

---

## STATISTICAL ASSESSMENT OF CAREERS OF DOCTORATE HOLDERS SURVEY BY CORRESPONDENCE ANALYSIS: THE CASE OF TURKEY

Zerrin AŞAN GREENACRE \*

Department of Statistics, Faculty of Science, Anadolu University, 26470, Eskişehir, Turkey

### ABSTRACT

Nowadays, doctorate holders are not only academicians and researchers, but they also play roles across a broad ground of working areas within society. These areas are versatile and increasing in way. The aim of this study is to make statistical evaluation of careers of doctorate holders survey using the correspondence analysis with supplementary variables. The focus of this paper is doctorate holders in Turkey. A survey regarding this matter was carried out by Turkish Statistical Institute. This survey includes various questions such as gender, age, employment situation, field of science, etc. Using correspondence analysis we investigated doctorate holders in relation to their gender and employment situation and employment sector. How the employment situation of doctorate holders changes according to gender and year of doctorate award; and doctorate holders' satisfaction level on their principal job by selected title have been shown.

**Keywords:** Doctorate holders, Correspondence analysis, Supplementary variable

---

### 1. INTRODUCTION

Science, technology and innovation are important drivers of economic development and social welfare at both national and global levels. It is widely recognized that the education and qualification of the labor force are crucial for scientific achievements, technological breakthroughs, and innovation excellence [1].

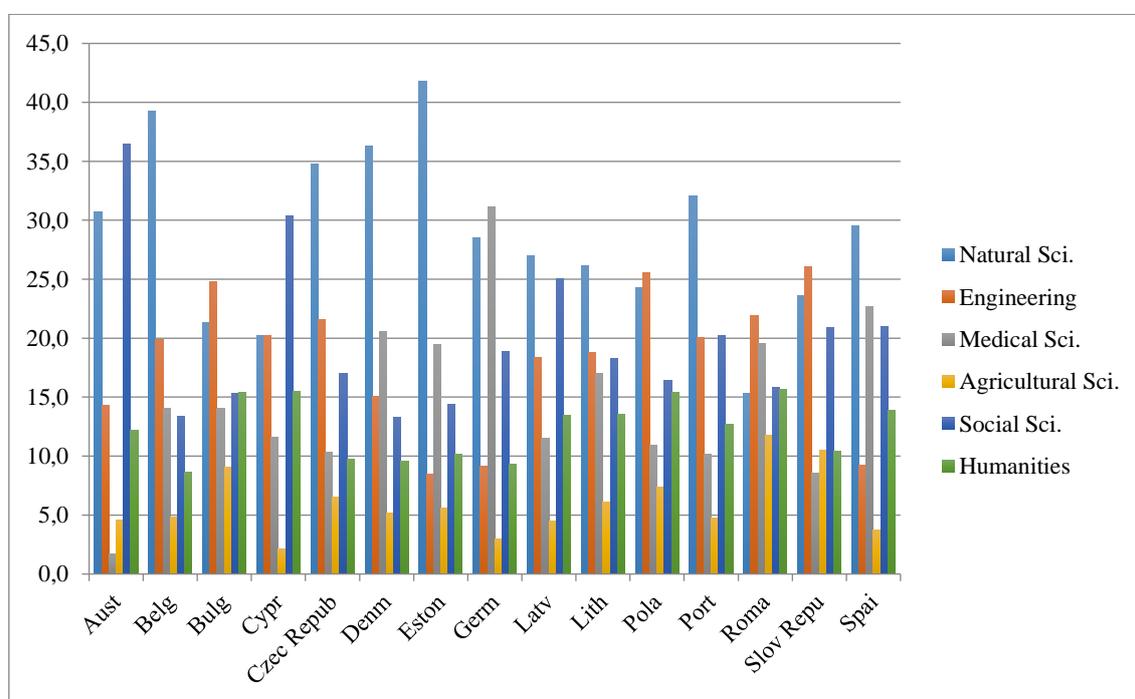
The creation, use and spread of innovation today depends on human resources [2]. The fascination of a country's research jobs is related to its capacity to put human resources into innovation. Among these resources, doctorate holders have the ability to conduct a research. Also they are the main part of a research system [3]. They are the most qualified in accordance with their educational status. Moreover, doctorate holders are specially trained to conduct research. In a knowledge-based society valuing higher education and advanced skills, there is a strong connection between graduate education, research and economic growth. Doctorate graduates play a major role in such a society because of their extremely specialized education and wide training in research. Statistics are gathered to measure the characteristics of careers doctorate holders (CDH) such as demographics, career, salary characteristics, and working levels [2].

The aim of this study is to examine the doctorate holders involved in such an important educational step. The aim is to make a statistical evaluation of CDH survey using the correspondence analysis with supplementary variable. CDH survey data in Turkey has been used for the application. One of the reasons why Turkey was taken as a case study in this study is the lack of any direct and detailed research in literature regarding this field in Turkey. Inanc and Ozcan, 2016, studied the relationship patterns among gender, family and academic careers in Turkey. They also used CDH survey data in their studies. They have shown how many doctoral graduates are involved in academic careers [4].

Firstly, the importance of CDH is explained. Secondly, data and methodology is described. Correspondence analysis techniques are used with supplementary variables in this section. Many variables such as gender, age, employment situation, field of science, graduate years have been included in the analysis. The results of the CDH survey show in detail the graphs of frequency distribution and correspondence analysis. The results obtained are discussed in the last section.

## 2. THE IMPORTANCE OF CAREERS OF DOCTORATE HOLDERS

Throughout the world, doctoral education systems are increasingly evolving with various constructive policies. Governments and some organizations such as the European Union and the Organization for Economic Co-operation places more emphasis on doctorate education. They are trying to expand and develop doctorate education in their countries and territories. Especially in the last two decades, doctoral education has increased in many countries with advanced higher education systems [5]. The total number of doctorate holders for The Organisation for Economics Co-operation and Development (OECD) countries between 1990-2009 are shown in Figure 1 and 2.

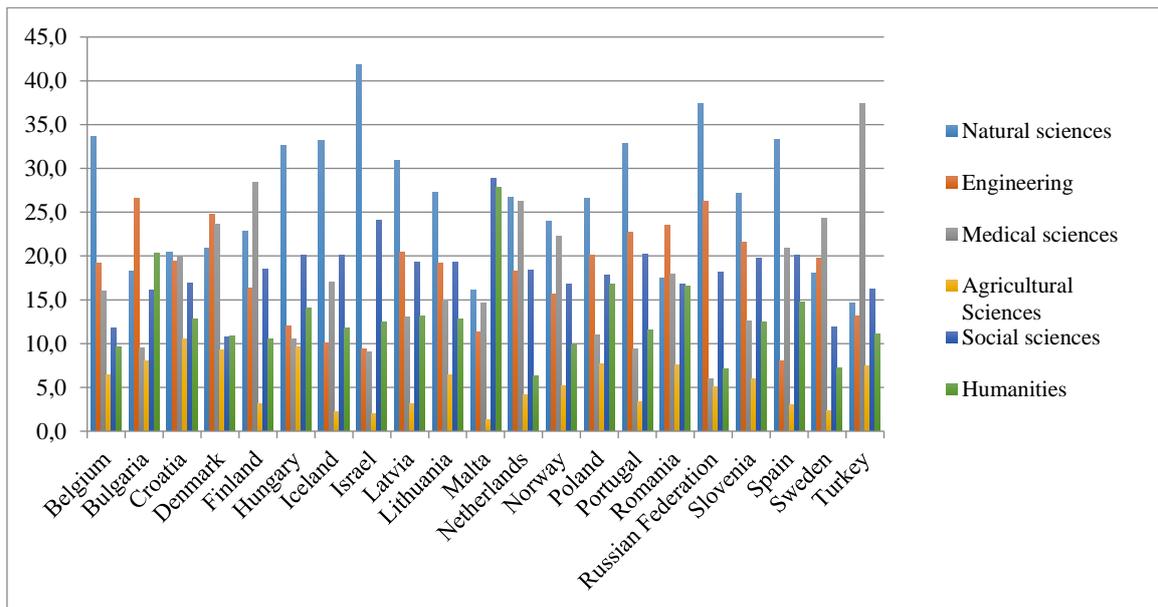


Source: Indicators for CDH, www.oecd.org.

**Figure 1.** Number of doctoral degrees received under 70 years of age between 1990-2006

As is seen in Figure 1, the rates are as follows: The highest number of doctorate holders is in Estonia, Belgium and Denmark in Natural Sciences, followed by Slovak Republic, Poland and Bulgaria in Engineering, Germany in Medical Sciences, Romania in Agricultural Science, Austria, Cyprus and Latvia in Social Sciences, and Cyprus, Romania, Bulgaria and Poland in Humanities.

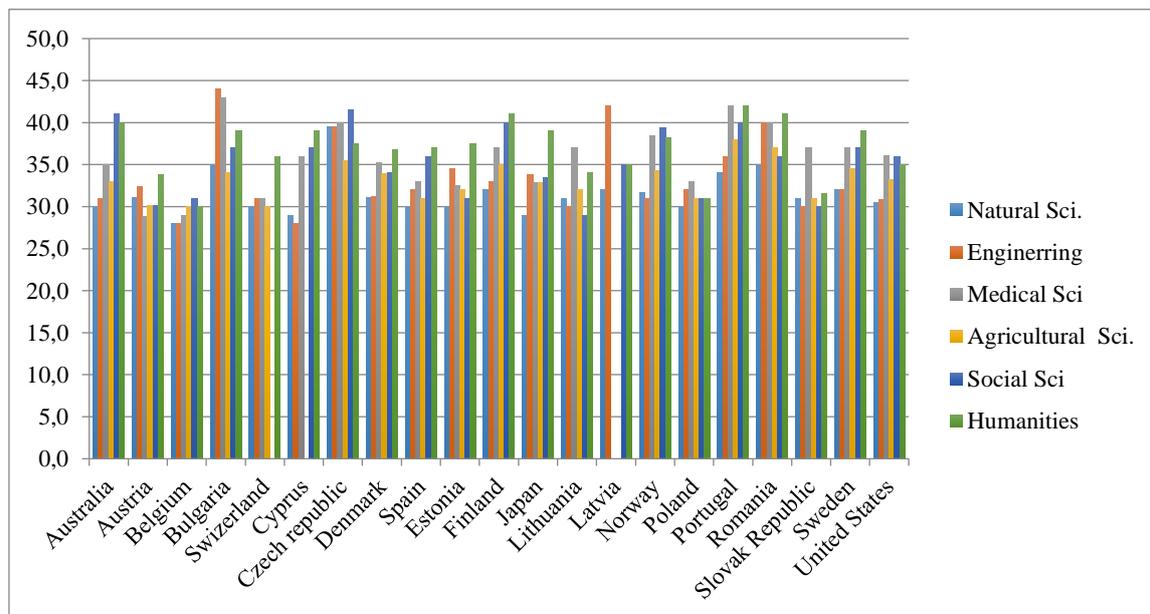
Some rates for 2009 are shown in Figure 2. This figure shows the total number of doctorate holders by field of study in 2009. The highest number of doctorate holders is in Israel in Natural Sciences, followed by Russian Federation and Bulgaria in Engineering and Technology, Turkey in Medical Sciences, Croatia in Agricultural Science, Malta in Social Sciences and Humanities. Looking at the economic data, it is seen that in the higher education sector, doctorate holders have more annual income than those who do not work as researchers [6]. It is evident that the amount of knowledge and innovation increases in proportion to the employment of doctoral holders in a country [3].



Source: Indicators for CDH, www.oecd.org.

**Figure 2.** Total doctorate holders according to field of study for 2009

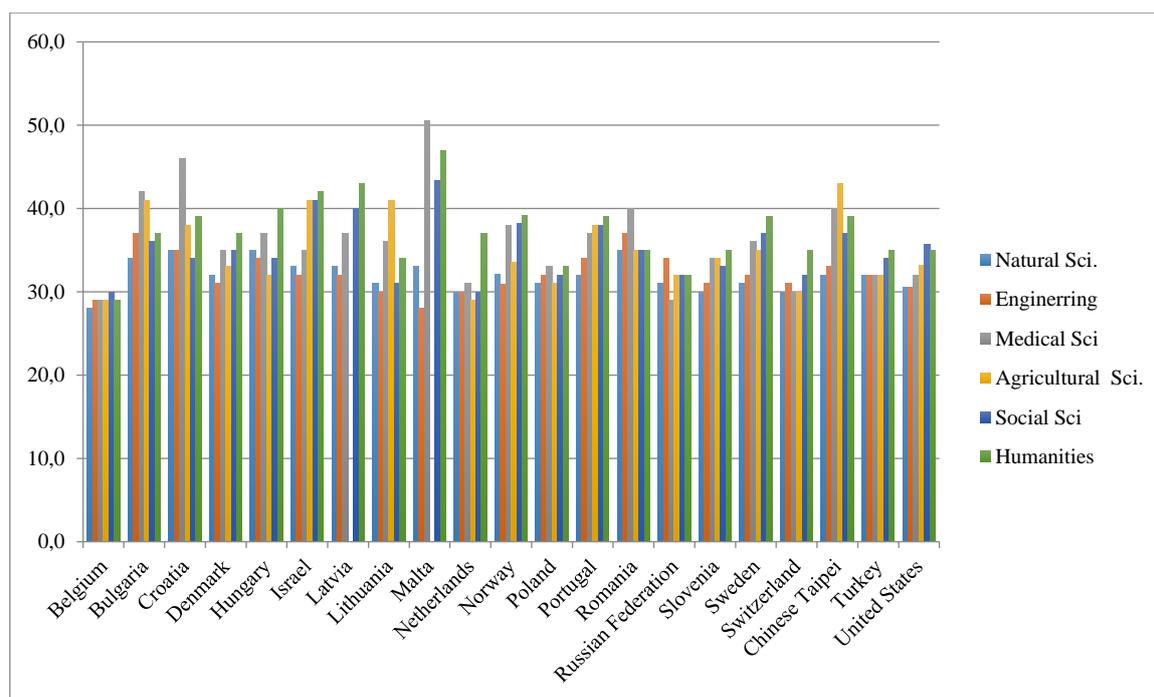
In Figure 3, median age defined the age of the person in the middle when all the persons in the population are arranged in ascending order at doctoral graduation by fields of study according to gender for 2006 is shown [7]. Median age at graduation is generally contained approximately between 29 and 34 years old. While Australia, Bulgaria, Czech Republic, Finland and Portugal have the highest median age (40 and above), Belgium has the lowest median age (28 years old). The lowest median age is generally seen in the engineering and natural sciences and the highest median age in the engineering and medical science compared with the other field of study. Within countries, Belgium has the lowest median age in engineering and natural science and Bulgaria has highest median age in engineering and medical science.



Source: Indicators for CDH, www.oecd.org.

**Figure 3.** Media age distribution of recent doctorate holders at graduation according to gender and field of study for 2006

Figure 4 shows median age at doctoral graduation by fields of study according to gender for 2009. Median age at graduation is generally contained between 30 and 35 years old. It can be seen that the lowest median age is 28 years. The highest median age is about 50 years. As in 2006, the lowest median age in 2009 is 28 years old in Belgium. Besides, the highest median age is 40 and above years in Bulgaria, Croatia, Israel, Latvia, Lithuania, Malta and Chiese Taipei. The lowest median age is generally seen in the natural sciences and engineering, and the highest median age in the humanities. Within countries, Belgium and Malta has the lowest median age in engineering and natural science and Malta has highest median age in medical science.



Source: Indicators for CDH, www.oecd.org.

**Figure 4.** Median age distribution of recent doctorate holders at graduation according to gender and field of study for 2009

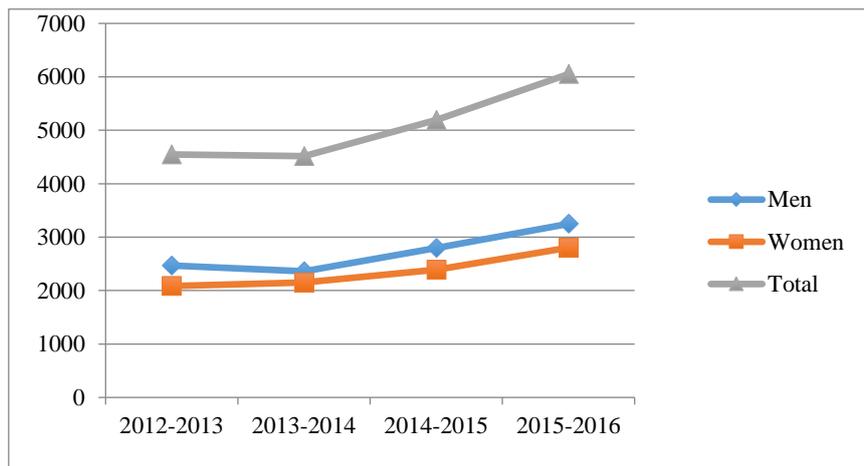
The training process is one of the most important steps of getting a doctorate degree. When we look at the work we do in this world, the most comprehensive study is the doctorate holders projects carried out by the OECD. These projects provide the views and some characteristics of those who have received doctoral studies in OECD countries [8]. Some other examples of studies in this field around the world can be listed as follows:

Gu et al. (2011) researched young doctorate holders in science in China [9]. Wildy et al. (2015) demonstrated a rise in the number of profession doctorate holders with case studies in China, Iceland, and Australia [10]. They demonstrated the important factors about doctorate holders. A survey called Careers of Doctorate Holders Survey was carried out in Belgium, 2010. This survey showed that nearly one in three doctorates awarded in Belgium is in natural sciences. Next, medical and health sciences take up a share of engineering and technology. These are followed by the social sciences, agricultural sciences and humanities. The number of males is considerably higher than the number of female doctorate holders in the natural sciences and engineering [11]. Mertens and Röbbken (2013) did research on German doctorate holders. Whether possible differences exist between the fields of German doctorate holder's studies was analysed [12]. Borrego et al. (2010) investigated postdoctoral scientists in Spain with a gender perspective [13]. Statistics Netherlands, a Dutch governmental institution, carried out the Dutch CDH Survey in 2014. According to this survey, the most popular field is medical and health sciences. The natural sciences are at second place. In every age group the

number of male doctorate holders is higher than the number of female doctorate holders, except for ages younger than 35 years in Dutch CDH Survey. There are slightly more women than men with a PhD in this age class [14].

Bin et al. (2015) studied the experiences and perspectives of Brazilian doctorate holders [15]. Waaijer et al. (2016) investigated whether the longevity of doctorate recipients' publication careers has changed in the United States. It is emphasized that the number of PhD students has increased since the 1950s in the USA in this paper [16].

The Doctorate program in Turkey aims to provide the students with the skills necessary to conduct independent research regarding scientific problems, to interpret the data in a broad and deep perspective, to analyse and reach new synthesis. In order to apply to doctoral programs (in other words, to do PhD), you must have a bachelor's or graduate degree in Turkey. One cannot apply for a doctorate study without at least twenty-one credits for at least seven courses, proficiency exam, thesis proposal and thesis work [17]. Figure 5 shows the number of doctoral students who graduated between 2012-2016 in Turkey. The number of male doctoral students graduated in Turkey is higher than the number of female doctoral students graduated according to years. After 2013, the number of doctoral graduates has shown an increase every year. The number of doctorate students registered in the 2016-2017 academic year was 91267 in total. About 59 percent of these doctoral students are male students and about 41 percent are female doctoral students [18].



Source: The Council of Higher Education's Statistics.

**Figure 5.** The number of doctoral students who graduated across the years in Turkey

### 3. THE DATA SOURCE AND METHODOLOGY

The statistics of the doctorate holders have an important place within the education statistics. It is important to evaluate and interpret the statistics obtained regarding this subject. For this reason, the CDH survey was used in this research. The doctorate holders' survey conducted by the Turkish Statistical Institution in Turkey, which is an OECD country where the number of doctoral graduate students increases each year, is covered in the application part of this study. It is aimed to statistically evaluate the data obtained to be used in this study. In this study, CDH survey carried out by regional statistical offices in 2009 is used. According to this survey, the data sources, type of data and sampling method is determined as follows.

**The data sources:** Turkish citizens under the age of 70 residing in Turkey and foreign nationals were covered on the reference date. Sampling method was applied for the citizens of Republic of Turkey and full counting method was applied for foreign nationals.

**Type of data collected:** Regional statistical offices gathered data. In this survey, some of them preferred to conduct face-to-face interviews, while others used the mixed method (postal and face to face). Stratified random sampling method was used in this study. The total sample size consisted of 77424 doctorate holders. The number of recent doctorate recipients was 7438 from the total sample size [19]. In this study, frequency distribution was used to show how doctorate holders’ data was distributed, and correspondence analysis was used to reveal the relationship between the related variables for statistical evaluation by R package program.

### 3.1. Correspondence Analysis

Correspondence analysis (CA) is a statistical technique that provides a graphical representation of contingency tables. All the data must be measured on the same scale so that it makes sense to compute row sums and column sums in CA. This technique analyses the row profiles and/or the column profiles of the data matrix in Table 1. These are the rows of data divided by their respective row sums or the columns divided by their respective column sums [20]. Each row and column is weighted by its own mass. The sums of the rows and columns related to the grand total of the data.

$r_i = n_{i.} / n_{..}$ ,  $c_j$  : the mass of the  $j$ th column  $c_j = n_{.j} / n_{..}$ ,  $a_{ij}$ : the  $j$ th element of the profile of row  $a_{ij} = n_{ij} / n_{i.}$ ,  $b_{ij}$  : the  $i$ th element of the profile of column  $b_{ij} = n_{ij} / n_{.j}$ .

Table 1. Contingency Table 2x2

1. Variable (row)	2. Variable (column)						Total	
	1	2	3	..	j	..		c
1	$n_{11}$	$n_{12}$	$n_{13}$	..	$n_{1j}$	..	$n_{1c}$	$n_{1.}$
2	$n_{21}$	$n_{22}$	$n_{23}$	..	$n_{2j}$	..	$n_{2c}$	$n_{2.}$
:	:	:	:	..	:	..	:	:
i	$n_{i1}$	$n_{i2}$	$n_{i3}$	..	$n_{ij}$	..	$n_{ic}$	$n_{i.}$
:	:	:	:	..	:	..	:	:
r	$n_{r1}$	$n_{r2}$	$n_{r3}$	..	$n_{rj}$	..	$n_{rc}$	$n_{r.}$
<b>Total</b>	$n_{.1}$	$n_{.2}$	$n_{.3}$	..	$n_{.j}$	..	$n_{.c}$	$n_{..} = n$

Distances between row profiles or between column profiles are defined by the chi-square distance. The distance between the points of the data in the multi-dimensional space can be defined. In Euclidean space, the distance between two points is defined by Euclidean distance. The distance can be computed by a function known as  $\chi^2$  distance. They are expressed in relation to the row and column profiles.

$\chi^2$  distance between the  $i^{\text{th}}$  and  $i' ^{\text{th}}$  row profiles is shown by

$$d_c(i, i')^2 = \|a_i - a_{i'}\|_c^2 = \sum_{j=1}^c \frac{(n_{ij} - \frac{n_{i.}n_{.j}}{n})^2}{n_{.j}} \quad (1)$$

CA is the most versatile of the methods based on singular value decomposition for visualizing data. The total inertia of a contingency table is the  $\chi^2$  statistic divided by the total of the table. It can be written as

$$in(r) = in(c) = \sum_{i=1}^r \sum_{j=1}^c \frac{(P_{ij} - r_i c_j)^2}{r_i c_j} = C^2 / n \quad (2)$$

which is the sum of the squares of singular values [21].

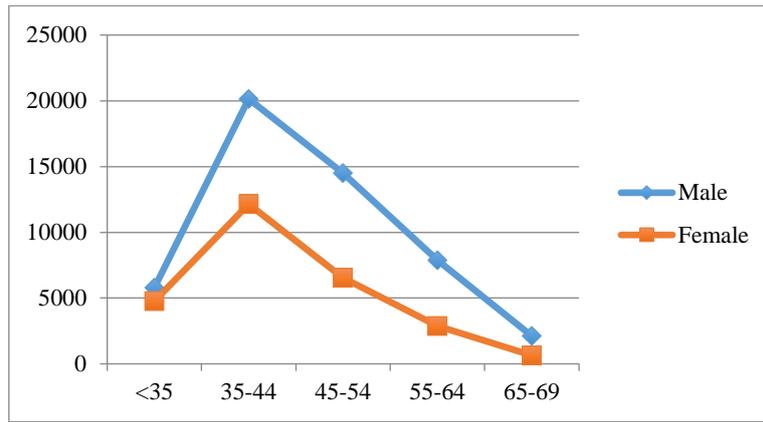
A graphical representation is provided for data analysis with CA. It is a generalization of a simple graphical concept with which we are all familiar, namely the scatterplot. The scatterplot is the representation of data as a set of points with respect to two perpendicular coordinate axes: The horizontal axis often referred to as the x-axis and the vertical one as the y-axis. When CA is used for data analysis, all data used should be nonnegative and on the same scale. Also this technique influences rows and columns equivalently. CA is generally used in contingency tables. Correspondence analysis decomposes the chi-squared statistics associated with this table into orthogonal factors. Because CA is a descriptive technique, it can be applied to tables whether or not the statistic is appropriate [22].

There are many variants of correspondence analysis such as multiple CA, Subset CA and canonical CA [19]. Demographic variables such as education level, age etc. can be used in CA as supplementary variables. Supplementary row or column points are mostly calculated as weighted averages of the column or row standard coordinates respectively in simple CA: for example, to position a supplementary column in principal coordinates, the profile of the supplementary column is used to calculate a weighted average of the row standard coordinates [23].

#### 4. RESULTS

CDH survey for Turkey was statistically analysed in this study. The statistical analysis of this study included gender, age, doctorate graduation year, the field of science, job qualification, employment status of doctorate holders and doctorate holders satisfaction level on their principal job variables.

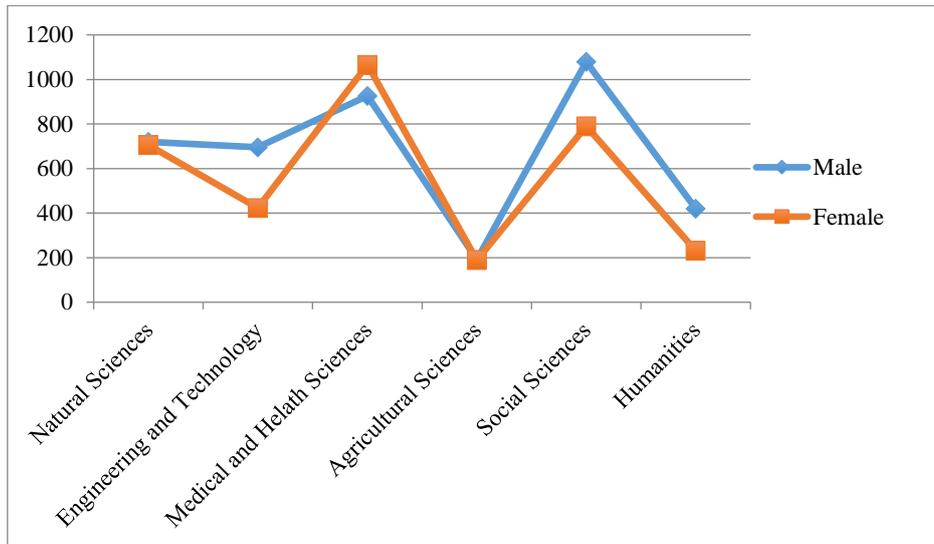
Firstly, the frequency distribution of CDH according to gender, age and the field of graduation can be seen. According to the results of this survey: One out of three doctoral holders in the survey was female. 65.1% of doctorate holders were male while 34.9% of those were female in Turkey. The gap between genders decreases as age group decreases in Figure 6. This figure shows that average age at graduation was higher for males, with 33.2, than it was for females, with 32.4.



Source: Doctorate holders Survey, 2009 by Turkish Statistical Institute.

**Figure 6.** Doctorate holders by age group and gender

In Figure 7, doctorate holders by the field of science group and gender are shown. It can be seen in Figure 7 that males prefer social sciences more than females. While the highest proportion of female doctorate holders was in medical and health sciences with 31%, the lowest proportion was in agricultural sciences with 6% on the field of scientific classification. While the highest proportion of male doctorate holders was in social sciences with 27%, the lowest proportion was in agricultural sciences with 5% on the field of scientific classification.



Source: Doctorate holders Survey, 2009 by Turkish Statistical Institute.

**Figure 7.** Doctorate holders by the field of science group and gender

Secondly, correspondence analysis was used to evaluate CDH. Therefore, some variables such as gender, graduation year, the field of science, job qualification, employment status and satisfaction level on the principal job were chosen. The relationship between variables was demonstrated by correspondence analysis. In this study, correspondence analysis was carried out under three headings;

- Gender-Job qualification
- Gender- Employment status and Employment by graduation years
- Gender- Satisfaction level on the principal job by selected title

#### 4.1. The Relationship Between Gender, Job Qualification and Field of Doctorate Degree

According to gender (Male- Female), the relationship between job qualification (Related- Partly related- Unrelated- Unknown) and field of doctorate degree (Natural Sciences, Engineering, and Technology, Medical and Health Sciences, Agricultural Sciences, Social Sciences and Humanities) variables was analysed by correspondence analysis.

This relationship was shown by a plot in Figure 8 and was found to be 93.67% in total. Doctorate holders who work in jobs related to their fields are females in humanities, males and females in engineering & technology, and males in natural sciences and social sciences. It is apparent in Figure 8 that doctorate holders who work in jobs related to their fields are males in humanities, females in social sciences and males in medical & health sciences. Male and female doctorate holders who work in jobs unrelated to their fields are in agricultural sciences.

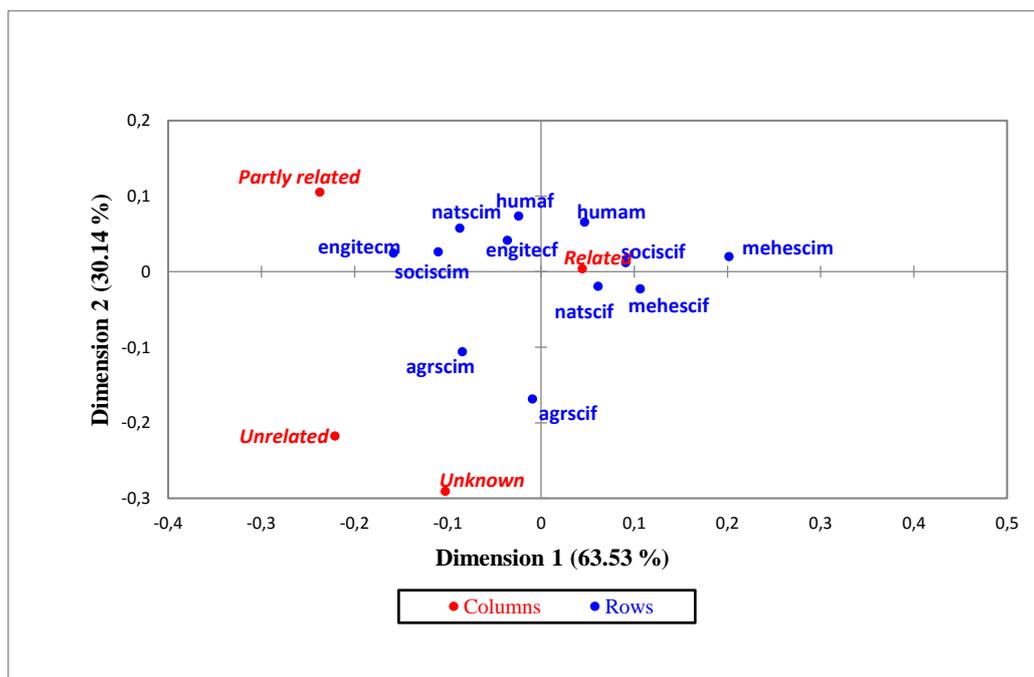


Figure 8. The relationship between job qualification and field of doctorate degree according to gender

#### 4.2. The Relationship Between Gender, Employment Status and Employment by Graduation Years

The relationship between gender (Male- Female), employment status (Employment - Unemployed – Inactive – Unspecified employment) and employment by graduation years (Before 1990, 1990-1999, 1999-2009) variables was analysed by correspondence analysis. This relationship was shown by a plot in Figure 9 and was found to be 99.65% in total. This plot shows a plot according to the employment status of doctorate holders: unemployed, inactive, unspecified employment status, employment by graduation year: before 1990, 1990-1999, 1999-2009 and male&female. We can see female and male doctorate holders in 1990-1999 and before 1990 have the same employment status.

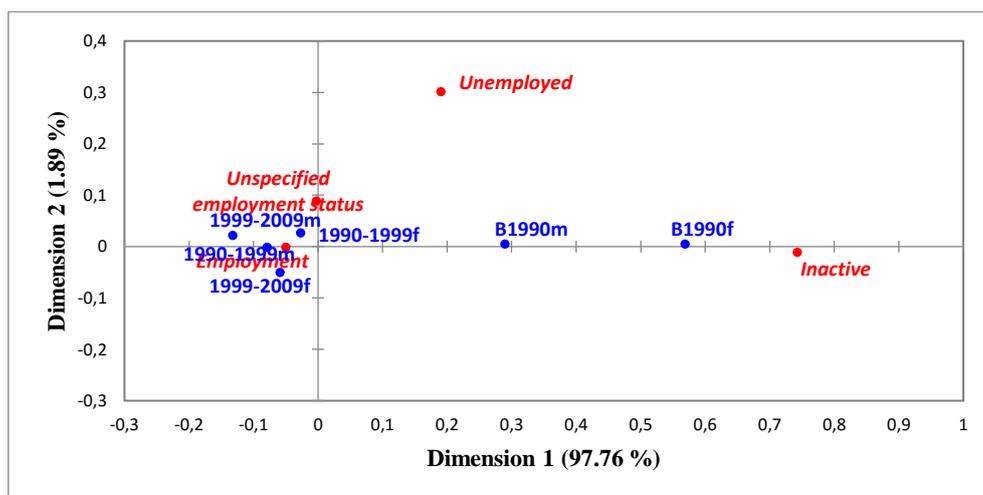


Figure 9. Employment status of doctorate holders by graduation year and gender

### 4.3. The Relationship Between Gender and Satisfaction Level on Principal Job by Selected Title

The relationship between gender, satisfaction level on principal job by selected title variables was analysed by correspondence analysis.

The data matrix in Table 2 is designed regarding job qualification: **very satisfied, somewhat satisfied, somewhat dissatisfied, very dissatisfied** and selected title: **salary, benefits, job security, location, working conditions, opportunities, intellectual challenge, level of responsibility, degree of independence, contribution to society and social status according to male and female**. New-old doctorate students are also shown in the data matrix. In Table 2, data was organized according to gender and principal job by selected title under the new-old doctorate students. New-old doctorate students were taken as the supplementary variable in correspondence analysis. The relationship between gender and satisfaction level on principal job by selected title was shown by a plot in Figure 10 and was found to be 99.65% in total.

Table 2. Data Matrix: Doctorate holders satisfaction level on their principal job by selected title according to gender

	Very satisfied (Vs)	Somewhat satisfied (Ssat)	Somewhat dissatisfied (Sdis)	Very dissatisfied (Vdis)
old students-salary-male (oS <sub>m</sub> )	4.4	39.8	37.9	17.9
old students-salary-female (oS <sub>f</sub> )	2.9	34	43.4	19.7
old students-benefits-male (oB <sub>m</sub> )	7.6	59	27.1	6.3
old students-benefits-female (oB <sub>f</sub> )	5.5	56	31.3	7.2
old students-job security-male (oJ <sub>S<sub>m</sub></sub> )	18.9	59.8	16.3	5
.	.	.	.	.
old students-degree of independence -female (nD <sub>I<sub>f</sub></sub> )	9.4	57.4	27.4	5.8
old students- contribution to society -male (nCS <sub>m</sub> )	18.7	58.9	18.5	3.9
old students-contribution to society-female (nCS <sub>f</sub> )	17.5	67.2	12.8	2.5
old students-social status -male (nSS <sub>m</sub> )	17	64.9	14.2	3.8

This plot shows a plot according to doctorate holders satisfaction level on their principal job by selected title. We can see similarities regarding job security, level of responsibility, contribution to society, location and social status both for old female and male doctorate holders. Recent female and male doctorate holders are similar in their level of responsibility, location and job security as is seen in Figure 10. It can be seen that old male doctorate holders and new female doctorate holders is same about salary at left part. The new male and female holders show no difference in opportunities and



1990-1999 and female doctors holders between 1999- 2009. Before 1990, the male and female doctorate holders were seen as inactive in the position. According to the latest correspondence analysis, it seemed that there was a high correlation between gender-satisfaction level on the principal job by selected title. Although all doctorate holders are very dissatisfied with salary, but all old doctorate holders are very satisfied with location, social status, and contribution to society.

It is observed that the number of doctorate is increasing by year, it is revealed that men are more likely to get a doctorate than women. The most preferred science for doing a doctorate is medical and health and social sciences. The least preferred science for doctorate is agricultural science. In the results obtained for the careers doctorate holders survey, gender was found to be related to job qualification, employment status, employment by graduation years, and satisfaction level on the principal job by selected title.

The statistical results of this study are expected to be useful in the future careers doctorate holder surveys. We can compare this survey with other survey data from European countries in future study. Thus, the difference between European countries and Turkey about careers of doctorate holders can be examined as a result.

## **REFERENCES**

- [1] Gokhberg L, Shmatko N, Auriol L. The Science and Technology Labor Force. Springer 2016.
- [2] Auriol L. Careers of Doctorate Holders, Employment and Mobility Pattern. OECD, Science, Technology and Industry. Working Paper. OECD Publishing 2010.
- [3] OECD. Science, Technology and Industry Scoreboard, Innovation and Growth in Knowledge Economies. OECD Publishing 2009.
- [4] Inanc H, Ozcan B. Gender, family and academic careers in Turkey, Advances in Life course Research 2016; 29, pp. 52-65.
- [5] Nerad M, Evans B. Globalization and Its Impacts on the Quality of PhD Education. Sense Publishers 2014.
- [6] OECD. Science, Technology and Industry Scoreboard, Innovation and Growth in Knowledge Economies. OECD Publishing 2011.
- [7] Turkish Statistical Institute. <http://tuikapp.tuik.gov.tr/nufusmenuapp/menu.zul>.
- [8] Auriol L, Misu M and Freeman RA. Careers of Doctorate Holders, Analysis of Labour Market and Mobility Indicators. OECD, Science, Technology and Industry. Working Paper. OECD Publishing 2013.
- [9] Gu J, Lin Y, Vogel D, Tian W. What are the major impact factors on research performance of young doctorate holders in science in China, High Education 2011; 62, pp. 483-502.
- [10] Wildy H, Peden S, Chan K. The rise of profession doctorates: case studies of the doctorate in Education in China, Iceland And Australia, Studies in Higher Education 2015; 40, 5, pp. 761-774.

- [11] Boosten K, Vandeveld K, Derycke H., Kaat A, Rossem RV. Careers of Doctorate Holders survey 2010, R&D innovation in Belgium, The Belgian Science Policy Office ( [www.belspo.be](http://www.belspo.be)) 2010.
- [12] Mertens A, Röbbken H. Does a doctoral degree pay off? An empirical analysis of rates of return of German doctorate holders, *High Education* 2013; 6, 2, pp. 217-231.
- [13] Borrego A, Barrios M, Villarroya A, Olle C. Scientific output and impact of postdoctoral scientists: a gender perspective. *Scientometrics* 2010; 83, pp. 93-101.
- [14] Maas B, Korvost M, Mooren FVD. Careers of Doctorate Holders in the Netherlands, Statistics Netherlands ([www.cbs.nl](http://www.cbs.nl)) 2014.
- [15] Bin A, Carneiro, AM, Salles-Filho S, Colugnati, FAB. Employment, research performance and decentralization: The experience and perspectives of doctorate holders in Brazil. *Science and Public Policy* 2015; 42, pp. 646-660.
- [16] Waaijer CJF, Macaluso B, Sugimoto CR, Lariviere V. Stability and longevity in the publication careers of U.S doctorate recipients. *Plos, One* 2016.
- [17] Council of Higher Education. [www.yok.gov.tr](http://www.yok.gov.tr).
- [18] Council of Higher Education. <https://istatistik.yok.gov.tr>.
- [19] Careers of Doctorate Holders Surveys. [www.tuik.gov.tr](http://www.tuik.gov.tr).
- [20] Beh EJ, Lombardo R. *Correspondence Analysis Theory, Practice and New Strategies*. Wiley 2014.
- [21] Greenacre, M. *Correspondence Analysis in Practice, Third Edition*. London: Chapman & Hall/CRC (2017).
- [22] Greenacre M. *Theory and Applications of Correspondence Analysis*. London: Academic Press 2013.
- [23] Nenadic O, Greenacre M. *Computation of Multiple Correspondence Analysis, with code in R*. Pompeu Fabra University, Working Paper 2005; No:887.