

Work Accidents and Occupational Diseases Cluster Analysis with respect to the Classification of the Provinces in Turkey

Türkiye'deki illerin Kümeleme Analizi İş Kazaları ve Meslek Hastalıkları Göre Sınıflandırılması

Gülcan GENCER¹, Kerem GENCER²

¹ Karamanoglu Mehmetbey University,, Karaman, TURKEY, ORCID: [0000-0002-3543-041X](https://orcid.org/0000-0002-3543-041X),

e posta: gulcangencer@kmu.edu.tr

² Karamanoglu Mehmetbey University, Department of Computer Programming, , Karaman, TURKEY,

ORCID: [0000-0002-2914-1056](https://orcid.org/0000-0002-2914-1056), e posta: keremgen@kmu.edu.tr

Geliş tarihi/Received: 01.04.2021

Kabul tarihi/Accepted: 04.10.2021

Yayın tarihi/Online published: 15.12.2021

DOI: 10.54862/pashid.907854

ABSTRACT

Clustering analysis is a multivariate statistical method used to determine whether a data set contains different groups, and if so, to identify these groups. The clusters that can be obtained as a result of cluster analysis are expected to be as similar within themselves and as different among themselves. Clustering analysis is a method used in almost all scientific fields. It is a method that is used more extensively in science fields such as Medicine, Biology, Psychology, Sociology, Archeology where uncertainty conditions and complex formations exist. In this study, using 19 variables related to occupational accidents and diseases, provincial groups showing the same structures in 81 provinces were tried to be determined. Using the 2019 data published by the

Social Security Institution (SSI), 81 provinces are divided into clusters according to occupational accidents and diseases. As a result of the analysis, the provinces with the best and worst conditions in terms of occupational accidents and diseases were revealed. It is aimed to clearly demonstrate the necessity of examining the source of the risk by determining the provinces with high risk in terms of work accidents and occupational diseases.

Keywords: Clustering, Work Accidents, Occupational Diseases, Turkey.

ÖZET

Kümeleme analizi, bir veri kümesinin farklı gruplar içerip içermediğini belirlemek ve varsa bu grupları tanımlamak için kullanılan çok değişkenli bir istatistiksel yöntemdir. Kümeleme analizi sonucunda elde edilebilecek kümelerin kendi içlerinde benzer (homojen) ve kendi aralarında farklı (heterojen) olması beklenmektedir. Kümeleme analizi, hemen hemen tüm bilimsel alanlarda kullanılan bir yöntemdir. Belirsizlik durumlarının ve karmaşık oluşumların olduğu Tıp, Biyoloji, Psikoloji, Sosyoloji, Arkeoloji gibi bilim alanlarında daha yaygın olarak kullanılan bir yöntemdir. Bu çalışmada iş kazaları ve meslek

hastalıklarına ilişkin 19 değişken kullanılarak 81 ilde aynı yapıları gösteren il grupları belirlenmeye çalışılmıştır. Sosyal Güvenlik Kurumu (SGK) tarafından yayınlanan 2019 verileri kullanılarak 81 il, iş kazaları ve meslek hastalıklarına göre kümelere ayrılmıştır. Yapılan analizler sonucunda iş kazası ve meslek hastalıkları açısından en iyi ve en kötü koşullara sahip iller ortaya çıkmıştır. İş kazaları ve meslek hastalıkları açısından yüksek riskli iller belirlenerek riskin kaynağının incelenmesi gerekliliğinin net bir şekilde ortaya konulması amaçlanmaktadır.

Anahtar Kelimeler: Kümeleme, İş Kazaları, Meslek Hastalıkları, Türkiye.

INTRODUCTION

Traditionally occupational health and safety; It is defined as systematic work carried out in order to provide a better working environment by keeping the workplaces away from the dangers caused by the execution of the business and purifying the conditions that may harm health accidents at work (Öztürk, D., 2018). Occupational health and safety is of great importance for a country like Turkey where occupational accidents and diseases are very high.

Because it is important to compensate for the damage that occurs in terms of people who suffer from work accidents or occupational diseases, and at this point, encounter work accident and occupational disease insurance and the benefits provided from this insurance branch. Work accident

and occupational disease insurance is one of the first insurance branches in the world. Creating a systematic way the social security system in Turkey has been realized since 1945 (Okur et.al, 2020).

Occupational accidents and occupational diseases are among the major social problems in all countries where the industrialization process continues (Okur et al., 2009). Turkey is one of the places where these accidents occurred as intense. The statistics on occupational accidents and diseases, which are among the statistics published every year by the SSI, reveal this situation. In Figure 1, Turkey's work accident and occupational diseases statistics are given by using the 2019 data published by the SSI (SGK, 2019).

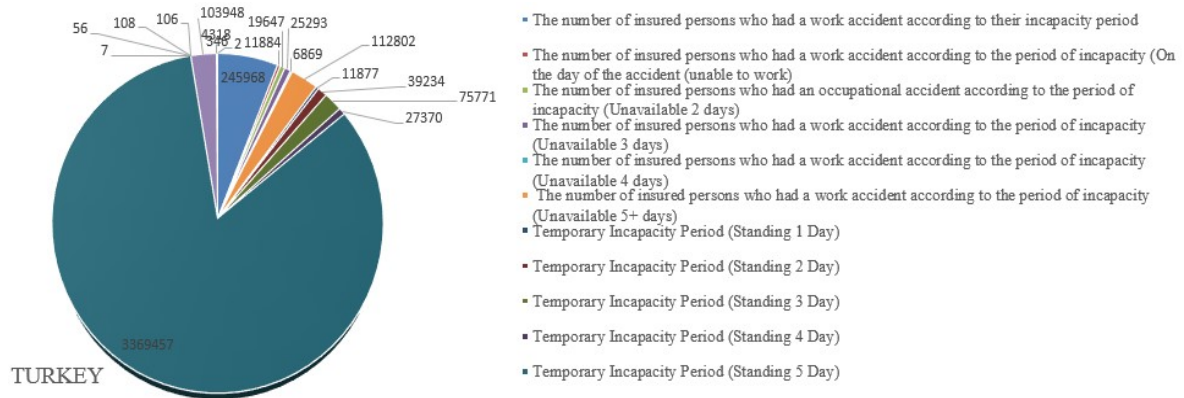


Figure 1. Statistics Of Occupational Accidents And Diseases Related To Turkey

When the results of work accidents or occupational diseases occur in Turkey, the result of this accident is often seen as a temporary incapacity period (5 days standing) In this study, the statistics for 2019 published by the SSI will be used. Considering that the number of occupational accidents and occupational diseases caused by our social security system is very high, the best and worst provinces groups will be determined by

grouping the provinces of Turkey according to occupational health and occupational diseases. It is aimed to clearly demonstrate the necessity of examining the source of the risk by determining provinces with high risk in terms of occupational accidents and diseases. The necessity of examining the source of the risk on a provincial or cluster basis will be clearly stated and it will be recommended to prepare a guide in this direction for the concerned.

LITERATURE REVIEW

In the literature, there are various studies on work accidents and occupational diseases recently. Ocal and Flowers (2017) evaluated through the concept of occupational accidents in work health in Turkey and safety practices, occupational accident data were compared with the European Union, and these assessments have been made in work accidents in Turkey in the light of the constituent problems.

Cerev and Yıldırım (2018) conducted a study on the effects of personal characteristics of employees on occupational accidents and diseases. In the study, it was observed that age, gender and education status are an important factor in occupational accidents and diseases, and suggestions were made to minimize occupational accidents and diseases.

Öztürk and Akın (2021) analyzed 2018 occupational accidents belonging to all sectors in Turkey. In addition, he evaluated the loss of life as a result of the accident in terms of the age of the victim. Caner (2021) researched how industry 4.0 approach and practices contribute to the prevention of accidents and diseases and to raise the health and safety standards of people in working life in the prevention of risks arising from hazards in occupational health and safety processes. Kıpçak et. al. (2021) revealed that testing the common factors affecting the occurrence of occupational accidents and occupational diseases according to two different businesses and comparing the effectiveness of International Organization for Standardization (ISO) and Occupational Health and Safety Assessment Series (OHSAS) applications used in only one enterprise in preventing the factors

that affect the occurrence of occupational accidents and occupational diseases. Bayyurt and Ekşi (2021) evaluated the impact of risk and safety culture on employee behavior in terms of occupational health and safety in the highway sector. Karaahmetoğlu (2021) although the situations considered as work accidents within the scope of article 13 of the SSI numbered 5510 are examined, a research has been carried out on which incidents can be accepted as work accidents or not. In this way, it aimed to outline the events that can be considered as work accidents. Umutlu and Karcıoğlu (2021) revealed that there is a relationship between occupational health and safety and job satisfaction and between the sub-dimensions of occupational health and safety and job satisfaction, and that occupational health and safety, occupational health and safety sub-dimensions and job satisfaction scores differ according to the demographic characteristics of the employees. Çımrın (2021) conducted a study on occupational disease as the reason for social visibility. Gül et al., (2021) has drawn a roadmap by making comparisons using data from other

countries in order to take necessary and up-to-date measures against occupational accidents and injuries. Durmuş (2020) made a study on why Covid-19 should be considered a work accident, focusing on the concepts of work accident and occupational disease. Göksu et al., (2020) made evaluations on the occupational accidents and occupational diseases experienced not only by the workers working in the production of wood raw materials, but also by the technical staff, trying to eliminate the lack of data, which is an important problem, and then examining the activities at the time of the accident and determining the measures to be taken has made suggestions for. Soykan et al., (2020) aimed to determine the attitudes and behaviors, safety awareness, work accident and occupational diseases experiences of aquaculture engineers diving with scuba diving in Turkey on occupational health and safety. Bekar et al., (2020) It aims to calculate the visible and invisible costs of occupational accidents and occupational diseases that occurred in Turkey between 2005 and 2014, and to reveal the factors affecting the costs.

MATERIAL AND METHOD

Model of the Research

In the study, the k-mean technique, one of the non-hierarchical clustering methods, was used to determine the distance matrix. In addition, in the k-mean technique, the significance levels of variables in various number of sets were analyzed by analysis of variance.

19 variables were considered in the study. Clustering analysis is an analysis that aims to bring similar units together. It is very difficult to give a single definition for cluster analysis due to the differences in usage areas and purposes. If a definition is given; Gan et al., (2007) define cluster analysis as methods of grouping units such that

similar units are in the same cluster and different units are in different clusters. Under these definitions, the most general definition of cluster analysis in the field of statistics is, "The $n \times p$ dimensional X data matrix, which is formed by measuring the units without a prior knowledge about the cluster structure, in terms of p variables, bringing together different units in different clusters and similar units in the same cluster. It is a multivariate statistical analysis consisting of methods that reveal hidden clusters and enable estimates about the population by determining the properties of these clusters" Gan et al., (2007). There are many algorithms used to combine similar

units in cluster analysis. Although these algorithms have different classifications, they are generally divided into hierarchical and nonhierarchical clustering methods according to the method they follow in bringing together similar units (Everitt et al., 2001). The material of this research is the most up-to-date 2019 occupational accident and occupational diseases among the statistics published by the SSI every year. The variables belonging to 81 used in this research are as follows. In the study, the k-mean technique, one of the non-hierarchical clustering methods, was used to determine the distance matrix. 19 variables were added to the study. These variables are explained below.

Inability to work: The period during which the insured is unable to work due to work accident.

Occupational Disease: It is the temporary or permanent illness, physical or mental disability that the insured suffers due to a recurring reason or the conditions of the work execution.

Outpatient Treatment: Providing healthcare services without hospitalization of patients in health institutions and organizations or where they are located is considered as outpatient treatment.

Income: Refers to the continuous payment made to the insured in case of work accident or occupational disease or to the right holders in the event of death of the insured.

Sickness: Diseases that cause incapacity of the insured, excluding work accident and occupational disease.

Permanent Incapacity: It refers to the fact that the earning power of the health service providers authorized by the Institution due to the illness and disability caused by the occupational accident has decreased by at least 10% by the Health Board of the Institution

Inpatient Treatment: It is the treatment applied in health institutions from the date of hospitalization to discharge (Law No. 5510).

X_1 : The number of insured persons who had a work accident according to their incapacity period

X_2 : The number of insured persons who had a work accident according to the period of incapacity (On the day of the accident (unable to work))

X_3 : The number of insured persons who had an occupational accident according to the period of incapacity (Unavailable 2 days)

X_4 : The number of insured persons who had a work accident according to the period of incapacity (Unavailable 3 days)

X_5 : The number of insured persons who had a work accident according to the period of incapacity (Unavailable 4 days)

X_6 : The number of insured persons who had a work accident according to the period of incapacity (Unavailable 5+ days)

X_7 : Temporary Incapacity Period (Standing 1 Day)

X_8 : Temporary Incapacity Period (Standing 2 Day)

X_9 : Temporary Incapacity Period (Standing 3 Day)

X_{10} : Temporary Incapacity Period (Standing 4 Day)

X_{11} : Temporary Incapacity Period (Standing 5 Day)

X_{12} : Temporary Incapacity Period (1 Day in Bed)

X_{13} : Temporary Incapacity Period (2 Day in Bed)

X_{14} : Temporary Incapacity Period (3 Day in Bed)

X_{15} : Temporary Incapacity Period (4 Day in Bed)

X_{16} : Temporary Incapacity Period (5 Day in Bed)

X_{17} : Among the insured who have had a work accident or occupational disease in the past years, who have earned income in 2019 Number of occupational accidents

X_{18} : The number of occupational diseases among the insured who had an occupational accident or

occupational disease in the previous years, whose income in 2019

X_{19} : The number of insured persons who died as a result of a work accident

Two factors should be considered in order to reach the most accurate results in clustering study. One of them is to choose the correct variable, the other is to determine the number of sets appropriately. In this study, two approaches were used to determine the number of clusters. One is principal component analysis and the other is the following approach given by (Tatlıdil, 1996).

$$k \cong \sqrt{\frac{n}{2}}$$

As a result, the number of clusters was selected as 7 by working according to both techniques.

Place and Time of Research

In this study, it has been tried to determine the provincial groups that 81 provinces have been clustered according to occupational accidents and diseases by using the data of 2019, the year the SSI published it.

Data Collection Tools

2019 data published by SSI was used.

Evaluation of Data

The data were analyzed with the SPSS 26.0 package program. Cluster analysis was used in the statistical evaluation.

Ethical Aspect of Research

There is no need for an ethics committee for this research. Research and Publication Ethics has been taken into account in the article.

RESULTS

The classification of provinces according to occupational accidents and diseases was made by cluster analysis, which is one of the multivariate statistical analysis techniques. According to the non-hierarchical k-means technique, the results in the same cluster, Antalya, Manisa and Tekirdağ are in the same cluster. Afyonkarahisar, Bilecik, Bolu, Çanakkale, Diyarbakır, Hatay, Adana, Aydın, Balıkesir, Bilecik, Eskişehir,

Table 1 were reached and the provinces were examined in 7 clusters. According to these results, it is seen that the leading industrial cities such as Kocaeli, Istanbul, Izmir and Manisa are located in a separate cluster alone, while Ankara and Bursa are Kırklareli, Kütahya, Kahramanmaraş, Samsun, Kayseri, Muğla, Sakarya, Konya, Zonguldak are in the

Table 1. Classification Of Provinces By Clusters With The Non-Hierarchical K-Means Technique

Cluster	Provinces
1	Kocaeli Afyonkarahisar, Bolu, Çanakkale, Diyarbakır, Hatay, Kırklareli, Kütahya, Kahramanmaraş, Samsun, Trabzon, Uşak, Karaman, Yalova, Karabük, Osmaniye, Düzce, Adıyaman, Ağrı, Amasya,
2	Artvin, Bingöl, Bitlis, Burdur, Çankırı, Çorum, Edirne, Elazığ, Erzincan, Erzurum, Giresun, Gümüşhane, Hakkari, Isparta, Kars, Kastamonu, Kırşehir, Malatya, Mardin, Muş, Nevşehir, Niğde, Ordu, Rize, Siirt, Sinop, Sivas, Tokat, Tunceli, Şanlıurfa, Van, Yozgat, Aksaray, Bayburt, Kırıkkale, Batman, Şırnak, Bartın, Ardahan, Iğdır, Kilis
3	İzmir
4	Antalya, Tekirdağ, Manisa
5	İstanbul
6	Ankara, Bursa
7	Adana, Aydın, Balıkesir, Bilecik, Eskişehir, Denizli, Mersin, Gaziantep, Kayseri, Muğla, Sakarya, Konya, Zonguldak

Denizli, Mersin, Gaziantep, same cluster, Trabzon, Uşak. It was observed that 60 provinces including Karaman, Yalova, Karabük, Osmaniye and Düzce

were gathered in a cluster with similar characteristic. The provinces most and least similar to each other are given in Tables 2 and 3.

Table 2. Provinces That Are Most Similar To Each Other

Cluster Combined		Coefficients
Cluster 1	Cluster 2	
Tunceli	Bayburt	1993
Hakkari	Sinop	4804
Adıyaman	Van	8949
Kars	Iğdır	9033
Tunceli	Ardahan	10828,5
Gümüşhane	Hakkari	11988

Table 3. Provinces That Are Least Similar To Each Other

Cluster Combined		Coefficients
Cluster 1	Cluster 2	
İzmir	Kocaeli	1819588957
Adana	Adıyaman	2835853368
Ankara	Antalya	7011808397
Adana	Ankara	21507584666
Adana	İzmir	75456500386
Adana	İstanbul	3,78295E+11

As seen as Table 4, when the k-means technique is applied, according to the results obtained, the significance (p) values of all variables are significant since the significance level is $p < 0.05$. The most influential variables are the number of insured persons who had an occupational accident

(accident day (working)) according to the Temporary Incapacity Period (Standing 5 Days) and the duration of incapacity, respectively.

Table 4. The Most Influential Variables According To The K-Means Technique

Variables	Cluster Mean Square	df	Error Mean Square	df	F	Sig.
X ₁	641951980,469	9	1570275,189	71	408,815	0,000
X ₂	1333717,195	9	3479,996	71	383,253	0,000
X ₃	3868762,943	9	7249,796	71	533,637	0,000
X ₄	5559117,991	9	9638,586	71	576,757	0,000
X ₅	406023,782	9	554,603	71	732,098	0,000
X ₆	83735248,009	9	55073,705	71	1520,422	0,000
X ₇	1332048,824	9	3486,331	71	382,078	0,000
X ₈	15458502,555	9	28858,280	71	535,670	0,000
X ₉	49901949,483	9	86337,784	71	577,985	0,000
X ₁₀	6470674,355	9	8830,141	71	732,794	0,000
X ₁₁	70685551505,839	9	24455244,789	71	2890,405	0,000
X ₁₂	0,693	9	0,030	71	22,813	0,000
X ₁₃	18,848	9	1,685	71	11,185	0,000
X ₁₄	155,767	9	2,283	71	68,228	0,000
X ₁₅	38,224	9	4,328	71	8,832	0,000
X ₁₆	44640330,085	9	216623,021	71	206,074	0,000
X ₁₇	61298,810	9	561,972	71	109,078	0,000
X ₁₈	606,659	9	30,875	71	19,649	0,000
X ₁₉	0,106	9	0,014	71	7,499	0,000

DISCUSSION

In the research, 81 provinces were classified according to the year 2019 work accident and occupational diseases. The data of 2019 published by SSI was used. 81 provinces are grouped according to occupational accidents and diseases using 19 variables. As a result of the cluster analysis, the provinces with the best and worst

conditions in terms of occupational accidents and diseases were revealed. The provinces with the highest risk and the provinces with the lowest risk in terms of occupational accidents and diseases were determined and the necessity of examining the source of the risk on a provincial basis was clearly demonstrated. With the re-evaluation of the

following years, the necessity of creating a guide on work accidents and occupational diseases on a

cluster basis will be fully revealed.

CONCLUSION

In this study, non-staged clustering techniques was used. According to the findings obtained, it was observed that metropolitan provinces such as Kocaeli, Istanbul, Izmir and Manisa formed clusters individually in k-means clustering technique. With non-hierarchical kluster analysis method with uncovered this cluster structures, in the provinces of development that developed in Turkey, business as accidents and occupational diseases to be sufficient measures taken on and it was determined that further development required.

On the other hand, another method that needs to be done to reduce occupational accidents and

diseases is to increase the number of inspections. Especially in provinces such as Ankara, Istanbul, Izmir, Kocaeli, Bursa and Antalya, which are located in separate clusters, with an effective inspection, the number of labor inspectors and the frequency of inspections can be increased and a healthy and safe work environment can be achieved. In addition, a separate inspection and precaution guideline can be issued for each group divided into clusters by the authorities.

REFERENCES

Arslan Durmuş, S. (2020). COVID-19'un iş kazası ve meslek hastalığı bakımından değerlendirilmesi. İstanbul Hukuk Mecmuası, 78 (2), 363-393.

Bayyurt, D. and Ekşi, A. (2021). Karayolları sektörü iş sağlığı ve güvenliği uygulamalarında risk ve güvenlik kültürü etkisinin değerlendirilmesi. Hastane Öncesi Dergisi, 6 (1), 83-98.

Bekar, İ., Oruç, D. and Bekar, E. (2017). İş kazası ve meslek hastalıklarının maliyeti (2005-2014). Uluslararası Ekonomik Araştırmalar Dergisi, 3 (3), 479-489.

Caner, V. (2021). Fiziksel risk etmenleri maruziyetine bağlı iş kazası ve meslek hastalıklarının önlenmesinde endüstri 4.0 yaklaşımının değerlendirilmesi. Ohs Academy, 4(1), 55-61.

Cerev, G. and Yıldırım S. (2018). Çalışanların kişisel özelliklerinin iş kazası ve meslek hastalıklarına etkisi üzerine bir inceleme. Fırat Üniversitesi İİBF Uluslararası İktisadi ve İdari Bilimler Dergisi, 2(1), 53-71.

Çımrın, A.H. (2021). Sosyal görünürlüğü'nün sebebi olarak meslek hastalığı. Olgu Sunumu, 819-829.

Everitt, B.S., Landau, S. and Leese, M. (2001). Cluster Analysis. Oxford University Press Inc., New York, 237p.

Gan, G., Ma, C. and Wu J., (2007). Data clustering: theory, algorithms, and applications. Society For Industrial & Applied Mathematics, U.S. 466p

Göksu, E., Başar, H. and Erol, M. (2020). Odun hammaddesi üretiminde çalışanların geçirdikleri iş

kazası ve meslek hastalıkları yönünden değerlendirilmesi. *Ormancılık Araştırma Dergisi*, 7(2), 120-130.

Gül, A.F., Özçoban, N.O. and Aydoğan, F. (2021). Türkiye'deki iş kazaları ile ilgili verilerin karşılaştırması, varolan sorunlar ve çözüm önerileri. *Sağlık Okuryazarlığı Dergisi*, 2(1), 50-63.

Gülşahin, A., Cerim, H. and Soykan, O. (2020). Su ürünleri mühendisliği'nde donanımlı dalışın iş sağlığı ve güvenliği açısından değerlendirilmesi. *Düzce Üniversitesi Bilim ve Teknoloji Dergisi*, 8(1), 94-101.

Umutlu, S. and Karcıoğlu, F. (2021). İş sağlığı ve güvenliğinin iş tatmini üzerine etkisi: çimento sektöründe bir uygulama. *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 26(1), 65-81.

Karaahmetoğlu, A. (2021). Sosyal güvenlik hukuku açısından iş kazası sayılan haller. *İş ve Hayat*, 67-88.

Kıpçak, E., Çalık, A., Mehri, H. (2021). İş kazaları ve meslek hastalıklarının oluşumuna etki eden faktörler: iso ve ohsas uygulamalarının etkileri. *Pamukkale Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, (43), 1-26.

Okur, A.R., Güzel, A. and Caniklioğlu, N. (2020). *Sosyal Güvenlik Hukuku*. B.12, Beta Yayınları, İstanbul.

Öçal, M. And Çiçek, Ö. (2017). Türkiye ve avrupa birliği'nde iş kazası verilerinin karşılaştırmalı analizi. *Hak İş Uluslararası Emek ve Toplum Dergisi*, 6 (16), 616-637.

Öztürk, D. (2018). Türkiye'de ve dünya'da iş sağlığı ve güvenliği, <http://www.bilgiosgb.com/haber-detay/turkiye-de-ve-dunya-da-is-sagligi-ve-guvenligi.html>, [Accessed date: 01.03.2021]

Özturk, T. and Caner Akın, G. (2021). Türkiye'de tüm sektörlere ait 2018 yılı iş kazalarının ve kaza sonucu yaşam kaybının kazazede yaş bakımından değerlendirilmesi. *Avrupa Bilim ve Teknoloji Dergisi*, Ejosat 2021 Ocak, 410-415.

SGK: İş kazası ve meslek hastalıkları ile ilgili SGK İstatistikleri, (2019). http://www.sgk.gov.tr/wps/portal/sgk/tr/kurumsal/istatistik/sgk_istatistik_yilliklari, [Accessed date: 01.03.2021].

Tatlıdil, H., *Uygulamalı Çok Değişkenli İstatistiksel Analiz*, Ankara, 1996.

5510 Sayı ve 31.05.2016 Tarihli Sosyal Sigortalar ve Genel Sağlık Sigortası Kanun, <http://www.mevzuat.gov.tr/MevzuatMetin/1.5.5510.doc>, [Accessed date: 20.03.2016].