO DE DADOS PARA ENTENDER AS MULHERES EMPREENDEDORAS EM TODO O MUNDO

RESUMO

Objetivo: O objetivo deste estudo é examinar o padrão das empresas de propriedade de mulheres, especialmente se o país, a idade, a educação, a cidadania, a classe social, a renda e a religião estão associados ao empreendedorismo feminino.

Estrutura Teórica: O estudo desenvolve um modelo demográfico global de empreendedorismo feminino usando técnicas de mineração de dados.

Projeto/Metodologia/Abordagem: O estudo emprega uma abordagem de mineração de dados nos dados da World Values Survey Wave 7 2017-2020, que inclui 1.575 mulheres trabalhadoras em 56 países.

Conclusões: Os resultados mostram que os antecedentes das mulheres empreendedoras são amplamente heterogêneos de um país para outro. O fator mais importante que influencia o empreendedorismo é o nível educacional, enquanto o menos importante é o status de cidadania.

^A PhD in Management Science. Faculty of Economics and Management, Universiti Kebangsaan Malaysia. Bangi, Malaysia. E-mail: abuhanifah.ayob@ukm.edu.my Orcid: <https://orcid.org/0000-0001-9759-7381>

^B PhD in Software Technology. Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia. Bangi, Malaysia. E-mail: zao@ukm.edu.my Orcid: <https://orcid.org/0000-0002-4238-5266>

^C PhD in Computer Science. Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia. Bangi, Malaysia. E-mail: sabrinatiun@ukm.edu.my Orcid: <https://orcid.org/0000-0002-1134-973X>

^D PhD in Computer Science. Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia. Bangi, Malaysia. E-mail: zurinam@ukm.edu.my Orcid: <https://orcid.org/0000-0003-4699-9828>

Implicações Sociais, Práticas e de Pesquisa: O estudo contribui para modelar um conjunto específico de atributos de mulheres empreendedoras para cada país. Os governos poderiam estabelecer políticas apropriadas para incentivar o empreendedorismo feminino que se adaptem ao ambiente local e sejam exclusivas de cada país.

Originalidade/Valor: O estudo contribui para o uso de técnicas de mineração de dados para produzir descobertas mais abrangentes do que as abordagens estatísticas que são comumente empregadas em pesquisas de economia empresarial.

Palavras-chave: Mulheres, Trabalho Autônomo, Mineração de Dados, Classificação, Pesquisa de Valores Mundiais.

UN ENFOQUE DE MINERÍA DE DATOS PARA COMPRENDER A LAS EMPRESARIAS DE TODO EL MUNDO

RESUMEN

Propósito: El objetivo de este estudio es examinar el patrón de las empresas propiedad de mujeres, en particular si el país, la edad, la educación, la ciudadanía, la clase social, los ingresos y la religión están asociados con la iniciativa empresarial femenina.

Marco Teórico: El estudio desarrolla un modelo demográfico global de la iniciativa empresarial femenina utilizando técnicas de minería de datos.

Diseño/Metodología/Enfoque: El estudio emplea un enfoque de minería de datos sobre los datos de la Encuesta Mundial de Valores Wave 7 2017-2020 que incluye 19575 mujeres trabajadoras en 56 países.

Hallazgos: Los resultados muestran que los antecedentes de las mujeres emprendedoras son en gran medida heterogéneos de un país a otro. El factor más significativo que influye en el emprendimiento es el nivel educativo, mientras que el menos es el estatus de ciudadanía.

Implicaciones Sociales, Prácticas y de Investigación: El estudio contribuye a modelizar un conjunto específico de atributos de las empresarias para cada país. Los gobiernos podrían establecer políticas adecuadas para fomentar el espíritu empresarial femenino que se ajusten al entorno local y sean propias de cada país.

Originalidad/Valor: El estudio contribuye al uso de técnicas de minería de datos para obtener resultados más completos que los enfoques estadísticos que se emplean habitualmente en la investigación sobre economía de la empresa.

Palabras clave: Mujeres, Autoempleo, Minería de Datos, Clasificación, Encuesta Mundial de Valores.

1 INTRODUCTION

Academic research on female entrepreneurs did not emerge until the 1980s and was initially concentrated in only two developed countries: the United States and the United Kingdom (Jennings & Brush, 2013). Prior to that, the success stories of women-owned businesses received limited coverage due to the historically dominant role of men in the economic sector.

Today, a paradigm shift is taking place where women are recognized as a subgroup in society that plays an important role in the socio-economic progress of a country (Hechevarria et al., 2019). Their contributions are widely acknowledged in boosting economic development locally and internationally. In particular, women involvement in business activities contributes to national prosperity through tax revenue, employment opportunities, and poverty reduction.

For that, recent research on female entrepreneurship focuses on their employment status: self-employed (entrepreneurs) or paid-employed.

Economic theory posits that the decision for self-employment is made upon a comprehensive consideration for a better return between the two career options even without complete information (Saridakis et al., 2014). According to the employment category approach by Gorgievski and Stephan (2016), those who are self-employed are synonymously defined as entrepreneurs. Although this simple definition is criticized (Henrekson & Sanandaji, 2014) because entrepreneurs rely heavily on innovation which does not always hold for all self-employed individuals, the general understanding that self-employment means entrepreneurship remains commonly accepted.

Following that, prior research examines the factors associated with the propensity of entrepreneurial ventures among women by comparing them with male entrepreneurs. For example, Saridakis et al. (2014) found that women are influenced by social demands such as marriage and family while men are driven by economic pursuits. However, such studies neglect the demographic heterogeneity across countries. In other words, women entrepreneurs are not only different than men but also different from each other based on the context of their home country.

To fill the gaps, this study attempts to develop a global demographic model of female entrepreneurship using data mining techniques. Specifically, we analyze the World Values Survey Wave 7 from 2017 to 2020 data that includes 19575 working women in 56 countries and compared self-employees and paid-employees on seven demographic characteristics: age, education, income, social class, religion, citizenship, and country.

The study contributes to the use of data mining techniques for producing more comprehensive findings than statistical approaches that are commonly employed in business economics research. For example, level of education is found to be positively related with the intention of being an entrepreneur by Drinkwater (2016) but Estrin et al. (2016) determined the opposite. This inconsistency may be due to the heterogenous demographic patterns of women entrepreneurship across economies, thus restricting the generalizability of the results. Instead, a classification method in data mining used in this study produces a unique demographic model for each country. Based on our findings, governments could establish appropriate policies to encourage female entrepreneurship that fits with the local environment and is unique to each country.

2 LITERATURE REVIEW

2.1 A COMPARISON OF FEMALE AND MALE ENTREPRENEURSHIP

Although the involvement of women in entrepreneurship has increased over the last few decades, their participation still trails behind men. For example, the latest data from the Global Entrepreneurship Monitor in 2018/2019 found that the number of women-owned new businesses in some countries such as Greece, Slovenia, Sweden, Switzerland, Turkey, and the United Kingdom is less than half of those started by men. A similar trend is observed in North America where the rate of women-owned start-ups remains below 5 percent. This figure has not changed much since 2009 when the Organization for Economic Cooperation and Development reported that the growth of female entrepreneurship is two times lower than for men.

The huge discrepancy in the number of female and male entrepreneurs motivated researchers to investigate if the entrepreneurial phenomenon is different between genders. Indeed, some of the results have improved our understanding on the subject. First, women are found to be less interested in entrepreneurship because it requires a high level of commitment to be successful. Compared to a paid job that guarantees a monthly wage, owning a business requires a lot of energy, time, and money. Therefore, many women find it difficult to divert their attention from their spouse, children, and family to business operations (Allen & Curington, 2014). Additionally, psychological studies suggest that women are strongly influenced by the opinions of their family members or friends. In the case of an unfavorable career choice, women would feel unsupported socially which in turn hinders the intention to become entrepreneurs.

On the other hand, men are aggressive in pursuing unlimited financial success as entrepreneurs (Dileo & Pereiro, 2019). They usually see the ownership of an enterprise as an opportunity to acquire maximum self-actualization as opposed to a paid job that is highly dependent on the company's achievements. Psychologically, men want to be seen as role models to families and society. Hence, they venture into business in part to impress their close circle of friends and family and become an example to the community.

Secondly, entrepreneurship is described as a competitive and high-risk profession that do not entirely fit with a woman's personality (Bird & Brush, 2002). Business strategies often create intense rivalries with various parties such as competitors, investors, and suppliers. Those

dealings contradict the general traits of women who are tolerant, sympathetic, and agreeable, thus reducing the competitiveness of their business (Dileo & Pereiro, 2019).

Finally, limited credit access also contributes to a low percentage of female participation in entrepreneurship. According to Guzman and Kacperczyk (2019), the gender gap in external financing is still apparent, women are 63 percent less likely to acquire it than men. Thus, women are restricted from starting a business if they rely solely on internal sources of capital such as their own savings or loans from family members. In short, both intrinsic and extrinsic factors remain major challenges for women to participate in entrepreneurship.

2.2 DEMOGRAPHIC CHARACTERISTICS

2.2.1 Age

Empirical works offer mixed results regarding the effect of age on entrepreneurial ventures (Vutsova et al., 2023). On the one hand, older people are found to be more likely to become entrepreneurs. They have stronger social capital such as experience and skills compared to young people (Leoni & Falk, 2010). Also, they have stronger financial resources and can reduce their reliance on external loans from financial institutions. Finally, they have wider social networks which is important for support and assistance.

On the other hand, new businesses are found more prevalent among those aged 25 to 44 (Hundt & Sternberg, 2014). Young people are healthier and more energetic allowing them to more easily face huge barriers in business; physical and mental strength are important for staying focused on businesses that commonly require working overtime and in high-stress environments. They are also more courageous, take risks, and try new challenges to acquire life experience.

2.2.2 Education

Entrepreneurship requires a lot of strategic decision-making (Farro-Ruiz et al., 2023). Thus, one's level of education is a decisive factor influencing an individual to become an entrepreneur (Dileo & Pereiro, 2019), because being highly knowledgeable helps them to leverage their innovation skills in addition to other management aspects such as accounting and

marketing. Therefore, the level of education is proposed to have a positive effect on the intention of starting a business (Parker, 2018).

Alternately, those who are highly qualified are more likely to build a career through paid employment in large companies. In other words, being highly educated qualifies a person to get a decent salary thus reducing their interest in self-employment (Estrin et al., 2016). Instead, they will opt for running a business only if the expected return exceeds the risk of becoming an entrepreneur.

2.2.3 Income

Self-employment is associated with a higher level of income because being an entrepreneur provides a greater opportunity for building wealth compared to receiving a monthly salary (Sorgner et al., 2017). Utility theory posits that entrepreneurs are individuals who choose to work for themselves based on the expectation of higher returns compared to paid employment. However, not all entrepreneurial endeavors end in success. In fact, data shows that the percentage of failed entrepreneurs far exceeds those who succeed (Henrekson & Sanandaji, 2014). Running a business is high risk where a failure would weaken the entrepreneur's financial position. Thus, there is little consensus from prior research whether entrepreneurs have high incomes (Hundt & Sternberg, 2014).

2.2.4 Social class

The social class of entrepreneurs varies following the local culture (Kalden et al., 2017). This informal institution refers to the social perception and acceptance towards entrepreneurship as a career (Fritsch & Wyrwich, 2014). For example, the Global Entrepreneurship Monitor 2021 report defines social class as the extent to which a country's social and cultural norms encourage individuals to start new businesses as a source for accumulating personal wealth. On the one hand, countries like the Netherlands, Iceland, and Indonesia strongly celebrate members of society who engage in entrepreneurial behavior. However, other nations such as Senegal, Croatia, and the Czech Republic generally do not favor entrepreneurship and consider it a poor career choice. To summarize, entrepreneurs enjoy a high social class in countries that revere their contribution to the country's economic and social well-being but belong to a low social class in areas that do not value their efforts to start a business.

2.2.5 Religion

Religion is a belief system in a formal structure that determines doctrine and observes specific rituals. Recent studies acknowledge that religion greatly influences an individual's decision to become an entrepreneur (Ayob & Mohd Nor, 2022). Therefore, a person's career choice can be a manifestation of their adherence to their religion.

Past research shows that the teachings of most of the world's major religions uphold principles that are compatible with entrepreneurship. For example, Christian Protestant ethics emphasize aspects of individualism and self-confidence, while Christian Catholic teachings encourage broad social relationships. Similarly, Islam encourages its people to be active and hardworking, Buddhism celebrates risk-taking behavior, and Hinduism supports free competition and ownership of wealth. All these values contribute positively to business ventures (Ayob & Saiyed, 2020).

However, empirical evidence in previous studies is rather mixed. For example, Zelekha et al. (2014) found that followers of Judaism are more likely to become entrepreneurs compared to other religions including Catholicism and Islam. In contrast, Audretsch et al. (2013) said that Islamic teachings are more closely related to entrepreneurial activities compared to Hinduism and Christianity. Regardless, the view that all religions allow their followers to own private businesses is still generally accepted.

2.2.6 Citizenship

Studies on the effect of nationality on entrepreneurship began in the 1970s during the migration wave into the United States. This issue remains relevant today as the world witnesses many workers migrating to foreign countries either because of war or to acquire a better quality of life (Castles, 2013).

However, it is still unclear whether local residents or immigrants are more likely to become entrepreneurs. One view believes that the probability of immigrants being self-employed is higher because they generally do not have the academic qualifications and skills to apply for paid jobs. This includes the inability to speak the local language fluently or lack of identification documents which make it difficult for them to be hired by a company. In addition, they are also commonly discriminated against in the labor market as many employers are

interested in hiring local people. In fact, even if hired, immigrants often take lower salaries. For that, the only option for them is to become entrepreneurs (Clark & Drinkwater, 2000).

An alternative opinion holds that opportunities for immigrants to become entrepreneurs are more limited because they usually experience financial constraints and lack of social support (Fairlie and Lofstrom, 2015). Many immigrants who have just arrived in a foreign country do not have a banking record, making it difficult for financial institutions to evaluate and provide capital loans. They also do not have the social network to acquire assistance and support. Additionally, there is a view that states that consumers are less interested in buying products from immigrants who are seen as competitors to local businesses. All of these challenges explain the low involvement of immigrants in entrepreneurship.

2.2.7 Country

Inconclusive findings regarding the factors above lead to a conclusion that the demographic pattern of female entrepreneurship is different in one country to another. For example, Fairlie et al. (2020) argues that the characteristics of entrepreneurs can be derived from their motives for becoming one. In general, there are two reasons why a person ventures into entrepreneurship: to grab business opportunities or a lack of paid jobs (Kwon & Sohn, 2021). In other words, the former focuses on innovation, while the latter merely replicates for profit. Due to the higher percentage of innovative entrepreneurs in developed countries compared to developing countries (Margolis, 2014), this study proposes that entrepreneurs in developed countries are more educated, earn higher incomes, and reside in the upper social class.

3 DATA MINING TECHNIQUES

Data mining is an analytical technique using artificial intelligence. Artificial intelligence uses algorithms that have intelligent features for specific problem-solving purposes. For example, using neural network techniques adapted from human brain neuron processes to make decisions in the form of formulas that have the nature of learning. There are many other intelligent techniques capable of analyzing complex data.

Data mining is also known as data analytics and uses various approaches such as classification, clustering, correlation, prediction, and diagnosis from past data, depending on the type of patterns to be acquired. These mining techniques are proven to be accurate when

deployed in various tasks such as diagnosing diseases from x-ray images, predicting stock exchange prices, and forecasting rainfall. Data mining can derive useful patterns from raw data for exhibiting high-frequency relationships with the level of confidence that exceeds 80 percent.

In this study, we use data mining to obtain patterns of female entrepreneurship across the globe. To attain that, we employ many classification techniques such as J48, Naive Bayes, Decision Tree, Multinomial, Multilayer Perceptron, Random Forest, and Logistic Regression.

4 METHODOLOGY

4.1 DATA PREPARATION

This study draws data from the World Values Survey Wave 7 (WVS) 2017-2020. The WVS is one of the most widely used datasets in cross-disciplinary research that covers a wide range of topics including politics, science, and religion. It was founded on the European Values Study by Jan Kerkhofs and Ruud de Moor from Tilburg University, Netherlands. Each wave of WVS data is collected over a period of 5 years starting with Wave 1 1981-1984.

WVS data collection employs strict sampling rules including a minimum sample size of 1200 participants over 18 years old in each country. Additionally, the WVS uses a probability procedure or a combination of probability and stratified sampling. WVS data is collected through face-to-face interviews by appointed academic social scientists. To ensure the accuracy of the data, questionnaires are translated into the local language.

4.2 PRE-PROCESSING

During this stage, we select from the total sample only working women and eight attributes for analysis. A data cleaning process is performed to improve accuracy when executing data classification. From the total of 22,609 respondents, 3,034 samples have incomplete data and were removed leaving only 19,575 for analysis in the next process. Table 1 below shows the description of the final data after pre-processing.

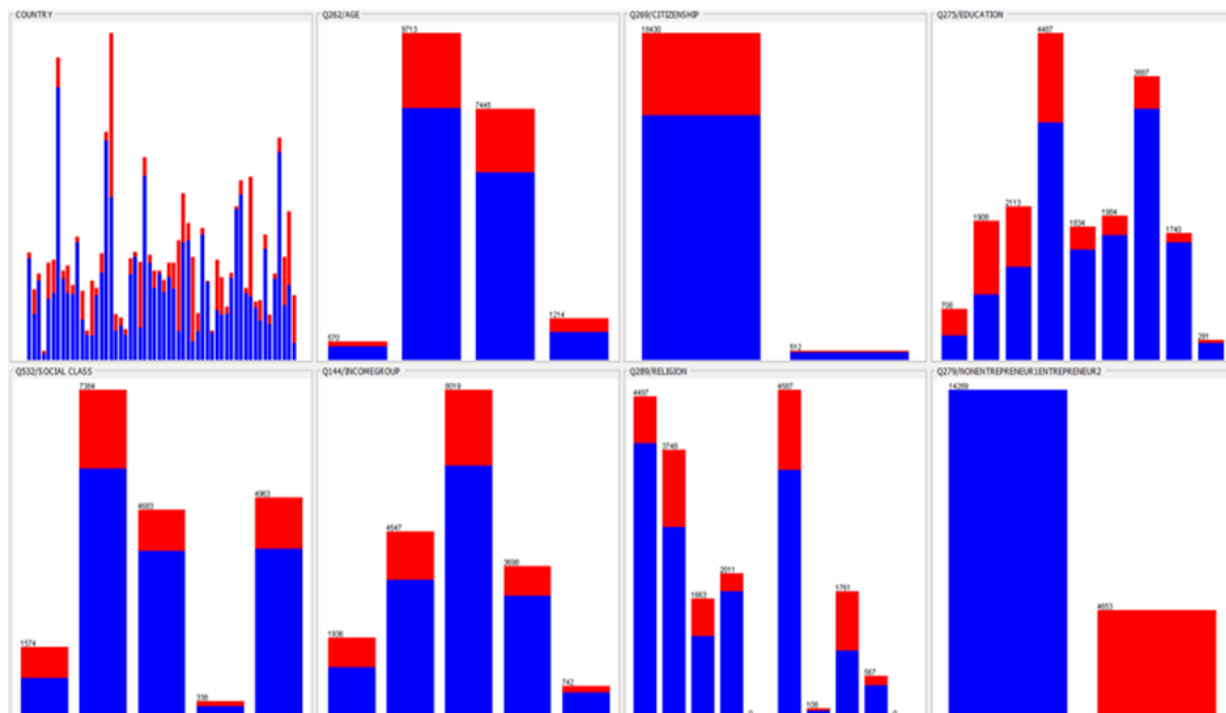
Table 1*Data description*

Attribute	Value
Country	Andorra, Armenia, Bolivia ... (56 countries)
Age	1 – less than 20, 2 – 21 to 40, 3 – 41 to 60, 4 – 61 and above
Education	0 - No education, 1 - Primary education, 2 - Lower secondary education, 3 - Upper secondary education, 4 – Post secondary education, 5 - Short higher education, 6 - Bachelor's degree, 7 - Master's degree, 8 - Doctoral degree
Income	1 – Bottom, 2 – Bottom middle, 3 – Middle, 4 - Top middle, 5 – Top
Social Class	1 – Bottom, 2 – Bottom middle, 3 – Middle, 4 - Top middle, 5 – Top
Religion	0 – No denomination 1 – Roman Catholic 2 – Protestant 3 – Orthodox 4 – Jew 5 – Muslim 6 – Hindu 7 - Buddhist 8 - Other
Citizen	1 – Yes 2 – No
Entrepreneur Status	1 – No entrepreneur 2 – Entrepreneur

Figure 1 shows the distribution of female entrepreneurs (red) and non-entrepreneurs (blue) for each attribute. In general, the figure exhibits that the ratio between the two for each attribute is rather unbalanced because entrepreneurs only represent around 25 percent of participants. However, data mining is still capable for producing model accuracy of over 80 percent.

Figure 1

Distribution of female entrepreneurs (red) and non-entrepreneurs (blue) for each attribute



4.2.1 Model development

The classification method is developed using the WEKA (Waikato Environment for Knowledge Analysis) application with 10 fold-cross validation. There are 8 classification methods used to compare and acquire the best classification method: Random Forest, J48, Naïve Bayes, Decision Tree, Logistic Regression, Multinomial, Multilayer Perceptron, and Random Forest.

Table 2 shows the experimental models conducted using all classification methods with evaluation parameters such as Accuracy, Mean Absolute Error (MAE), Root Mean Square Error (RMSE), F-Measure, Precision, Recall and ROC. From all the methods, the J48 model produces the best model based on all parameters. Figure 2 shows the pattern of differentiation between entrepreneurs and non-entrepreneurs based on the J48 algorithm.

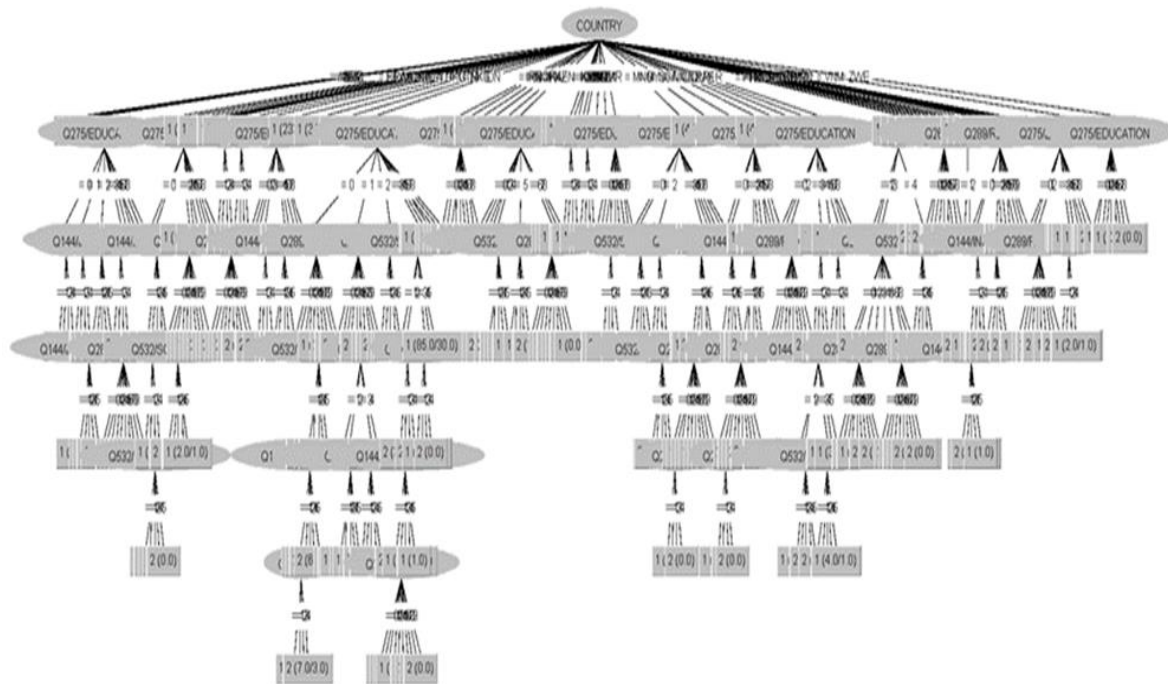
Table 2

List of tested classification models

Model	Accuracy	MAE	RMSE	F-Measure	Precision	Recall
J48	82.288	0.2518	0.3645	0.814	0.813	0.823
Decision Tree	81.8551	0.2716	0.3643	0.811	0.809	0.819
Logistic Regression	81.7601	0.2566	0.3583	0.808	0.807	0.818
Naïve Bayes	80.5723	0.2437	0.3779	0.803	0.802	0.806
Random Forest	80.0549	0.2454	0.3827	0.793	0.790	0.801
Multilayer Perceptron	79.0835	0.2195	0.4093	0.790	0.790	0.791
Random Tree	78.8776	0.2436	0.4268	0.779	0.776	0.789
Multinomial	75.4355	0.3706	0.4305	-	-	0.754

Figure 2

Pattern of differentiation between entrepreneurs and non-entrepreneurs based on the J48 algorithm



5 RESULTS AND DISCUSSION

The J48 algorithm is the best classification model with 82.288 percent accuracy in producing a list of the most important demographic characteristics in the global female

entrepreneurship model. The most dominant feature that distinguishes female entrepreneurs from non-entrepreneurs based on the largest index value is country (0.50), education level (0.39), social class (0.35), income group (0.34), age (0.33), religion (0.31), and citizenship (0.22).

Therefore, our foremost finding in this study is that the demographic pattern of global female entrepreneurs is largely heterogenous across nations. However, a demographic model can only be constructed for 20 out of the 56 countries analyzed: Bolivia, Brazil, Colombia, Ecuador, Ethiopia, Philippines, Indonesia, Iran, Kenya, Kyrgyzstan, Morocco, Myanmar, Mongolia, Nicaragua, Peru, Thailand, Tunisia, Venezuela, Vietnam, and Zimbabwe. It is observed that most of these countries have almost the same ratio of entrepreneurs and non-entrepreneurs which allows for more accurate classification. Table 3 shows specifically the demographic pattern of female entrepreneurs in those 20 countries.

Table 3

Demographic model of female entrepreneurs in each country

<p>IF Country = 'BOL' AND IF Age = 1, 2, 3 AND Education = 0 OR IF Income Group = 3 AND Age = 4 AND Education = 1 OR IF Age = 1, 3 AND Education = 1 OR IF Income Group = 1, 2 AND Education = 2 OR IF Religion = 0 AND Age = 3 AND Education = 3 OR IF Age = 2, 4 AND Education = 3 THEN Entrepreneur ELSE No Entrepreneur</p>
<p>IF Country = 'BRA' AND IF Education = 0 OR IF Social Class = 1, 3, 4, 5 AND Age = 3 AND Income Group = 2 AND Education = 1 OR IF Income Group = 3 AND Education = 1 OR IF Social Class = 2 AND Religion = 2 AND Education = 2 AND IF Religion = 0, 8 AND Education = 2 THEN Entrepreneur ELSE No Entrepreneur</p>
<p>IF Country = 'COL' AND IF Religion = 1, 2, 3, 4, 5, 6, 7, 8, 9 AND Age = 4 THEN Entrepreneur ELSE No Entrepreneur</p>
<p>IF Country = 'ECU' AND IF Age = 3, 4 THEN Entrepreneur ELSE No Entrepreneur</p>
<p>IF Country = 'ETH' AND IF Age = 2, 3, 4 AND Education = 3 OR IF Income Group = 2, 4 AND Education = 4 OR IF Education = 0, 1, 2, 8 THEN Entrepreneur ELSE No Entrepreneur</p>
<p>IF Country = 'IDN' AND IF Age = 2, 4 AND Income Group = 1 AND Social Class = 1 AND Religion = 5 AND Education = 0 OR IF Religion = 0 - 9 AND Income Group = 2 AND Age = 2 AND Social Class = 2 AND Education = 3 OR IF Income Group = 4, 5 AND Social Class = 1 AND Religion = 5 AND Education = 0 OR IF Income Group = 2, 5 AND Age = 2 AND Religion = 5 AND Education = 1 OR IF Social Class = 2, 3, 4, 5 AND Age = 3 AND Religion = 5 AND Education = 1 OR IF Income Group = 3 AND Age = 2 AND Social Class = 2 AND Education = 3 OR IF Social Class = 3, 4 AND Religion = 5 AND Education = 0 OR</p>

<p>IF Age = 4 AND Religion = 5 AND Education = 1 OR IF Age = 3 AND Social Class = 2 AND Education = 3 OR IF Age = 2, 4 AND Social Class = 3 AND Education = 3 OR IF Religion = 0, 1, 2, 3, 4, 7, 8, 9 AND Education = 0 OR IF Religion = 0, 1, 2, 3, 4, 6, 7, 8, 9 AND Education = 1 OR IF Income Group = 1, 3, 4, 5 AND Education = 2 OR IF Social Class = 1, 4 AND Education = 3 THEN Entrepreneur ELSE No Entrepreneur</p>
<p>IF Country = 'IRN' AND IF Education = 2, 4, 5 THEN Entrepreneur ELSE No Entrepreneur</p>
<p>IF Country = 'KEN' AND IF Social Class = 1, 2, 3, 4 AND Education = 4 OR IF Social Class = 1 AND Education = 5 OR IF Religion = 2 AND Education = 6 OR IF Education = 0, 1, 2, 3, 8 THEN Entrepreneur ELSE No Entrepreneur</p>
<p>IF Country = 'KGZ' AND IF Age = 1 THEN Entrepreneur ELSE No Entrepreneur</p>
<p>IF Country = 'MAR' AND IF Age = 4 THEN Entrepreneur ELSE No Entrepreneur</p>
<p>IF Country = 'MMR' AND IF Age = 2, 3, 4 AND Education = 3 OR IF Education = 0, 1, 2, 4, 5, 8 THEN Entrepreneur ELSE No Entrepreneur</p>
<p>IF Country = 'MNG' AND IF Age = 1, 2, 4 AND Religion = 0 AND Income Group = 1 AND Education = 3 OR IF Social Class = 1, 3, 4 AND Age = 3 AND Education = 2 OR IF Social Class = 1, 2, 3, 4 AND Education = 1 OR IF Age = 2, 4 AND Education = 2 OR IF Income Group = 2, 4 AND Education = 3 THEN Entrepreneur ELSE No Entrepreneur</p>
<p>IF Country = 'NIC' AND IF Age = 1, 3, 4 AND Religion = 1 AND Religion = 1 AND Social Class = 5 AND Education = 1 OR IF Religion = 2, 3, 4, 5, 6, 7, 8, 9 AND Social Class = 5 AND Education = 1 OR IF Social Class = 2, 4 AND Education = 1 OR IF Income Group = 3 AND Education = 2 OR IF Education = 0, 7 THEN Entrepreneur ELSE No Entrepreneur</p>
<p>IF Country = 'PER' AND IF Social Class = 2 AND Income Group = 2 AND Age = 2 AND Education = 3 OR IF Social Class = 1, 2, 3, 4 AND Income Group = 3 AND Age = 2 AND Education = 3 OR IF Income Group = 1, 5 AND Age = 2 AND Education = 3 OR IF Religion = 0, 1, 3, 4, 5, 6, 7, 9 AND Education = 2 OR IF Age = 3, 4 AND Education = 3 OR IF Age = 3 AND Education = 5 OR IF Education = 0, 1, 8 THEN Entrepreneur ELSE No Entrepreneur</p>
<p>IF Country = 'PHL' AND IF Religion = 8 AND Education = 1 AND Age = 3 OR IF Religion = 1, 2, 3, 4, 6, 7, 8, 9 AND Education = 6 AND Age = 3 OR IF Education = 2, 3, 4, 8 AND Age = 3 OR IF Social Class = 1, 4 AND Age = 4 THEN Entrepreneur ELSE No Entrepreneur</p>
<p>IF Country = 'THA' IF Education = 0, 1, 2, 3, 4, 8 THEN Entrepreneur ELSE No Entrepreneur</p>

IF Country = 'TUN' IF Citizenship = 2 THEN Entrepreneur ELSE No Entrepreneur
IF Country = 'VEN' IF Income Group = 1, 3, 4 AND Age = 1 AND Religion = 1 OR IF Age = 2 AND Religion = 1 OR IF Income Group = 1, 3, 4 AND Religion = 2 OR IF Religion = 8 THEN Entrepreneur ELSE No Entrepreneur
IF Country = 'VNM' AND IF Religion = 0, 3, 4, 5, 6, 7, 8, 9 AND Education = 2 OR IF Age = 1, 3 AND Education = 3 OR IF Education = 0, 1, 4 THEN Entrepreneur ELSE No Entrepreneur
IF Country = 'ZWE' AND IF Education = 0, 1, 2, 3, 7, 8 THEN Entrepreneur ELSE No Entrepreneur

5.1 COUNTRY

In general, there are two distinct patterns by country. First, 9 countries exhibit simple characteristics with only one or two dominant demographic factors. This includes level of education in Iran, Thailand, and Zimbabwe; age in Ecuador, Kyrgyzstan, and Morocco; and citizenship in Tunisia. Age and religion are important demographics in Colombia while age and education level are significant in Myanmar.

For example, the specific demographic characteristics of female entrepreneurs compared to non-entrepreneurs in Ecuador are only among those over 40 years old, while in Kyrgyzstan they are under 20 years old, and in Morocco they are over 60 years old. In addition, female entrepreneurs in Tunisia are only non-citizen women. Finally, female entrepreneurs in Colombia differ from wage earners as they are mostly over 60 years old and follow a variety of religions.

Second, another 11 countries show a more complex pattern especially in Indonesia, Peru, Bolivia, Brazil, Mongolia, and Nicaragua. This is due to a larger number of respondents and the ratio of entrepreneurs and non-entrepreneurs.

5.2 AGE

In terms of age, 8 countries, Bolivia, Ethiopia, Indonesia, Myanmar, Mongolia, Nicaragua, Peru, and Vietnam show female entrepreneurs among both the young and old. Female entrepreneurs in Kyrgyzstan and Venezuela are dominated by young people under the age of 40. Conversely, women over 40 in Brazil, Ecuador, the Philippines, and Morocco are more likely to be self-employed.

5.2.4 Education

The level of education of female entrepreneurs also varies across nations. In 6 countries: Bolivia, Brazil, Indonesia, Mongolia, Nicaragua, and Vietnam, most women have low education up until upper secondary level only. Next, the educational achievement of female entrepreneurs in Iran is slightly higher up until the bachelor's degree. Finally, the level of education among female entrepreneurs in Kenya, Myanmar, Thailand, and Zimbabwe is diverse.

5.3 INCOME

Although previous studies argue that owning a business offers entrepreneurs the opportunity to accumulate unlimited wealth, the analysis found a balanced pattern between high- and low-income in 6 countries namely Brazil, Ethiopia, Indonesia, Mongolia, Peru, Venezuela. Most female entrepreneurs in Bolivia are found to have a low-income level at only at the middle class.

5.4 SOCIAL CLASS

The social class characteristics of female entrepreneurs are also mixed between high and low social hierarchy in 7 countries: Brazil, Philippines, Indonesia, Kenya, Mongolia, Nicaragua, and Peru. The findings suggest that the social position of entrepreneurs in a society is not fixed but rather depends on the local culture.

5.5 RELIGION

Religion also has a varied effect on female entrepreneurs. Followers of the world's main religions are less likely to choose to be self-employed. Instead, non-religious women are more likely to become entrepreneurs in Bolivia, Brazil, Mongolia, and Venezuela. Members of religious minorities in Brazil and Venezuela are also more likely to become entrepreneurs. In addition, many Protestants in Kenya are also entrepreneurs. Finally, female entrepreneurs in 6 countries, namely Colombia, the Philippines, Indonesia, Nicaragua, Peru, and Vietnam, are adherents of different religions.

5.6 NATIONALITY

In general, our analysis found that citizenship or immigration status does not determine the probability of becoming an entrepreneur among women in all countries. On the other hand, the analysis confirms that immigrants are more likely to be self-employed than locals, but only in Tunisia. This finding may be driven by immigration trends from North Africa and sub-Saharan Africa into the country since the early 2000s. This large number of immigrants makes it difficult for them to find work, thus necessitating self-employment.

6 CONCLUSION

This study aims to produce a global demographic model of female entrepreneurs. From all the classification models tested, the J48 algorithm produced the most accurate model, showing that the most dominant characteristic is country, followed by education, social class, income, age, religion, and citizenship. Thus, the most important finding in this study is that the demographic pattern of global female entrepreneurs is very different from one country to another.

Despite that, demographic models are constructed for only 20 out of 56 countries analyzed. For example, women over 40 tend to be entrepreneurs in Ecuador, they are under 20 in Kyrgyzstan, and over 60 in Morocco. The demographic characteristics of social class are also diverse across countries. In addition, the level of education of female entrepreneurs in many countries is also low to moderate which leads to low incomes like in Bolivia. Next, non-religious groups or religious minorities are more likely to become entrepreneurs in most countries. Finally, nationality is largely insignificant in all countries except Tunisia.

Taken together, these findings are important to stakeholders, especially government agencies, to ensure that programs to encourage entrepreneurship are targeted effectively. This includes increasing campaigns to groups that are still less involved in business activities. For example, the education system can be improved by strengthening entrepreneurship education to boost more highly educated women to get involved.

In conclusion, this study has successfully applied data mining techniques to develop a model of global female entrepreneurs based on demographic characteristics. This contribution proves that data science has the potential to provide more detailed knowledge than existing analytical techniques. However, there are caveats in this study, mainly the number of samples across countries is not equal and the ratio of entrepreneurs to non-entrepreneurs is also

unbalanced. Therefore, further research is recommended to examine other factors such as the type of entrepreneurship and the level of well-being of entrepreneurs. In addition, analysis can also be done using other data sources such as the Global Entrepreneurship Monitor and the World Bank.

ACKNOWLEDGEMENT

This work partially contributes to the research funded by LRGS/1/2020/UKM/01/5/2 grant under the Ministry of Higher Education, Malaysia.

REFERENCES

- Allen, W. D. & Curington, W. P. (2014). The self-employment of men and women: What are their motivations? *Journal of Labor Research*, 35(2), 143-161. <https://doi.org/10.1007/s12122-014-9176-6>
- Audretsch, D. B., Boente, W. & Tamvada, J. P. (2013). Religion, social class, and entrepreneurial choice. *Journal of Business Venturing*, 28(6), 774-789. <https://doi.org/10.1016/j.jbusvent.2013.06.002>
- Ayob, A. H. & Mohd Nor, S. (2022). Individual religiosity and career choice: does cultural religiosity moderate the relationship? *Cross-Cultural Research*, 56(4), 323-344. <https://doi.org/10.1177/106939712211080622>
- Ayob, A. H. & Saiyed, A. A. (2020). Islam, institutions and entrepreneurship: evidence from Muslim populations across nations. *International Journal of Islamic and Middle Eastern Finance and Management*, 13(4), 635-653. <https://doi.org/10.1108/IMEFM-11-2019-0472>
- Bird, B. & Brush, C. (2002). A gendered perspective on organizational creation. *Entrepreneurship Theory and Practice*, 26(3), 41-65. <https://doi.org/10.1177/104225870202600303>
- Castles, S. (2013). The forces driving global migration. *Journal of Intercultural Studies*, 34(2), 122-140. <https://doi.org/10.1080/07256868.2013.781916>
- Clark, K. & Drinkwater, S. (2000). Pushed out or pulled in? Self-employment among ethnic minorities in England and Wales. *Labour Economics*, 7(5), 603-628. [https://doi.org/10.1016/S0927-5371\(00\)00015-4](https://doi.org/10.1016/S0927-5371(00)00015-4)
- Dileo, I. & García Pereiro, T. (2019). Assessing the impact of individual and context factors on the entrepreneurial process. A cross-country multilevel approach. *International Entrepreneurship and Management Journal*, 15(4), 1393-1441. <https://doi.org/10.1007/s11365-018-0528-1>

- Drinkwater, S., 2018. Migrant Entrepreneurship. In *The SAGE Handbook of Small Business and Entrepreneurship* (pp. 94-106). SAGE Publications Sage UK: London, England. <https://doi.org/10.4135/9781473984080>
- Estrin, S., Meyer, K. E., Nielsen, B. B. & Nielsen, S. (2016). Home country institutions and the internationalization of state owned enterprises: A cross-country analysis. *Journal of World Business*, 51(2), 294-307. <https://doi.org/10.1016/j.jwb.2015.11.002>
- Fairlie, R. W. & Lofstrom, M. (2015). Immigration and entrepreneurship. In *Handbook of the Economics of International Migration* (Vol. 1, pp. 877-911). North-Holland. <https://doi.org/10.1016/B978-0-444-53768-3.00017-5>
- Farro-Ruiz, L. Z., Figueroa, I. K. G., Cardenas, M. B., Ramon, E. F. C. & Chiparra, W.E.M. 2023. Education in entrepreneurship: systematic review of publications in Spanish from 2013 to 2023. *International Journal of Professional Business Review*, 8(8), e03096-e03096. <https://doi.org/10.26668/businessreview/2023.v8i8.3096>
- Fritsch, M. & Wyrwich, M. (2014). *The effect of regional entrepreneurship culture on economic development: Evidence for Germany* (No. 2014-014). Jena Economic Research Papers. <http://hdl.handle.net/10419/98454>
- Gorgievski, M. J. & Stephan, U. (2016). Advancing the psychology of entrepreneurship: A review of the psychological literature and an introduction. *Applied Psychology*, 65(3), 437-468. <https://doi.org/10.1111/apps.12073>
- Guzman, J. & Kacperczyk, A. O. (2019). Gender gap in entrepreneurship. *Research Policy*, 48(7), pp.1666-1680. <https://doi.org/10.1016/j.respol.2019.03.012>
- Hechavarria, D., Bullough, A., Brush, C. & Edelman, L. (2019). High-growth women's entrepreneurship: Fueling social and economic development. *Journal of Small Business Management*, 57(1), 5-13. <https://doi.org/10.1111/jsbm.12503>
- Henrekson, M. & Sanandaji, T. (2014). Small business activity does not measure entrepreneurship. *Proceedings of the National Academy of Sciences*, 111(5), 1760-1765. <https://doi.org/10.1073/pnas.1307204111>
- Hundt, C. & Sternberg, R. (2014). How did the economic crisis influence new firm creation? *Jahrbücher für Nationalökonomie und Statistik*, 234(6), 722-756. <https://doi.org/10.1515/jbnst-2014-0605>
- Jennings, J. E. & Brush, C. G. (2013). Research on women entrepreneurs: challenges to (and from) the broader entrepreneurship literature? *Academy of Management Annals*, 7(1), 663-715. <https://doi.org/10.1080/19416520.2013.782190>
- Kalden, J. N., Cunningham, J. & Anderson, A. R. (2017). The social status of entrepreneurs: Contrasting German perspectives. *The International Journal of Entrepreneurship and Innovation*, 18(2), 91-104. <https://doi.org/10.1177/1465750317706439>
- Kwon, I. & Sohn, K. (2021). Trust or distrust: entrepreneurs vs. self-employed. *Small Business Economics*, 56(4), 1553-1570. <https://doi.org/10.1007/s11187-019-00278-y>

- Leoni, T. & Falk, M. (2010). Gender and field of study as determinants of self-employment. *Small Business Economics*, 34(2), 167-185. <https://doi.org/10.1007/s11187-008-9114-1>
- Margolis, D. N. (2014). By choice and by necessity: Entrepreneurship and self-employment in the developing world. *The European Journal of Development Research*, 26(4), 419-436. <https://doi.org/10.1057/ejdr.2014.25>
- Parker, S. C. (2018). *The economics of entrepreneurship*. Cambridge University Press.
- Saridakis, G., Marlow, S. & Storey, D. J. (2014). Do different factors explain male and female self-employment rates? *Journal of Business Venturing*, 29(3), 345-362. <https://doi.org/10.1016/j.jbusvent.2013.04.004>
- Sorgner, A., Fritsch, M. & Kritikos, A. (2017). Do entrepreneurs really earn less? *Small Business Economics*, 49(2), 251-272. <https://doi.org/10.1007/s11187-017-9874-6>
- Vutsova, A., Arabadzhieva, M. & Angelova, R. (2023). The youth entrepreneurship as response to the youth unemployment-examples of Western Balkan region. *International Journal of Professional Business Review*, 8(6), 21. <https://doi.org/10.26668/businessreview/2023.v8i6.2299>
- Zelekha, Y., Avnimelech, G. & Sharabi, E. (2014). Religious institutions and entrepreneurship. *Small Business Economics*, 42(4), 747-767. <https://doi.org/10.1108/IJEER-10-2012-0121>