

**Proceedings of the**  
**International Conference on**  
**Gender Research**  
**ICGR 2018**

**Hosted by**  
**ISCAP, P.Porto**  
**Porto, Portugal**

**12 - 13 April 2018**

**Edited by**  
**Dr Ana Azevedo**  
ISCAP/ Polytechnic of Porto, Portugal  
**Dr Anabela Mesquita**  
Porto Business and School / Polytechnic of Porto, Portugal

Copyright The Authors, 2018. All Rights Reserved.

No reproduction, copy or transmission may be made without written permission from the individual authors.

### **Review Process**

Papers submitted to this conference have been double-blind peer reviewed before final acceptance to the conference. Initially, abstracts were reviewed for relevance and accessibility and successful authors were invited to submit full papers. Many thanks to the reviewers who helped ensure the quality of all the submissions.

### **Ethics and Publication Malpractice Policy**

ACPIIL adheres to a strict ethics and publication malpractice policy for all publications – details of which can be found here:

<http://www.academic-conferences.org/policies/ethics-policy-for-publishing-in-the-conference-proceedings-of-academic-conferences-and-publishing-international-limited/>

### **Conference Proceedings**

The Conference Proceedings is a book published with an ISBN and ISSN. The proceedings have been submitted to a number of accreditation, citation and indexing bodies including Thomson ISI Web of Science and Elsevier Scopus.

Author affiliation details in these proceedings have been reproduced as supplied by the authors themselves.

The Electronic version of the Conference Proceedings is available to download from DROPBOX <http://tinyurl.com/ICGR2018> Select Download and then Direct Download to access the Pdf file. Free download is available for conference participants for a period of 2 weeks after the conference.

The Conference Proceedings for this year and previous years can be purchased from <http://academic-bookshop.com>

E-Book ISBN: 978-1-911218-78-4

E-Book ISSN: 2516-2810

Book version ISBN: 978-1-911218-77-7

Book Version ISSN: 2516-2802

Published by Academic Conferences and Publishing International Limited

Reading

UK

Tel: +44-118-972-4148

[www.academic-conferences.org](http://www.academic-conferences.org)

# Clustering the Portuguese Employee Population Regarding Gender: A NUTIII Cluster Analysis

Alcina Nunes<sup>1</sup>, Sofia Cardim<sup>2</sup> and Carlos Balsa<sup>2</sup>

<sup>1</sup>Applied Management Research Unit (UNIAG), Institute Polytechnic of Bragança, Portugal

<sup>2</sup>Institute Polytechnic of Bragança, Portugal

[alcina@ipb.pt](mailto:alcina@ipb.pt)

[sofiacardim@ipb.pt](mailto:sofiacardim@ipb.pt)

[balsa@ipb.pt](mailto:balsa@ipb.pt)

**Abstract:** Throughout the last century and the last decade of the XXI century, where the issue is being discussed in several social, political and scientific forums, women have achieved unquestionable labour market rights increasing their presence and visibility. Nevertheless, they continue to be discriminated in modern labour market environments and gender segregation continues to be a feature common to many markets, namely the ones of more developed economies such as the Portuguese. The phenomenon is undesirable and has been fought by a number of public measures, not just those addressed directly at the labour market but also those which aim to combat gender discrimination in societies as a whole. In order these measures could be effective and produce the desired effects, it is necessary to know and understand the reality in order to act on it. In particular, it is necessary to understand if there are regional differences that can be considered in specific public policies directed to well-targeted populations. The availability of regional statistical information on the labour market, which provides information by gender, for variables such as staff employed by companies regarding professions, duration and type of work or earnings and remuneration earned by men and women in different economic activities, allows also the use of increasingly sophisticated statistical techniques and thus to develop and present research that characterizes the female and male NUTIII employee population in mainland Portugal. Considering the new potential of information and statistical methodologies, this paper applies the exploratory analysis of clusters to develop hypotheses about the existence of regional differences related to men and women employee in the national labour market. The latest data for the desired regional division, which match the intermunicipal communities, made available by the National Statistical Institution refer to 2013, a period not too far away from the present time to identify patterns on which the public power may act developing the most adequate and effective measures.

**Keywords:** labour market, employee population, gender, cluster analysis, Portugal

---

## 1. Introduction

It is unquestionable that, in the last decades, women have achieved (a never seen before) a highest participation in the paid labour market (Coelho, 2010; Drutarovská, et al, 2016; Instituto Nacional de Estatística, 2012) as in other spheres of their active life (Abrantes, 2013), such as political and public governance structures participation. Several recent research investigations support this idea. “The promotion of women’s economic independence has long been identified as pivotal for advancing gender equality in Europe. Accordingly, the EU’s main economic and social policy instruments have stressed the need to bring more women into the labour market and to increase their participation in economic activity” (Addabbo et al., 2015, p. 451).

However, and apart from these figures and facts which are unquestionable, there is also evidence that discrimination, considering a gender approach, continues to be a constant at several and distinct levels (Carmo and Cantante, 2015; Coelho, 2010; Instituto Nacional de Estatística, 2012; Jaba et al, 2017; Périvier, 2014), even in more developed contemporary societies. Portugal is not unaware to this phenomenon (Perista, 2016).

Effectively, even before the entry of Portugal into the European Union (at that time European Economic Community) in 1986, was already abundant the existence of legislation that sought to promote women's access and integration into education and, subsequently, into the labour market, which meant that Portugal, notwithstanding a slight more than a decade after the end of a 40-year dictatorship, was a pioneer country on this subject (Ferreira, 2010). Inherently, while there is a continued concern about the discrimination phenomenon by the competent authorities, particularly, discrimination by gender, on access and permanency on the labour market by women, this is a reality that tends to perpetuate itself, and even showing some setbacks in times of crisis, as it was the case of the post-2008 financial crisis (Pivoňka and Löster, 2013). Despite the country's small size, Portugal has many regional disparities that determine the existence and the persistence of diverse asymmetries, namely in the labour market and regarding gender issues, that also change in time of crisis (Nunes and Cardim, 2017).

Also, if, on the one hand, and as already mentioned, there is an abundance of recent research that seeks to analyse and to understand the reason for the persistence and maintenance of inequalities, on the other hand, studies with a regional character are scarce. Therefore, it is important to know and to understand these regional particularities, to define policies and measures, for an effective transformation. Even though the development of global policies, that combat inequalities, is desirable, it is also important to know the specificity of each region, to legislate and define appropriate measures for the disadvantaged populations, specifically referring to the case of the women's collective and their position in the labour market. The adequacy of the policies developed is the first step towards its success in combating inequalities (Carmo and Cantante, 2015). Indeed, Pereira and Galego (2015, p. 81) argue how important is to understand intra-regional dimension of wage inequality "in order to improve the effectiveness of policies and to design inequality correction programmes when needed. In fact, intraregional inequality dynamics as well as the underlying causes of inequality may differ across regions".

There are several researches that analyse the presence of the female collective in the labour market. In a generalist way, two distinct types of investigations can be identified: the research works developed to assess wage differentials as an explained variable (González, Santos and Santos, 2005; Pereira and Galego, 2015; Perista, 2004; Vieira, Couto and Tiago, 2005), and the researches which aim to explain the reason for gender inequalities in the labour market, having as explained variable an index of segregation, for example. These models are still rare (in Portuguese reality) due to unavailability of data, but can be mentioned the investigation of Dueñas, Iglesias and Llorente (2014) and García, García and Montuenga (2011), both applied to Spanish labour market.

All this models have as explained variables: (i) the specific and inherent characteristics of labour like tenure, type of employment contract, professional experience, type of position held, skills associated with the position or number of hours worked (Finzi, 2007; Frederiksen, 2008; Pascual, 2010; Iglesias and Llorente, 2008); (ii) company characteristics like dimension, wage or the duration of the working period; and (iii) personal characteristics like age, status, educational level or nationality (Gelves, 2011; Finzi 2007), which have as main objective to understand the reason for the maintenance of the asymmetries due to gender. Furthermore, and with a regional character, such investigations are not so common, since data are not always available, with the desired disaggregation which allows to conduct detailed regional analyses.

One type of analysis that is usually conducted nowadays, and allows to make spatial analysis, is the clusters analysis, which consists in a statistical approach, which intends to evaluate a group of observations (groups or clusters), and whose main objectives are: data exploration, grouping and consequent data reduction and hypothesis generation. Indeed, the statistical method is being applied as an effective tool, in a set of problems as social ones, to generate hypothesis about the classification and categorization of a structure of data. As Anderberg (1973, p. 4) mentions "an algorithm can assemble observations into groups which prior misconceptions and ignorance would otherwise preclude". Being the objective of this research work to identify the existence of regional differences related to men and women employee in the Portuguese labour market, to develop a proper and specific regional approach to the public labour market policies, the application of the exploratory cluster methodology serves such objective. The application of the cluster methodology to the latest data for several labour market indicators - labour earnings and remuneration, labour duration and type of work developed by men and women - in different intermunicipal communities, made available by the National Statistical Institution, allows to identify patterns on which the public power may act developing and putting on practice the most adequate and efficient measures. This call of attention has also been made for the case of Portugal by Albuquerque and Ferreira (2015) who claim that a global phenomenon may hide different regional dynamics. They authors do not doubt national performance is the result of regional performances but believe a more detail regional perspective allows a differentiate spatial treatment able to promote regional and consequent national development.

The present paper is divided in six sections. In the next section, the methodology of cluster analysis is presented. After, the data is defined and described. Follows a section which presents the results of the cluster analysis and a section where the discussion of results is made. The final section, presents the main conclusions.

## **2. Methodology**

Clustering is an important process for data classification that consists in organising a set of data observations into groups, called clusters. A cluster is a subset of an original set of data observations that are close together in some distance measure. In other words, given a data matrix containing multivariate measurements on a large number of individuals (observations), the aim of the cluster analysis is to build up some natural groups (clusters) with homogeneous properties out of heterogeneous large samples. Indeed, groups are based on similarities and this is the main focus of the problem solving. As mentioned by Jain (2010, p. 651), “organizing data into sensible groupings is one of the most fundamental modes of understanding and learning” about a problem in order to formulate a solution.

The similarity depends on the distance between the data for observations and a reduced distance among them indicates that they such observations are more similar (Balsa, Nunes and Barros, 2015). So, to identify the similarities (or differences) is important to quantify them. Several distinct methods can be used to measure the (smallest or biggest) distance among the observations of a data set. One of the most popular and common ways to compute such a distance (namely do to its simplicity) is the traditional Euclidian distance (Duran and Ordell, 1974), that will be used also in this empirical work.

Another choice that has to be made regards the cluster technique. There are two main classes of clustering techniques: hierarchical and optimization (or partitional) methods (Jain, 2010). In hierarchical clustering is not necessary to know in advance the number of subsets in which we want to divide the data. The observations are successively included in groups of different dimensions depending on the level of clustering. The result is a set of nested partitions. In each step of the process, two groups are either merged (agglomerative methods) or divided (divisive or top-down methods) according to some criteria (Martinez; Martinez and Solka, 2010). In the agglomerative approach, single-members clusters (clusters with only one observation) are increasingly fused until all observations are in only one cluster. The divisive approach starts with a single set containing all points. This group will be increasingly divided as the distance between points is reduced. The set of nested partitions is represented graphically by a dendrogram that has a tree shape indicating the distances hierarchical dependence. In this work we use methods from both classes. We use the Complete Linkage and Ward hierarchical methods and the k-means most popular optimization method. “Compared to hierarchical clustering algorithms, the optimization clustering algorithms find all the clusters simultaneously as a partition of the data and do not impose a hierarchical structure” (Jain, 2010, p. 653).

The k-means (MacQueen, 1967) is an optimization method that partitions the data in exactly k clusters, previously determine. This is achieved in a sequence of steps which begins, for instance, with an initial partition randomly generated. In each step the cluster’s arithmetic vector mean (centroid) is computed. The minimum distance between each data point and the clusters’ different centroids will decide the formation of new clusters. The formation of a new cluster implies assigning each observation to the cluster which presents the lowest distance. After that the centroids are (re)calculated and the former step is repeated until the moment each individual belongs to a stable cluster, i.e., when the sum of the squared distances to the centroid of all data point over all the clusters is minimized (Eldén, 2007). According with Jain (2010) the ease of implementation, the simplicity, the efficiency and the empirical success of a method that has more than 50 years are the main reasons for its popularity and continuous application.

## **3. Data description**

The dataset here used comprises 23 intermunicipal communities – each one comprising a set of Portuguese municipalities geographically adjacent – called the Portuguese administrative NUTIII regions. Each intermunicipal community is an observation in the data set. The 23 NUTIII regions are: Alto Minho, Cávado, Ave, Área Metropolitana do Porto, Alto Tâmega, Tâmega e Sousa, Douro, Terras de Trás-os-Montes, Oeste, Região de Aveiro, Região de Coimbra, Região de Leiria, Viseu Dão Lafões, Beira Baixa, Médio Tejo, Beiras e Serra da Estrela, Área Metropolitana de Lisboa, Alentejo Litoral, Baixo Alentejo, Lezíria do Tejo, Alto Alentejo, Alentejo Central and Algarve.

For each NUTIII region are published several indicators that show how the presence of women and men is distributed in the labour market. The latest year for which such indicators are available is 2013, a period not too far away from the present time to identify patterns that may be found in the Portuguese labour market. For the present research work were collected from Pordata (2017), a database of contemporary Portugal, a set of 8

indicators that describe some employment characteristics of women and men. The 8 employment characteristics used in this study are identified and described in Table 1. As in the studies mentioned in the first section the characteristics selected reflect labour characteristics: (1) monthly basic remuneration; (2) monthly earnings; the situation in the main occupation, this is, if the individual is (3) an employer or (4) an employee; the duration of work, this is, if the individuals has (5) a full-time contract or (6) a part time contract and, finally, the type of contract, this is, if the contract is a (7) fixed-term contract or (8) a permanent contract. In Table 2 is possible to observe the differences between men and women, for each labour market characteristic, and by NUTIII region.

**Table 1:** Employment characteristics and definitions

Characteristic		Definition
1	Monthly basic remuneration	The gross amount, before deduction of taxes and social security contributions, in cash or in kind, paid regularly in the reference period and corresponding to the normal working period.
2	Monthly earnings	The gross amount (before deduction of any discounts) in cash and/or in kind paid monthly, on a regular basis, for the hours of work performed, as well as for the hours paid but not actually worked. This includes, in addition to the basic pay, all premiums and regular grants, as well as payment for overtime.
3	Situation in the main occupation Employers	An employer is a person who, working on his/her own account and with one or a few partners, obtains remuneration depending directly on (realised or potential) profits resulting from goods or services produced and therefore has engaged one or more persons to work for him/her in his/her business as employees.
4	Situation in the main occupation Employees	An employer is a person on those jobs where the incumbents hold explicit (written or oral) or implicit employment contracts which give them a basic remuneration not directly dependent upon the revenue of the unit for which they work.
5	Duration of work Full time	A full time worker is a worker whose normal hours of work are equal or greater than the normal hours of work prevailing on the enterprise/institution, for the respective professional category or in the respective profession.
6	Duration Work Part time	A part time worker is an employed person whose normal hours of work are less than the normal hours of work prevailing on the enterprise/institution, for the respective professional category or in the respective profession.
7	Type of contract Fixed-term contract	A person who is linked to the enterprise/institution by a fixed-term written contract, containing a specific reference to: (a) fixed term or a varying duration.
8	Type of contract Permanent contract	A person with a contract or a work contract with undetermined duration.

Source: Self elaboration using as source the metadata of Pordata (2017)

**Table 2:** Labour market differences between men and women, by characteristic and region, in 2013

NUTIII regions	Monthly basic remuneration	Monthly earnings	Situation in the main occupation Employers	Situation in the main occupation Employees	Duration of work Full time	Duration Work Part time	Type of contract Fixed-term contract	Type of contract Permanent contract
	Values in %							
Alto Minho	14.93	18.91	6.12	-5.57	6.12	-6.12	0.78	-0.56
Cávado	17.11	20.32	5.36	-5.38	3.04	-3.04	1.63	-2.37
Ave	16.07	21.01	4.73	-4.76	0.91	-0.91	-0.12	-0.15
Área Metropolitana do Porto	19.21	22.58	3.97	-3.97	8.21	-8.21	-0.99	0.93
Alto Tâmega	9.22	13.54	6.61	-6.29	3.77	-3.77	-2.63	2.09
Tâmega e Sousa	13.66	16.74	4.21	-4.23	2.71	-2.71	2.69	-4.69
Douro	11.96	17.14	5.00	-4.99	4.23	-4.23	-0.92	0.13
Terras de Trás-os-Montes	12.68	15.53	7.05	-7.34	3.03	-3.03	-1.81	1.85
Oeste	18.71	21.29	5.13	-5.16	3.85	-3.85	1.11	-0.88
Região de Aveiro	23.20	25.89	4.79	-4.81	5.09	-5.09	0.97	-1.70
Região de Coimbra	19.01	24.61	5.13	-5.13	5.67	-5.67	3.03	-3.60
Região de Leiria	22.41	26.41	3.86	-3.77	4.41	-4.41	2.65	-1.49
Viseu Dão Lafões	15.05	19.81	5.23	-5.32	4.64	-4.64	0.76	-0.75
Beira Baixa	15.48	21.33	4.67	-4.77	3.46	-3.46	-1.27	0.87
Médio Tejo	16.11	23.07	4.60	-4.60	4.15	-4.15	2.40	-1.69
Beiras e Serra da Estrela	11.34	17.40	6.31	-6.24	2.51	-2.51	3.36	-3.63
Área Metropolitana de Lisboa	21.78	24.37	3.33	-3.44	8.68	-8.68	-0.08	-1.08
Alentejo Litoral	33.22	43.31	2.44	-2.39	4.79	-4.79	0.54	-5.72
Baixo Alentejo	14.50	27.93	3.10	-3.25	4.08	-4.08	2.61	-2.40
Lezíria do Tejo	20.02	22.72	4.39	-4.45	3.55	-3.55	2.47	-2.23
Alto Alentejo	19.93	23.21	4.18	-4.46	3.46	-3.46	1.20	-1.52
Alentejo Central	18.50	22.13	4.54	-4.34	3.51	-3.51	0.88	-0.45
Algarve	15.08	18.11	5.32	-5.27	3.27	-3.27	-0.07	-0.04
<b>Total (average)</b>	<b>17.35</b>	<b>22.06</b>	<b>4.78</b>	<b>-4.78</b>	<b>4.22</b>	<b>-4.22</b>	<b>0.83</b>	<b>-1.26</b>

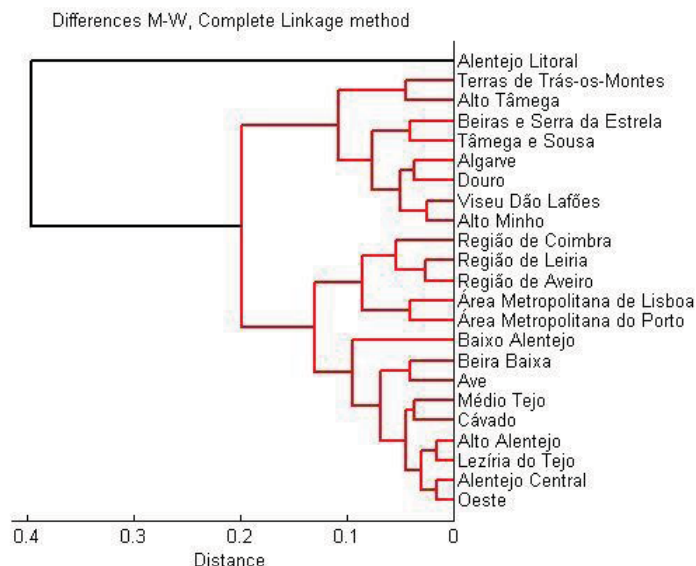
Source: Self elaboration using the information collected in Pordata (2017)

The value of each of the variables corresponds to the difference between the values observed for male employees and female employees, in the year 2013. So, a positive difference corresponds to a discrimination in favour of men. The original values for male and female employees were previously normalised – the original values for men and women are relative values regarding the value found for the all employed population independently of the gender. The last row in the table shows the average value found for the all territory.

It is possible to observe that the main differences are the ones on the monthly remuneration (the official remuneration of employees) and the monthly earnings (the real earning of employees). In Portugal, men can have a remuneration, in average, approximately 17% higher than women. In some regions the difference is much smaller (for instance, around 9% in Alto Tâmega) but in other the differences are significantly higher (for instance, around 33% in Alentejo Litoral). When the effective earnings of male and female employees are analysed, the gender discrimination is even more obvious. In average, the effective earnings could reach 22%. In all regions the difference is always above 13% with women in a clear disadvantage. The differences' range among regions remains. For the other six employment characteristics the differences are not so large. Still, women continue to be in disadvantage. There are more male than female employers, in all 23 regions. By opposition, the relative number of women that work for others is bigger than the number of men (also for all the 23 regions). The male employees are also the ones that relatively more have full time and permanent contracts (which explains the negative differences that appear for the variables part-time and fixed-term contracts). Note, however, that regarding the variable that shows the relative difference between men and women in terms of fixed-term contracts, seems to exist differences among regions. If the overall result is positive for men, that do not happen in all regions. Note that Addabbo et al (2015) found out that in the period between 2007 and 2012, the gender gap in temporary employment decreased in Portugal due to the deterioration of men's labour market conditions. However the previous authors just analysed the full Portuguese economy. This particular fact together, with the observation of a big range of differences among the 23 regions for the first two variables in analysis, makes relevant the application of a cluster analysis to check if exists a unique pattern of differences among men and women, in the Portuguese labour market or if some regions share patterns that demand specific public measures. The cluster methodology will allow to gain insight into data, generate hypotheses and identify perceptible features.

#### 4. Results

The classification of the 23 regional divisions in function of the 8 measures presented in Table 1 is achieved by means of the clustering methodology. Firstly, hierarchical methods are applied. The dendrogram obtained with the complete linkage method is presented in Figure 1.



**Figure 1:** Dendrogram obtained with the Complete Linkage cluster method

In Figure 1 is possible to observe two main clusters. One that includes a single a region, Alentejo Litoral, and another with all the remaining NUTIII. The distance between these two clusters is close to 0.4. Still, inside the

larger cluster it is also possible to distinguish two main groups distanced from near 0.2. These results indicate that the 23 NUTIII regions could be classified in two or three groups.

The dendrogram obtained with the ward method is presented in Figure 2. This dendrogram shows more crystal clear that there are three main groups. This is, is possible to join the several 23 regions in 3 distinct groups with different labour profiles, regarding the difference between men and women. As it was observed previously, with the Complete Linkage method, one group comprises exclusively the Alentejo Litoral region, whose distance from the other two groups is near 0.4. The other two groups present a distance between them close to 0.3. We observe also that the two other groups are very similar with the two groups resulting from the decomposition of the main cluster obtained with the Complete Linkage method. The single difference in the classifications are due to Beira Baixa and Ave regions that moves to a different cluster.

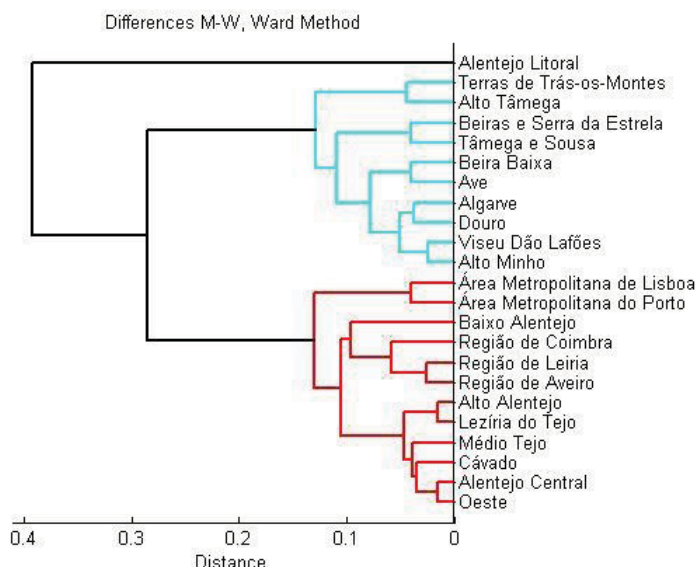


Figure 2: Dendrogram obtained with the Ward cluster method

The results obtained with the complete linkage and Ward hierarchical methods indicates that the optimal number of classes in which the 23 NUTIII regions is three. Consequently, we apply to the k-means the data in three classes. The results obtained with this method are presented in Table 3.

Table 3: Classification obtained with the k-means cluster method

Cluster #	NUT III
1	Alentejo Litoral.
2	Alto Minho, Cávado, Ave, Alto Tâmega, Tâmega e Sousa, Douro, Terras de Trás-os-Montes, Viseu Dão Lafões, Beira Baixa, Beiras e Serra da Estrela, Algarve.
3	Área Metropolitana do Porto, Oeste, Região de Aveiro, Região de Coimbra, Região de Leiria, Médio Tejo, Área Metropolitana de Lisboa, Baixo Alentejo, Lezíria do Tejo, Alto Alentejo, Alentejo Central.

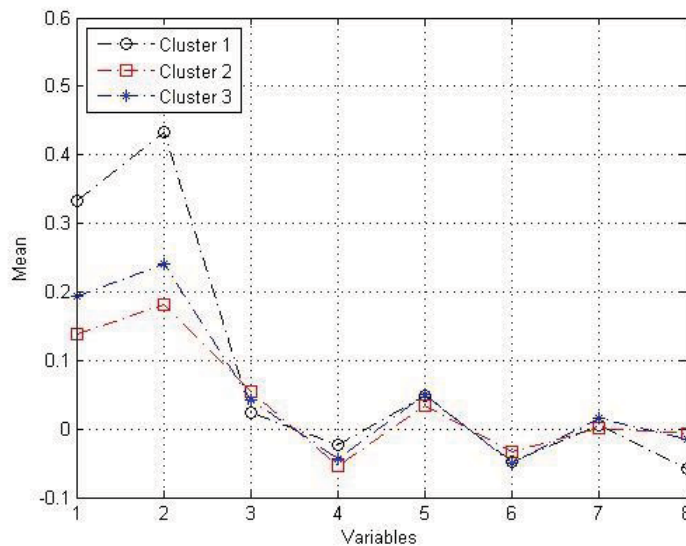
The results obtained with the k-means method are very similar to the ones obtained with the Wards cluster hierarchical method. As it was seen with the previous methods, the Alentejo Litoral is alone in a single group what shows this region present a pattern of differences between men and women in the labour market completely different from the rest of the economy. This result is robust and is obtained with all the methods applied. Regarding the hierarchical methods, the unique difference in the results of the K-means methods is due to the Cávado region that changes cluster.

## 5. Discussion

The results obtained with the different clustering method indicate that the NUTIII regions can be classified in three main groups concerning the differences between men and women in the Portuguese labour market. In the Alentejo Litoral regions these differences are important compared with the differences found for other regions. As already mentioned, for all the methods applied the regions appear always alone.

The composition of the other two groups of regions varies slightly with the classification method chosen. It was verified that only the Beira Baixa, Ave and Cávado regions alternate their classification group in function of the method. These results indicate that for these regions the differences between men and women in the labour market are equidistant of the mean of these two groups and, consequentially, they could be assigned to any of the two groups.

For a better understanding of the reasons for the groups (regions) classification it is necessary to compare the average value (the mean) for each variable in each cluster obtained using the k-means method (see Table 2). The results are present in Figure 3.



**Figure 3:** Average value (mean) of each variable for each cluster

Figure 3 allows to observe that the average values of the first and second variables are the ones that differ more among the three clusters. From these clear results is possible to conclude that the monthly real and effective earnings are responsible for three main regional profiles regarding the participation of men and women in the Portuguese labour market. For all the other variables, the average differences observed for each cluster do not show relevant differences in the participation of men and women in the labour market. The average differences are clearly different for cluster 1 which includes Alentejo Litoral. The difference in earnings, are bigger than 30%. The smaller differences in the earnings received by men and women are observed in cluster 2 which comprises the regions of: Alto Minho, Cávado, Ave, Alto Tâmega, Tâmega e Sousa, Douro, Terras de Trás-os-Montes, Viseu Dão Lafões, Beira Baixa, Beiras e Serra da Estrela and Algarve. In this cluster the differences are between 10 and 20% for real and effective earnings, respectively.

Between the previous two clusters is cluster 3 which includes Área Metropolitana do Porto, Oeste, Região de Aveiro, Região de Coimbra, Região de Leiria, Médio Tejo, Área Metropolitana de Lisboa, Baixo Alentejo, Lezíria do Tejo, Alto Alentejo and Alentejo Central. In these regions the difference in earnings are equal to 20% or slightly higher than this value.

## 6. Conclusions

Portugal is a small economy divided in small administrative regions. In these regions the differences between men and women in the labour market are a reality with men presenting a clear advantage. The distribution of women regarding the occupation, type and duration of a contract reflects some discrimination. However, the biggest discrimination is observed in terms of what the two genders receive at the end of each month. Such official and effective different earnings allow to divide the country 23 NUTIII regions in three different groups/clusters, that demand specific public policy measures. Among all regions is clear the difference observed in Alentejo Litoral, a depressed region in south of Portugal, yet all the other regions present high levels of earning discrimination by gender. The cluster analysis, as mentioned, does not permit explain the results obtained but generates hypothesis that further research studies may test and explain. For now, this research work shows the existence of different pattern of differences in the Portuguese labour market - even if more variables would be

important to include in the analysis, namely company and personal variables – and the need to better target the labour market and other social measures to overcome such differences. The addition of another dimension in the analysis would be important to understand such division – the regions are identified by different productive structures and, therefore, the analysis of gender discrimination among different economic activities is an important aspect to include in the previous analysis.

## Acknowledgements

The preparation of the paper was supported by UNIAG, R&D unit funded by the FCT – Portuguese Foundation for the Development of Science and Technology, Ministry of Science, Technology and Higher Education; “Project Code Reference UID/GES/4752/2016”.

## References

- Abrantes, M. (2013) “Fortalezas e Masmorras: a Persistência da Divisão Sexual das Profissões na Sociedade Contemporânea”, *Ex aequo*, Vol 27, pp 113-127.
- Addabbo, T., Bastos, A., Casaca, S. F., Duvvury, N., and Ni Leime, A. (2015) “Gender and labour in times of austerity: Ireland, Italy and Portugal in comparative perspective. *International Labour Review*”, Vol 154, No. 4, pp 449-473.
- Albuquerque, P., and Ferreira, J. (2015) “Envelhecimento, emprego e remunerações nas regiões portuguesas: uma análise shift-share”, *EURE (Santiago)*, Vol 41, No. 122, pp 239-260.
- Anderberg, M. R. (1973) *Cluster Analysis for Applications*, Probability and Mathematical Statistics, Academic Press, New York.
- Balsa C., Nunes A. and Barros E. (2015) “Optimization Clustering Techniques on Register Unemployment Data”, In Almeida J., Oliveira J., Pinto A. (Eds), *Operational Research* (pp 19-35), CIM Series in Mathematical Sciences, Springer, Cham.
- Carmo, R. M and Cantante, F. (2015) “Desigualdades, Redistribuição e o Impacto do Desemprego: Tendências Recentes e Efeitos da Crise Económico-Financeira”, *Sociologia, Problemas e Práticas*, No. 77, pp 33-51.
- Coelho, L. (2010) “Mulheres, família e desigualdade em Portugal” [online], PhD Thesis, Faculty of Economics, University of Coimbra, Coimbra,  
[https://estudogeral.sib.uc.pt/jspui/bitstream/10316/13833/1/Disserta%C3%A7%C3%A3o%20Dout\\_Lina%20Coelho.pdf](https://estudogeral.sib.uc.pt/jspui/bitstream/10316/13833/1/Disserta%C3%A7%C3%A3o%20Dout_Lina%20Coelho.pdf).
- Drutarovská, J., Kováčová, J., Pechová, H. and Podmajerská, K. (2016) “Analysis of Women’s Status in the Labor Markets of Countries in the European Union”, *Journal of International Women’s Studies*, Vol 17, no. 1, pp 30-38.
- Dueñas, D., Iglesias, F. and Llorente, H. (2014) “Segregación Ocupacional por Género en España: ¿Exclusión o Confinamiento?”, *Revista Internacional del Trabajo*, Vol 133, No. 2, pp 345-371.
- Duran, B. S. and Odell, P. L. (1974) *Cluster Analysis: a Survey* (Vol. 100), Springer Science & Business Media.
- Eldén, L. (2007) *Matrix Methods in Data Mining and Pattern Recognition*, SIAM.
- Ferreira, V. (2010) *A Igualdade de Mulheres e Homens no Trabalho e no Emprego em Portugal: Políticas e Circunstâncias*, CITE, Lisboa.
- Finzi, I. (2007) “Occupational Gender Segregation and Gender Wage Gap in Switzerland”, PhD Thesis, Faculty of Economics, University of Lugano, Switzerland.
- Frederiksen, A. (2008) “Gender Differences in Job Separation Rates and Employment Stability: New evidence from Employer-Employee data”, *Labour Economics*, Vol 15, No. 5, pp 915-937.
- García, I., García, G. and Montuenga, V. M. (2011) “Segregación Ocupacional por Género en Aragón”, *Documento de trabajo, Nº 56/2011* [online], Universidad Zaragoza, <http://www.fundear.es>.
- Gelves, L. F. A. (2011) “Segregación Laboral y Discriminación Salarial de Género en Colombia: El caso de las Trece Áreas Metropolitanas, 2001, 2005 y 2009”, *Revista de Economía del Caribe*, No. 8, 202-231.
- González, P., Santos, M. and Santos, L. (2005) “The Gender Wage Gap in Portugal: Recent Evolution and Decomposition”, *Discussion Papers No. 505, CETE – Centro de Estudos de Economia Industrial, do Trabalho e da Empresa*, [online], <https://www.fep.up.pt/investigacao/cete/papers/dp0505.pdf>.
- Iglesias, C. and Llorente, R. (2008) “Evolución Reciente de la Segregación Laboral por Género en España”, *Documento de Trabajo del Instituto de Análisis Económico y Social (IAES) No. 13*, [online], <http://uvadoc.uva.es/handle/10324/11395>.
- Instituto Nacional de Estatística (2012) “Estatísticas no Feminino: Ser Mulher em Portugal 2001-201” [online], [https://www.ine.pt/ngt\\_server/attachfileu.jsp?look\\_parentBoui=136477029&att\\_display=n&att\\_download=y](https://www.ine.pt/ngt_server/attachfileu.jsp?look_parentBoui=136477029&att_display=n&att_download=y).
- Jaba, E., Sandu, C., Ploeanu, A., Robu, I., and Istrate, M. (2017). “The Statistical Analysis of Labor Market and Female Labor Force Characteristics in Central and Eastern European Countries”, *Studies and Scientific Researches*, No. 25, pp 19-33.
- Jain, A. K. (2010) “Data Clustering: 50 Years Beyond K-means”, *Pattern Recognition Letters*, Vol 31, No. 8, pp 651-666.
- MacQueen, J. B. (1967) “Some Methods for Classification and Analysis of Multivariate Observations”, *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability* (pp 281-297), University of California Press.
- Martinez, W. L., Martinez, A. R. and Solka J. L. (2010) “Exploratory Data Analysis with MATLAB”, CRC Press.
- Nunes, A. and Cardim, S. (2017) “Discriminação salarial por sexo e por setores de atividade em Portugal: evidências regionais em período de crise”, *Journal of Studies on Citizenship and Sustainability*, No. 3, pp 137-157.

- Pascual, M. I. (2010) "Al Otro Lado de la Segregación Ocupacional por Sexo. Hombres en ocupaciones Femeninas y Mujeres en Ocupaciones Masculinas", *Revista Internacional de Sociología*, Vol 68, No. 1, pp 145-164.
- Pereira, J. and Galego, A. (2015) "Intra-regional Wage Inequality in Portugal", *Spatial Economic Analysis*, Vol 10, No. 1, pp 79-101.
- Perista, H. (2004) "O contexto: análise das desigualdades salariais de género em Portugal", In Salada, A. (Comp.) *Compilação de elementos para uma consulta especializada sobre igualdade de remuneração entre mulheres e homens* (pp 43-59), Direção Geral Estudos, Estatística e Planeamento, Lisboa.
- Perista, H., Cardoso, A., Brázia, A., Abrantes, M., Perista, P., and Quintal, E. (2016) *Os Usos do Tempo de Homens e de Mulheres em Portugal*, CESIS – Centro de Estudos para a Intervenção Social e CITE – Comissão para a Igualdade no Trabalho e no Emprego, Lisboa.
- Périvier, H. (2014) "Men and women during the economic crisis", *Revue de l'OFCE*, No. 2, pp 41-84.
- Pivoňka, T. & Löster, T. (2013) "Cluster Analysis as a Tool of Evaluating Clusters of the EU Countries before and during Global Financial Crisis from the Perspective of the Labor Market", *Intellectual Economics*, Vol 7, No 4, pp 411-425.
- Pordata (2017) "Employment and Labour Market Indicators" [online], <http://www.pordata.pt/en/Portugal>.
- Vieira, J. Couto, J and Tiago, T. (2005) "Inter-regional Wage Dispersion in Portugal", Proceedings of the 45th Congress of the European Regional Science Association "Land Use and Water Management in a Sustainable Network Society" (pp 1-24), European Regional Science Association (ERSA), Amsterdam.