

T.C.
BANDIRMA ONYEDİ EYLÜL ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

**BİR OTOMOTİV ANA SANAYİ İŞLETMESİNDE KALİTEYİ ETKİLEYEN
FAKTÖRLERİN VERİ MADENCİLİĞİ İLE İNCELENMESİ**

Yücel KURTULUŞ

Akıllı Ulaşım Sistemleri ve Teknolojileri Anabilim Dalı

Akıllı Ulaşım Sistemleri ve Teknolojileri Programı

TEMMUZ 2022

T.C.
BANDIRMA ONYEDİ EYLÜL ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

**BİR OTOMOTİV ANA SANAYİ İŞLETMESİNDE KALİTEYİ ETKİLEYEN
FAKTÖRLERİN VERİ MADENCİLİĞİ İLE İNCELENMESİ**

Yücel KURTULUŞ

DANIŞMAN

Prof. Dr. Mehmet TEKTAŞ

II. DANIŞMAN

Doç. Dr. Hasan ŞAHİN

Akıllı Ulaşım Sistemleri ve Teknolojileri Anabilim Dalı

Akıllı Ulaşım Sistemleri ve Teknolojileri Programı

TEMMUZ 2022

ONAY

Yücel KURTULUŞ tarafından hazırlanan “**Bir Otomotiv Ana Sanayi İşletmesinde Kaliteyi Etkileyen Faktörlerin Veri Madenciliği ile İncelenmesi**” adlı tez çalışması .../.../... tarihinde yapılan sınavla aşağıdaki jüri tarafından oybirliği/oyçokluğu ile Bandırma Onyedli Eylül Üniversitesi Fen Bilimleri Enstitüsü Akıllı Ulaşım Sistemleri ve Teknolojileri Anabilim Dalı’nda YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Jüri Üyeleri

İmza

Ünvanı, Adı ve Soyadı
(Danışman)

Ünvanı, Adı ve Soyadı
(Üye)

Ünvanı, Adı ve Soyadı
(Üye)

Ünvanı, Adı ve Soyadı
(Üye) (varsa)

Ünvanı, Adı ve Soyadı
(Üye) (varsa)

Bu tezin kabulü, Fen Bilimleri Enstitüsü Yönetim Kurulu’nun .../.../...gün ve .../... sayılı kararı ile onaylanmıştır.

Doç. Dr. Abdullah YEŞİL

Enstitü Müdürü

BANDIRMA ONYEDİ EYLÜL ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ
YÜKSEK LİSANS / DOKTORA TEZ ÇALIŞMASI
ETİK BEYANI

Bandırma Onyedi Eylül Üniversitesi Fen Bilimleri Enstitüsü tez yazım kılavuzuna göre hazırlamış olduğum “**Bir Otomotiv Ana Sanayi İşletmesinde Kaliteyi Etkileyen Faktörlerin Veri Madenciliği ile İncelenmesi**” adlı tezimin özgün bir çalışma olduğunu, tez hazırlanırken tüm aşamalarda bilimsel etik ilkelerine uygun davrandığımı, tez kapsamında sunulan tüm verileri bilimsel etik ilkelerine uygun elde ettiğimi, tezde faydalandığım tüm eserlere atıf yaptığımı ve kaynaklar kısmında bu eserleri gösterdiğimi beyan ederim. / / 20...

Öğrencinin Adı Soyadı

Yücel KURTULUŞ

İmza

ÖNSÖZ

Küreselleşmeyle birlikte, müşteri beklentileri değişmiş ve bu durum işletmeleri kaliteyi arttırmaya ve oluşabilecek hataları önlemeye yönelik araştırma geliştirme faaliyetleri yapmaya sevk etmiştir. Ayrıca birçok işletme, rekabetçi piyasa koşullarında, kalitenin arttırılmasının yanı sıra maliyetlerin azaltılması ve verimliliğin yükseltilmesine yönelik değişik sistemler uygulamaktadır. Bu sistemlerden biri de özellikle kalitenin geliştirilmesi konusunda son yıllarda önemi gittikçe artan ve geleneksel yöntemlere kıyasla bu alanda verimliliğin arttırılması ve maliyetlerin azaltılması konusunda önemli katkıları bulunan veri madenciliğidir. Veri madenciliği birçok alanda uygulanmasına karşın otomotiv sektöründe bu konuda çok fazla çalışmanın yapılmamış olması konun önemini bir kez daha gözler önüne sermektedir.

Bu çalışmam esnasında, tez konumun belirlenmesinden başlayarak en son aşamasına kadar her konuda benden yardımlarını esirgemeyen danışman hocam Sayın Prof. Dr. Mehmet TEKTAŞ ve ikinci danışmanım Sayın Doç. Dr. Hasan ŞAHİN'e teşekkürlerimi sunarım.

Bu çalışmayı hazırladığım günlerde, yoğun çalışma temposuna rağmen, benden maddi ve manevi desteğini esirgemeyen eşim Filiz'e, teşekkür ederim.

ÖZET

YÜKSEK LİSANS TEZİ

BİR OTOMOTİV ANA SANAYİ İŞLETMESİNDE KALİTEYİ ETKİLEYEN FAKTÖRLERİN VERİ MADENCİLİĞİ İLE İNCELENMESİ

Yücel KURTULUŞ

Bandırma Onyediy Eylül Üniversitesi

Fen Bilimleri Enstitüsü

Akıllı Ulaşım Sistemleri ve Teknolojileri Anabilim Dalı

Danışman: Prof. Dr. Mehmet TEKTAŞ

İkinci Danışman: Doç. Dr. Hasan ŞAHİN

Temmuz 2022, 77 sayfa

Bu çalışmada, bir otomotiv ana sanayi işletmesinde zamana bağlı olarak oluşan değişimler sonucu meydana gelen hatalar analiz edilmiştir. Son muayene ve deney kontrollerinde tespit edilen hatalar incelenerek, sorumlularının tespiti, hata oluşumlarının nedenleri, hata oluşumuna sebep olan dış etkenlerin sonuçlarının analiz edilmesi tezin temel amacını oluşturmaktadır. Analizler yapılırken veri madenciliği teknikleri kullanılmıştır. Yapılan araştırmalar veri madenciliğinin kalite kontrol konusunda önemli yararlar ortaya koyduğunu göstermektedir. Veri madenciliği teknikleri içerisinde birçok algoritma kullanılmakta ve bunların içerisinde Apriori Algoritmasının birliktelik analizi konusunda ön plana çıktığı görülmektedir. Bu çalışmada, hataların oluşumuna etki eden faktörlerin bir birliktelik halinde hatayı oluşturup oluşturmadığının tespiti amaçlandığından, kullanılacak veri madenciliği tekniği olarak Apriori Algoritması seçilmiştir. Bu çalışmada, bir otomotiv ana sanayi işletmesinde yapılması planlanan analizler sonucunda elde edilecek verilerin hata oluşumunun engellenmesinde ve oluşmuş hataların da ürün müşteriye sunulmadan tespit edilmesine katkıda bulunacağı düşünülmektedir.

Anahtar kelimeler: Otomotiv, Kalite, Veri Madenciliği, Apriori.

ABSTRACT

M.Sc. THESIS

INVESTIGATION OF FACTORS AFFECTING QUALITY IN AN AUTOMOTIVE MAIN INDUSTRY BUSINESS WITH DATA MINING

Yücel KURTULUŞ

**Bandırma Onyedi Eylül University
Graduate School of Natural and Applied Sciences
Department of Intelligent Transportation Systems and Technologies**

**Supervisor: Prof. Dr. Mehmet TEKTAŞ
Co-Supervisor: Assoc. Prof. Dr. Hasan ŞAHİN**

July , 2022, 77 pages

In this study, the errors that occur as a result of time-dependent changes in an automotive industry enterprise will be analyzed. The main subject of the thesis will be to examine these errors detected in the final examination and test controls, to determine the responsible ones, to analyze the results of the external factors that cause error occurrence. Data mining techniques will be used while analyzing. Studies have shown that data mining has significant benefits in quality control. Many algorithms are used in data mining techniques and it is seen that Apriori Algorithm comes to the fore in association analysis. In the thesis, the data mining technique to be used is the Apriori Algorithm, since it is aimed to determine whether the factors that affect the occurrence of errors together create the error or not. In this thesis, it is thought that the data obtained as a result of the analysis planned to be carried out in an automotive main industry enterprise will contribute to preventing the occurrence of errors and detecting the errors before the product is presented to the customer.

Keywords: Automotive, Quality, Data Mining, Apriori.

İÇİNDEKİLER

	Sayfa
ÖZET	I
ABSTRACT	II
İÇİNDEKİLER.....	III
ŞEKİL LİSTESİ	V
TABLO LİSTESİ.....	VI
1. GİRİŞ.....	1
1.1 VERİ MADENCİLİĞİNİN TARİHİ GELİŞİMİ.....	2
1.2 LİTERATÜR ARAŞTIRMASI.....	3
2. MATERYAL VE YÖNTEM	19
2.1 MATERYAL.....	19
2.1.1 Uygulama Sahası Hakkında Genel Bilgi	19
2.2 YÖNTEM	22
2.2.1 İşi Anlamak	23
2.2.2 Veriyi Anlamak.....	23
2.2.3 Veriyi Hazırlamak.....	24
2.2.3.1 Gürültülü Veri	25
2.2.4 Modelleme	27
2.2.4.1 Birliktelik Kuralları ve Apriori Algoritması	30
2.2.4.1.1 Birliktelik Kuralı	30
2.2.4.1.2 Apriori Algoritması	34
2.2.4.2 Pareto Analizi.....	41
2.2.4.2.1 Pareto Diyagramı Kullanım Alanları	42
2.2.4.2.2 Pareto Diyagramı Çizilmesi.....	43

2.2.4.3 Apriori Algoritması Adımlarının Veri Setine Uygulanması	46
2.2.5 Değerlendirme.....	49
2.2.6 Yayma.....	49
3. BULGULAR.....	50
3.1 VERİ SETİNİN EXCEL TABLOLARI İLE YORUMLANMASI.....	50
3.2 VERİ SETİNİN APRIORİ ALGORİTMASI İLE YORUMLANMASI.....	62
3.2.1. Web Grafiği.....	66
4. TARTIŞMA VE SONUÇ	68
5. KAYNAKLAR	70
ÖZGEÇMİŞ	77

ŞEKİL LİSTESİ

<u>Sekil</u>	<u>Sayfa</u>
Şekil 2-1: Veri toplamada izlenen iş akışı.....	21
Şekil 2-2: The Cross-Industry Standard Proces for Data Mining süreç şeması	23
Şekil 2-3: Microsoft Excel formatında düzenlenen veri seti alıntısı	25
Şekil 2-4: Veri toplama ve işleme aşamaları	38
Şekil 2-5: Hata Türleri	40
Şekil 2-6: Hata sayıları için Pareto diagramı.....	45
Şekil 2-7: Microsoft Excel formatında düzenlenen arındırılmış veri seti alıntısı.....	46
Şekil 2-8: Clementine programı açılış arayüzü görüntüsü	47
Şekil 2-9: Clementine programı açılış arayüzü görüntüsü	48
Şekil 2-10: Clementine programı veri tipi belirleme ekranı alıntısı	48
Şekil 3-1: Tespit edilen kusurların dağılımı.....	50
Şekil 3-2: 2010 yılı hata sayıları	55
Şekil 3-3: 2011 yılı hata sayıları	55
Şekil 3-4: 2012 yılı hata sayıları	56
Şekil 3-5: 2013 yılı hata sayıları	57
Şekil 3-6: 2014 yılı hata sayıları	57
Şekil 3-7: 2015 yılı hata sayıları	58
Şekil 3-8: 2016 yılı hata sayıları	59
Şekil 3-9: 2017 yılı hata sayıları	59
Şekil 3-10: Hata sayılarının yıllara göre dağılımı.....	60
Şekil 3-11: Yıllara göre üretilen ve kontrol edilen araç sayıları ile bulunan hata sayılarının gösterimi.....	61
Şekil 3-12: Web Grafiği.....	67

TABLO LİSTESİ

<u>Tablo</u>	<u>Sayfa</u>
Tablo 1-1: Veri Madenciliğinin Tarihçesi	2
Tablo 2-1: Veri madenciliği için kullanılan araç çeşitleri ve örnekleri.	26
Tablo 2-2: Verilerin büyüktten küçüğe doğru kümülatif dağılımları	44
Tablo 3-1: Yıllara göre toplam hata adetleri	51
Tablo 3-2: Montaj Hattı yıllara göre hata adetleri.....	52
Tablo 3-3: Boya Hattı yıllara göre hata adetleri.....	52
Tablo 3-4: Parça Hattı yıllara göre hata adetleri	53
Tablo 3-5: Gövde Hattı yıllara göre hata adetleri	53
Tablo 3-6: Plastik Hattı yıllara göre hata adetleri.....	54
Tablo 3-7: Diğer Hataların yıllara göre hata adetleri	54
Tablo 3-8: Test edilecek Güven ve Destek değerleri	62
Tablo 3-9: Veri setinde %50 minimum Güven ve %5 minimum Destek değerleri için bulunan kurallar	63
Tablo 3-10: Veri setinde %50 minimum Güven ve %10 minimum Destek değerleri için bulunan kurallar	64
Tablo 3-11: Veri setinde %59 minimum Güven ve %5 minimum Destek değerleri için bulunan kurallar	65
Tablo 3-12: Veri setinde %59 minimum Güven ve %10 minimum Destek değerleri için bulunan kurallar	66

1. GİRİŞ

İşletmeler; müşteri memnuniyetini sağlamak, pazar payını artırmak, daha kaliteli ürünler üretmek ve üretilen bu kaliteli ürünlerin devamlılığını sağlamak zorundadır. Küreselleşmeyle birlikte, müşteri beklentileri değişmiş, bu durum işletmeleri kaliteyi arttırmaya ve oluşabilecek hataları önlemeye yönelik araştırma geliştirme faaliyetleri yapmaya sevk etmiştir. Ayrıca birçok işletme, rekabetçi piyasa koşullarında, kalitenin arttırılmasının yanı sıra maliyetlerin azaltılmasıyla birlikte verimliliğinde yükseltilmesine yönelik değişik sistemler uygulamaktadır. Bu sistemlerden biri de özellikle kalitenin geliştirilmesi konusunda son yıllarda gittikçe artan önemi ve geleneksel yöntemlere kıyasla bu alanda verimliliğin arttırılması, maliyetlerin azaltılması konusunda önemli katkıları bulunan veri madenciliğidir. Veri madenciliği birçok alanda uygulanmasına karşın otomotiv sektöründe bu konuda çok fazla çalışmanın yapılmamış olması konunun önemini bir kez daha gözler önüne sermektedir.

Günümüz teknolojisinin gelişmesiyle birlikte ortaya çıkan Akıllı Ulaşım Sistemleri, ülkemiz açısından da karşı konulamaz bir ihtiyaçtır. AUS ile büyük veri kavramlarının entegrasyonundan elde edilecek verilerin analiz sonuçlarının işlenerek hayata geçirilmesiyle insan karar mekanizması en aza indirilerek hataların önlenmesi dolayısıyla daha hızlı ve güvenli bir ulaşım organizasyonunun sağlanmasına yardımcı olacaktır. Bunun olabilmesi için elde edilen trafik bilgilerinin yanı sıra bunları değerlendirerek en anlamlı veriyi çıkartacak sistemlerde gereklidir. Bu sistemler ise veri madenciliği sayesinde olacaktır [82].

Bu çalışmada, bir otomotiv ana sanayi işletmesinde zamana bağlı olarak oluşan değişimler sonucu meydana gelen hatalar analiz edilecektir. Son muayene ve deney kontrollerinde tespit edilen bu hataların incelenerek, sorumlularının tespiti, hata oluşumlarının nedenleri, hata oluşumuna sebep olan dış etkenlerin sonuçlarının analiz edilmesi tezin temel konusudur. Analizler yapılırken veri madenciliği tekniklerinden yararlanılacaktır. Yapılan araştırmalar veri madenciliğinin kalite kontrol konusunda önemli yararlar ortaya koyduğunu göstermiştir. Veri madenciliği teknikleri içerisinde birçok algoritma kullanılmaktadır ve bunların içerisinde Apriori Algoritmasının birliktelik analizi konusunda ön plana çıktığı görülmektedir. Çalışmada, hataların oluşumuna etki eden faktörlerin bir birliktelik halinde hatayı oluşturup oluşturmadığının tespiti amaçlandığından, çalışmada veri madenciliği tekniği olan Apriori Algoritması kullanılacaktır. Bu çalışmada, bir otomotiv ana sanayi işletmesinde yapılması planlanan analizler sonucunda elde edilecek verilerin hata oluşumunun engellenmesinde ve oluşmuş hataların da ürün müşteriye sunulmadan tespit edilmesine katkıda bulunacağı düşünülmektedir.

1.1 VERİ MADENCİLİĞİNİN TARİHİ GELİŞİMİ

“1950’li yıllarda matematikçiler veri madenciliği teknikleri üzerine çalışarak mantık ve bilgisayar bilimleri alanlarında yapay zekâ ve makina öğrenme alanlarını geliştirmişlerdir. Bunun yanında 1960’lı yıllarda istatistikçiler ise yeni bir algoritma keşfetmişlerdir. Veri madenciliğinin ilk adımlarını oluşturan bu algoritmalar regresyon analizi ve en büyük olabilirlik kestirimidir. Daha sonraki 20 yıllık süreçte önce verilerin sınıflara ayrılması ardından bu sınıflar arasında ilişkisel bağlantıların kurulması ile veri tabanı kavramı ortaya çıkarılmıştır. 1990’lı yıllara gelindiğinde ise veri tabanında bilgi keşfinin ilk adımları oluşturulmuş ve bununla birlikte büyük veri tabanları için veri ambarı geliştirilmiştir. Aynı zamanlarda yeni teknolojilerle beraber veri madenciliği yaygın olarak kullanılmaya başlanmıştır” [1].

Veri madenciliğinin gelişimi aşağıdaki tabloda daha net bir şekilde görülmektedir:

Tablo 1-1: Veri Madenciliğinin Tarihçesi

Gelişim Adımları	Cevaplanan Karar Problemi	Kullanılabilen Teknolojiler	Ürün Sağlayıcılar	Karakteristikler
Veri toplama (1960’lar)	Benim toplam karım geçen beş yılda ne kadardı?	Bilgisayarlar, teypler, diskler	IBM, CDC	Geriye dönük, statik veri dağıtımı
Veri erişimi (1980’ler)	İngiltere’de geçen mart ayında birim satışları ne kadardı?	İlişkisel veri tabanları, SQL, ODBC	Oracle, Sybase, Informix, IBM, Microsoft	Kayıt düzeyinde geri dönük, dinamik veri dağıtımı
Veri ambarlama ve karar destek sistemleri (1990’lar)	İngiltere’de geçen mart ayında birim satışları ne kadardı?	OLAP, çok boyutlu veri tabanı sistemi, veri ambarları	Pilot, Comshare, Arbor, Cognos, Microstrategy	Çoklu düzeylerde, geriye dönük, dinamik veri dağıtımı

Veri madenciliği (bugün)	Gelecek ay Bostondaki birim satışlar muhtemelen ne olabilir, niçin?	İleri düzeyde algoritmalar, çok işlemcili bilgisayarlar, büyük veri tabanları	Pilot, Lockheed, IBM, SGI, SPSS, SAS, Microsoft vs.	Geleceğe dönük, proaktif, enformasyon dağıtımı
--------------------------------	---	--	--	---

1.2 LİTERATÜR ARAŞTIRMASI

Çok geniş bir alana yayılmış olan veri madenciliği teknikleri üzerinde birçok çalışma yapılmıştır. Veri tabanları taranarak konu ile ilgili ulusal ve uluslararası makaleler, lisansüstü araştırmalar incelenmiştir.

Yapılan taramalar sonucunda;

- Veri Madenciliği

- Otomotiv ve Kalite

- Kalite ve Veri Madenciliği konuları hakkındaki araştırmalar derlenmiştir. Bu derlemelere aşağıda yer verilmiştir.

Veri madenciliğine ilişkin literatür aşağıdaki gibidir:

Akküçük [2], “Veri Madenciliği Kümeleme ve Sınıflama Algoritmaları” adlı çalışmasında da belirttiği üzere, tüm kurum ve kuruluşlarda toplanan veri miktarları arttıkça bu verilerin anlamlandırılması isteği, çok çeşitli matematiksel yöntemlerin bir araya gelerek “Veri Madenciliği” çatısı altında toplanmaları sonucunu doğurmuştur. Veri madenciliği, iki önemli amaç için kullanılabilir. İlk sırada “sınıflama” bulunurken ikinci sıra da ise “kümeleme” yer almaktadır. Bu durumda firmalar elde ettiği verileri yorumlarken, değişken gözetmeden demografik sıralama yapabileceği gibi segmentlere de ayırarak gruplandırabilir. Yalnızca tek amaca ulaşmak için çalışan algoritmaların yanı sıra, bu verileri sınıflandıran, kümeleyen ve gruplayan algoritmalar da bulunmaktadır. Çalışmanın ilk iki bölümünde Veri Madenciliğinin temel kavramlarından ve verilerin ön işlem uygulamalarından bahsedilmektedir. Daha sonraki diğer iki bölümde ise Veri Madenciliği kullanımında uygulanan temel istatistikî teknikleri açıklamıştır. Sonraki bölümde ise Veri Madenciliği amaçları olan “sınıflama” ile “kümeleme” konularına değinmiştir. Son bölümde ise hayali bir çalışma ile hayal edilen şirketin kayıp müşterilerinin tespitinin yapıldığı vaka analizi çalışması gerçekleştirilmiştir.

Doğan ve Büyükberber [3], veri madenciliği konusunda bu alanda yapılan lisansüstü tezler hakkında bir araştırma çalışması yapmışlardır. Son zamanlarda veri madenciliği teknikleri, kurumların ve şahısların farklı yapılar ve amaçlardaki çözümlemesi gereken problemlerine yanıt olmaktadır. Bu sebeple veri madenciliği, akademik alanda çalışma yapanları ve araştırmacıları kendine çekmektedir. Bu çalışmalarda, teknolojik ve bilimsel yenilikler kendini göstermektedir. Bu nedenle lisansüstü tezlerin içerikleri önem arz etmektedir. Bu araştırmada amaç; ülkemiz üniversitelerinde erişim imkânı bulunan ve “veri madenciliği” sorgusu ile saptanan lisansüstü tüm tezlerin farklı açılardan incelenmesidir. Bu çalışmada veri madenciliği ifadesi kullanılarak yapılan sorgular ile sosyal bilimler alanında 116 lisansüstü tez çalışması olduğu gözlemlenmiştir. Bu çalışmalardan 22 tanesi yapılan analizlerde kullanılamamıştır. Ulaşılan diğer 94 adet lisansüstü tez çalışmasının içeriği, üniversiteleri ve ana bilim dalları, çalışmaların uygulandığı alanlar incelenerek yorumlanmıştır. Ayrıca bu çalışmalar ek olarak, çalışmaların uygulama alanlarının ve ana bilim dallarının yazılan yıllara göre değişimleri de incelenmiştir.

Aydoğan ve Gencer [4], yaptıkları çalışmada toplanmış olan büyük verilerin gizlemiş olduğu anlamları ortaya çıkartmak için veri madenciliği yöntemini kullanmışlardır. Veri madenciliğin en önemli kavramlarından biri de sınıflandırmadır. Eksik ve büyük miktarlardaki veri kümelerinden bilgi toplamanın en güçlü aracı Kaba Küme Teorisidir. Bu teorinin temelinde ise veri madenciliğinin en önemli kavramları arasında yer alan sınıflandırmaya hizmet eden birçok sistem ve algoritma gerçekleştirilmiştir. Bu kapsamda çalışma içeriğinde Kaba Küme Teorisi işleyiş basamaklarını sunmuşlar ve sınıflandırma kavramına ait çalışmalar incelenerek bilgi vermişlerdir.

Emre ve Selçukcan Erol [5], yaptıkları çalışma ile, veri madenciliğinin artan önemine vurgu yapmışlardır. Çalışmaya göre, günümüzde gelişen teknoloji beraberinde veri çeşitliliğinde artma ve veri boyutlarında çoğalma meydana getirmiştir. Bu durum bilinen istatistiksel analiz yöntemlerinin yetersiz kalabileceğini göstermektedir. Günümüzde, yıllardır analiz için kullanılan istatistik teknikleri, veri madenciliği teknikleriyle beraber kullanılmaktadır. Bu yöntemler farklılık gösteren birçok araştırma alanında birlikte kullanılabilir. Genellikle sağlık alanında veri analizi istatistiksel yöntemlerle yapılmakta iken son yıllarda veri madenciliği araştırmacılar tarafından keşfedildiği gözlemlenmektedir. Bu çalışmada bu iki alan incelenerek mevcut benzerlikler ve farklılıkların ortaya konması amaçlanmıştır. Sağlık alanında çalışma yapacakların genellikle istatistiksel yöntemleri seçmeleri ve veri madenciliği yöntemlerini bilmiyor olmaları onları eski yönteme bağlı kılmaktadır. Çalışmada öncelikli olarak istatistiksel yöntemler ve veri madenciliği hakkında literatür çalışması gerçekleştirilmiştir. Sonuç kısmında ise bu iki kavram

genel olarak incelenmiş olup birbirleri arasındaki benzerlikler ve farklılıklar ana çizgileriyle ortaya konmuştur.

Veri madenciliğinin artan önemine ve istatistik ile karşılaştırmasına yer veren bir diğer çalışma, Tüzüntürk [6], tarafından 2010 yılında yapılmıştır. Bu çalışmada veri madenciliği ile istatistik konularını ele almışlardır. Bu konular arasındaki ilişkinin ortaya konması çalışmanın amacını oluşturmaktadır. Bu çalışma sayesinde istatistiksel yöntemleri ile veri madenciliği yönteminin ayrı düşünölmeyecek kadar içi içe olduđu belirtilmiştir. Yine çalışma içerisinde veri madenciliği literatür bilgilerinin yanı sıra OLAP küpleri kullanılarak bir veri madenciliği çalışması uygulanmıştır.

Doğan ve arkadaşları [7], tarafından yapılan ve 2016 yılını kapsayan çalışma, büyük verinin önemi, değeri, işleme teknikleri, veri ve metin madenciliği ve günümüzdeki durumuna ilişkindir. Çalışmaya göre, büyük veri, son günlerde yeni bir devir başlangıcı olarak yorumlanmakla beraber, çok yaygın ve herkesçe tanınan bir görüş halini almıştır. Bu sayede tüm alanlarda gerçek bir dönüşüm sağlanırken, ellerinde verileri bulunan birçok işletmenin ve kurumun toplanan verilere bakış açılarında değişme ve bu bilgilerden elde edilecek faydaların farkındalıkları keşfedilmiştir. İşletmelerin ve kurumların bu değişim ve dönüşümü görmemezlikten gelmeleri düşünölememektedir. Her ne kadar büyük veri, bu kadar tanınmış ve yaygın bir yöntem olmuş olsa da ulusal literatürümüze bakıldığında kurumsal anlamda alışılmışın dışında araştırmalarla çok fazla karşılaşılmamaktadır. Her ne kadar kurumsal yönleri işleyen yayın sayısı az olsa da teknoloji konusunda kendini geliştirmek isteyen ve bu konuyla ilgilenen birçok kuruluş ve işletme büyük veriye en üst düzeyde önem vermiş olup buna istinaden de büyük yatırımlar yapmışlar ve yapmaya devam etmektedirler. Bu çalışmada, literatür olarak büyük veri incelenmiş olup, birçok alanla olan benzerlikleri, büyük veriyi işleme yöntemleri incelenmiştir. Aynı zamanda dünya üzerinde büyük veriyi kullanarak yapılan farklı çalışmalardan örneklerle çeşitlendirilmiştir.

Doğan [8], 2017 yılında yaptığı çalışma ile ücretsiz veri madenciliği araçları ve Türkiye’de bilinirlikleri üzerine bir araştırma yapmıştır. Çalışmaya göre, veri madenciliğinin tanımı, istatistiksel temelli yöntemler kullanılarak ve bilgisayar teknolojisinden faydalanarak mevcut verilerden bilgiye ulaşmak olarak yapılmıştır. Her geçen gün kullanılan veri madenciliği analiz araçlarının sayısı artmaktadır. Bu çalışmada amaç, ücret ödemededen elde edilebilecek veri madenciliği yöntem ve araçlarını bir araya getirerek tanınmasını sağlamaktır. Bu amaç doğrultusunda ücretsiz olarak her türlü amaç için en sık kullanılan (akademik, ticari vb.) veri madenciliği yöntemleri belirlenmiştir. Veri madenciliği analizi hakkında kullanılan yazılımların geçmişi, kullanılan

alanları hakkında genel bilgilendirme yapılmıştır. Buna istinaden ülkemizde bu konu hakkında tez, proje, araştırma vb. çalışmalarında veri madenciliği yöntemini kullanan kişilerle kartopu örneklem metodu vasıtasıyla ulaşılarak bu yöntemin yazılımlarının ülkemizdeki yaygın olma durumu ve tanınırlık düzeyi tespit edilmeye çalışılmıştır. Araştırmacıların, bu çalışmada işlenen toplam 38 adet yazılımın sadece 5' i üzerinde odaklandığı diğer 33 tanesi hakkında bilgi sahibi olmamakla beraber hiçbir şekilde kullanmadıkları gözlemlenmiştir. Bu çalışma ülkemizdeki araştırmacıların veri madenciliği analizinde kullanacakları veri setlerinin oluşturulması ve diğer veri madenciliği yöntemlerinin tanınması konusunda öncülük edeceği düşünülmektedir.

Karacan ve Yeşilbudak [9], 2010 yılında yaptıkları çalışmada, kullanıcı merkezli interaktif veri madenciliği üzerine bir literatür taraması gerçekleştirmişlerdir. Çalışmaya göre, son zamanlarda tüm veri madenciliği modelleri, verilerin arkasında gizli kalmış olan belirli bir yapıyı otomatik olarak gün yüzüne çıkartmasını ve bu analiz süreçlerinde elde edilen verimliliğin önemine değinmektedir. Veride bulunan bilgi ağlarının yalnızca veri madenciliği araçları ile tespit edilmesi çok onaylanan bir yöntem değildir. Bunun sebebi elde edilen bilgiler fazla olmakla beraber veriyi kullanacak olan araştırmacılar için faydalı olmayabilir. Aynı zamanda ulaşılan bu öz bilgiler bir araştırmacı için yararlı olmayabilirken, diğer bir araştırmacı için yararlı olabilmektedir. Bu nedenle, veri madenciliği sürecinde araştırmacı odaklı ve etkileşimli bir analiz süreci geçirilmesi gerekmektedir. Araştırmacı odaklı ve etkileşimli bir analiz süreci gerçekleştiren araştırmacılar, bilgisayar sistemleri arasında etkin iletişim ağları oluşturarak amacına en uygun veri madenciliği algoritmalarını belirleme fırsatı bulmaktadır. Böylelikle araştırmacılar ve kullanıcılar için veri madenciliği analizi zor bir süreç olmaktan çıkarak araştırmacıların ulaşmak istedikleri ana bilgiye kolaylıkla ulaşma kolaylığı sağlamaktadır. Bu çalışmanın içeriğinde, araştırmacı odaklı ve etkileşimli bir veri madenciliği yöntemi hakkında literatür taramasına yer verilmiştir. Aynı zamanda, literatürde yapılan çalışmaların ve bu çalışmalardan çıkarılan sonuçların değerlendirmeleri ile bu değerlendirmeler hakkında yapılan önerilere yer verilmiştir. Bu analizlerde ulaşılan ana bilginin ortaya çıkarılması için yapılan çalışmanın, etkileşimli bir yapıya sahip olmasının yanında, en üst düzeyde görsellerle sunulması gerektiği konularına vurgu yapılmıştır.

Alaeddinoğlu ve arkadaşlarının [10], 2017 yılında yaptığı çalışmada, birliktelik kuralları ile mekânsal-zamansal veri madenciliğini ele almışlardır. Çalışmaya göre, araştırmacılar için öz bilgiye göre karar vermek oldukça büyük önem taşımaktadır. Bu kararın doğru olabilmesi için ise verilerle ilgili geçmişe dönük olarak bilgi birikimine ihtiyaç duyulmaktadır. Bu konuda bilgi birikimine sahip araştırmacıları bulmak ise zor, meşakkatli ve maliyeti açısından külfetli

olmaktadır. Gelişmekte olan teknolojiler vasıtası ve Veri Tabanı Yönetim Sistemlerindeki gelişimler verilecek olan kararların öz bilgilerine ulaşımı kolaylaştıracaktır. Bu durum karşısında araştırmacılar öz bilginin çıkarılmasının otomatik olarak yapılabilmesini ve bu iş ile uğraşan uzman kişilere bağımlılığı azaltmak için farklı yöntemler geliştirmişlerdir. Bu yöntemlerin en bilineni ise Veri Madenciliği yöntemidir. Bu analiz yöntemi de kendi içinde kümeleri oluşturma, birliktelikleri saptama ve sınıflandırma gibi alt gruplardan meydana gelmektedir. Birliktelikleri saptamada kullanılan birliktelik kuralı ise literatürde en çok kullanılmakta olan algoritma olarak Apriori Algoritması göze çarpmaktadır [66]. Veri madenciliği analizi alanında en fazla kullanılan bu algoritma değişkenlik gösteren mekân ve zaman kavramlarından oluşan veri kümelemelerinde pek kullanılmadığı gözlemlenmiştir. Bu verilerin diğer bilinen veri kümelerinden farklı yanı mevcut verilerin bilinen veri niteliklerine konum ve zaman niteliklerinin de eklenebilmesidir. Bu çalışmada zaman, mekân ve konum özelliklerinin de eklenebiliyor olmasından dolayı eldeki veriler Apriori Algoritmasına uyarlanarak Van Gölü'ne ait bir veri kümesi üzerinde çalışılmıştır.

Durmuşoğlu [11], 2017 yılında yaptığı çalışmada, veri madenciliği çalışmaları üzerine bir analiz yapmış ve Türkiye adresli yayınları incelemiştir. Veri Madenciliği analizi yöntemi elde ki mevcut büyük veri içerisindeki direkt olarak tespit edilemeyen öz bilgilerin tespit edilmesi için birtakım metotların uyarlandığı bir araştırma yöntemidir. Son dönemde gelişen mevcut teknolojilerin ucuzlaması ve aynı zamanda toplanan verilerin büyüklüğü tüm dünyanın gözünü veri madenciliği uygulamalarına çevirdiğini göstermektedir. Bu konulardaki artışa paralel olarak konu hakkındaki bilimsel çalışma literatürünün de arttığı gözlemlenmektedir. Bu çalışma kapsamında son on yıllık bir süreçte dünya genelindeki indeksli dergilerde yer alan ve aynı düzeyde konferanslarda sunulan veri madenciliği çalışmaları ile bu çalışmalardaki yöntemler, yazarlar ve ülkeler incelenerek dağılımları analiz edilmiştir. Yine ülkemizde yayınlanan bu konu hakkındaki çalışmalar ayrı bir başlık ile ele alınarak irdelenmiştir.

Tektaş ve Çiçek [78], yaptıkları çalışmada “veri zarflama analizi, benzer mal veya hizmet üreten ekonomik karar verme birimlerinin göreceli etkinliklerinin ölçülmesi amacı ile geliştirilmiş doğrusal programlama esaslı bir yöntemdir. Karar verme ile ilgili problemlerin çözümünde birçok analiz bulunmasına rağmen veri zarflama analizi en çok tercih edilen ve parametrik olmayan bir analiz yöntemidir. Tektaş ve Çiçek yaptıkları çalışmada, veri zarflama analizi ile ilgili Türkiye’de 1994-2019 yılları arasında ulusal tez merkezinde yayınlanmış olan lisansüstü tezlerin bibliyometrik analizini yapmıştır. Çalışmada 2006 yılından sonra veri zarflama ile ilgili yapılan çalışmalarda önemli bir artış meydana gelmekle birlikte 2014 yılında en yüksek seviyeye ulaşmıştır. Yapılan çalışmaların üniversite türü bakımından değerlendirildiğinde devlet

üniversitelerinin özel üniversitelere göre daha fazla çalışma yaptıkları görülmektedir. Çalışmada ayrıca son yıllarda artan veri zarflama analizi ile ilgili tezlerin uygulama alanları ile ilgili değerlendirmeler yapılmıştır.”

Şahin [79] çalışmasında Dumlupınar Üniversitesi Meslek Yüksekokulları performanslarını değerlendirmiştir. “Bu amaçla farklı ölçüm birimleri ile ifade edilebilen girdi ve çıktıların bir arada değerlendirilmesine imkân veren Veri Zarflama Analizi (VZA) kullanmıştır. Problemin çözümü LINGO programında yapmıştır. VZA’nın girdisi olarak öğrenci sayısı, akademik personel sayısı, idari personel sayısı, bölüm sayısını almış, çıktı olarak mezun sayısını kullanmıştır. Çalışmada toplam 14 Meslek Yüksekokulunu analiz dahil etmiştir. Yapılan analiz sonucunda; 5 Meslek Yüksekokulunu etkin olarak bulmuştur”.

Otomotiv sektöründe kaliteye ilişkin literatür aşağıdaki gibidir:

Fritzsche ve arkadaşları [12], çalışmalarında otomotiv sektöründe kalite analizi yapılmasını destekleyen interaktif ve web tabanlı bir veri madenciliği uygulaması sunmuşlardır. Bu uygulamanın, otomotiv mühendislerine, kalite ile ilgili kök sebepleri tanımlama konusundaki görevlerini yerine getirirken yardımcı olması için dizayn edildiğini ifade etmektedirler. Çalışmaya göre, tam olarak neyin probleme neden olduğunun bilinmesi ve konuya göre etkilenme ihtimali olan araçların tanımlanması, etkili hizmet aksiyonlarının planlanmasına ve uygulanmasına yardımcı olmaktadır.

Shankar ve Sundararajan [13], kalitenin artırılmasında, geleneksel kalite ölçümüne karşı veri madenciliğinde uç değerlerin kullanılması ile ilgili bir çalışma yapmışlardır. Çalışmada kullanılan veri, otomotiv motor montajı ve motor testleri sonrasında elde edilmiştir. Çalışmanın bulgularına göre, motorların geleneksel tarzda kontrolü, özgül yakıt tüketim değer kalıbına dayanmaktadır. Bu yöntem, ürün kalitesinin dizayn, güvenlik ve güvenilirlik açısından genellikle minimum düzeyde kalmasına neden olmaktadır. Son zamanların rekabetçi otomotiv sektörü, ürün kalitesinin geliştirilmesinde değişik bir yaklaşımın izlenmesini zorunlu kılmaktadır. Bu yaklaşım şüphesiz veri madenciliğidir. Uç/aykırı değerlerin analizi 500 motorun ele alındığı örnek verinin 1/4’ü, geleneksel yaklaşımla elde edilen milyonda 400 parçaya karşı, milyonda 15000 parçaya ulaşıldığını ortaya çıkarmıştır.

Kruse ve arkadaşlarının [14], yaptığı çalışmaya göre, otomobillerin dizaynı ve montajı, karmaşık bir görevdir ve kısa döngüler halinde gerçekleştirilmesi gerekmektedir. Bununla birlikte, müşterilerin güvenilirlik, dayanıklılık ve konfor gibi artan beklentilerinin de karşılanması gerekmektedir. Bu zorunluluklarla başa çıkmak için, aracın tüm yaşam döngüsü aşamalarında

toplanan verilerin analizini büyük ölçüde kolaylaştırmaya yönelik birtakım araçların geliştirilmesi önem taşımaktadır. Çalışmada, geliştirme, üretim ve planlama aşamalarında olduğu kadar ana ve tamamlayıcı bakım aşamalarındaki kalıp keşif görevleri için bazı metotlar ortaya konulmuştur. Bu konuda ilk yaklaşım, sonradan ilginç kalıplar bulmak için zihinde canlandırılabilen birliktelik kurallarını destekleyen Bayesian ağının yorumlanmasıdır. İkinci kısımda ise Markow ağları, bir aracın montajı yapılırken yürütülen planlama faaliyeti ile ilgili dayanışmayı modellemek için kullanılmıştır. Son kısım da ise zaman serisinde, simülasyon parametrelerinin ayarlanmasında kullanılan tekrar eden kalıpların bulunması ile ilgilidir.

Baysal ve arkadaşları [15], yaptıkları çalışmada, otomotiv yan sanayinde hata türü ve etkileri analizi gerçekleştirmişlerdir. Çalışmaya göre, Hata Türü ve Etkileri Analizi sistem içerisinde mevcut olan problemleri tanımlayarak bu problemleri elemek için kullanılan bir mühendislik alanına dair bir tekniktir. Hata Türü ve Etkileri Analizi ürünün tasarlanma sürecinde başlayarak daha sonrasında da sürekliliği olan bir serüvendir. Hata Türü ve Etkileri Analizi, ürün süreci ve bu süreçte oluşabilecek sorunları oluşmadan önlemek için geliştirilen sistematik bir yöntemdir. Bu çalışmada, Hata Türü ve Etkileri Analizi çalışması örneklendirilerek nasıl gerçekleştirileceği gösterilmiş olup otomotiv sanayisinde bir uygulama yapılarak bir takım çalışmasıyla örneklendirilmiştir

Liang, yaptığı çalışmanın konusu, otomotiv endüstrisinde müşteri değer analizinin veri madenciliği teknolojileri aracılığıyla yapılmasıdır. Çalışmada öncelikle müşterinin öneminden ve müşterinin değerinin işletmeler açısından anlaşılmasının gerekliliğinden ve etkili bir müşteri ilişki yönetiminin kurulmasının yararlarından bahsedilmiştir. Çalışmanın uygulama konusu, Tayvan'da faaliyet gösteren bir otomotiv ana işletmesidir. Çalışmada söz konusu işletmede veri madenciliği teknikleri kullanılarak müşteri değerinin analizi ve artırılmasının sistematik yolları aranmıştır. Bu bağlamda müşteri değerinin analizi için öncelikle K-means ve Kendini Düzenleyen Haritalar (SOM) metotları uygulanmıştır. Analiz sonuçlarına göre müşteriler değer grupları açısından düşük, orta ve yüksek olmak üzere üç sınıfa ayrılmıştır [16].

Sönmez ve Unğan [17], 2017 yılında yaptıkları çalışmada, hata türü etkileri analizinin otomotiv parçaları üretiminde uygulanmasına yer vermişlerdir. Çalışmada kullanılan Hata Türü ve Etkileri Analizi, oluşabilecek olan hataların öncesinde tespit edilerek oluşabilecek etkileri en aza indirmeye yöntemidir. Son yıllarda çok sayıda modern işletme tarafından yaygın olarak kullanıldığı görülmektedir. Bu çalışmada otomotiv endüstrisinde faal durumda olan ve metal işleme faaliyeti ile uğraşan bir işletmede Hata Türü ve Etkileri Analizi çalışması yapılmıştır. Çalışma süreci firmanın ürettiği ürünler ve bu ürünlerin üretildiği alt süreçlerin ham madde ve

yardımcı madde ihtiyaçlarının kabul edilmesi, stoklanması, proseslere sevkiyatı ve kalite kontrol aşamaları ile paketleme proseslerini içermektedir. Bu çalışma sonucunda olası Hata Türü ve Etkileri Analizi yapılarak uygunsuzlukların önceden tespiti sağlanmış olup ham madde ve yardımcı madde ihtiyaçlarının kabul edilmesi, stoklanması, proseslere sevkiyatı ve kalite kontrol aşamalarında %52,6'lık, paketleme prosesinde ise %54,3'lük bir iyileşme gerçekleştirilmiştir.

Teo [18], Victoria Üniversitesinde yaptığı yüksek lisans tezinde, otomotiv sektöründe garanti konusunda gerçek veriye dayanan örnekler üzerinden edinilen fikirleri veri madenciliği yöntemi ile yansıtmıştır. Çalışmanın sınırlılıkları yaş ve kilometre performansı şeklinde belirtmiş ve bir dönem aralığını baz almamıştır. Çalışmada non-parametrik çerçevede lineer yaklaşım ile parçalı lineer yaklaşımı izlemiştir. Tek değişkenli ve iki değişkenli olguların araştırıldığı çalışmada, tek değişkenli olgular ile yaşa, kilometre performansına ve gerçek takvim zamanına bağlı olarak elde edilen ortalama kümülatif maliyet garantileri ve standart hataları incelemiş, iki değişkenli olgular ile yaş ve kilometre performansının fonksiyonu olarak ortalama kümülatif garanti maliyetini ele almıştır.

Kargari ve Sepehri, otomotiv yedek parçalarının taşıma maliyetlerinin azaltılması için veri madenciliği tekniklerini kullanarak bir çalışma yapmışlardır. Çalışmaya göre, perakendecilerin bir dağıtım ağı içerisinde gruplandırılması hem dağıtımda hem de taşımada maliyetlerin indirilmesi açısından önemli katkı sağlamaktadır. Çalışmada bu kapsamda belirli otomotiv yedek parçalarının dağıtımı ve satış sonrası servis hizmetlerini yapan bir şirketten alınan üç yıllık veriyi analiz etmişlerdir. Çalışmada K-Means algoritmasını ve veri madenciliği uygulaması için ise R yazılımı kullanmışlardır. Elde edilen sonuçlar ışığında, her küme için uygun bir dağıtım politikası geliştirmişlerdir. Bu çalışmada elde edilen sonuçlar, otomotiv yedek parçası dağıtımı ve taşımacılığı maliyetlerinde %32 oranında bir azalma olduğunu göstermiştir [19].

Karabay [20], 1998 yılında yaptığı çalışma ile, ISO 9001-QS 9000 kalite sistem standartları kapsamında geometrik ölçülendirme ve toleranslama üzerine çalışmıştır. Geometrik ölçülendirme ve toleranslama, bilinen ölçülendirme ve toleranslamaya göre avantajlar sağladığından mekanik parça imalatında daha fazla uygulanabilirliği bulunmaktadır. Ayrıca ucuz ve işlevsel ekipman tasarımıyla birlikte bunların üretimine de olanak sağlamaktadır. Bu çalışmada ISO 9001 ile QS 9000 sistemlerini karşılaştırarak mevcut ve oluşabilecek gereksinimler üzerinde durulmuştur.

Aydoğan ve Semiz [21], 2005 yılında yaptıkları bir çalışmada, işletmelerde kullanılan ileri yönetim teknolojilerinin kullanım nedenlerini araştırmış ve otomotiv sektöründe bir alan araştırması yapmışlardır. Çalışmada işletmelerin günümüz şartlarında devamlılıklarını sürdürebilmeleri için teknolojik gelişmeleri benimsemeleri gerekmektedir. Bu uyum sürecinde ise

mevcut kaynakları en optimum şekilde kullanmaları gerekmektedir. Çalışmada gelişmekte olan üretime dair teknolojik yeniliklerin incelenmesi ve bu süreçte kullanılmakta olan mevcut yönetim teknolojileri ile uyumu incelenmiş olup, ülkemiz otomotiv ana sanayi işletmelerinde yapılmış bir çalışma ile örneklendirilmiştir.

Atmaca ve Keskin [22], 2007 yılında yaptıkları bir çalışma da TS 16949 Kalite Yönetim Sisteminin otomotiv sektöründeki yeri üzerine bir değerlendirme yapmışlardır. Bu çalışmada Bursa ili sınırları içinde üretim yapan otomotiv yan sanayi işletmelerinin katıldığı TS16949 Kalite Yönetim Sistemini kapsayan bir saha çalışması yapmışlardır. Bu konu hakkında oluşturulan anket sorularını SPSS programı ile analiz edilerek değerlendirmişlerdir. Değerlendirme sonuçlarına göre işletmelerin beklentilerin ne olduğu ve ilgili sertifikalandırma sürecinde sonra oluşan rekabet ve bu rekabetin avantaj ve dezavantajlarını açıklamışlardır.

Ertuğrul ve Aytaç [23], 2014 yılında yaptıkları bir çalışmada, kalite fonksiyon göçerimin de Markov zincirlerini analiz etmişlerdir. Kalite fonksiyon göçerimi, müşterinin beklentilerini karşılayacak sonuçlara ve bu sonuçlara ulaşırken müşterinin tatminini sağlayacak şekilde üretilecek ürünün imalatı sırasında kullanılacak kalite kontrol noktaları haline getirmek için tasarlanmış olan tasarım aşamasındaki kaliteyi artırmak amacı ile geliştirilmiş bir yöntem olarak bilinmektedir. Bu sayede müşterilerin talepleri ile mevcut üretim şartlarının birbiri ile olan ilişkisini analiz etmişlerdir. Bu araştırmada Markov zincirleme yöntemi ile elde edilen ilişkinin yorumlanması ve araç satın alan müşterilerin talepleri doğrultusunda oluşan araç tasarımı için kalite fonksiyonu yöntemi ile bir kişiden diğerine aktarılmaktadır. Bunun sonucunda Markov zincirleme yöntemi müşterilerin talepleri ile mevcut teknik şartlar arasındaki bağıntıları değerlendirerek mevcut teknik şartların gereksinimlerini ne şekilde değerlendirilmesi gerektiğini belirlemişlerdir.

Kesici ve Yıldız [24], 2016 yılında, kalite kontrol faaliyetlerinde yapay zekâ kullanımının bir otomotiv yan sanayisinde uygulanması üzerinde bir araştırma yapmışlardır. Çalışmaya göre, günümüz değişen dünya şartlarında işletmeler, rekabet koşullarının artması sebebiyle sıfır hata ve maliyet azaltma konularında fazlasıyla hassas davranmaktadırlar. Bu duruma istinaden bilimsel araştırma tekniklerinde ve işletmelerde karşılaşılan her türlü karmaşık problem de yapay zekâ kullanarak çözüme ulaşma düşüncesi ağırlık kazanmıştır. Bu elde edilen süreçler her sektörde olduğu gibi otomotiv endüstrisinde de kullanılmaya başlanmıştır. Bu yöntemlerle sıfır hata, karşılaşılan karmaşık problemlere etkin çözümler üretilmesinin yanı sıra, üretim süreçlerindeki kalite kontrol aşamalarına da destek olduğu gözlemlenmektedir. Bu araştırmada bir otomotiv yan sanayi üretim şirketindeki, üretim bantlarında personel kaynaklı hataların minimize edilmesi

kapsamında yapay zekâ çalışması gerçekleştirilmiştir. Gerçekleştirilen bu çalışma ile üretilmiş olan yapay zekâ yöntemi sayesinde kullanıcılarda olumsuz görüşe sebebiyet verecek olan kusurlar işletme sahasında üretim bandında tespit edilerek nihai kullanıcıya gitmemiş ve müşteri şikâyeti oluşmadan engellenmiştir.

Baray ve Sarı [25], kalite geliştirmede deney tasarımı yöntemi üzerine çalışmışlar ve otomotiv sektöründe bir uygulama gerçekleştirmişlerdir. Çalışmada, üretimi yapılan ürün kalitesinin yükseltilerek kalite kapsamındaki hataların çözümlenmesi için son dönemde etkin bir şekilde kullanılan deney tasarım yönteminin olumlu yönleri ele alınıp, bu yöntemin aktif bir şekilde kullanılarak yaygınlaştırılmasında katkı sağlanması hedeflenmiştir. Bu çalışmanın literatür bölümünde, deney tasarım yöntemi hakkında ana prensiplere değinilerek açıklanmış olup, uygulama kısmında ise otomotiv yan sanayi işletmesinde bir prosesteki parçanın dayanımı ele alınarak, deney tasarım yöntemi ile ulaşılabilecek çözümler hakkında önerilerde bulunmuşlardır. Aynı zamanda analiz yöntemi olarak verilerin işlenmesinde ANOVA ve bu sonuçlardan optimum sonucun çıkarılması için ise Yanıt Yüzeyi Yöntemini kullanmışlardır. Bu çalışma sonucunda ise kalite verilerinde büyük bir ölçüde iyileşme gerçekleştiğini gözlemlemişlerdir.

Arıkan Kargı ve Aydın [26], 2018 yılında yaptıkları çalışmada, istatistiksel kalite kontrol teknikleri ile otomotiv sektöründe bir uygulama gerçekleştirmişlerdir. Günümüz şartlarında firmalar arasında yoğun rekabet yaşandığı gözlemlenmektedir. İşletmeler bu rekabetçi şartlarda yarışabilmeleri için ürettiği ürünleri müşteri istekleri doğrultusunda gerçekleştirirken aynı zamanda kaliteyi de göz ardı etmemek gerektiğinin bilincinde olmalıdırlar. Üretilen mamulün kalitesinin artırılması, uluslararası standartlarda ürün üretilmesi, kalitesinin ne düzeyde olduğunun tespiti ve araştırılması istatistiksel yöntemlerle yapılabilmektedir. Yine bu sürecin de kontrolünün sağlanıp sağlanmadığı elde edilen verilerden oluşturulan grafiksel gösterimlerle gözlemlenebilmektedir. Bu çalışmada, Bursa ilinde faaliyet gösteren bir otomotiv ana sanayinde, üretilmiş olan otomobillerin ön kapılarının kapanma hızlarının istatistiksel olarak kontrol altında olup olmadığı yapılan analiz sonuçlarının grafiklerle gösterilerek belirlenmiştir. Bu bağlamda, konuya ait olan 2017 yılı ocak ayı verileri SPSS 17 analiz programı vasıtası ile incelenerek sonuçların istenen düzeylerde olup olmadığı, süreçlerin ve makinelerin yeterliliklerinin hesaplanması gerçekleştirilmiştir. Çalışma sonucunda ise tüm kontrol sürecinin istenilen standartlarda olduğunu tespit etmişlerdir.

Surinova ve Skurkova [27], 2013 yılında yaptıkları çalışma ile, otomotiv üretiminde kalite denetiminde Alman standartlarının kısa bir görünümünü vermeye çalışmışlardır. Otomotiv endüstrisinde, kalite yönetim sistemlerinin, diğer endüstriyel fabrikalar ile kıyaslandığında, birkaç

noktada farklılaştığı görülmektedir. Her müşterinin kendi özel gereksinimi bulunmaktadır. Denetimler, kalite yönetiminin önemli/öz araçlarından birisidir ve sürecin sağlıklı bir şekilde işleminde rolleri büyüktür. Açıktır ki, ISO/TS 16949: 2009 gereksinimi, otomotiv endüstrisini bu açıdan destekleyici bir niteliğe sahiptir. Fakat resmi ISO9001 Q-yönetim sistemi standartları ve ISO/TS16949'la birlikte birtakım standartlar da bulunmaktadır. Çalışmanın ana sorusu, otomotiv endüstrisinde bu özel standartların neler olduğu, hangi standardın neden ve nasıl kullanıldığıdır.

Luczak ve Wolniak [28], 2015 yılında bir çalışma yaparak otomotiv sanayinde problem çözme ve kalite yönetimi metotları geliştirme üzerine çalışmışlardır. Kalite yönetim teknik ve metotlarının etkin bir şekilde kullanımına dair bilgi, yüksek örgütsel kültürün bir göstergesi sayılmaktadır. Bu tür metot ve tekniklerin etkili ve verimli bir şekilde kullanılması, örgütün kalite yönetim sisteminde ulaşılmış olan olgunluk düzeyinin de bir göstergesidir. Çalışma, Polonya'da otomotiv sektöründe, kalite yönetimi sürecinde uygulanan problem çözme metot ve tekniklerinin bir analizi yer almaktadır. Alan araştırması, ISO/TS16949 sertifikasına sahip olan sanayi işletmelerini kapsamaktadır. Çalışmanın sonuçları ve yazarın önerileri, etkili kalite yönetim sistemlerinin geliştirilmesinde hangi stratejiler doğrultusunda yön verilmesi gerektiği konusunda aydınlatıcı olmayı amaçlamaktadır. Çalışmada elde edilen sonuçların, toplam kalite yönetimi stratejisinin tüm dünya tarafından fark edilmesi açısından da önemli olduğu vurgulanmıştır.

Özçalıcı [29], otomotiv sektöründe veri madenciliği yöntemi ile yaptığı çalışmaya göre, otomotiv piyasasındaki el değiştiren ikinci el araç sayısı her geçen gün daha da artmaktadır. Bu durumda ikinci el otomobil piyasasının istatistiksel olarak detaylıca incelenmesi konusunu gün yüzüne çıkartmaktadır. Gelişen teknolojiyle birlikte geliştirilen analiz teknikleri ile toplanan büyük veri setlerinin de desteği ile bu sektörün işleyiş yöntemleri hakkında bilgi sahibi olunması kolaylaşacaktır. Bu çalışmada amaç, ikinci el araç piyasasına ait olan verilerden oluşan bir veri seti oluşturularak birliktelik kuralı yöntemi ile durumun irdelenmesidir. Aynı anda gözlemlenen özelliklerin belirlenmesi ise birliktelik kuralı vasıtası ile gerçekleştirilmektedir. Bu çalışmada 2016 yılı temmuz ayı ağustos ayının üç haftalık bir dönemi ele alınarak, internet üzerinde ikinci el otomobil satışı gerçekleştiren bir internet sitesinden 211109 adet otomobil için 73 adet değişken saptanarak bir araya toplanmıştır. Birliktelik kuralının meydan getirilebilmesi için mevcut verinin mantıksal değişkenlerden bir araya getirilmesi gerekmektedir. Elde edilen sayısal değişkenler, mantıksal yapıya çevrilerek dağılımlar analiz edilmiştir. Elde edilen veriler Apriori algoritması yardımı ile ve birliktelik kuralı uygulanıp ağ grafiğine dönüştürülerek görsel olarak verilmiştir. Bu

çalışmanın sonucunda, “dizel araçlar az yakar, fazla hız yapmazlar, torkları yüksektir; yeni ve pahalı araçların vergisi yüksektir.” Sonuçlarına ulaşılmıştır.

Kalite ve veri madenciliğine ilişkin literatür aşağıdaki gibidir:

Köksal vd. [30], 2010 tarihinde yaptıkları çalışmada günümüz dünyasının en rekabetçi unsurunu kalite olarak görüldüğünü belirtmişlerdir. Firmalar, ürünlerin üretim süreçleri, bu süreçlerin kalite kontrol aşamaları, ürün tasarımı ve kalite iyileştirme çalışmaları için daha fazla emek vermekte ve iyi bir seviyeye gelip sürekliliğinin sağlanması için bu süreçlere tüm çalışanlarında dahil edilmesi gerektiğinin bilincinde olmaları gerekmektedir. Tüm bu çalışmaların sonucunda sıfır hatalı ürün üreterek müşteri memnuniyeti sağlayıp, harcanan maliyetlerde minimum seviyelere düşürülmesi bunlara paralel olarak da verimin ve kârlılığın maksimum seviyede artırılması planlanmaktadır. Kaliteyi iyileştirme çalışmaları, öncelikle üretim sahasından toplanan verilerin incelenmesi daha sonrasında ise müşteri geri dönüşlerinden elde edilen verilerin analizi ile gerçekleştirilebilmektedir. Bu tür analiz çalışmaları için, çok karmaşık yapıya sahip olan ve çok fazla girdiyi içeren değişkenlerin bulunduğu büyük verilerin analizleri için veri madenciliği yöntemlerine başvurulduğu görülmektedir. Bu araştırmada, veri madenciliği süreci anlatılarak, 1997 ile 2007 yıllarını kapsayan on yıllık bir süre içerisindeki imalat sektöründe kullanılan kalite iyileştirme faaliyetlerinde uygulanmış olan veri madenciliği çalışmalarını kapsayan literatürden faydalanılmıştır. Bu süreçte, kalite süreci ve üretilen ürünün kalitesinin tanımlanması, mevcut ve elde edilen kalite sınıflandırılması çalışması ve kalite parametrelerini oluşturan değişkenlerin optimizasyonu hakkında bilgi verilmiştir. Ayrıca çalışma içeriğinde, veri madenciliği yöntemleri arasında yer alan karar ağaçları uygulaması kullanılarak bir döküm işletmesinde tespit edilen döküm proses hatalarına sebebiyet veren değişkenler ile bu değişkenlerin seviyelerini tespit etmeye çalışılmışlardır.

Reuter ve arkadaşları [31], veri madenciliği algoritmalarının üretim sürecine uyarlanmasına ilişkin bir çalışma yapmışlardır. Çalışmaya göre, üretim işletmeleri gün geçtikçe artan bir rekabet ve karmaşa halinde bulunan bir çevre içerisinde faaliyetlerini sürdürmek durumundadırlar. Bu çevre içinde işletmelerin varlıklarını sürdürebilmeleri, söz verilen teslimat sürelerinde ürünlerin yetişmesi işletmelere rekabet üstünlüğü sağlayacaktır. Çünkü, dinamik bir doğası olan ve sürekli değişen üretim şartlarında ve yüksek çeşitlilik içeren üretim programlarında faaliyet göstermek zorunda olan üretim işletmeleri, lojistik hedef konusunda başarısız olabilmektedirler. Bu çalışma, veri madenciliği algoritmaları kullanılarak üretim süreçlerine dair verinin kaliteli bir şekilde elde edilmesine yardımcı olacak bir model önermektedir. Bu bağlamda, Almanya’da faaliyet gösteren orta büyüklükteki üretim işletmelerinden gerçek algoritmaları kıyaslamaya çalışmışlardır.

Köksal ve arkadaşları [32], yaptıkları çalışmada, kalite geliştirme programlarından bahsetmişler ve kalite geliştirme sürecinde veri madenciliği kullanımına yoğunlaşmışlardır. Çalışma, işletmelerde kalite geliştirme için kullanılan altı sigma ve kaizen gibi programların yanı sıra, veri toplamak ve analiz etmek açısından sayısız avantajları bulunan veri madenciliği tekniklerinin de yaygınlaştığını ortaya koymaktadır. Çalışmada, 1997-2007 yılları arasında üretim endüstrisinde kaliteyi hedefleyen veri madenciliği uygulamaları ile ilgili olarak yapılmış olan literatür çalışmaları ve çeşitli analizler bir araya getirilmiştir. Bunlardan kalite ile ilgili olanlar arasında ürün-süreç kalitesi, kalite öngörüsü, kalite sınıflandırması ve parametre optimizasyonu yer almaktadır.

Çınar ve Silahtaroglu [33], 2012 yılında yaptıkları çalışmada, veri madenciliği teknikleri ile müşteri memnuniyetine etki eden gizli nedenleri ortaya çıkarmayı amaçlamışlardır. Bu çalışmayla birlikte 301 kişiden oluşan müşteri memnuniyet anket sonuçlarını irdelemiş olup müşteri cevaplarına göre oluşan memnuniyet veya memnuniyetsizlik sebepleri araştırarak, bunların altında yatan gizli nedenlere ulaşmaya çalışmışlardır. İncelenen verileri grafiksel analiz gösteriminin yanı sıra K-Means ve C5.0 algoritmaları gibi veri madenciliğinin sınıflandırma ve kümeleme algoritmalarından da faydalanmışlardır.

Acar ve Çakırkaya [34], yaptıkları çalışmada bir üretim hattında meydana gelen hataların önem derecelerini Pareto analizi ile tespit etmişlerdir. Çalışmaya göre, firmalar içinde bulundukları rekabet mücadelesinde müşterilerin talep ettikleri ürünleri istenilen kalite de üretebilmek ve pazarda ayakta kalabilmek için üretim hatlarındaki tüm hata gruplarını sıfırlamak veya minimum düzeylere indirebilmek için birçok analiz ve bunların sonuçları için birçok çalışma gerçekleştirmektedir. Bu tür çalışmaların başında proses kalite kontrol uygulamaları ve bu uygulamalarda tespit edilen hataların iyileştirme aktiviteleri, aynı hatalarla karşılaşılmasında yapılan aksiyonların sürekliliğinin kontrollerini sağlamak firmalar için oldukça önemlidir. Bu sayede tüm uygunsuzluklar tespit edilerek oluşum sebepleri araştırılarak tekrarlamamaları sağlanabilir. Çalışmada literatür olarak istatistiksel proses kontrol teknikleri irdelenmiş, Pareto Analizi grafiği uygulanarak tespit edilen hataların önem sırası belirlenmiştir. Bu bağlamda Pareto Analizi, basit bir yöntem olmasına karşın sonucu göstermede önemli bir analiz tekniğidir. Uygulama aşamasında ise işletme üretim bandında oluşan hataların tespiti yapılmış olup Pareto Analizi grafiği yardımı ile gösterimi gerçekleştirilmiştir.

Liu ve Sun [35], 2009 yılında yaptıkları çalışmada, üretim kalitesi yönetiminde veri madenciliği uygulamasını ele almışlardır. Çalışmada öncelikle üretim işletmelerinin kalite yönetimi sürecinde veri madenciliğinin kullanımına yönelik literatüre yer verilmiştir. Ayrıca,

kalite faktör analizinin, kalite kontrolü ve ürün yönetimi açısından çok önemli olduğu vurgulanmıştır. Ardından, ürün kalitesini etkileyen değişik faktörler arasındaki ilişkilerin ortaya konması amacıyla Apriori algoritmasından yararlanılmıştır.

Yalçın Pirinçciler ve Şen [36], veri madenciliği kullanımının, süreç iyileştirme aktiviteleri üzerinde bir çalışma yapmışlardır. Bu çalışmaya göre, günümüz teknolojilerindeki gelenen aşamalarda, şirketlerin ellerinde bulunan çok sayıdaki büyük veri topluluklarını arşivlemeleri ve bu verilerden anlaşılabilir veriye dönüştürmeleri mümkün olmaktadır. Veri madenciliği teknikleri, eldeki bu büyük veri topluluklarını anlamlı birer bilgi kümesine dönüştürmek amacı ile kullanılan tekniklerin başında gelmektedir. Bu verilerden elde edilen sonuçların yorumları ise müşteri odaklı çalışan firmalar için veri girişi olarak değerlendirilmektedir. Çalışmada, günümüz rekabet koşullarında işletmelerin ayakta kalabilmeleri adına müşteri şikayetlerini sıfıra indirerek üretim süreçlerini oluşan veya oluşabilecek hatalar doğrultusunda sürekli iyileştirme faaliyetleri kapsamında Altı Sigma aktivitelerini incelemişlerdir. Ayrıca Altı Sigma uygulamaları kapsamında kullanılmakta olan DMAIC döngüsü aşamalarına değinmişlerdir. Bu çalışma için verilerin derlenmesinde, analizinin yapılması ve ölçülebilir bir duruma gelmesinde Veri Madenciliği tekniklerini kullanmışlardır.

Wallis ve arkadaşları [37], montaj sürecinde veri madenciliği tekniklerinin kullanımı ile ilgili bir çalışma yapmışlardır. Son zamanlarda, dijital montaj araçlarının kullanımı, üretim sürecinin karmaşıklığına bir çözüm yolu olarak gittikçe artan oranda yaygınlaşmaktadır. Bunun bir sonucu olarak, montaj sürecine dair önemli oranda veri, değişik sistemlerde birikmektedir. Bu veri, veri madenciliği teknikleri kullanılarak, gelecekteki montaj planlamaları açısından yararlı hale getirilebilir. Bu çalışma, montaj planlamasının, planlama sürecinde etkinlik ve kalitenin sağlanması için, veri madenciliği teknikleri ile yapılması ile ilgili bir yaklaşım sunmaktadır. Bu yaklaşım, üretim ve sürece dair verilerin sınıflandırılmasına ve aralarındaki bağlantının tanımlanmasına dayanmaktadır.

Yılmaz [38], hata türü ve etki analizi ile ilgili yaptığı çalışmasında üretim bantlarındaki ürünlerin ve bu ürünlerin üretildiği prosesler de oluşabilecek risklerin oluşmaması ve bu tür çalışmaların belgelenebilmesini gerçekleştirmiştir. Bu analiz türü, tüm müşteri odaklı konularda mamul, hizmet vb. oluşabilecek hata türü ve konularının tespiti ve oluşum adedi ile oluşum sıklığı gibi önem arz eden başlıkların gruplandırılması amacı ile kullanılmaktadır. Kaliteyi arttırarak ürün ve üretim hattında meydana gelebilecek uygunsuzluklarla ilgilenmektedir. Yine bu analiz türü oluşabilecek hata gruplarını tespit ederek, bu hata türlerinin doğurabileceği yeni sonuçların

etkilerinin büyüklüğünü tespit etmeyi hedeflemektedir. Aynı zamanda hata çeşitlerinin oluşmaması adına yapılan düzenleyici önleyici faaliyetleri de içermektedir.

Asilkan, yaptığı çalışmada, ikinci el otomobillerin güncel pazar fiyatlarını veri madenciliği yöntemi ile modellemiştir. Çalışmaya göre, “otomobiller, tüketicilerin en çok ilgi gösterdiği dayanıklı ürünlerdendir. Çok sayıda teknik ve opsiyonel özelliğe sahip olan otomobillerin ikinci el pazarındaki fiyatlarının matematiksel bir modelini oluşturmak oldukça güçtür. Çalışmada iki farklı veri madenciliği yöntemi kullanılarak ikinci el otomobillerin pazardaki güncel fiyatlarını modellenmeye (tahmin edilmeye) çalışmıştır. Girdi verisi olarak Avrupa merkezli çok sayıda ikinci el otomobil web sitesinden elde edilmiş olan ilanlar kullanılmıştır. Regresyon Analizi ve Yapay Sinir Ağları uygulaması ile modeller oluşturulduktan sonra bu modeller karşılaştırılmıştır. Elde edilen sonuçlar, veri madenciliğinin ikinci el otomobillerin güncel pazar fiyatlarını modellemede başarılı olarak kullanılabileceğini ortaya koymuştur” [39].

Vazan ve arkadaşları, yaptıkları çalışmada, modern üretim sistemlerinin toplanması, saklanması, yönetilmesi ve analiz edilmesi gereken önemli miktarda veri yarattıklarını ifade etmişler ve çalışmalarını üretilen veriye göre üretim süreci davranışının önceden tespit edilmesi üzerine yoğunlaştırmışlardır. Çalışma veri madenciliği yöntemi ile üretim sisteminin gelecekteki davranışı hakkında bilgi edinme yollarının ortaya konması ile ilgilidir. Kontrol sürecinde gerekli olan bilginin toplanması için gerçek bir üretim sisteminin simülasyonu dizayn etmişlerdir. Üretim süreci ile ilgili öngörüler, veri madenciliğinin bazı tekniklerini kullanarak yapmışlardır. Yazarlar değişik metodlar arasından nöral ağ metodunu seçmişlerdir. Araştırmanın amacı, üretim sisteminin kontrolünü kolaylaştıracak ve amaca ulaşmayı hızlandırarak karar verme sürecini destekleyecek veri madenciliği araçları dizayn etmektir [40].

Akteke-Öztürk ve arkadaşları, kalite iyileştirmede veri madenciliği kullanımı ile ilgili bir çalışma yapmışlardır. Bu çalışmaya göre, “üretim sürecinde hatalı ürünlerin oluşma nedenleri, bir üründe hatanın oluşup oluşmayacağını tahmin edilebilmesi ve hata oluşumunu en çok etkileyen üretim süreci parametrelerinin belirlenmesi için veri analizi yapmışlardır. Bu nedenle özellikle birçok değişken tipinin büyük hacimli olarak saklandığı karmaşık verilerin analizinde başarılı bir şekilde uygulanan veri madenciliği teknikleri kullanılarak kalite verisi etkin bir şekilde izlenebilir, hatta toplanan veride gizli olan bilgi ortaya çıkartılabilir ve süreçte meydana gelecek hatalar hızlı bir şekilde saptanabilir. Veri madenciliği öğrenmeye dayalı bir yöntemle mevcut veride gizli olan örüntülerin keşfedilerek bilgi elde edilmesini amaçlamaktadır. Endüstriyel süreçler genellikle büyük hacimli, sürekli veya kesikli veriyi saklayan veri toplama sistemleri ile desteklenmektedirler. Karar (sınıflandırma ve regresyon) ağaçları, yapay sinir ağları, destek vektör

makinaları, k-ortalamalar gibi Veri Madenciliği araçları veriden bilgi çıkarımı sağlamakta etkin ve hızlı çözümler sunabilmektedir” [41].



2. MATERYAL VE YÖNTEM

Bu bölümde öncelikle çalışmanın amacına, önemine ve sonucuna ulaşmak için izlenecek yöntem açıklanacaktır. Ayrıca çalışmanın kapsamına ve uygulama metotlarına değinilecektir. Daha sonra ise elde edilen gerçek veriler kullanılarak, profesyonel bir paket program vasıtası ile bu veriler analiz edilecektir.

2.1 MATERYAL

Bu çalışmanın temel materyalini, Türkiye'de otomotiv sektöründe faaliyet gösteren bir ana sanayi işletmesinden temin edilen ve belirli bir dönemi kapsayan gerçek veriler oluşturmaktadır. Kullanılacak veriler 8 yıllık bir zaman dilimini kapsamakla birlikte üretilmiş olan ürünler üzerinde yapılan son kontrollerde tespit edilmiş olan kalite hatalarından oluşmaktadır.

Her yıl tespit edilen hatalar üretici firma veri tabanında arşivlenmektedir. Bu hata grupları öncelikle sorumlusu olduğu departmanlar üzerinden gruplandırılmaktadır. Daha sonra tespit edilen hatalar, hata gruplarına göre kategorilere ayrılmaktadır.

2.1.1 Uygulama Sahası Hakkında Genel Bilgi

Bu teze konu olan üretici firma dünyaca ünlü bir otomotiv firması olup şirket yönetimi tarafından ve şirket politikası nedeniyle, işletme bilgilerinin gizli tutulması istenmiştir.

İşletme 90'lı yılların başlarında Türkiye'de kurulmuş olup birkaç yıl sonra da seri üretime başlamıştır. İşletme bugün ortalama 6.000- 6.500 kişi istihdam ederek yıllık yaklaşık 300.000 adet araç üretmektedir. Üretilen araçlar 5 kıtada yaklaşık 100 ülkeye ihraç edilip ülke ekonomisine de büyük bir katkı sağlamaktadır.

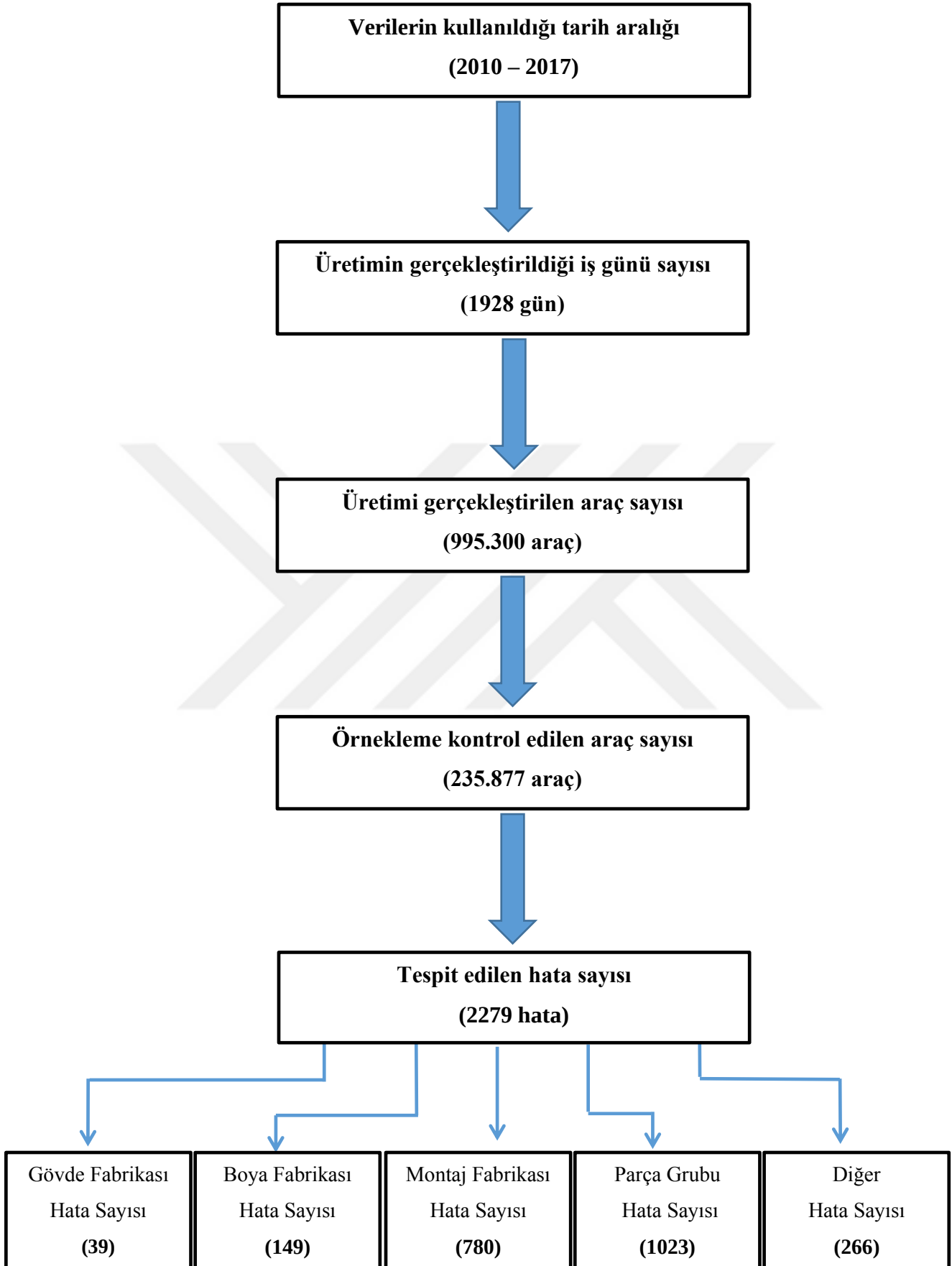
İşletme pres, kaynak, boya ve montaj fabrikalarından oluşup seri üretim mantığı ile çalışmaktadır. Saç levhalar pres fabrikasında uygun formlarda şekillendirilmektedir. Şekillenen bu saç malzemeler kaynak fabrikasında birleştirilerek aracın iskeleti oluşturulur. Bu iskelet yapı üretim bantları vasıtası ile boya fabrikasına yönlendirilerek gerekli ses yalıtım malzemeleri ve dış boya rengi ile kaplanır. Boyanmış olan araç iskeleti montaj fabrikasında müşteri standart ve opsiyonlarına göre araç son şeklini almaktadır.

Üretimin gerçekleştiği her sahanın sonunda ve üretimi tamamlanmış her araç için, belirlenen standartlar içerisinde olup olmadığı kontrol edilmektedir. Bu kalite kontrol yöntemi

%100 her araca uygulanmaktadır. Üretimi tamamlanmış olan araçlar müşteri gözüyle kontrol etmek için son muayene kalite kontrol grubu tarafından üretim yoğunluğuna göre örnekleme yöntemi ile kontrolü gerçekleştirilmektedir. Aynı zamanda dış tedarikçilerden gelen ürünler giriş kalite kontrol proseslerinde kontrol edilerek uygunsuz olan ürünlerin fabrika içerisine girişi engellenmektedir. Karşılaşılan uygunsuzluklar ise şirket veri tabanında raporlanarak arşivlenmektedir.

Teze konu olan veriler bu veri tabanında depolanmış olan ve son muayene kalite kontrol grubunun tespit etmiş olduğu hata raporlarından oluşmaktadır.

Çalışmanın verileri 2010-2017 yılları arasında, 1928 çalışma gününde, üretici firma genelinde gerçekleşen 995.300 adet arasında örnekleme kontrolü yapılan 235.877 adet araçtan tespit edilen 2279 adet hata sayısını ifade etmektedir. Aşağıdaki şekilde hata sorumlusu grup dağılımları da görülmektedir (Şekil 2-1).

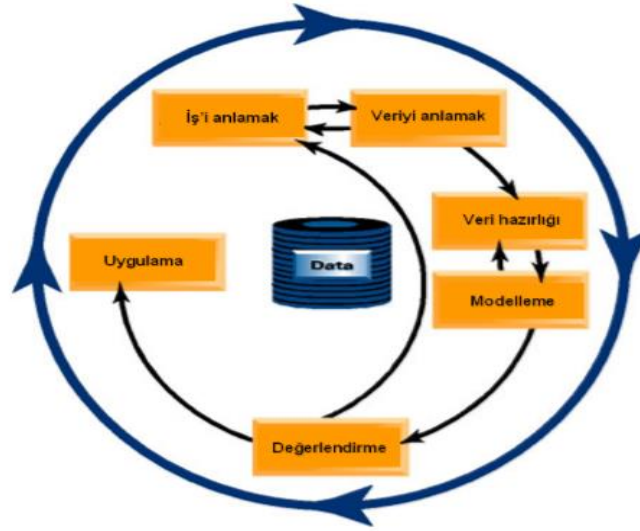


Şekil 2-1: Veri toplamada izlenen iş akışı

2.2 YÖNTEM

Veri madenciliği, araştırma yapılarak ve bunun tekrarlamasına dayalı bir yöntem olmasının yanında elde edilen verilerin veri tabanlarındaki gizli bilgilerin keşfedilmeleriyle genişleyerek anlam bulan bir süreci ifade etmektedir. Bu süreç dahilinde yeni yeni bilgilerle karşılaşılabilir ve bu durumları analiz edebilmek için yeni hipotezlerle konular tekrardan irdelenebilir. Bu durum farklı konularında düşünülerek irdelenmesine yol açabilmektedir. Bu tür verilerin özündeki bilgilere ulaşılabilmesi adına, analiz yapan kullanıcının konu hakkındaki bilgi birikimini de işin içine katması gerekmektedir. Analiz aşamaları sürecinde elde edilen veriler araştırmacıya farklı konularda git-gel ler yaşatabilir [42].

Veri madenciliği, büyük veri içine saklanmış olan ana fikirlerin ortaya çıkartılma sürecidir. Bu analizleri bir süreç olarak değerlendirebilmemiz için gerçekleştirilen her bir işlem basamağının tek tek ve dikkatlice takip edilmesi gereklidir. Elde edilen her veri, işlenen her prosesin çıktısı olarak değerlendirileceği gibi bir sonraki prosesinde girdisi olacaktır. Bundan dolayı veri madenciliği analizin de takip edilmesi ve adımların terkedilmemesi gereken standart iş uygulaması gerekmektedir. Bu standart iş süreci, 1996 yılı sonlarında veri madenciliği alanında çalışmalar yürüten üç firma tarafından kurulmuş olup sürecin adını The Cross-Industry Standard Proces for Data Mining olarak isimlendirmişlerdir. Bu üç firmanın ilki Daimler Chrysler'dir. Veri madenciliği analizleri konusunda birçok endüstriyel organizasyona öncülük etmişlerdir. Statistical Package for the Social Sciences işletmesi 1990 yılından bugüne veri madenciliği analiz hakkında birçok hizmet sağlayıcı uygulamayı aktifleştirerek 1994 yılında Clementine'i hayat geçirmiştir. National Cash Register, firması ise veri madenciliği hakkında danışmalık yaparak ve ilgili teknolojileri de kullanarak müşterilerin talepleri ve beklentilerini tespit edilebilmesi için takımlar kurmuştur. Bu süreç altı aşamadan oluşmakta ve en meşakkatli süreci de veri hazırlama süreci olarak göze çarpmaktadır. Bu süreci de karar verilen problemin saptanması, elde edilen verilerin analiz edilmesi ve elde edilen sonuçların doğru bir biçimde yorumlanması süreçleri takip etmektedir [43].



Şekil 2-2: The Cross-Industry Standard Proces for Data Mining süreç şeması .

2.2.1 İş'i Anlamak

Veri madenciliğinin ilk basamağını iş ve iş ortamını anlamak oluşturmaktadır. İş akışını algılayarak girdilerin ve çıktıların işe olan etkisinin analizinin yapılması gerekmektedir.

The Cross-Industry Standard Proces for Data Mining yönteminin birinci aşamasında, araştırmanın amacının ve ihtiyaçlarının yine aynı açıdan gözlemlenerek anlaşılması ve daha sonra elde edilen veriyi veri madenciliği analizi yöntemi ile anlayarak, amaçlar doğrultusunda sonuca ulaşabilecek yönde işlenebilecek bir hale dönüştürmeyi amaçlamaktadır. Veri madenciliği analizi yapacak olan araştırmacının önceliği tam anlamıyla nereye veya neye ulaşmak istediğini tam olarak belirlemektir. Veri sonuçlarında ve verilerin sahip olduğu işletmelerde dengelenmesi gerekli olan birçok amaç mümkün olabilir. Araştırmacının asıl ulaşması gereken nokta ise sonuca şimdiden etki edebilecek olan ana etkenleri tespit ederek başlangıç seviyesinde ortaya koymaktır. İş'i anlama aşamasına gerekli özen ve hassasiyetin gösterilmiyor olması, işletme için, yanlış yönde kürek çekmekten başka bir çabadan öte gitmeyecektir [45].

2.2.2 Veriyi Anlamak

Veri Madenciliğinde ikinci aşama veriyi kavramaktır. Bu aşama verinin toplanması ile başlayarak veri içindeki sorunların tanımlanması, verilerin keşfedilmesi ve elde edilen verilerdeki gizlenmiş olan bilgilerin ortaya çıkartılması için sınıflandırılması ve kümelenmesi aktiviteleri ile

devam etmektedir. Yine bu aşamada verinin kalitesi ve veri içindeki bilgilerin neyi ifade ettiği tespit edilir. İşi ve veriyi anlamak birbirinin ayrılmaz bir parçasıdır [46]. İşletme düzenli olarak her günkü üretim sonuçlarına ilişkin raporları haftalık, aylık ve yıllık olarak çeşitli gruplar altında veri tabanında bulundurmaktadır. Çalışmanın başlangıç aşamasında 2010-2017 yılları arasında bitmiş olan araçların son kontrolünü yapan kalite kontrol grubunun raporları incelenerek tespit edilmiş olan hatalar üzerinde çeşitli gruplandırmalar yapılacaktır.

Bu aşamada öncelik verinin elde edilmesidir. Farklı yerlerden toplanarak elde edilen veriler tek bir tabloda düzenlenmelidir. Toplanan ve tek bir tablo halini getirilen bu veri kümesi genel olarak neyi ifade ettiği konusunda inceleme yapılarak anlamaya çalışılmalıdır. Bu aşamanın sonraki aşamadan farkı ise, veri üzerinde hiçbir şekilde değişiklik yapılmadan olduğu gibi kabul edilerek inceleniyor olmasıdır. Yine bu aşamada incelenmiş olan veri kümesi raporlanarak sonraki aşamada nasıl hazırlanması gerektiği ve hangi noktaların temizlenerek arınması gerektiği konusunda fikir vermiş olacaktır [45].

Veri madenciliği analizi yöntemi ile kullanılmakta olan klasik istatistiksel analiz yöntemi arasındaki en önemli fark elde edilen veri topluluğunun büyüklüğüdür. Bir veri madenciliği araştırmacısının karşılaşacağı veri miktarı milyon veya milyar seviyesindeki bir büyük veri topluluğu iken bu rakamlar klasik yöntemle istatistiksel veri analizi yapan araştırmacılarda elde edilen mevcut kümenin büyüklüğünün yüzölçümü veya binli rakamları geçmediği görülmektedir. Teknolojinin de gelişmesiyle işletmelerin büyümeyle de bağlantılı olarak büyük veri gruplarıyla çok sık karşılaşmaktadır. Elde edilen verilerin değişkenliklerinin çok fazla farklılık gösteriyor olması, ilerleyen aşamalarda zorluklarla karşılaşılacağına bir göstergesidir. Elde edilen değişkenlerin fazlalığı ve farklılığı yine bu duruma etki eden hücrelerin birim bazda üstel olarak artışa sebep olacaktır. Elimizdeki bir tek ikili değişken ile elde edilecek olan hücre sayısı ile iki tane ikili değişkenden elde edilen hücre sayısı arasında iki kat fark gözlemlenmektedir. Örnekler arttırılarak sayıların arttığı görülebilir [47].

2.2.3 Veriyi Hazırlamak

Bu aşamada farklı dönemlerin verileri, değişik kaynaklardan toplanarak tek bir veri tabanı oluşturulması sağlanmaktadır. Kolay yollarla ve üstünkörü yapılan iyileştirme faaliyetlerinin ilerleyen süreçlerde daha büyük ölçekli sorunlara sebebiyet vereceği için verilerin hazırlanması oldukça önemlidir.

Çalışmanın amacını gerçekleştirmek için işletmenin çeşitli aksiyonları aldığı zaman aralığı tercih edilmiş olup gürültülü veriler temizlenmiştir. Bu zaman aralığında işletme tekli, ikili ve üçlü vardiya olarak çalışmakta olup aynı zamanda üretilen araç modellerinde de bir, iki ve üç modeli aynı anda üretmektedir.

Çalışmanın verileri 2010-2017 yılları arasında, 1928 çalışma gününde, fabrika genelinde üretimi gerçekleşen 995.300 araç arasında örnekleme kontrolü yapılan 235.877 araçtan tespit edilen 2290 hata sayısından oluşturmaktadır. Bu veriler Şekil 2-3’de görüldüğü üzere Microsoft Excel formatında düzenlenerek bir tablo haline getirilmiştir.

The screenshot shows a Microsoft Excel spreadsheet with a data table. The table has columns for dates (from 4.1.2010 to 22.2.2010), shifts (1, 2, 3), models (1, 2, 3), and various error types (Hata Aray Model, Hata Bulun Vardiyası, 1. Vardiyası Kontrol Edilen Hata Aray, etc.). The data is organized into a grid with rows representing dates and columns representing different categories of errors and shifts.

Şekil 2-3: Microsoft Excel formatında düzenlenen veri seti alıntısı

2.2.3.1 Gürültülü Veri

Büyük verilerin analizlerinde, veri girişlerinin tüm aşamalarında kullanıcı kaynaklı hataların yanı sıra ölçülen değerlerin hatalı olması gibi veri sistemine dahil olmayan hatalardan dolayı oluşan veri hatalarına gürültü denmektedir. Bunun yanında gürültü, ölçülen değerlerdeki değişiklik veya rastgele hata olarak da adlandırılır. Günümüzdeki veri tabanı yönetim sistemleri çok gelişmiş olsalar da bu tür hatalar gözden kaçabilmektedir. Böyle bir durum karşısında çalışılan veri madenciliği analiz yöntemi veri kümesi içerisindeki gürültülü verilere karşı hassasiyet

göstermelidir. Eğer veriler içerisinde gürültülü veri tespit edilirse bu verileri tanımlayarak analiz dışı bırakabilmelidir [48-49]. Uygulama için toplanan verilerin yeterli sayıda, tutarlı ve gürültülü verilerden ayıklanmış olmaları gerekmektedir [75]. Bunun yanında veri toplanmasında olabilecek hataların sebepleri ve muhtemel sonuçları ise şu şekilde sınıflandırılabilir.

- A. Hata sebepleri
 - Hatalı veri toplama gereçleri
 - Veri giriş problemleri
 - Veri girişi sırasında kullanıcıların hatalı yorumları
 - Veri iletim hataları
 - Teknolojik sınırlamalar
 - Veri isimlendirmede veya yapısında uyumsuzluk
- B. Muhtemel sonuçları
 - Çelişkili (tutarsız) veriler
 - Tekrar eden (çift) kayıtlar
 - Yetersiz (eksik) veriler

Veri madenciliği analizi yöntemlerinin uygulanabilirliğini kolaylaştırmak için bir takım bilgisayar yazılım programları ve araçlar geliştirilmiştir [50]. Veri Madenciliği analizi yöntemleri için kullanılan çeşitli yazılım ve araç çeşitlerine ait örnekler Tablo 2-1’de verilmiştir [51].

Tablo 2-1: Veri madenciliği için kullanılan araç çeşitleri ve örnekleri.

Kullanılan Araç	Örnekler
Programlama Dilleri	Perl, Python, SQL, XML
Genel Amaçlı	Angoss, Clustran, Cross-Graph, Cross-z, Daisy, Data Distilleries, Database Marksman, DataMind, GVA, IBM Intelligent, Miner, Insightful Miner, Integral Solutions, KXEN, Magnify, MatLab, NeoVista Solutions, Oracle Darwin, Quadstone, SAS, Spotfire, SPSS Clementine, StatPac, Syllogic, WEKA hinkAnalytics, Thinking Machines,
Biyoenformatik Amaçlı	MEME, PIMA, Pratt, PrattWWW, SPEXS

Veri Madenciliği analiz yöntemlerini uygulayabilmek için; WEKA, Keel, R, ARTool, RapidMiner, C4.5, Orange, KNIME vb. açık yazılım kodu kullanılan programlar ile SPSS, SPSS Clementine, SAS, MATLAB, Angoss, SQL Server ve KXEN vb. ticari amaçlı programlar da kullanılmaktadır.

Birçok veri madenciliği bilgisayar yazılımları, OLAP, geleneksel istatistiksel metotlarını (kümeleme analizi, diskriminant analizi ve regresyon analizi gibi) ve geleneksel olmayan istatistiksel metotlarını (sinir ağları, karar ağaçları, birliktelik kuralları gibi) kapsamaktadır. Veri madenciliği analiz yönteminin bu kadar popüler olup ilgi çekmesinin sebebi bilgisayar bilimi, istatistik ve veri tabanı yönetimi gibi disiplinler arası bir çalışmaya imkân sağlamasından oluşmaktadır [42].

Veri madenciliği yazılımları sektörlere göre büyük farklılık göstermese de kurumsal bir yapı için kullanılacak olan bir yazılım seçerken aşağıdaki hususlara dikkat edilmesi gerekmektedir [52-53]:

- Yazılımın uygulayabildiği veri madenciliği teknikleri,
- Yazılımın veri boyutu, kullanıcı sayısı, veri kümesi içindeki özellik sayısı, uygun donanım tipi gibi kapasite kısıtları,
- Veri tabanlarına ve dosyalara erişim imkânı,
- Oluşturulan modellere ilişkin detaylı açıklama desteği,
- Görsel sunum, raporlama formatlarının yetersizliği,
- Farklı yazılımlar ile etkileşim uyumu,
- Danışma ve eğitim desteğidir.

2.2.4 Modelleme

Verilerin modellenmesi aşamasına gelindiğinde, en optimum değerlerin elde edilebilmesi adına farklı farklı modelleme yöntemleri seçilerek modellemeye uygun olan analiz parametreleri kullanılmaktadır. Artık bu noktaya gelindiğinde uygulanması için seçilmiş olan yöntem mevcut veri kümeleri üzerinden çalıştırılarak elde edilen değerler analiz edilmektedir. Bu analiz süreçleri

birbirini tekrarladıkça elde edilen sonuçlar olumlu yöne evrilerek süreçten elde edilecek olan nihai verinin güvenilirliği de artacaktır. Bu noktada analiz için seçilmiş olan veri madenciliği yönteminde kullanılacak olan algoritmaların ve ilgili tekniklerin mevcut model için uygunluğunun tespiti konusuna karar verilecektir. Genelde birbirine benzer veri madenciliği analizlerinde aynı problem analizi için birçok farklı yöntem bulunduğu için yine bu tekniklerin bazıları da özel bir düzenleme istiyor olmasından dolayı eldeki verinin düzenlenmesi gerekebilir. Bu nedenle bu noktalara gelinmiş olmasına rağmen oluşabilecek problemlerde çoğunlukla veri hazırlama aşamasına geri dönüşlerin yapıldığı gözlemlenmektedir [59].

Modelleme yapılırken meydana gelecek olan problemler, sıklıkla geri dönüp düzeltmeler yapılmasına sebebiyet vereceği gibi tüm verilerin en baştan düzenlenmesi de gerekebilir. Bu nedenle veri hazırlığı ve ilgili model kurulumu araştırmacı için tüm analiz sürecinin %50'si ile %85'i gibi bir zaman aralığının harcanmasına sebep olabilmektedir.

Çalışmada kullanılacak olan veriler aşağıdaki aşamalardan geçtikten sonra analiz sürecine geçilebilir.

1. Toplama (Collection)

Tespit edilen sorun için ihtiyaç duyulan veriler ile bu verilerin toplanacağı veri kaynaklarının tespit edilmesi aşamasıdır. Verilerin elde edilmesinde işletmenin veri tabanının dışında aynı alanda çalışmalar yapan diğer kuruluşların verilerinden de yararlanılabilir.

2. Değer Biçme (Assessment)

Analizi yapılacak olan veri topluluğunun farklı birimlerden toplanıyor olması, beraberinde uyumsuzlukları da getirecektir. Bu uyumsuzlukların başlıcaları, ölçüm birimi farklılıkları, hatalara departmanlarca verilen hata kodu gruplarının farklılığı ve veriler üzerindeki birim zaman farklılıkları olabilmektedir. Bundan dolayı mükemmel sonuca ulaşmak için kurulacak olan modellemeler sadece düzgün uyarlanmış veri kümelerinden elde edilebileceği için, modelleme için kullanılacak olan veri topluluklarının ne derece birbirleri ile uyumlu olduklarının bu aşamada değerlendirilmesi gerekmektedir.

3. Birleştirme ve Temizleme (Consolidation and Cleaning)

Bu aşamada farklı birimlerden elde edilen veri topluluklarının bir önceki adımda bahsedilen uyumsuzluklardan arındırılmış bir şekilde ele alınarak derlenmiş ve düzenlenmiş bir veri kümesi halinde tek bir veri tabanında bir araya getirilmiş olması istenmektedir.

4. Seçim (Selection)

Bu aşamada modelleme için uygun olan veri tercihleri gerçekleştirilir. Örneğin müşteri şikayetlerini azaltmak için müşterilerden geri dönüş yapılan veriler incelenmesinin yanı sıra üretim bantlarının da incelenerek sonraki proses bir önceki prosesin müşterisidir mantığı ile hataların müşterilere ulaşmadan firmada tespiti sağlanabilir.

5. Dönüştürme (Transformation)

Büyük veri içerisinde birbirleri ile ilişkili olan veri kümeleri incelenerek tek bir anlam ifade eden bir oluşuma evrilebilirler. Örneğin, yapay sinir ağları ile yapılan bir algoritma uygulamasında değişken kategorisel değerlerin evet ve hayır olması veya yine bir diğer algoritma çeşidi olan karar ağacı uygulamasında kullanılmasının durumundaysa gelir değişkenlerinin değerlerini düşük-orta-yüksek şeklinde gruplandırılıyor olması bu modelin etkinliğini arttıracaktır [76].

Analizleri gerçekleştirmek amacıyla kullanılacak olan veri kümesinin oluşturulmasına ilişkin tüm faaliyetler modelleme aşamasında değerlendirilmektedir. Ayrıca çalışmadaki tablo, değişken ve kayıt seçimi, birleştirme ve temizleme işlemleri gibi faaliyetler de yine bu aşamada gerçekleştirilmektedir. Modelleme aşamasının içeriğinde, modelin gelişim süreci kapsamında kullanılan veri kümesinin oluşturulması için gerekli olan tüm faaliyetler yer almaktadır. Yine bu aşama için önceden belirlenmiş bir kalıp bulunmamakta, mevcut veri kümesi üzerinde tablo, alan ve kayıt seçimlerinin yanı sıra modelleme araçları için veri temizlemesi ve veri dönüşümü gibi işlemler yürütülmektedir. Bu aşama sonucunda, modelleme ya da çalışmanın temel analizleri için kullanılacak veri kümesinin hem oluşturulması hem de tanımlanması mümkün olmaktadır [54].

Veri toplama aşamasında karşılaşılan sorunların yanı sıra, büyük verilerle analiz işlemleri gerçekleşirken öngörülemeyen yeni sorunlarla da karşı karşıya kalınabilir. Bu büyük veri kümelerinin içerisinde kirli veri, gürültülü veri, hatalı veri ve eksik veriler bulunabilmektedir. Çok az da olsa tüm bu uygunsuzluklardan arınmış büyük veri kümeleri ile de karşılaşılabilir. Bu durumda Expectation Maximization algoritması ve bu prensipte çalışmakta olan algoritma çeşitleri vasıtası ile eksik olan veriler için tahminleme veya yerini doldurma yöntemleri

kullanılarak modelin özelliklerini taşıyan yapay veriler üretilerek model tamamlanabilir. Klasik istatistiksel analiz yöntemlerinde de bu tür uygunsuzluklarla karşılaşmakla beraber verilerin az olması problem oluşturmazken veri madenciliği analiz yöntemlerinde büyük veri topluluklarında büyük sorun oluşturabilmektedir. Kısaca veri madenciliği analizinde veri ambarlarının büyüklüğü sebebi ile öngörülemeyen birçok yeni problem ile karşılaşmak bu tür büyük veri analizinin doğasında yer almaktadır [55].

Veri madenciliği analizi yönteminde büyük veri işleminin en çok zaman alan kısmı, mevcut verilerin işlenmeden önce gözden geçirilmesi dönemidir. Bu analiz yönteminde, toplamda harcanacak olan zamanın %80'lik büyük bir kısmı, büyük veri kümesindeki temizlenme aşamaları ve modellemeden önce ön işlemlerden geçirilmesi süreçlerinde harcanmaktadır [56].

Verinin hazırlanma aşamasında veriler veri madenciliği için hazırlanmaktadır. Verinin hazırlanması süreci oldukça uzun sürmektedir. Bu süreç, veri içerisindeki gürültülü verinin temizlenmesi, dönüştürülmesi ve birleştirilmesi aşamalarını da kapsamaktadır [57].

Gelişen teknoloji ve büyüyen firmalar sayesinde büyük veri tabanları oluşmakta ve bu büyük verilerin incelenmesi ve hazırlanması sürecinde şu hususlara dikkat edilmesi gerekmektedir [58].

- Amaca hizmet etmeyecek değişkenler varsa çıkarılmalı,
- Eksik ve hatalı veri girişi yapılmış ise bunlar ayıklanmalı,
- Eksik verilerin sistematik bir hataya yol açıp açmayacağı kontrol edilmeli,
- Birbirine eşdeğer tekrar niteliğinde olan veri alanları temizlenmeli,
- Ekleniecek yeni değişkenin verisini hazırlamak için harcanacak çabaya değer olup olmadığı gözden geçirilmelidir.

2.2.4.1 Birliktelik Kuralları ve Apriori Algoritması

2.2.4.1.1 Birliktelik Kuralı

Birliktelik kuralı, eldeki tespit edilmiş olan veri topluluğu içindeki birlikteliklerin analizi yapılarak gelecek planlamaları için öngörülerde bulunabilme analizidir. 1990 yıllara kadar mevcut satış rakamlarında satılan ürün ve alım yapan müşteri bilgilerinin çok az sayıda ulaşılabilirken, mali yönden gelir verilerinin arşivlenmesi gerçekleştiriliyordu. 1990'lı yıllarla birlikte gelişen teknoloji ile verilerin arşivlenmesi çalışmalarındaki uygulamaların gelişmesi ile birlikte işletmelerin satış yapılan alanlarda kullandıkları son teknoloji ürünler vasıtası ile müşterinin ve ürünlerin tanınması kolaylaşmıştır. Teknolojideki bu tür gelişmeler ile satış anından itibaren satışı gerçekleştirilen ürünler hakkında bilgi toplanması mümkün olmuş ve aynı verilerin elektronik

ortama aktarılması ile şirketlere ait veri tabanları da oluşmaya başlamıştır. Günümüze gelindiğinde ise büyük alışveriş merkezlerinde, süper ve zincir marketlerde ve yine orta ve büyük kapsamlı alışveriş mağazalarında bu tür teknolojik ürünlerin yaygınlaştığı görülmektedir. Tüm bu alışveriş hareketlerinden elde edilen bilgiler, işlemin gerçekleştirildiği saat, işlemin gerçekleştirildiği tarih, alımı yapılmış olan ürün bilgileri yer almakta ve bu işleme ait tanımlanmış olan sipariş numarası o işlem için üretilmiş ve o işleme özel olarak tek üretilmiştir. Bu konulara özen gösteren bazı işletmeler tüm bu bilgilerin bulunduğu veri ambarlarını kullanma yönünde çalışmalar yapmaktadırlar [60].

“Birliktelik kuralları verideki güçlü birliktelik özelliklerini tanımlayan örüntüleri keşfetmek için kullanılan bir analiz yöntemidir. Keşfedilmiş olan örüntüler, özel olarak çıkarılan kurallar veya özellik alt grupları şeklinde temsil edilebilir. Araştırma uzayının üssel büyüklüğünden dolayı birliktelik analizinin amacı önemli örüntülerin (affinity) analizi olarak da anılır”. Bu teknikle birlikte mevcut olayın ortaya çıkışının yanı sıra bu olaya etki eden diğer faaliyetlerin olma olasılıklarının veya ihtimallerinin ölçümü söz konusu olabilmektedir. Bu yöntemle birlikte veri tabanı içerisindeki tüm oluşabilme ihtimali olan örüntülere ulaşılabilmesi de sağlanmaktadır. Kuşkusuz tüm olasılıkların incelenebiliyor olması bu yöntemi avantajlı kılarken, yine büyük ölçekli verilerin analiz edilmesi sürecini ise zorlaştırmanın yanı sıra zaman sürecinin de uzamasına sebep olması dezavantajlı bir yönüdür. Bu yöntem genellikle araştırmacılar tarafından veri içerisindeki birliktelikleri ortaya koymak ve veriler arasındaki ilişkilerin tespiti amacıyla kullanılmaktadır [61].

“Birliktelik kuralları, bir A hareketi (öncül olay) meydana geldiğinde B hareketinin (sonuç) de meydana gelme olasılığını bularak iki hareket arasındaki ilişkiyi tespit etmeye çalışmaktadır. A değerini içeren olayların, B değerini de içermeye durumu, yani birliktelik kuralı, $A \Rightarrow B$ biçiminde gösterilmektedir. Bir birliktelik kuralının kural desteği (s), $A \cup B$ 'yi kapsayan hareketlerin sayısının, veri tabanındaki toplam hareketlerin sayısına (N) yüzde cinsinden oranıdır. Kural desteği (s) = $\frac{A \cap B}{A \cup B}$ Güven (α) ise, $A \cup B$ 'yi kapsayan hareketlerin sayısının, A'yı kapsayan hareketlerin sayısına yüzde cinsinden oranı yani A'yı kapsayan hareketlerin yüzde kaçının B'yi de kapsadığının ifadesidir. Kural desteği (s) ne kadar yüksekse kural o kadar sık tekrar etmektedir; güven (α) ne kadar yüksekse kuralın istisnası o kadar az demektir. Veri setlerinden birliktelik kurallarını çıkaran basit ve iyi bilinen bir algoritma olan Apriori algoritması, Agrawal ve diğerleri tarafından 1994 yılında geliştirilmiş, veri madenciliği tarihinde birliktelik kurallarının çıkarılması konusunda en çok uygulanan algoritma olmuştur. Apriori algoritması, belirlenen bir eşik değerin üzerindeki destek ve güvene sahip birliktelik kurallarının çıkarılmasını sağlar. Tekrarlayan bir niteliğe sahiptir

ve veri tabanlarında sık geçen veri kümelerini bulmak için birçok kez veri tabanını taramak gerekmektedir” [62].

Birliktelik kurallarına ilişkin modelde $I = \{i_1, i_2, \dots, i_m\}$ kümesine nesneler kümesi denmekte ve i 'ler nesneleri oluşturmaktadır. İşlemler kümesi olarak adlandırılan D veri tabanındaki tüm hareketleri gösterirken bir işlemdeki ürünleri gösteren T sembolü de nesnelerin yapmış olduğu hareketleri sembolize etmektedir. “TID her hareketi ifade eden belirteçtir. I kümesindeki bazı iş ya da nesneler seti olan A için bir T işlemler kümesi $A \subseteq T$ ise T , A 'yı kapsıyor denilir. Birliktelik kuralı $A \Rightarrow B$ biçiminde tanımlanabilir ve burada $A \subset I$, $B \subset I$ ve $A \cap B = \emptyset$ olmaktadır. Birliktelik kuralları oluşturulurken kuralların ilginçliğini ve ilgililiğini ifade eden destek ve güven ölçütleri belirlenir. $A \Rightarrow B$ Kuralının destek değeri $A \cup B$ 'nin D işlemler kümesinde bulunma olasılığı iken $A \Rightarrow B$ kuralının güven değeri D işlemler kümesinde A 'yı içeren işlemlerin B 'yi de içermesi olasılığıdır. Başka bir ifade ile $A \Rightarrow B$ kuralı için destek ve güvenilirlik değerlerine yönelik formüller aşağıdaki biçimde gösterilebilir” [61]:

Destek değeri: $P(A \text{ ve } B) = A \text{ ve } B \text{ ürünlerini alan müşteri adedi} / \text{Toplam müşteri adedi}$

Güvenilirlik değeri: $P(A/B) = P(A \text{ ve } B) / P(A) = A \text{ ve } B \text{ ürünlerini alan müşteri adedi} / A \text{ ürününü satın alan müşterilerin adedi.}$

Birliktelik kurallarının oluşturulmasında izlenen yöntemin çeşitliliğine göre her bir sıklıktaki eleman kümeleri için I , hiç elemanı olmayan boş alt kümeleri içinse I_a kullanılırken, destek ve güvenilirlik değerlerinin minimum değerini gösteren sınır değerini sağlayan $I_a \Rightarrow I - I_a$ biçiminde olası kurallar oluşturulur. Güvenirlilik değer oranının %100 olması halinde uygulanan kurallar kesin kural adını alırken, aynı zamanda yapılan tüm veri analizlerinin de doğru oldukları kabul edilmektedir. Birliktelik kuralı veri madenciliğine yönelik geliştirilen algoritmalara örnek olarak AIS, SETM, Apriori, Partition, Rapid Association Rule Mining (RARM), CHARM verilebilir. Bunlar arasında en yaygın kullanılan ise önsel (Apriori) algoritması olmaktadır [63].

Agrawal ve arkadaşları birliktelik kuralının matematiksel model çalışmalarını 1994 yılında ortaya koymuşlardır. Bu modellemede $I = \{i_1, i_2, \dots, i_m\}$ kümesine “ürünler” denmektedir.

Birliktelik kuralı şu şekilde tanımlanabilir;

$$A_1, A_2, \dots, A_m \Rightarrow B_1, B_2, \dots, B_n$$

Bu ifadede yer alan, A_i ve B_j , yapılan iş veya nesnelerdir. Bu kural, genellikle “ $A_1, A_2, \dots, A_m$ ” hareketleri oluştuğunda, sıklıkla “ $B_1, B_2, \dots, B_n$ ” hareketlerinin aynı olayın ve hareketin içerisinde bulundukları anlaşılmaktadır [77]. Birliktelik kuralları analizi araştırmacı tarafından belirlenmiş olan minimum değerlerin güvenilirlik ile destek değerlerinin sağlanacağı bir şekilde üretilmektedir. Analiz aşamasındaki bir modelde işlenen veri kümesinin içindeki destek göstergesi, D harfi ile ifade edilmekte olan bütün oluşumları meydana getiren sonuçlar kümesini oluşturan bütün hareketlerin yüzdesel oranıdır. Eğer A ve B kümelerini oluşturan materyallerin birliktelik kuralı, “ $A \Rightarrow B$ ” şeklinde gösterilirse, kümelerle ilişkin olan, destek kümesinin gösterimi aşağıdaki gibi olacaktır.

$$\text{Destek } (A \Rightarrow B) = (\text{A ve B'nin bulunduğu satır sayısı}) / (\text{toplam satır sayısı})$$

Açıklayacak olursak, $A \Rightarrow B$ birliktelik kuralında güven değerini ifade etmekteyse, A’yı meydana getiren tüm oluşumların aynı zamanda B’yi de meydana getirme olasılığının yüzdesidir. Örnek olarak sonuçlanmış olan bir model de güven değeri %85 olarak bulunduyorsa A’yı oluşturan tüm değerlerin hepsi %85 oranında B’yi de oluşturmaktadır.

Oluşturulan model sonucunda güven değeri oranı %100 ise, yazılan kuralın tüm veri analizi için doğru olduğu kabul edilir ve yazılan bu kurallar için “kesin” denilir. Büyük veri tabanlarından verilerin sağlanması konusunda kullanılan algoritmaların birçoğu art arda sıralanacağı gibi birbirine paralel olarak da kullanılabilir. Art arda kullanılmakta olan algoritmalarda içerisinde ürünlerin küme olarak meydana getirildiği ve sayılarak mantıksal ifadelerle dönüştürüldüğü ardışık sistemlerdir. Paralel şekilde kullanılan algoritmaların yapısında ise büyük veri tabanlarının paralellik oluşturularak meydana getirilmesi amaçlayan bir yapıya sahiptir. Birliktelik kuralının iletilmesi amacıyla geliştirilmiş ve kullanılmakta olan bazı algoritma çeşitleri aşağıdakilerdir;

AIS, SETM, Apriori, Partition, RARM - Rapid Association Rule Mining, CHARM. Bu algoritmalarından ilk olarak kullanıma çıkan AIS ve yine bu algoritmalarından en çok bilinen ve kullanılan algoritmayla Apriori algoritmasıdır [63].

2.2.4.1.2 Apriori Algoritması

Veri madenciliği analizi için veri kümelerinin ve bu kümeler arasındaki bağıntının ortaya konulması için birçok algoritmaya ihtiyaç duyulmakla birlikte günümüzde bu amaca hizmet edecek algoritmalar geliştirilmiş ve geliştirilmeye devam etmektedir. Geliştirilen teknolojiler de verilerin analiz edilmeleri ve işlenerek sonuca ulaşılma süreci araştırmacıların çok zamanını almakta ve aynı zamanda birçok veri kümelerine ihtiyaç duyulmaktadır. 1994 yılında Agrawal ve Srikant veri analizi konuları için Apriori Algoritması geliştirmişlerdir. Veri madenciliği analizi çalışmalarındaki birliktelik kurallarının bulunması için bilinen yöntem Apriori Algoritmasıdır. Bu algoritma ile çok büyük veri tabanları üzerinde yapılmakta olan veri madenciliği analizleri gelişme göstermiştir. Genellikle birliktelik kuralının çıkarım yapılması alanında çok kullanılmaktadır. Bu algoritmada amaç ise veri kümeleri içerisinde bulunan veri toplulukları arasındaki bağlantıları tespit ederek ortaya çıkarmaktır. Algoritma da ismini, kendinden önce tespit edilen sonuçlara bağlı olarak çalıştığı için Latince “önce” anlamına karşılık gelen “prior” kelimesinden almaktadır. Algoritmada yapısal çalışma aşağıdan yukarıya doğru mantığı ile çalışmakta olup her aşamada bir tek element incelenmekte ve bu elementin diğer veriler ile olan ilişkilerinin ortaya konması için çalışılmaktadır [64].

Bu algoritmanın diğer algoritmalarından farklı yönleri model için aday olan nesne kümelerinin üretilmesinde ve yine sayılacak olan nesne gruplarının seçilmesindeki farklılıklardır. Daha önce de belirtildiği gibi hem AIS hem de SETM algoritmalarında önceki sonuca bağlı olarak yaygın nesne kümelerinin birbirleri arasındaki benzer nesne kümeleri ve aynı hareketi gösteren diğer nesneler tespit edilmektedir. Bu benzer nesne kümeleri hareket halinde olan diğer serbest nesneler ile büyümeleri sağlanarak yeni aday nesne kümeleri oluşturulmaktadır. Yine de serbest olan nesnelerin yaygın olmadıkları gözlemlenmiştir. Bilindiği üzere genişlemiş bir nesne kümesinin süper kümesiyle ufak bir nesne kümesi, ufak bir nesne kümesi sonucunu oluşturur. Bu yöntemlerle ufak ölçekli yeni çok fazla aday nesne kümesi meydana getirilebilir. Apriori algoritması da bu tespit edilen değerli alan üzerinde çalışmaktadır. Bu bağlamda mevcut algoritma, önceki geçişten elde edilmiş olan geniş nesne kümelerini bir araya getirerek oluşan aday nesne kümelerini meydana getirir ve veri tabanında oluşan oluşumları göz ardı ederek, önceki geçişte meydana gelen küçük alt kümeleri silerek yoluna devam eder. Önceki geçişlerin oluşturduğu geniş

nesne kümelerinin irdelenmesiyle geniş aday nesne kümelerinin sayısının da anlamlı bir azalma olduğu gözlenmektedir. İlk aşamada yalnızca bir nesne içeren nesne kümeleri dikkate alınır. İlk aşamada tespit edilen geniş nesne kümelerinden `apriori_gen ()` fonksiyonu kullanılarak, diğer aşamada kullanılacak olan aday kümeler oluşturulmaktadır. Aday nesne kümeleri elde edildiğinde veri tabanı işlenerek iki uzunluklu geniş nesne kümelerini tespit etmek için destek değerleri sayılır. İkinci aşamadaki geniş yaygın nesne kümeleri, sonraki aşamadaki geniş nesne kümelerini meydana getirecek olan aday kümeler olarak incelenir. Bu işlem yeni bir geniş küme elde edilemeyinceye kadar devam ettirilir. Algoritma çalışırken her bir i . geçişi yaptığında veri tabanında tarama yaparak i uzunluğundaki geniş nesne kümelerini saptar. L_i ; i uzunluğundaki yaygın nesne kümelerini, C_i de i uzunluğundaki adayları ifade eder. Apriori de seviye mantığı araması olarak da bilinmekte olan tekrarlamalı bir yöntem kullanılmaktadır. Bu yöntemde k -nesne kümeleri, $(k+1)$ şeklinde ifade edilmekte olan nesne gruplarının analiz edilmesi amacıyla kullanılmaktadır. Çalışmalarda öncelikle, 1-nesne gruplarının kümeleri bulunur. Tespit edilen bu kümeye L_1 adı verilir. Burada L_1, L_2 'nin yani 2-nesne gruplarının kümelerinin bulunabilmesi için kullanılmaktadır. Yine aynı şekilde L_2 burada L_3 'ün bulunabilmesi amacıyla kullanılırken algoritmanın çalışma mantığı birbirini tekrarlayarak daha fazla k -nesne kümelerinin tespitinin sağlanmasını amaçlamaktadır. Elde edilecek olan her bir L_k verisi mevcuttaki tüm veri tabanının taranarak elde edilmiş olduğunu gösterecektir [65].

“Ayrıca algoritmanın her eleman için çalışmasını, bir arama algoritmasına benzetmek mümkündür. Algoritma, bu anlamda sığ öncelikli arama (breadth first search) yapısında olup, sanki adayları birer ağaç (tree) gibi düşünerek bu ağaç üzerinde arıyor kabul edilebilir. Ağaç yapısında, k elemanlı bir aday listesinden $k-1$ elemana baktıktan sonra, alt frekans örüntüsü yetersiz olan elemanları budamakta ve kalan elemanların üzerinden arama yapmaya devam etmektedir. Klasik apriori algoritması aşağıda yazıldığı gibidir” [63]:

- 1) $L_1 = \{\text{sık geçen 1-öge kümesi}\};$
- 2) for ($k=2$; $L_{k-1} \neq \emptyset$; $k++$) do begin
- 3) $C_k = \text{apriori-gen}(L_{k-1});$ // Yeni adaylar
- 4) for all transactions-hareketler $t \in D$ do begin
- 5) $C_t = \text{subset}(C_k, t);$ // Adaylar t içindedir
- 6) for all candidates – adaylar $c \in C_t$ do
- 7) $c.\text{count}++;$
- 8) end

- 9) $L_k = \{c \in C_k \mid c.count \geq minsup\}$
- 10) end
- 11) Answer = $\cup_k L_k$;

“Aşağıdaki şekilde yer alan apriori-gen fonksiyonu (k-1) adet öğeye sahip L_{k-1} sık geçen öge kümesini kullanarak k adet öğeye sahip aday kümeleri oluşturur. Bu fonksiyon ile, ilk önce, L_{k-1} sık geçen öge kümesine kendisi ile birleştirme (join) işlemi uygulanır. Birleştirme işleminde L_{k-1} sık geçen öge kümesinin her satırında yer alan son öge haricinde diğer öğelerin çapraz olarak benzerliği aranır ve son öge haricinde diğer öğelerle yakalanan benzerliklerden yeni aday öge kümeleri oluşturulur. Oluşan kümeler budama (prune) adımı ile budanarak fonksiyondan dönülür. Apriori-Gen Fonksiyonu aşağıdaki gibidir” [63]:

```

insert into Ck
  select p.items1, p.items2,.....,p.itemsk-1, q.itemk-1
  from Lk-1 p, Lk-1 q
  where p.item1=q.item1, ..., p.itemk-2=q.itemk-2,
        p.itemk-1 < q.itemk-1;

forall itemsets c  $\in$  Ck do
  forall (k-1)-subsets s of c do
    if (s  $\notin$  Lk-1) then
      delete c from Ck

```

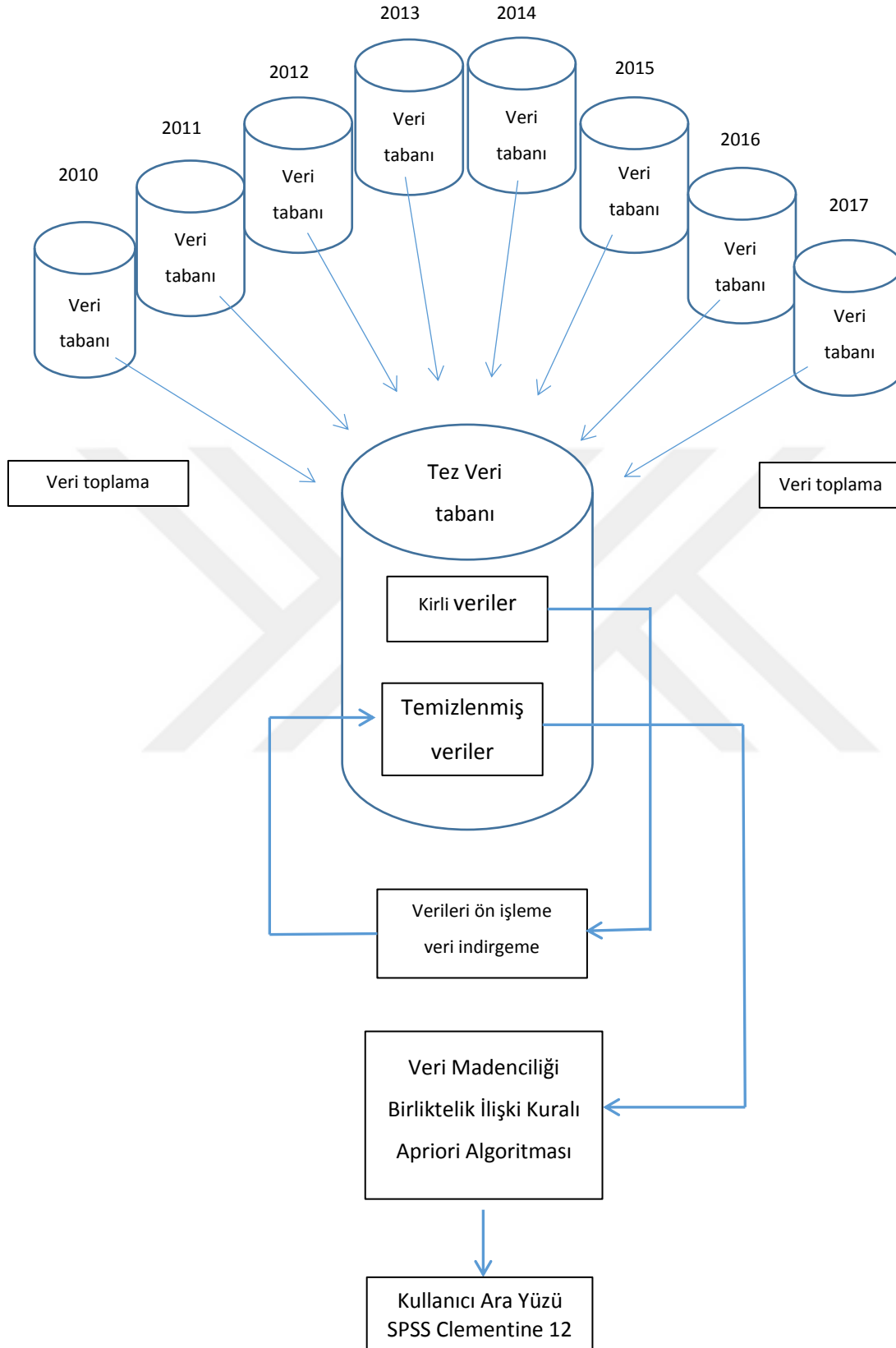
“Budama işleminde; c aday kümesinin (k-1) öğeye sahip alt kümelerinden L_{k-1} sık geçen öge kümesinde yer almayan tüm alt kümeler silinir. Farklı bir ifade ile budama, C_k aday öge kümesindeki öğelerin alt kümelerinin L_{k-1} sık geçen öge kümesindeki varlığı kontrol edilir, bir öğenin alt kümelerinden biri, L_{k-1} sık geçen öge kümesinde yer almıyorsa ilgili öge değerlendirme dışı kalır ve C_k aday öge kümesinden silinir” [63, 66].

Özel kural formundaki bu kurallar “sol ve sağ kısım olmak üzere birbiriyle bağlantılı iki kısımdan oluşur. Bu iki kısımda yapılan iş veya nesneler yer alır ve veriler arasındaki ilişkiler eğer-sonra ifadeleri vasıtasıyla gösterilir”. “Eğer kısmı ile ilgili durumlar öncül ve sonra kısmı ile ilgili durumlar sonuç olarak ifade edilir. Birliktelik kurallarında önemli olan konu gürültülü veriden

değerli bilgiyi ayırt etmeyi sağlayan eşik değerini bulmaktır”. Bu amaçla ilginç birliktelik kurallarından ilginç olmayanları ayırmakta destek (support) ve güven (confidence) denilen bu gruplardan yararlanılmaktadır. “Birliktelik kurallarının güçlü yönleri açık ve anlaşılır sonuçlar üretmesi, yönlendirilmemiş veri madenciliğini desteklemesi, çeşit sayısı fazla ve büyük miktardaki veri üzerinde çalışabilmesi, hesaplamaların anlaşılma kolaylığı olurken zayıf yönleriyse problem boyutunun hesaplama karmaşıklığı nedeniyle artması, doğru özellik sayısının bulunmasındaki güçlük, nadir görülen özelliklerin göz ardı edilmesi olarak sıralanmaktadır” [61].

Agrawal ve arkadaşları 1994 yılında Apriori Algoritması ile veri gruplarındaki birbirleri ile ilişkili durumları ortaya koyan ve kolayca kullanılabilen bir yöntemi geliştirmişlerdir. Bu yöntem tüm zamanlarda veri madenciliği kapsamında birliktelik ilişkisinin tespitinde kullanılan en yaygın algoritma olmuştur. Bu algoritma kullanıcı tarafından belirlenen bir sınır değer üzerinde oluşan güven ve desteğe sahip olan birbirleri ile ilişkili olan durumların ortaya koyulmasını sağlamaktadır. Kendini tekrarlayabilen bir özelliğe sahip olan bu algoritma ile büyük veri kümeleri içerisinde tekrarlayan kümeleri bulabilmek adına mevcut verileri tekrar tekrar analiz etmek gerekebilmektedir [62].

Tezde gerçekleştirilecek olan veri madenciliği aşamaları aşağıdaki şekilde görülmektedir.



Şekil 2-4: Veri toplama ve işleme aşamaları

Yukarıdaki şekilden de anlaşılacağı üzere üretici veri tabanında bulunan veriler yıllara göre düzenlenmiştir. Her bir yıl için o yıl içerisinde çıkmış olan hatalar gruplandırılarak sorumlu departmanlara göre analiz yapılmıştır. Düzenlenmiş olan bu veriler çalışma veri tabanı adı verilen ortak bir dosyada toplanmıştır. Toplanan bu veriler kirli verilerden ayıklanarak düzenlemeleri yapılmıştır.

Toplanan veriler şunlardır:

- Kontrol tarihi,
- Kontrol yapılan gün,
- Kontrol yapılan ay,
- Özel gün,
- Mevsim,
- Araç model adedi,
- Vardiya adedi,
- Çalışma şekli,
- Yeni personel,
- Ramazan ayı,
- Hatalı araç modeli,
- Hatayı bulan vardiya,
- Hatayı bulan vardiyanın çalışma şekli,

Bu toplanan veriler etkenlerine göre oluşturulacak excel tablosunda analizleri yapılmıştır.

Bulunmuş olan hata sayılarının hangi departmana göre oluştuğunun kırılımı yapıp daha sonra da genel hata adetlerinin oluşum sebepleri araştırılmıştır. Bu araştırma hataların oluşumunda bir birliktelik sebebinin var olup olmadığını varsa da hata oluşumuna etki eden faktörlerin ne derece güvenilir olduğunu ve bu hata oluşumunu ne derecede desteklediğinin analizini yapmamıza olanak sağlayacaktır.

Çalışma kapsamında incelenecek hata türleri aşağıdaki şekilde verilmiştir:



Şekil 2-5: Hata Türleri

Hata gruplarında meydana gelen hataların oluşumunda yukarıda sıraladığımız etkenlerin ne derecede birliktelik gösterdikleri sonucuna varılmaya çalışılacaktır. Böylece zaman içerisinde yaşanan değişimlerin hata oluşma olasılığına etkisi hakkındaki sonuçlar yorumlanacaktır.

Bu aşamada farklı dönemlerin verileri, farklı kaynaklardan toplanarak tek bir veri tabanı oluşturulması sağlanmıştır. Kolay yollarla ve baştan savarak yapılan iyileştirme faaliyetlerinin ilerleyen süreçlerde toparlanarak gelen daha büyük ölçekli sorunlara sebebiyet vereceği düşünülmüş ve veri tabanı hazırlanmıştır.

Çalışmanın amacı doğrultusunda gürültülü veriler temizlenirken, işletmenin çeşitli aksiyonları aldığı zaman aralığı seçilmiştir. Bu zaman aralığında işletme tekli vardiya, ikili vardiya ve üçlü vardiya olmak üzere 3 çalışma şekli sergilemiş, aynı zamanda üretilen araç modellerinde de bir model, iki model ve üç modeli aynı anda üretmiştir.

2.2.4.2 Pareto Analizi

Pareto analizi tekniđi; önem derecesi en büyük olan noktalara dikkat çekmek için kullanılan bir analiz tekniđidir. 1900’lü yıllardan sonra ünlü İtalyan iktisatçı Vilfredo Pareto adıyla anılmaya başlayan Pareto analizi, eldeki tüm vakaların (hata, talep, maliyet vb.) içindeki verilerin diğere verilere göre daha fazla orana sahip olanı göstermektedir. Bu analiz tekniđi, mevcut durumları önemine göre sıralayarak oranlamayı sağlamakla birlikte, oransal anlamda daha küçük orana sahip olan problemleri terk ederek en büyük orana sahip probleme odaklanarak çözüme kavuşturmayı hedeflemektedir [67].

Normal dağılımlar incelendiğinde “nedenlerin en önemli %20’si, sonuçların %80’ini; sonraki %30’u, sonuçların %15’ini ve kalan %50’si ise sonuçların yalnızca %5’ini ortaya çıkarmaktadır. Maliyetlerin neredeyse %80’ninin elemanların yalnız %20’sinden kaynaklandığı veya servetin yaklaşık %80’ninin nüfusun %20’sinin elinde olması bu duruma örnek gösterilebilir. Pareto prensibine söz konusu oranlar nedeniyle literatürde “80-20”, “90-10” kuralı veya “70-30” kuralı da denmektedir”. Literatürde ABC analiz tekniđi olarak da adlandırılmakta olan bu analiz tekniđi yapılacak olan analizin önceliklendirilmesi amacıyla kullanılmaktadır [68].

Hatalar, harcanacak olan iyileştirme masrafları ve üretilen ürünün performansına etki eden faktörler olarak gruplandırılabilir. Bu aşamada Pareto analizi devreye girerek, hangi kusur gurundan analize başlanacağı konusunda öncelik sırası belirleyerek yol gösterici bir durum sergilemektedir. Bu analiz yöntemi analizci ve araştırmacılar tarafından yapılacak olan faaliyetlerin en verimli bir şekilde hangi alan da olması gerektiği konusunda yol göstericidir. Yine aynı şekilde bu analiz yöntemi grafiksel olarak hataları büyük oranlardan küçük oranlara doğru sıralayarak aksiyonların alınması konusunda başlanacak noktayı işaret etmektedir [69].

Analizi yapılan hata gruplarının hepsi, probleme etki etmeleri açısından aynı oran dağılımına sahip değillerdir. Bu etki oranlarının ve bu oranların sayısal ölçeklendirilmesinin gösterimi konusunda pareto grafiđi oluşturulmaktadır. Bir anlamda Pareto grafiđi, önem bakımından küçük sayılardaki sorunları, önemsiz olarak analiz edilen çok sayıdaki sorunlardan ayırt etme tekniđi olarak da tanımlanmaktadır. Bir üründe ortaya çıkabilecek bütün hataların aynı öneme sahip olduğu söylenemez. Pareto analizinin amacı farklı sayıdaki birçok nedenden, daha yüksek öneme sahip nedenleri ayırmaktır. Pareto analizi ile bir olay grafik yardımıyla

gösterilebilmektedir. Böylelikle dikkati, yaşanan problemin en önemli nedeni üstüne toplamakta ve önceliklerin belirlenmesini sağlamaktadır. Bu özelliklerinden dolayı ekonomi dışındaki tüm alanlarda da kullanılabilir. Eldeki verilen analizinden sonra başlanacak kalite iyileştirme faaliyetlerine başlama noktasında, hangi hata grubunun ele alınacağı grafikten elde edilen büyük yüzde oranına sahip olan hata grubu olduğunun tespiti konusunda yol göstermektedir [70].

Hatalar A, B ve C olarak 3 gruba ayrılmaktadır. A grup hatalar yüksek değere sahip az sayıdaki problemlerdir. Toplam miktarın yalnızca %15-20'si olmalarına rağmen toplam değer %70-80 gibi önemli bir kısmını oluştururlar. B grubu hatalar hem değer hem miktar olarak ortalama gruptur. Toplam hataların %30-40; toplam değer ise %15-20'lik kısmıdır. C grubu hatalar ise değerce düşük çok sayıdaki problemten oluşur. Toplam miktarın %40-50'sini oluşturmalarına rağmen toplam değer sadece %5-10'luk kısmı olduklarından bu hataların ortaya çıkmaması için alınması gereken önlemler basit ve seyrek tutulabilir. Bu bağlamda, maliyet giderlerini düşürmek isteyen bir firmanın ekonomik anlam taşıyan A grubu hataların ortadan kaldırılmasına odaklanması gerektiği sonucu çıkarılabilir. B ve C gruplarının üzerindeki işlemler ise mümkün olduğunca yalınlaştırılmalıdır [71].

2.2.4.2.1 Pareto Diyagramı Kullanım Alanları

Pareto analizi araştırmacılar ve analizciler tarafından birçok konuda kullanılarak sonuç elde edilebileceği konusunda hatırı sayılır bir yanı olsa da genele bakıldığında Pareto analizinin yaygın bir şekilde kullanılmadığının da göz ardı edilmemesi gerekmektedir. İstatistiksel proses kontrolünün yaratıcısı sayılan Kaoru Ishikawa'ya göre; Pareto analizi hatalar içerisinde daha fazla zarar yol açan hataların kolayca tespitini sağlarken, düzeltici, önleyici faaliyetler programından alınan sonuçların takip edilmesinde de kullanılabilir. Örneğin ilk paretoda yer alan hatalar için çözüm üretildikçe hataların sıralaması da değişir. Bu şekilde kalite elemanı grafiğe baktığında iyileşmenin ne seviyede olduğunu rahatça görebilir [72].

Pareto analizi; işletme içerisinde enerji tasarrufu sağlamada, verimliliğin artırılması aşamasında, üretim hatlarında oluşan duruş sebeplerinin analiz edilmesinde, mevcut stokların kontrol ve takibinin yapılmasında, ürün satış ve pazarlamasının yapılmasında, işletmenin güvenliğinin sağlanması noktasında, satın alma süreçlerinin takibinde, ürünlerin lojistik ve dağıtımların takibinde, ürünlerin satışlarının takibinde, fire edilen ürün tabinde, ürün çeşitliliğinin

takibinde ve ürün temin etme süreci gibi bir çok konuda analiz yapabilme ve bu analiz sonuçlarına göre elde edilen faydalı durumların hayata geçirilmesi konusunda bilgi sahibi olmamızı sağlamaktadır [73].

2.2.4.2.2 Pareto Diyagramı Çizilmesi

Çeşitli kategoriler içeren bir süreçte yetersizlikler, maliyet ve hata çeşitleri gibi farklı kategorilerin göreceli önemini göstermek için bir Pareto diyagramı tasarlanmıştır. Pareto diyagramının çizimi altı aşamada özetlenebilir [72]:

Aşama 1: Tüm elemanların listelenmesi: İlk olarak sürece etki eden etkenler belirlenmelidir. Bu etkenler tüm sağlayan unsurlar araştırılır, toplanır ve listelenir. Etkenler detaylı şekilde incelenmeli ve tüm yönleri ile iyi tanınmalıdır.

Aşama 2: Elemanların ölçümü: Etkenlerin süre aralığında gösteren bir form hazırlanır. Belirlenen süre zarfındaki hata adetleri ve sorumluları tabloya işlenir.

Aşama 3: Elemanların sıralanması: Bu aşamada elde edilen veriler değer olarak büyükten küçüğe doğru sıralanır.

Aşama 4: Kümülatif dağılımların hesaplanması: Elemanlar sıralandıktan sonra bütün veriler toplanır. Ardından her verinin toplam veri içindeki yüzdesi hesaplanmaktadır. Sonrasında ise elde edilen yüzdelik değerlerin kümülatif toplam değerleri bulunarak elde edilmiş olan son verilerle birlikte Pareto grafiğinin çizimi gerçekleştirilir.

Analiz yapılan süre aralığında üretim hatlarına ait tespit edilen hataların kümülatif hesaplanması Tablo 2.2’de gösterilmektedir.

Tablo 2-2: Verilerin büyükten küçüğe doğru kümülatif dağılımları

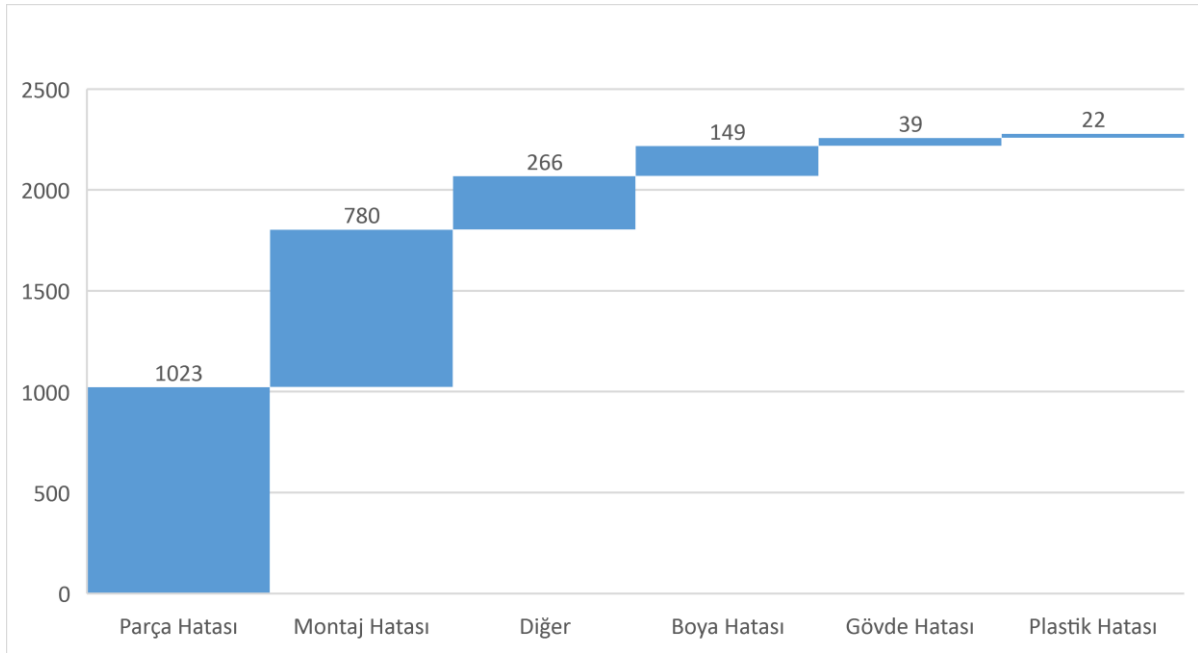
Sorumlu	Hata Sayısı	Yüzde	Kümülatif Yüzde
Parça Hatası	1023	45	45
Montaj Hatası	780	34	79
Diğer	266	12	91
Boya Hatası	149	6	97
Gövde Hatası	39	2	99
Plastik Hatası	22	1	100
Toplam	2279	100	

Aşama 5: Pareto grafiğinin çizimi: Ele alınan sorun için tespit edilen nedenler yatay eksen doğrultusunda önemlilik seviyelerine göre ve birbirine eş mesafelerle sütun şeklinde yazılır. Sorunların en büyüğü grafiğin en soluna eklenir. Grafiğin sağına doğru sorunun nedenleri önem derecelerine göre azalan sırayla konumlandırılır. Önemsiz sütunlar genellikle birleştirilip grafiğin en sağına “diğerleri” ifadesiyle yerleştirilir. Bu noktada önemsiz görünen fakat büyük bir öneme sahip sütunları gözden kaçırmamaya dikkat edilmelidir. Ayrıca hata çeşidi arttıkça elde edilecek fayda azalır. Kontrollerin yapılması zor hale gelebilir. Bir ürün için normal şartlarda 3-6 hata sayısı belirlemek yeterlidir. Yüzdelik değerler ile mevcut kusur adetleri dikey eksen üzerinde yer almaktadır. Her bir hatanın toplam sayısının bütün hataların toplam sayısına oranı sütunların uzunluğunu belirlemektedir. Pareto grafiğindeki toplam eğri grafiğin sol alt köşesinden başlayıp sütun boylarını toplayarak sağ üst köşedeki %100 değerine ulaştığında tamamlanır. Bu eğri genelde öncesini ve sonrasını karşılaştırmak amacıyla kullanılır [71].

Dolayısıyla her grafikte pareto grafiği çizilmeyebilir. En sağda hata yüzdelerini gösteren 0’dan 100’e olacak şekilde dik bir eksen çizgisi çizilir. Hata yüzdesi aşağıda verilen denklem kullanılarak kolaylıkla hesaplanabilir;

$$\text{Hata Yüzdesi} = \frac{\text{Sütundaki Hata Sayısı}}{\text{Toplam Hata Sayısı}} \times 100$$

Tespit edilen hata sayıları için örnek Pareto grafiği Şekil 2-6’da gösterilmektedir.



Şekil 2-6: Hata sayıları için Pareto diagramı

Aşama 6: Pareto grafiğinin yorumlanması: Pareto diagramın da amaç; hangi sorunlar üzerinde daha çok odaklanması gerektiğinin gösterilmesidir. Grafik üzerinde tepiş edilen en büyük hata oranının gösterildiği sütunun çözüme ulaştırılması, yine aynı grafik üzerindeki daha küçük kusur gruplarının çözüme ulaştırılmasından daha karlı olacağının göstergesidir. Grafik yorumlanırken araştırmacıların dikkat etmesi gerekli olan en önemli husus, grafik üzerindeki en çok hatayı gösteren sütunun bize her zaman büyük problemi vermediğini, bir sonraki sütunun çözümünün bizleri istenilen sonuca daha kolay ve zahmetsizce olabileceği ihtimalini de göz önünde bulundurmalarıdır. Pareto diagramının verdiği sonuçlar kalite kontrol uygulamaları arasında en kullanışlısı olmasına rağmen, grafiği oluşturan verilerin okunması ve yorumlanması analizi yapan analizcinin bilgi, tecrübe ve deneyimine kalıyor olması sistemin handikapları arasındadır [72].

Hata grupları yorumlandığında Parça Grubu hataların buradaki yüzdesi en büyük hata olduğu görülmektedir. Toplamın %45’ini oluşturan bu hatanın nedenleri araştırılıp başka bir grafik tasarlandığında belki de bu hatanın çözümü için gerekli olan maliyet ve zamanın çok büyük olduğu ve bunun sonucunda ikinci hata olarak grafikte görülen hatanın düşük bir bütçe ile çok daha az bir zaman diliminde çözümlenebileceği görülebilir.

2.2.4.3 Apriori Algoritması Adımlarının Veri Setine Uygulanması

Çalışmamızın bu aşamasında, birliktelik kurallarının tespit edilmesi işlemi için veri madenciliği uygulamalarının kullanılması amacı ile geliştirilmiş olan SPSS Clementine 12.0 paket programından faydalanılmıştır. Veri tabanının oluşturulmasında ise Microsoft Excel çalışma sayfası kullanılmıştır. Tesiste elde edilen veriler arşivlenmek suretiyle saklanmakta olup, düzenlemeler yapıldıktan sonra kullanılacak son hale getirilmiştir.

Elde edilen veriler üzerinde yapılan düzenlemelerle, Ocak 2010 ile Aralık 2017 yılları arasındaki sekiz yıl boyunca örnekleme kontrol yapılan kalite kontrol deparmanı verileri düzenlenerek 2279 satır kusur verisi elde edilmiştir.

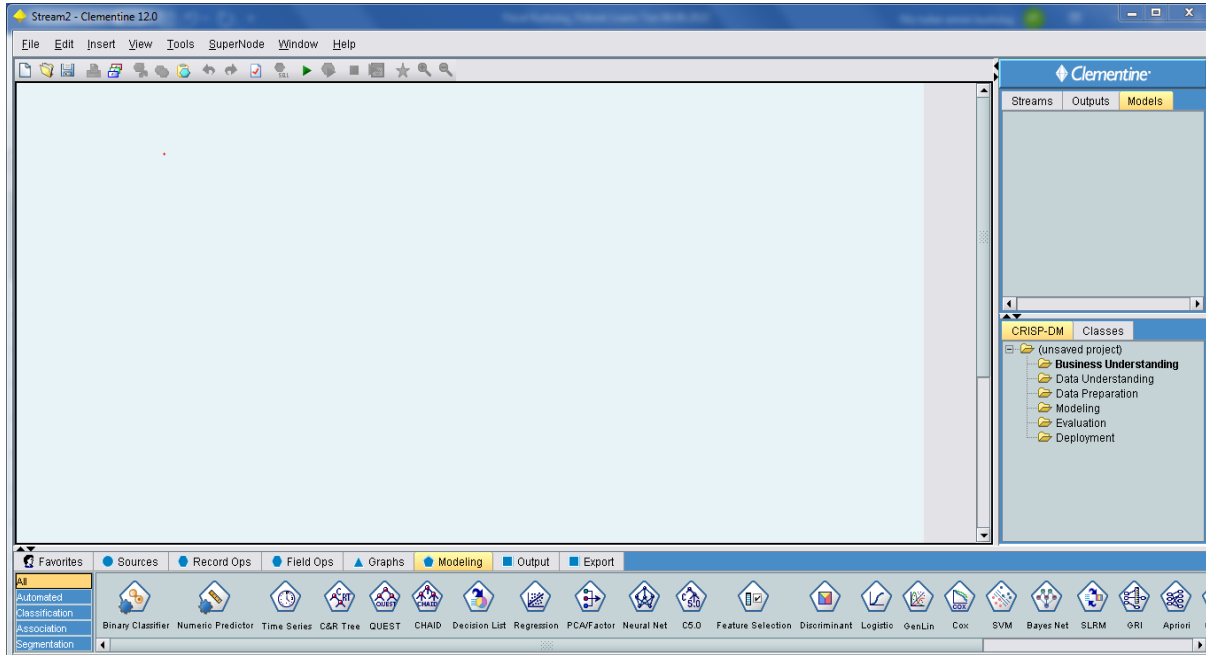
SPSS Clementine 12.0 programında birliktelik kurallarının tespiti için kullanılacak olan arındırılmış veri setinin kesiti Şekil 2.7’ de gösterilmiştir.

	A	B	C	D	E	F	G	H	I	J	K
1	Sorumlu	Kontrol Tarihi	Gün	Ay	Özel Gün	Model Adedi	Vardiya Adedi	Yeni Personel	Hata Bulan Vardiya	Hatalı Araç Modeli	Hata Bulan Vardiya Çalışma Saati
2	Boya	2010	Perş.	Temmuz	Normal	2	2	Yok	Kır.	AU	2
3	Boya	2010	Cuma	Temmuz	Normal	2	2	Yok	Kır.	AU	2
4	Boya	2010	Ptesi	Ağustos	Normal	2	2	Yok	Bey.	AU	2
5	Boya	2012	Ptesi	Ağustos	Normal	2	2	Yok	Bey.	AU	2
6	Boya	2010	Ptesi	Ağustos	Normal	2	2	Yok	Bey.	AU	2
7	Boya	2010	Cuma	Temmuz	Normal	2	2	Yok	Kır.	AU	2
8	Boya	2010	Ptesi	Ağustos	Normal	2	2	Yok	Bey.	AU	2
9	Boya	2010	Ptesi	Ağustos	Normal	2	2	Yok	Bey.	AU	2
10	Boya	2010	Ptesi	Ağustos	Normal	2	2	Yok	Bey.	AU	2
11	Boya	2012	Ptesi	Ağustos	Normal	2	2	Yok	Bey.	AU	2
12	Boya	2011	Cuma	Mart	Normal	2	2	Yok	Kır.	CO	3
13	Boya	2011	Cuma	Mart	Normal	2	2	Yok	Kır.	CO	3
14	Boya	2011	Cuma	Mart	Normal	2	2	Yok	Kır.	CO	3
15	Boya	2011	Cuma	Mart	Normal	2	2	Yok	Kır.	CO	3
16	Boya	2011	Cuma	Mart	Normal	2	2	Yok	Kır.	CO	3
17	Boya	2011	Çarş.	Mayıs	Normal	2	2	Yok	Bey.	AU	3
18	Boya	2011	Cuma	Ağustos	Tatil Son.	2	2	Yok	Kır.	CO	3
19	Boya	2011	Cuma	Ağustos	Tatil Son.	2	2	Yok	Kır.	CO	3
20	Boya	2011	Cuma	Ağustos	Tatil Son.	2	2	Yok	Kır.	CO	3
21	Boya	2013	Salı	Şubat	Normal	1	1	Yok	Kır.	AU	2
22	Boya	2013	Ptesi	Mart	Normal	1	1	Yok	Kır.	AU	2
23	Boya	2013	Ptesi	Mart	Normal	1	1	Yok	Bey.	VE	3
24	Boya	2013	Ptesi	Haziran	Normal	2	2	Yok	Kır.	VE	3

Şekil 2-7: Microsoft Excel formatında düzenlenen arındırılmış veri seti alıntısı

SPSS Clementine 12.0 paket programı üzerinde çalışmaların yapılacağı sayfa Stream Canvas (Akım Tualî) olarak adlandırılır. Bu alan görsel programlama metotları aracılığıyla veri

madenciliği yapılmasına imkân tanımaktadır. Clementine’in açılış arayüzü Şekil 2.8’de gösterilmiştir.

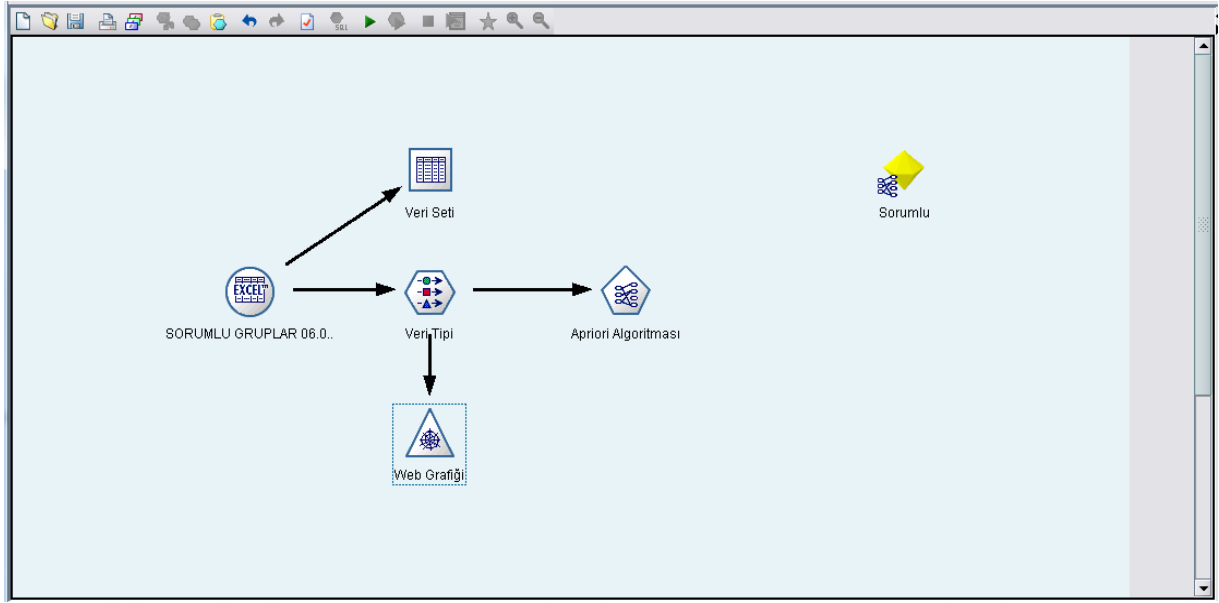


Şekil 2-8: Clementine programı açılış arayüzü görüntüsü

Görsel programa metoduunda yer alan düğümler veri üzerinde yapılacak işlemleri ifade etmekte ve her palet kendisiyle ilişkili birkaç düğüm içermektedir. Örnek olarak “Sources” paleti içerisinde, verilerin modele eklenmesi için ihtiyaç duyulan düğümler yer almaktadır. Düğümler, akım tualine yerleştirildikten sonra birbirine bağlanarak akımlar (streams) oluşturulur. Akımlar, düğümlerden veri akışını simgelemekte ve her akım bir çıktı (output) veya modelle sonlanmaktadır [80].

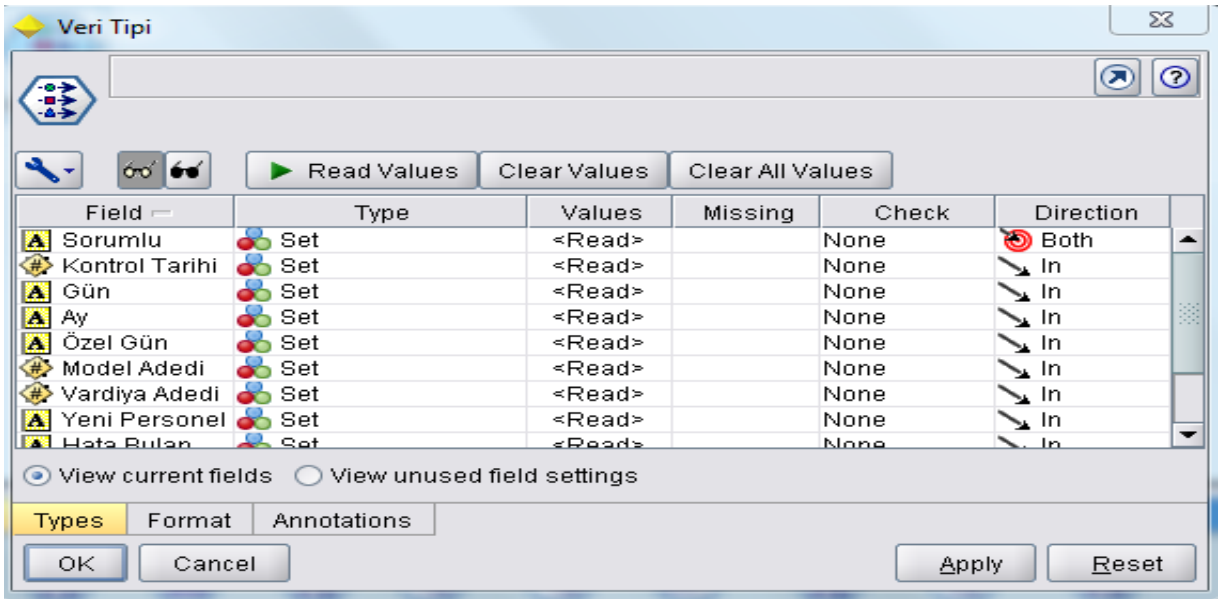
Elde edilen ve düzenlemesi yapılan veri setinin analiz edilecek son şekli Excel üzerinde düzenlendikten sonra Clementine programına veri kaynağı olarak tanımlaması yapılmıştır. Bu çalışmamızda veri analizi yöntemlerinden, birliktelik analizi algoritmaları yöntemleri arasında basit ve iyi bilinen algoritma olan Apriori algoritmasından faydalanılmıştır [81].

Apriori algoritması adımları veriler üzerinde işletilirken 2 adet minimum destek ve 2 adet güven değeri içeren 4 farklı kombinasyon denenmiştir. Şekil 2.9’da Clementine akım tualı üzerinde kurulan modelin son görüntüsü verilmiştir [71].



Şekil 2-9: Clementine programı açılış arayüzü görüntüsü

Excel kaynağından alınan tespit edilmiş olan kusur sorumluları verileri “Table” düğümü ile görüntülenmiş ve verilerin doğru şekilde aktarıldığından emin olunmuştur. “Type” düğümü kullanılarak tespit edilmiş olan kusur sorumluları verilerinin tipi belirlenmiştir. Elde edilen verinin oluşturduğu tüm kısıt alanları “Set” seçilmiştir. Oluşturulan modelde sorumlu grupları çıktı, veriyi oluşturan diğer kısıtların tamamı ise girdi olarak atanmıştır. Böylece girdi değişkenlerinin sorumlu prosese olan çıkışına etkilerinin incelenmesi amaçlanmıştır. Veri tipi belirleme düğümünün içeriği Şekil 2.10’da gösterilmektedir.



Şekil 2-10: Clementine programı veri tipi belirleme ekranı alıntısı

2.2.5 Değerlendirme

Değerlendirme aşamasında ise hazırlanan modele ait değerlendirmeler yapılmaktadır. Değerlendirme sürecinin amacı, verilerin sunulmasından öte, tasarımın en ince ayrıntısına kadar değerlendirilerek yola çıkılan hedeflerle elde edilen verilerin uyuşup uyuşmadığının kontrollerinin yapılmasını sağlamaktır. Bunun yanında elde edilen karşılaştırma sonuçlarına göre sürece dahil olmayan yeni verilerin ortaya çıkıp çıkmadığının değerlendirilmesi yapılmıştır [74].

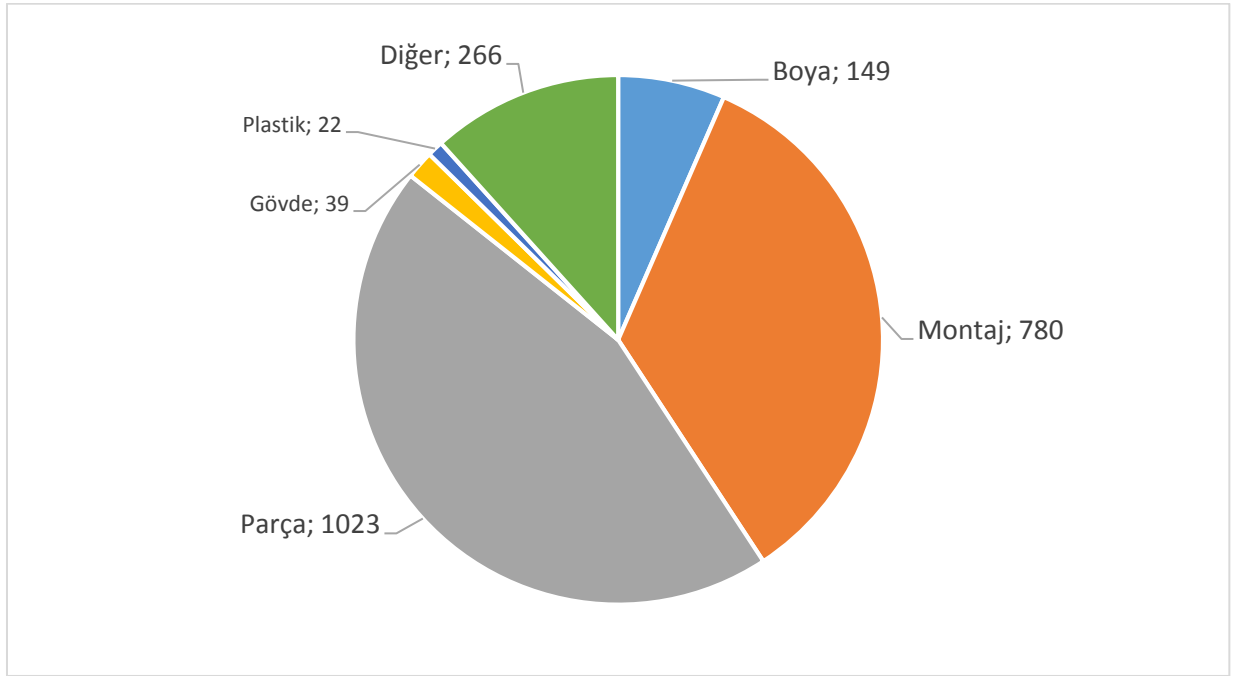
2.2.6 Yayma

Tamamlanmış olan model çalışması, düzenlenen tasarımın final sonucu olarak değerlendirilmemelidir. Amaçlanan hedef bilinen sonuçları sayılarını arttırmak olsa bile, ulaşılan veriler düzenlenerek kullanılacak hale getirilmeli ve sunumları gerçekleştirilmelidir [74].

3. BULGULAR

3.1 VERİ SETİNİN EXCEL TABLOLARI İLE YORUMLANMASI

Oluşturulan model çalışmada girilen veriler doğrultusunda elde edilen sonuçlar incelenmiştir. Veriyi topladığımız 8 yıllık çalışma diliminde üretici firma 1928 çalışma gününde, toplam 995.300 araç üretilmiştir. Kalite kontrol departmanının son kalite kontrol bölümü örnekleme olarak 235.877 aracın son kalite kontrolünü gerçekleştirmiştir. Bu kontroller sonucunda kontrol için seçilen araçlar üzerinde 2279 adet kusur tespit edilmiştir. Tespit edilen kusurların dağılımı aşağıdaki grafikte gösterilmiştir.



Şekil 3-1: Tespit edilen kusurların dağılımı

Tespit edilen hata gruplarının dağılımını incelediğimizde %45’lik oranla en büyük hata oranının parça grubunda olduğu, bunu montaj grubu %34, diğer %12, boya grubu %6, gövde grubu %2 ve son olarak plastik grubunun %1 ile takip ettiği tespit edilmiştir.

Grupların analiz yapıldığında üretim hatlarının gövde grubundan başlayıp sırası ile boya, montaj ve kalite grubu olarak devam ettiği, gövde grubundaki bir hatanın kalite grubundan önceki bölümlerde tespit edilmiş olması ihtimalinden kalite grubuna hata olarak yansması düşük olduğu tahmin edilmektedir. Buna rağmen 39 adet kusur kalite grubunda tespit edilmiştir. Araç üzerindeki, fabrika sınırları içinde üretilen plastik parçaların sayısının azlığı, plastik grubu hatalarının aynı oranda küçük kalmasına etken olduğu söylenebilir.

Boya grubunun %6'lık 149 adet kusuru gövde ve plastik grubuna göre yüksek olmasına karşın montaj ve parça grubunun oranları göz önüne alındığında iyileştirilebilir görünmektedir. Diğer olarak adlandırılmış olan kusur grupları ise kusurun sorumlusunun tam olarak tespit edilemediği hata grubudur.

780 adet kusur ile en çok ikinci kusur adedine sahip olan bölüm olan montaj grubu gerek çalışan sayısının çokluğu gerekse araç üzerine takılan parçaların fazla olması kusur sayısının bu rakamlara gelmesinde etkin rol oynamaktadır. Aynı zamanda montaj grubundan çıkan tamamlanmış aracın, kalite grubuna direkt olarak geçiyor olması, araç üzerinde oluşmuş olan bir kusurun direkt olarak tespit edilmesine olanak sağlamaktadır. Aynı zamanda parçayı araçla birleştiren montaj grubu çalışanı birleştirme esnasında, aracın boya yüzeyine zarar verdiği tespit edilirse bu kusurda montaj hattına kusur adedi olarak arşivlenmektedir.

%45 lik bir oranla neredeyse tespit edilen 2279 adet kusurun yarısını parça grubunun oluşturduğu veya sebep olduğu anlaşılmaktadır. Parça grubu olarak adlandırılan gruplar üretici firmaya dışarıdan tedarik yolu ile parça temin eden yan sanayilerdir. Ülkenin ve dünyanın çeşitli yerlerinden gelen parçalar yol şartlarından dolayı deforme olmuş olabiliyorlar. Her ne kadar parça grubu kalite kontrol ekipleri kontrollerini yapıyor olsalar da kırık, çatlak, çizik, fonksiyonunu yerine getirmeyen parçalar gözden kaçıp fark edilmeyerek üretime gönderilerek araç ile buluşmaktadır. Aynı zamanda üretici firmanın kalıplarında, preslerinde oluşan enjeksiyon hataları da her ışık ortamında görülmediği için araca montesi yapılarak kalite grupları ile buluşabilmektedir.

Tablo 3-1: Yıllara göre toplam hata adetleri

Yıl	Hata Sayısı
2010	237
2011	260
2012	215
2013	268
2014	249
2015	185
2016	497
2017	368
Toplam	2279

Yukarıdaki yıllara göre hata dağılımı tablosu incelendiğinde ise örnekleme kalite kontrol proseslerinde tespit edilen hata sayılarının ilk beş yıl ortalama rakamlarda seyrettiği bir anormallik görülmediği gözlemlenmiştir. Fakat 2015 yılında önceki yıllara göre hata sayısında bir miktarda olsa düşüş olduğu dikkatlerden kaçmamıştır. Yine bu yılı takip eden sonraki yılda önceki yılların neredeyse iki katı bir artış olduğu ve 2017 yılında bu anormal yükselişin hata adetleri bazında düşüş yönünde bir eylem gerçekleştirdiği görülmektedir.

Tablo 3-2: Montaj Hattı yıllara göre hata adetleri

Sorumlu	2010	2011	2012	2013	2014	2015	2016	2017
Montaj Hattı	50	40	65	65	83	76	236	165

Montaj hattının yıllar içinde hata kaçırma kırılımını incelediğimizde ilk altı yıl, ortalama değerler içinde gerçekleşmiş olup montaj hattı ortalamaları belirli bir düzeyde kalmıştır. Bu duruma karşın 2016 yılında anormal bir yükseliş gerçekleşirken geçmiş altı yıllık ortalama hata sayısının bu yıl içerisinde neredeyse %400'e yakın bir değerde arttığı gözlemlenmektedir. 2017 yılında ise bir önceki yılın anormal yükselişinin aksine, ilk altı yıla oranla yüksek bir hata sayısı olmasına karşın bir önceki yılın anormal yükselişine kıyasla %30 oranında bir iyileşme gerçekleşmiştir.

Tablo 3-3: Boya Hattı yıllara göre hata adetleri

Sorumlu	2010	2011	2012	2013	2014	2015	2016	2017
Boya Hattı	21	9	5	14	14	12	42	36

Boya hattının yıllar içinde hata kaçırma kırılımını incelediğimizde ilk altı yıl, ortalama değerler içinde gerçekleşmiş olup boya hattı ortalamaları belirli bir düzeyde kalmıştır. Bu duruma karşın 2016 yılında anormal bir yükseliş gerçekleşirken geçmiş altı yıllık ortalama hata sayısının bu yıl içerisinde neredeyse %350'ye yakın bir değerde arttığı gözlemlenmektedir. 2017 yılında ise bir önceki yılın anormal yükselişinin aksine, ilk altı yıla oranla yüksek bir hata sayısı olmasına karşın bir önceki yılın anormal yükselişine kıyasla %15 oranında bir iyileşme gerçekleşmiştir.

Tablo 3-4: Parça Hattı yıllara göre hata adetleri

Sorumlu	2010	2011	2012	2013	2014	2015	2016	2017
Parça Hattı	125	111	119	135	119	75	217	122

Parça hattının yıllar içinde hata kaçırma kırılımını incelediğimizde ilk beş yıl ortalama değerler içinde gerçekleşmiş olup parça hattı ortalamaları belirli bir düzeyde kalmıştır. Diğer gruplara göre parça hattında 2015 yılında bir iyileşme gözlemlenerek daha önceki yıllarda gerçekleşen üç basamaklı hata sayılarının yerini bu yılda iki basamaklı sayılara bıraktığı görülmektedir. Bu duruma karşın 2016 yılında anormal bir yükseliş gerçekleşirken geçmiş altı yıllık ortalama hata sayısının bu yıl içerisinde neredeyse %90'a yakın bir değerde arttığı gözlemlenmektedir. 2017 yılında ise bir önceki yılın anormal yükselişinin aksine, ilk altı yıla oranla yüksek bir hata sayısı olmasına karşın bir önceki yılın anormal yükselişine kıyasla %45 oranında bir iyileşme gerçekleşerek ilk altı yıllık ortalama değerlere dönmüştür.

Tablo 3-5: Gövde Hattı yıllara göre hata adetleri

Sorumlu	2010	2011	2012	2013	2014	2015	2016	2017
Gövde Hattı	3	4	5	5	3	3	10	6

Gövde hattının yıllar içinde hata kaçırma kırılımını incelediğimizde ilk altı yıl, ortalama değerler içinde gerçekleşmiş olup gövde hattı ortalamaları belirli bir düzeyde kalmıştır. Bu duruma karşın 2016 yılında anormal bir yükseliş gerçekleşirken geçmiş altı yıllık ortalama hata sayısının bu yıl içerisinde neredeyse %150'ye yakın bir değerde arttığı gözlemlenmektedir. 2017 yılında ise bir önceki yılın anormal yükselişinin aksine, ilk altı yıla oranla yüksek bir hata sayısı olmasına karşın bir önceki yılın anormal yükselişine kıyasla %15 oranında bir iyileşme gerçekleşmiştir.

Tablo 3-6: Plastik Hattı yıllara göre hata adetleri

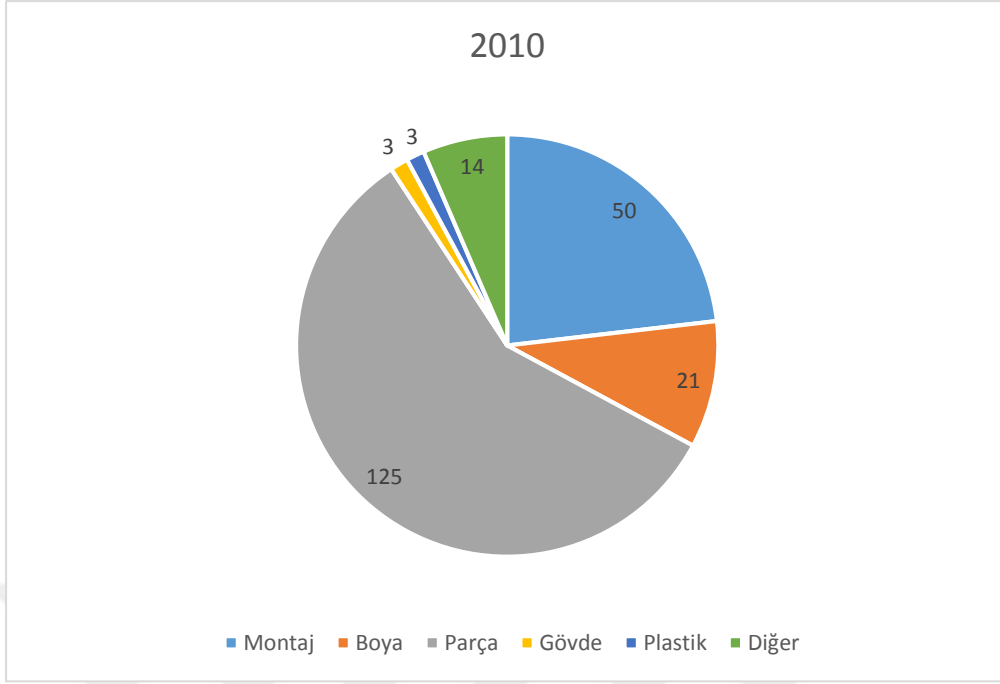
Sorumlu	2010	2011	2012	2013	2014	2015	2016	2017
Plastik Hattı	3	0	2	0	2	3	7	5

Plastik hattının yıllar içinde hata kaçırma kırılımını incelediğimizde ilk altı yıl içerisinde 2 ve 3 adet hata kaçırmanın yanı 2011 ve 2013 yıllarında örnekleme kontrol yapan kalite grubu tarafından kayıtlara geçen hatası bulunmamaktadır. Bu duruma karşın 2016 yılında anormal bir yükseliş gerçekleşirken geçmiş altı yıllık ortalama hata sayısının bu yıl içerisinde neredeyse %450'ye yakın bir değerde arttığı gözlemlenmektedir. 2017 yılında ise bir önceki yılın anormal yükselişinin aksine, ilk altı yıla oranla yüksek bir hata sayısı olmasına karşın bir önceki yılın anormal yükselişine kıyasla %30 oranında bir iyileşme gerçekleşmiştir.

Tablo 3-7: Diğer Hataların yıllara göre hata adetleri

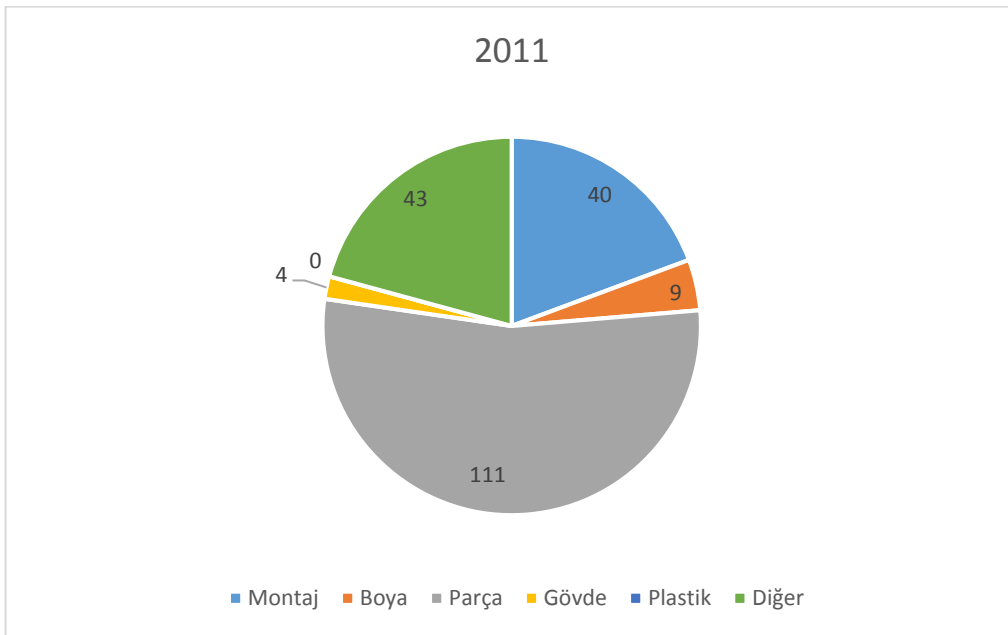
Sorumlu	2010	2011	2012	2013	2014	2015	2016	2017
Diğer Hatalar	14	43	22	30	27	12	67	51

Diğer hata oranlarının yıllar içinde hata kaçırma kırılımını incelediğimizde ilk altı yıl, ortalama değerler içinde gerçekleşmiş olup diğer hataların ortalamaları belirli bir düzeyde kalmıştır. Bu duruma karşın 2016 yılında anormal bir yükseliş gerçekleşirken geçmiş altı yıllık ortalama hata sayısının bu yıl içerisinde neredeyse %170'ye yakın bir değerde arttığı gözlemlenmektedir. 2017 yılında ise bir önceki yılın anormal yükselişinin aksine, ilk altı yıla oranla yüksek bir hata sayısı olmasına karşın bir önceki yılın anormal yükselişine kıyasla %25 oranında bir iyileşme gerçekleşmiştir.



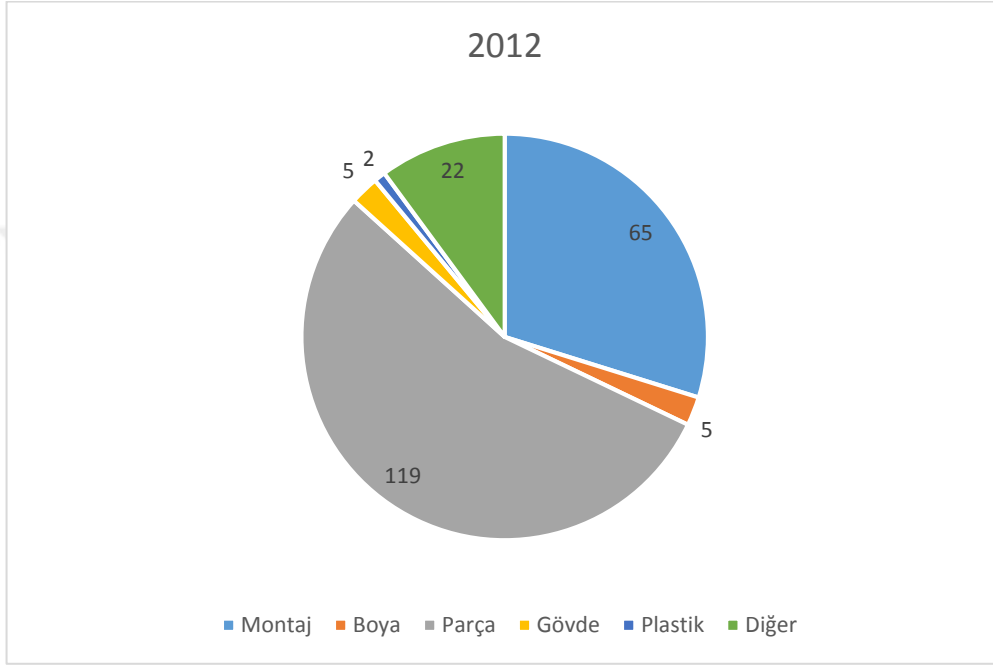
Şekil 3-2: 2010 yılı hata sayıları

2010 yılının hata kırılımına baktığımızda, yıl içerisinde tüm hatların kaçırdığı ve örnekleme kalite kontrol grubunda tespit edilen toplamda 214 adet hata grubu bulunmaktadır. Bu hataları en fazla kaçıran sorumlu hatların ilk üçüne baktığımızda ise, toplamda yıl içerisindeki tüm hataların neredeyse yarısını kaçırmış olan 125 hata ile parça grubu birinci sırada yer alırken, 50 adet hata kaçırmaya sayısı ile montaj hattı ikinci sırada yer almakta üçüncü sırayı ise 21 hata kaçıran boya hattının aldığı gözlemlenmektedir.



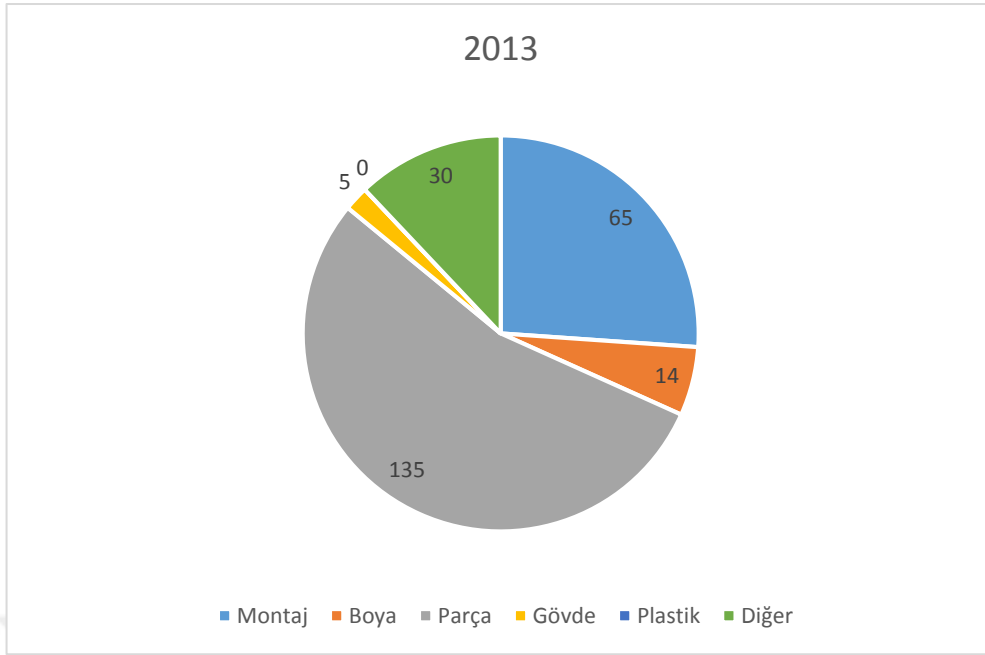
Şekil 3-3: 2011 yılı hata sayıları

2011 yılının hata kırılımına baktığımızda ise, yıl içerisinde tüm hatların kaçırıldığı ve örnekleme kalite kontrol grubunda tespit edilen toplamda 207 adet hata grubu bulunmaktadır. Bu hataları en fazla kaçırın sorumlu hatların ilk üçüne baktığımızda ise, toplamda yıl içerisindeki tüm hataların neredeyse yarısını kaçırmış olan 111 hata ile parça grubu birinci sırada yer alırken, 43 adet hata kaçırma sayısı ile plastik hattı ikinci sırada yer alırken üçüncü sırayı 40 hata kaçırın montaj hattının aldığı gözlemlenmektedir.



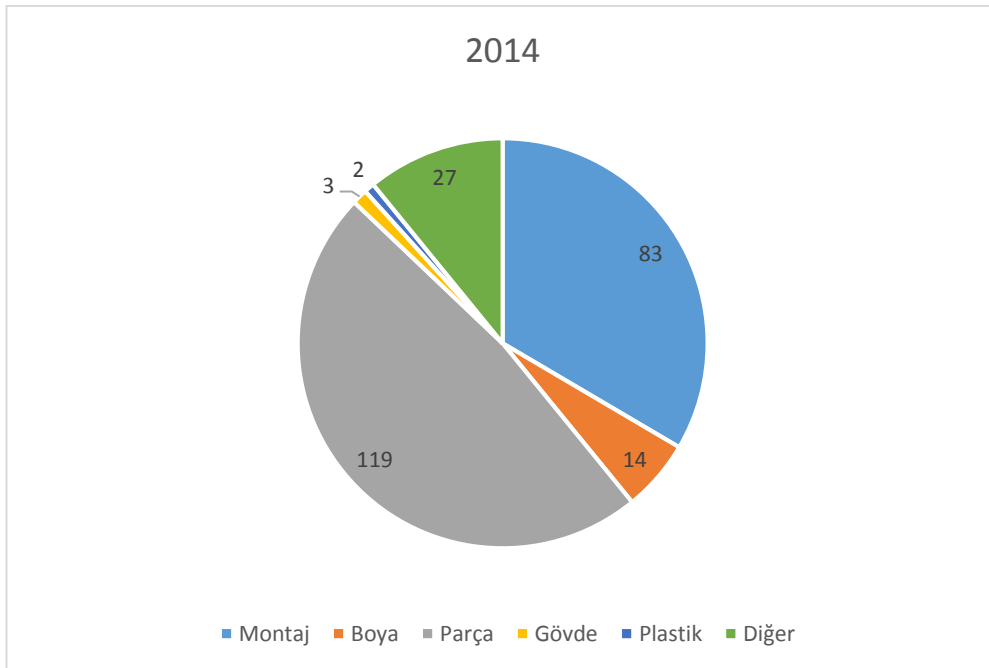
Şekil 3-4: 2012 yılı hata sayıları

2012 yılının hata kırılımına baktığımızda ise, yıl içerisinde tüm hatların kaçırıldığı ve örnekleme kalite kontrol grubunda tespit edilen toplamda 218 adet hata grubu bulunmaktadır. Bu hataları en fazla kaçırın sorumlu hatların ilk üçüne baktığımızda ise, toplamda yıl içerisindeki tüm hataların neredeyse yarısını kaçırmış olan 119 hata ile parça grubu birinci sırada yer alırken, 65 adet hata kaçırma sayısı ile montaj hattı ikinci sırada yer almakta ve üçüncü sırayı 22 hata kaçırın diğer hataların aldığı gözlemlenmektedir.



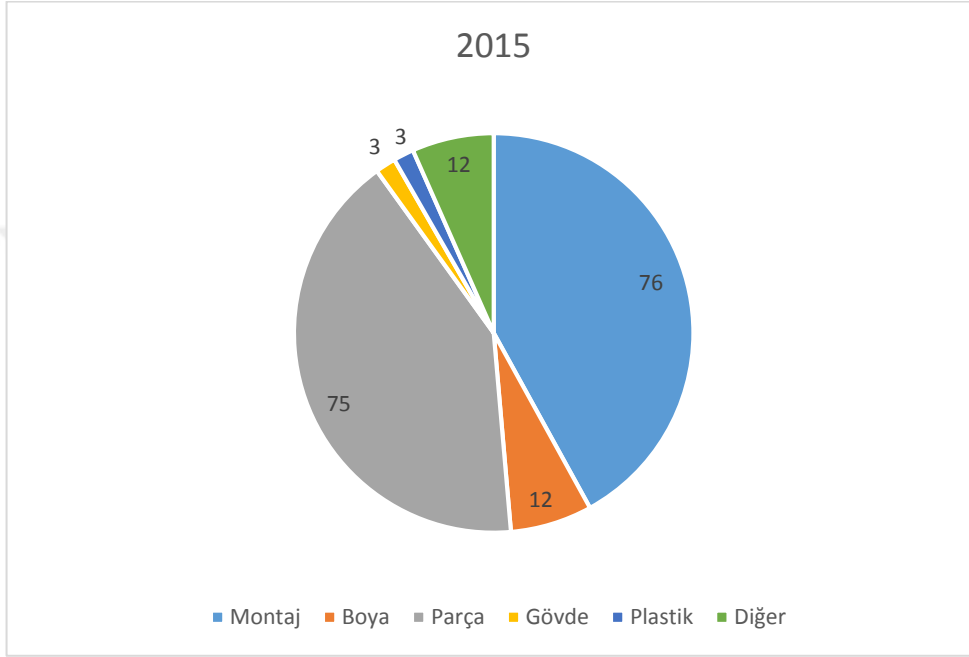
Şekil 3-5: 2013 yılı hata sayıları

2013 yılının hata kırılımına baktığımızda ise, yıl içerisinde tüm hatların kaçırıldığı ve örnekleme kalite kontrol grubunda tespit edilen toplamda 249 adet hata grubu bulunmaktadır. Bu hataları en fazla kaçırılan sorumlu hatların ilk üçüne baktığımızda ise, toplamda yıl içerisindeki tüm hataların neredeyse yarısını kaçırmış olan 135 hata ile parça grubu birinci sırada yer alırken, 65 adet hata kaçırma sayısı ile montaj hattı ikinci sırada yer almakta ve üçüncü sırayı 30 hata kaçırılan diğer hataların aldığı gözlemlenmektedir.



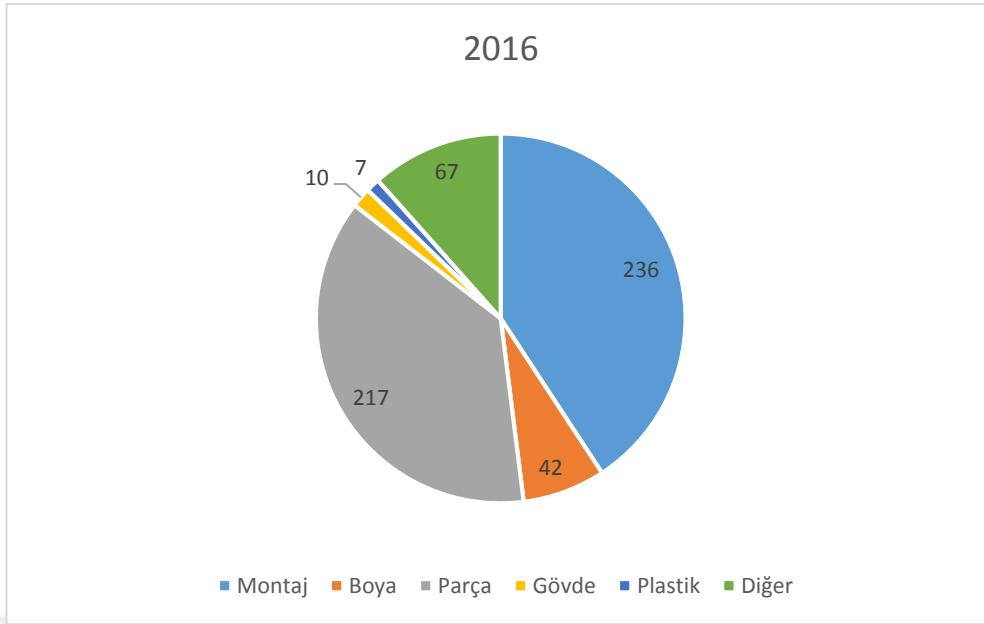
Şekil 3-6: 2014 yılı hata sayıları

2014 yılının hata kırılımına baktığımızda ise, yıl içerisinde tüm hatların kaçırdığı ve örnekleme kalite kontrol grubunda tespit edilen toplamda 248 adet hata grubu bulunmaktadır. Bu hataları en fazla kaçıran sorumlu hatların ilk üçüne baktığımızda ise, toplamda yıl içerisindeki tüm hataların neredeyse yarısını kaçırmış olan 119 hata ile parça grubu birinci sırada yer alırken, 83 adet hata kaçırma sayısı ile montaj hattı ikinci sırada yer almakta ve üçüncü sırayı 27 hata kaçıran diğer hataların aldığı gözlemlenmektedir.



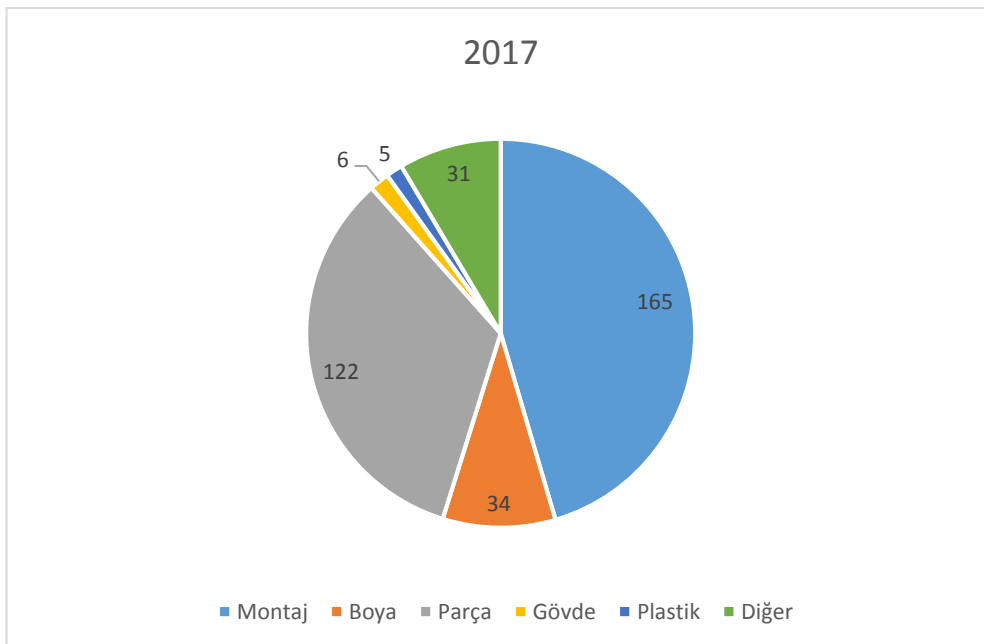
Şekil 3-7: 2015 yılı hata sayıları

2015 yılının hata kırılımına baktığımızda ise, yıl içerisinde tüm hatların kaçırdığı ve örnekleme kalite kontrol grubunda tespit edilen toplamda 184 adet hata grubu bulunmaktadır. Bu hataları en fazla kaçıran sorumlu hatların ilk üçüne baktığımızda ise, toplamda yıl içerisindeki tüm hataların neredeyse %80 'ini oluşturan 76 ve 75 hata adet ile montaj ve parça grubu birinci sırada yer alırken, yine 12 şer adet hata kaçırma sayıları ile boya hattı ve diğer hatalar grubunun ikinci sırada yer aldığı gözlemlenmektedir.



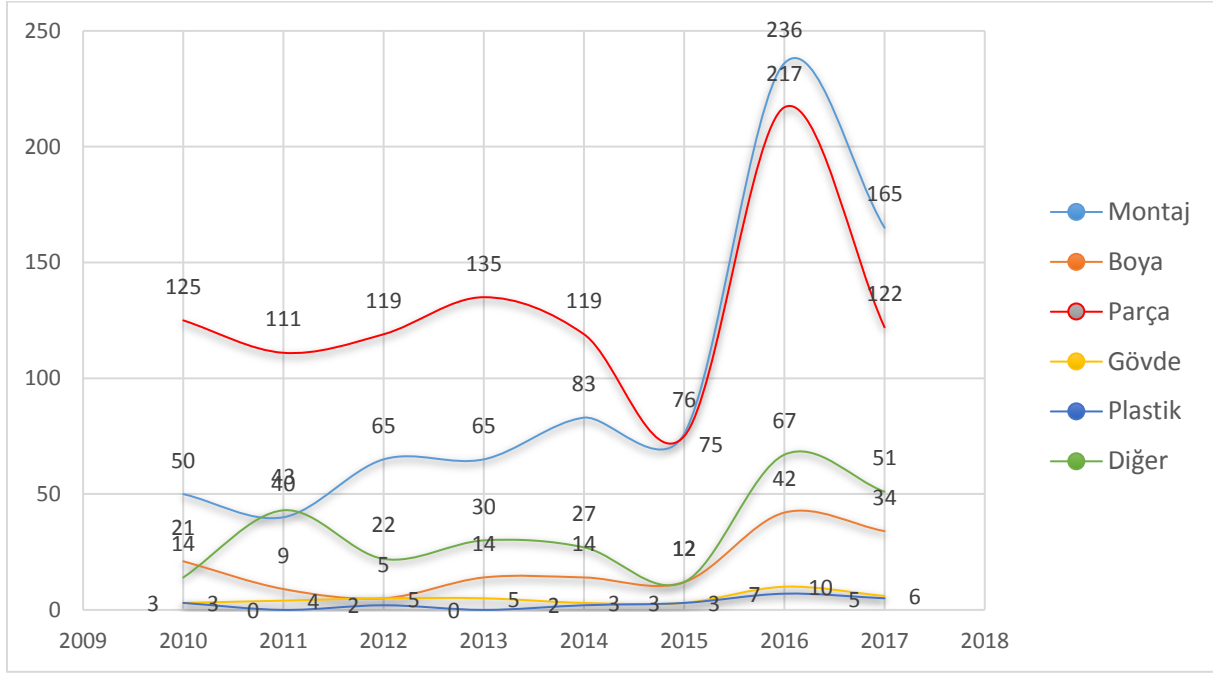
Şekil 3-8: 2016 yılı hata sayıları

2016 yılının hata kırılımına baktığımızda ise, yıl içerisinde tüm hatların kaçırıldığı ve örnekleme kalite kontrol grubunda tespit edilen toplamda 579 adet hata grubu bulunmaktadır. Bu hataları en fazla kaçırılan sorumlu hatların ilk üçüne baktığımızda ise, toplamda yıl içerisindeki tüm hataların neredeyse %40' ını kaçırmış olan 236 hata ile montaj grubu birinci sırada yer alırken, %37' sini kaçırmış olan 217 adet hata sayısı ile parça hattı ikinci sırada yer almakta ve üçüncü sırayı 67 hata kaçırılan diğer hataların aldığı gözlemlenmektedir.



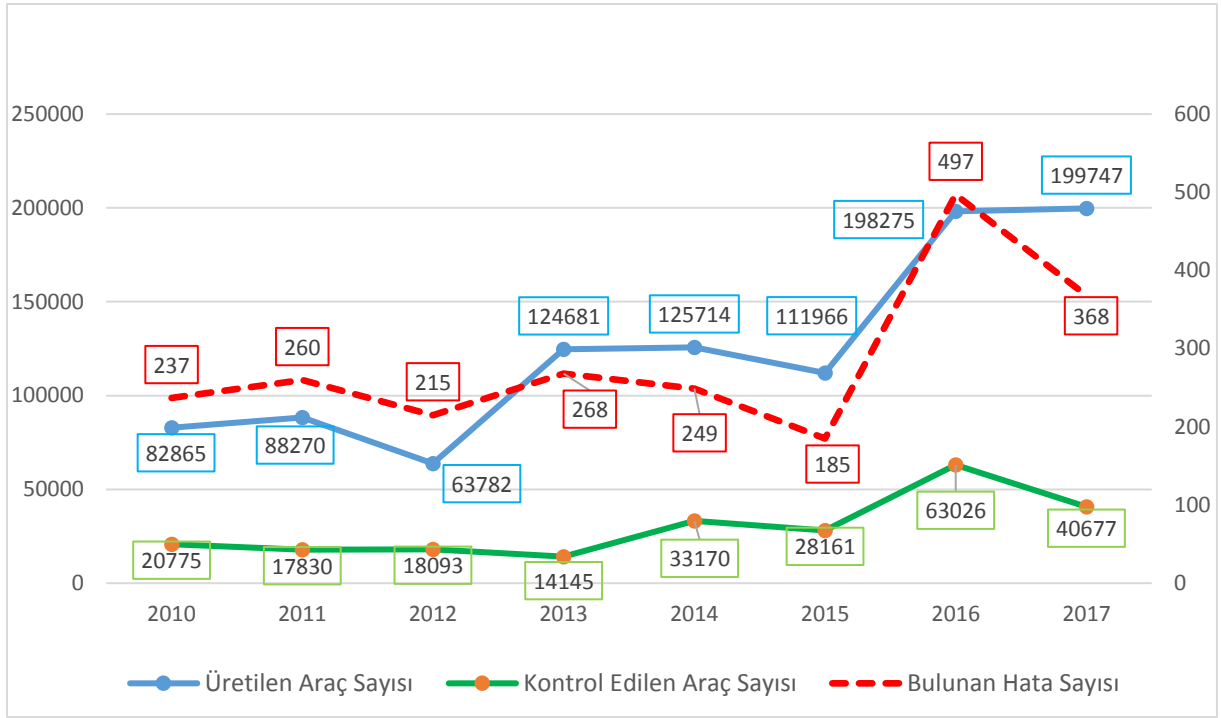
Şekil 3-9: 2017 yılı hata sayıları

2017 yılının hata kırılımına baktığımızda ise, yıl içerisinde tüm hatların kaçırdığı ve örnekleme kalite kontrol grubunda tespit edilen toplamda 383 adet hata grubu bulunmaktadır. Bu hataları en fazla kaçıran sorumlu hatların ilk üçüne baktığımızda ise, toplamda yıl içerisindeki tüm hataların neredeyse yarısını kaçırmış olan 165 hata ile montaj grubu birinci sırada yer alırken, 122 adet hata kaçırma sayısı ile parça hattı ikinci sırada yer almakta ve üçüncü sırayı 34 hata kaçıran boya hattının aldığı gözlemlenmektedir.



Şekil 3-10: Hata sayılarının yıllara göre dağılımı

Tüm yılları ve sorumluları bir arada incelediğimizde, tüm gruplar 2015 yılına kadar ki süreçte kendi ortalamalarında hata düzeylerini korurlarken, 2015 yılında tüm gruplarda düşüş gözlemlenmiştir. 2016 yılında yine aynı şekilde tüm gruplarda yine birlikte bir hareket söz konusu iken bu hareketin önceki yılın tersi yönünde aşırı bir yükseliş olarak gözlemlenmiştir. 2017 yılında ise bir önceki yıla göre bir şeylerin düzeldiği ve göstergelerin düşüş eğiliminde olduğu görülmektedir.



Şekil 3-11: Yıllara göre üretilen ve kontrol edilen araç sayıları ile bulunan hata sayılarının gösterimi

Şekil 3-11 de görüleceği üzere üretilen araç sayıları ile son kalite kontrol bölümünde kontrol edilen araç sayılarının birbirleri arasında genelde paralellik gösterdiği, tespit edilen hata sayılarının yıllara göre değişiklik gösterdiği gözlemlenmektedir. Bu durum 2015 yılından 2016 yılına geçişte, üretimi yapılan araçlarda yaklaşık %80 lik bir artış olarak gözlemlenirken buna paralel olarak kontrolü yapılan araçlarda artış oranı %120 olarak gözlemlenmiştir. Aynı dönem içinde bulunan hata oranına bakıldığında %170 lik ani bir yükseliş görülmektedir. Bu durum bize araç düşüşü ile aynı anda düşüşe geçen bir hata kaçırma oranı sergilerken, araç üretiminin artışına bağlı olarak yükselmesi beklenenden çok fazla sayıda hata kaçırma oranıyla karşı karşıya olduğumuzu göstermektedir.

3.2 VERİ SETİNİN APRIORİ ALGORİTMASI İLE YORUMLANMASI

Oluşturulan model çalıştırılmadan önce son aşama birliktelik kurallarının minimum destek ve güven değerlerinin belirlenmesidir. Bu değerlerin büyük seçilmesi az sayıda kuralın ortaya çıkmasına sebebiyet vermektedir. Küçük seçilmesi ise çok sayıda kuralın ortaya çıkmasına yol açacaktır. Diğer bir deyiş ile; minimum destek ve güven değerlerinin elde edilecek birliktelik kuralı sayısı ile doğru orantılı olduğunu söylemek doğru bir ifade olacaktır [71].

Bu çalışmada kurulan model 2 adet minimum destek ve 2 adet minimum güven değeri ile sınanacak ve 4 farklı kombinasyon ile denenmiş olacaktır. Böylece farklı destek ve güven değerleri seçilmesinin elde edilecek kuralları nasıl etkileyeceği analiz edilecektir.

Öncelikle belirlenen güven değeri sabit tutulacak ve sırasıyla %5 ve %10 destek değerleri ile model çalıştırılacaktır. Aynı işlem diğer güven değerleri için tekrarlanacaktır. Analizde kullanılacak güven değerleri sırasıyla %50 ve %59 olarak seçilmiştir. Sınanacak tüm güven ve destek değeri kombinasyonları Tablo 3.8’de özetlenmiştir.

Tablo 3-8: Test edilecek Güven ve Destek değerleri

Deneme Sayısı	Güven (%)	Destek (%)
1	50	5
2	50	10
3	59	5
4	59	10

Apriori algoritması model üzerinde ilk olarak %50 minimum güven ve %5 minimum destek değeri ile çalıştırılmıştır. Elde edilen 578 adet birliktelik kuralı Tablo 3.9’da gösterilmektedir. Kuralların sıralaması en yüksek güven değerinden en düşük güven değerine doğru yapılmıştır.

Tablo 3-9: Veri setinde %50 minimum Güven ve %5 minimum Destek değerleri için bulunan kurallar

No	Sorumlu	Öncül Değişken	Destek (%)	Güven (%)
1	Parça	Kontrol Tarihi: 2011 H. B. V. Çalışma Saati:2	5,26	65,83
2	Parça	Kontrol Tarihi: 2011 Yeni Personel: Yok H. B. V. Çalışma Saati:2	5,26	65,83
3	Parça	Kontrol Tarihi: 2011 Vardiya Adedi: 2 H. B. V. Çalışma Saati:2	5,26	65,83
4	Parça	Kontrol Tarihi: 2011 Model Adedi: 2 H. B. V. Çalışma Saati:2	5,26	65,83
5	Parça	Kontrol Tarihi: 2011 Hatalı Araç Modeli: AU	5,22	65,54
6	Parça	Kontrol Tarihi: 2011 Yeni Personel: Yok Hatalı Araç Modeli: AU	5,22	65,54
7	Parça	Kontrol Tarihi: 2011 Vardiya Adedi: 2 Hatalı Araç Modeli: AU	5,22	65,54
8	Parça	Kontrol Tarihi: 2011 Hatalı Araç Modeli: AU Model Adedi: 2	5,22	65,54
9	Parça	Gün: Cuma H. B. V. Çalışma Saati:2 Özel Gün: Normal	7,06	64,60
10	Parça	Gün: Cuma H. B. V. Çalışma Saati:2 Yeni Personel: Yok	5,92	64,44

Elde edilen ilk kuralı yorumladığımızda, kontrol tarihinde 2011 yılı, hatayı bulan vardiya çalışma saatinde 2 olması durumunda hata sorumlu grubunun parça grubu olma olasılığı %65,83 olduğu anlamına ulaşılmaktadır. Kuralların sıralaması da en yüksek güven değerinden en düşük güven değerine doğru yapılmıştır.

Bir sonraki aşamada ise algoritmayı model üzerinde %50 minimum güven ve %10 minimum destek değerleriyle tekrardan çalıştırılmıştır. Sonuçlarda 164 tane birliktelik kuralı ile karşılaşılmıştır. Elde edilen bu kuralların 10 tanesi en yüksek güven değerinden en düşük güven değerine doğru Tablo 3.10' da sıralanmıştır.

Tablo 3-10: Veri setinde %50 minimum Güven ve %10 minimum Destek değerleri için bulunan kurallar

No	Sorumlu	Öncül Değişken	Destek (%)	Güven (%)
1	Parça	Hatalı Araç Modeli: AU Hatayı Bulan Vardiya: Kır Model Adedi: 2	11,54	59,31
2	Parça	Hatalı Araç Modeli: AU Hatayı Bulan Vardiya Çal. S.: 2 Vardiya Adedi: 2	15,79	59,16
3	Parça	Hatalı Araç Modeli: AU Hatayı Bulan Vardiya: Kır Model Adedi: 2	11,49	59,16
4	Parça	Hatalı Araç Modeli: AU Hatayı Bulan Vardiya Çal. S.: 2 Vardiya Adedi: 2	16,06	59,01
5	Parça	Hatalı Araç Modeli: AU Yeni Personel: Yok Özel Gün: Normal	10,57	58,92
6	Parça	Hatalı Araç Modeli: AU Hatayı Bulan Vardiya Çal. S.: 2 Özel Gün: Normal	16,80	58,74
7	Parça	Hatayı Bulan Vardiya Çal. S.: 2 Yeni Personel: Yok Özel Gün: Normal	29,35	58,59
8	Parça	Hatalı Araç Modeli: AU Hatayı Bulan Vardiya: Kır	12,98	58,44
9	Parça	Hatalı Araç Modeli: AU Hatayı Bulan Vardiya: Kır Yeni Personel: Yok	12,54	58,39
10	Parça	Hatayı Bulan Vardiya: Kır Yeni Personel: Yok Model Adedi: 2	30,75	58,34

Elde edilen ilk kuralı yorumladığımızda, hatalı araç modelinin AU, hatayı bulan vardiyanın Kır, model adedinin de 2 olması durumunda hata sorumlu grubunun parça grubu olma olasılığı %59,31 olduğu anlamına ulaşılmaktadır. Kuralların sıralaması da en yüksek güven değerinden en düşük güven değerine doğru yapılmıştır.

%50 minimum güven değeri için %5 ve %10 minimum destek değerleri için herbiri kullanılarak oluşan kural çıktıları tamamlandıktan sonra bu işlemlerim aynısını %59 minimum güven değeri için tekrarı sağlanacaktır.

Çalışmanın bu sürecinde algoritma model üzerindeki verilerin %59 minimum güven ve %5 minimum destek değeri ile yeniden çalıştırılmıştır. Bunun sonucunda sistem 47 tane birliktelik kuralı oluşturmuş olup, elde edilen kuralların sıralaması en yüksek güven değerinden en düşük güven değerine doğru yapılmıştır. Bu durum Tablo 3.11’ de gösterilmiştir.

Tablo 3-11: Veri setinde %59 minimum Güven ve %5 minimum Destek değerleri için bulunan kurallar

No	Sorumlu	Öncül Değişken	Destek (%)	Güven (%)
1	Parça	Kontrol Tarihi: 2011 Hatayı Bulan Vardiya Çal. S.: 2	5,26	65,83
2	Parça	Kontrol Tarihi: 2011 Hatayı Bulan Vardiya Çal. S.: 2 Yeni Personel: Yok	5,26	65,83
3	Parça	Kontrol Tarihi: 2011 Hatayı Bulan Vardiya Çal. S.: 2 Vardiya Adedi: 2	5,26	65,83
4	Parça	Kontrol Tarihi: 2011 Hatayı Bulan Vardiya Çal. S.: 2 Model Adedi: 2	5,26	65,83
5	Parça	Hatalı Araç Modeli: AU Kontrol Tarihi: 2011	5,22	65,54
6	Parça	Hatalı Araç Modeli: AU Kontrol Tarihi: 2011 Yeni Personel: Yok	5,22	65,54
7	Parça	Hatalı Araç Modeli: AU Kontrol Tarihi: 2011 Vardiya Adedi: 2	5,22	65,54
8	Parça	Hatalı Araç Modeli: AU Kontrol Tarihi: 2011 Vardiya Adedi: 2	5,22	65,54
9	Parça	Gün: Cuma Hatayı Bulan Vardiya Çal. S.: 2 Özel Gün: Normal	7,06	64,59
10	Parça	Gün: Cuma Hatayı Bulan Vardiya Çal. S.: 2 Yeni Personel: Yok	5,92	64,44

Apriori algoritması model üzerinde %59 minimum güven ve %10 minimum destek değeri ile yeniden çalıştırılmıştır. Elde edilen 4 adet birliktelik kuralı Tablo 3.12’de gösterilmektedir.

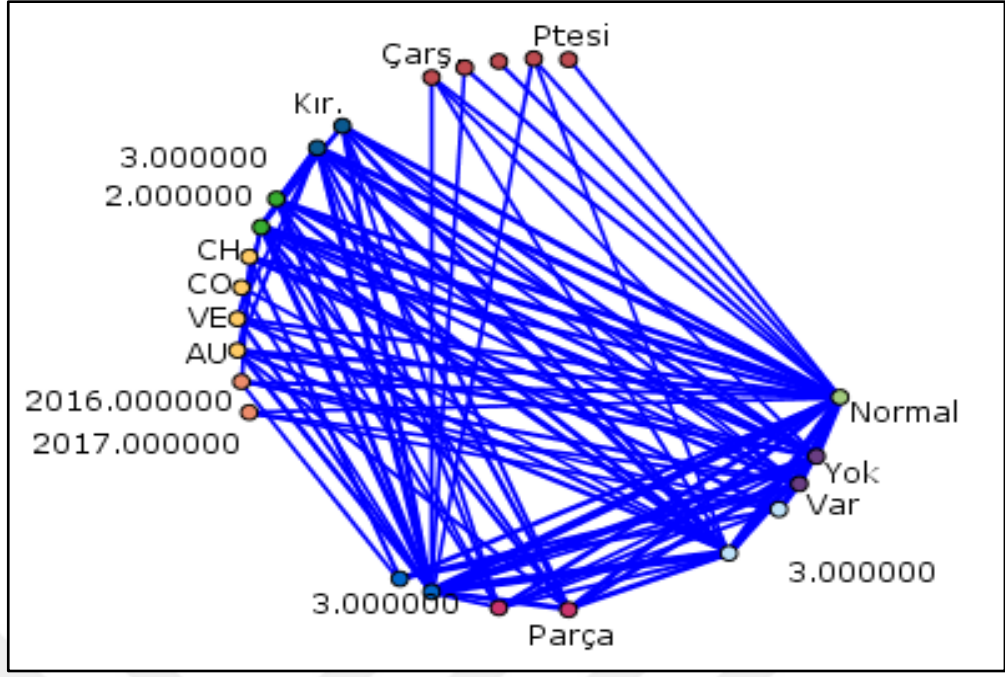
Tablo 3-12: Veri setinde %59 minimum Güven ve %10 minimum Destek değerleri için bulunan kurallar

No	Sorumlu	Öncül Değişken	Destek (%)	Güven (%)
1	Parça	Hatalı Araç Modeli: AU Hatayı Bulan Vardiya: Kır Model Adedi: 2	11,54	59,31
2	Parça	Hatalı Araç Modeli: AU Hatayı Bulan Vardiya Çal. S.: 2 Vardiya Adedi: 2	15,79	59,16
3	Parça	Hatalı Araç Modeli: AU Hatayı Bulan Vardiya: Kır Vardiya Adedi: 2	11,49	59,16
4	Parça	Hatalı Araç Modeli: AU Hatayı Bulan Vardiya Çal. S.: 2 Model Adedi: 2	16,06	59,01

Yapılan tüm analizler sonucunda, sorumlu grubu olarak parça grubu tespit edilmektedir. Bu durumu bize Pareto grafiğinde doğrulamaktadır. Pareto grafiği en çok hata grubunun parça grubunda olduğunu gösterirken yine tüm güven ve destek aralıklarında algoritma çalıştırılmış ve öncül sorumlunun parça grubu olduğunu sonuçlar bize göstermiştir.

3.2.1. Web Grafiği

Tüm verilerin yer aldığı veri setinde girdiler için kullanılan verilerin birbirleri arasındaki ilişkinin bağıntısını daha net görüntüleyebilmek ve ifade edebilmek için SPSS Clementine programının Web grafiği adından faydalanılmıştır. Bu grafikte yer alan her çizginin kendine has olan kalınlığı çizgileri birbirine bağlayan faktörler arasındaki baskının gücünü göstermektedir. İki nokta arasındaki çizgi kalınlığı artış gösterdikçe ilişkinin gücü de artmaktadır. Kalın çizgiler iki faktör arasında güçlü bir ilişkili olduğu ve incelenmesi gerektiği anlamına gelmektedir [71].



Şekil 3-12: Web Grafiği

4. TARTIŞMA VE SONUÇ

Üretim işletmelerinin günümüz değişen teknoloji ve rekabetçi koşulları yakalayıp ayakta kalabilmesi için sürekli kendilerini geliştirmeli ve yatırımlarını makine, donanım ve insan odaklı olarak yapmaları gerekmektedir. Üretim sistemi her ne kadar makine ağırlıklı olsa da ve bu ekipmanların görevlerini sorunsuz bir şekilde yerine getirmeleri beklense de insan kaynaklı hatalar makine kaynaklı hataların önüne geçmektedir.

Bu çalışma kapsamında 8 yıllık süreçte tespit edilen hata sayılarının yıllara ve gruplara göre veri madenciliği teknikleri kullanılarak analizleri yapılmıştır. Literatürde Apriori algoritmasının daha çok market sepet analizi konularında kullanıldığı gözlemlenmiştir. Bu bağlamda otomotiv endüstrisi alanındaki bu çalışmamızla farklı bir alandaki Apriori algoritması örneği ortaya konulmuştur.

Bu çalışmada verilere genel bir bakış yapılarak hangi yıllarda farklılıklar görüldüğü tespit edilmiştir. Veriler sorumlu hatlara ve yıllara göre tek tek analiz edilerek genele yansıyan verilerin doğruluğu incelenmiştir.

Yapılan analizler doğrultusunda üretici ana sanayi şirketin hata oranlarını yükselten sebebin genele bakıldığında parça grubundan kaynaklandığı saptanmıştır. 2016 yılına gelindiğinde üretici şirket ürettiği model sayısını 2 den 3 e, çalışma vardiya sayısını 2 den 3e çıkarttığını ve takt time süresini de 90 saniyeden 70 saniye ye indirdiğini gözlemlemekteyiz. Bu zaman diliminde ise montaj hattı hatalarının artışı gözlemlemekteyiz. Bu değişimlere tedarikçi yan sanayi firmaların ayak uydurabilmiş gibi görünse de ana sanayi üreticisinin uyum sağlamasında zorlanma yaşamış olma ihtimali söz konusu olabilir. Ana üretici firmanın tedarikçi yan sanayilerden parça talep oranının üretim sayısı oranında arttığı görülmektedir. Tedarikçi firmaların yine kendilerine yarı mamul malzeme tedarik eden firmalardan olan malzeme talebi ve üretimi arttıracak makine, ekipman ve tecrübeli işçi bulmaları bir diğer etken olabilir.

Aslında ana sanayi üreticisinin de montaj hattı hatalarının bu dönemde birinci sırada olduğuna dikkat edilirse, şirket üretim hattını ürün artırımına göre ayarlamış olmasına karşın tecrübeli personel tedarikinde sıkıntılar yaşamış olduğu gözlenmektedir. Tecrübeli olmayan ve deneyim kazanamaya zamanı olmayan yeni personel uyum sağlamada da zorlanmaktadır. Mevcut personelin eğitim eksikliği de söz konusu olabilir.

En önemli nokta ise yapılan analizler sonucunda görülmüştür ki, parça grubunun hata sayıları incelenen yıllarda diğer hatların hata sayılarının hep üzerinde olduğu yani bu durumun parça grubu için anormal olmadığı fakat, montaj hattının yukarıda bahsedilen 2016 ve 2017

yıllarında ki geiş srecinden kaynaklı uyumsuzluk nedeniyle hata sayısının para grubunu geerek, geiše uyum saėlayamadığı ve bu kritik dnemdeki hata sayılarının artışına sebep olduėudur.

Sonuç olarak; řirket bu tr geiş srelerinde, mevcut alıřanlarının eėitim faaliyetlerini daha geniř bir zaman dilimine yayarak ve alacaėı deneyimsiz personelin eėitim ve uyum srelerini daha n plana alarak montaj hattı hata sayılarını dřrebileceėi gibi, tedariki firmalarla da bu tr geiřleri daha optimum bir srete gerekleřtirmesi ve daha geniř bir zaman dilimine yayması olası hata oranlarının dřmesine yardımcı olacaėı dřnlmektedir.

alıřmanın en nemli sınırlılıėı bir otomotiv ana sanayi iřletmesi kapsamında gerekleřtirilmesidir. alıřmanın verileri 2010-2017 yılları arasında, 1928 alıřma gnnde, retici firma genelinde gerekleřen 995.300 adet arasında rnekleme kontrol yapılan 235.877 adet aratan tespit edilen 2279 adet hata sayısını ifade eden verilerden derlenebilmiřtir.

Arařtırma sonularının tm iřletmelere genellenmesi iin daha farklı sektrlere sahip, farklı illerdeki iřletmeler de rnekleme katılarak alıřmanın rnekleme geniřletilebilir. Bu alıřma otomotiv endstrisinde Birliktelik Kuralları ve Apriori Algoritması aracılıėıyla yapılan bir bařlangı alıřmasıdır. Daha kapsamlı bir alıřma ile elde edilen sonuları da kullanarak aıklama gc daha yksek modellere ulařılabileceėi dřnlmektedir.

Gelecekte yapılacak alıřmalarda seilmiř imalat endstrileri, rneėin gıda ya da tekstil gibi alanlarda karřılařtırmalı alıřmalar yapılabilir. Ayrıca arařtırma modelini basit tutmak iin deėiřken sayısı sınırlandırılmıřtır. Son olarak gelecekteki alıřmalarda ok kriterli karar verme yntemleri yanında yapay sinir aėları, bulanık mantık gibi yapay zeka teknolojilerinin de kullanılması ilgin olacaktır.

5. KAYNAKLAR

- [1] <https://medium.com/@ipekkrdmn/veri%CC%87-madenci%CC%87li%CC%87%C4%9Fi%CC%87-nedi%CC%87r-7a8d936eff95> (Eriřim Tarihi: 21.04.2021)
- [2] Akk, U. (2011). Veri Madencilięi Kmeleme ve Sınıflama Algoritmaları. Yalın Yayıncılık. İstanbul.
- [3] Doęan, O., & Bykberber, S. (2017). Trkiye’de Veri Madencilięi Konusunda Yapılan Lisansst Tezler zerine Bir Arařtırma. *Gazi niversitesi İktisadi ve İdari Bilimler Fakltesi Dergisi*, 19(3), 929-951.
- [4] Aydoęan, E. K. & Gencer, C. (2007). Kaba Kme Yaklařımı Kullanılarak Veri Madencilięi Problemlerinde Sınıflandırma Amalı Yapılmış Olan alıřmalar. *Savunma Bilimleri Dergisi*, 6(2), 17-32.
- [5] Emre, İ. E. & Selukcan, E. . (2017). Veri analizinde istatistik mi veri madencilięi mi. *Biliřim Teknolojileri Dergisi*, 10(2), 161-167.
- [6] Tzntrk, S. (2010). Veri Madencilięi ve İstatistik. *Uludaę niversitesi İktisadi ve İdari Bilimler Fakltesi Dergisi*, 29(1), 65-90.
- [7] Doęan, K., Arslantekin, S. & Grer, A. (2016). Byk Veri: nemi Yapısı ve Gnmzdeki Durum. *Ankara niversitesi Dil ve Tarih-Coęrafya Fakltesi Dergisi*, 56(1).
- [8] Doęan, O. (2017). cretsiz Veri Madencilięi Araları ve Trkiye’de Bilinirlikleri zerine Bir Arařtırma. *Ege Stratejik Arařtırmalar Dergisi*, 8(1), 77-93.
- [9] Karacan, H. & Yeřilbudak, M. (2010). Kullanıcı Merkezli İnteraktif Veri Madencilięi: Bir Literatr Taraması. *Biliřim Teknolojileri Dergisi*, 3(1), 17-22.
- [10] Alaeddinoęlu, M., Aydın, T. & Dal, D. (2014). Birliktelik Kuralları İle Meknsal-Zamansal Veri Madencilięi. *Erzincan niversitesi Fen Bilimleri Enstits Dergisi*, 5(2), 191-212.
- [11] Durmuřoęlu, A. (2017), Veri Madencilięi alıřmaları zerine Bir Analiz: Trkiye Adresli Yayınlar. *Elektronik Sosyal Bilimler Dergisi*, 16(62), 1034-1047.
- [12] Fritzsche, S., Mueller M. & Lanquillon C. (2009). Interactive Quality Analysis in The Automotive Industry: Concept and Design Of An Interactive. Web-Based Data Mining Application. *Enterprise Information Systems* 402-414.

- [13] Shankar, R. & Sundararajan, M. (2015). Manufacturing Quality Improvement With Data Mining Outlier Approach Against Conventional Quality Measurements. *Indian Journal of Science and Technology*, 8(15), 1-8.
- [14] Kruse, R., Steinbrecher, M. & Moewes, C. (2010). Data Mining Applications In The Automotive Industry. *Proceedings of the 4th International Workshop on Reliable Engineering Computing*, 23-40.
- [15] Baysal, M. E., Canyılmaz, E. & Eren, T. (2002). Otomotiv Yan Sanayinde Hata Türü ve Etkileri Analizi. *Z.K.Ü. Karabük Teknik Eğitim Fakültesi Dergisi*, 5(1-2), 83-90.
- [16] Liang, Y. H. (2010). Integration Of Data Mining Technologies To Analyze Customer Value For The Automotive Maintenance Industry. *Expert Systems with Applications*, 37(12), 7489-7496.
- [17] Sönmez, Y. & Unğan, M. C. (2017). Hata Türü Etkileri Analizi ve Otomotiv Parçaları Üretiminde Bir Uygulama. *İşletme Bilimi Dergisi*, 5(2), 217-245.
- [18] Teo, H. G. (2010). Data Mining In Automotive Warranty Analysis. A Thesis Submitted To The Victoria University Of Wellington.
- [19] Kargari, M. & Sepehri, M. M. (2012). Stores Clustering Using A Data Mining Approach For Distributing Automotive Spare-Parts To Reduce Transportation Costs. *Expert Systems with Applications*, 39(5), 4740-4748.
- [20] Karabay, M. (1998). ISO 9001-QS 9000 Kalite Sistem Standartları Kapsamında Bir Gereksinim: Geometrik Ölçülendirme ve Toleranslama. *Mühendis ve Makina*, 39(466), 34-40.
- [21] Aydoğan, E. & Semiz, S. (2005). İşletmelerde İleri Yönetim Teknolojilerinin Kullanım Nedenleri Ve Otomotiv Sektöründe Bir Alan Araştırması. *Politeknik Dergisi*, 8(4), 337-344.
- [22] Atmaca, E. & Keskin, H. (2007). TS 16949 Kalite Yönetim Sisteminin Otomotiv Sektöründeki Yeri. *Süleyman Demirel Üniversitesi İ.İ.B.F. Dergisi*, 12(2), 53-65.
- [23] Ertuğrul, İ. & Aytaç, E. (2014). Kalite Fonksiyon Göçeriminde Markov Zincirleri: Otomotiv Sektörü Örneği. *Dokuz Eylül Üniversitesi İşletme Fakültesi Dergisi*, 8(2), 181-200.
- [24] Kesici, B. & Yıldız, M. S. (2016). Kalite Kontrol Faaliyetlerinde Yapay Zeka Kullanımı ve Bir Otomotiv Yan Sanayisinde Uygulanması. *Yalova Üniversitesi Sosyal Bilimler Dergisi*, 6(12), 307-323.

- [25] Baray, A. & Sarı, T. (2006). Kalite Geliştirmede Deney Tasarımı Yöntemi ve Otomotiv Sektöründe Bir Uygulama. *İstanbul Üniversitesi İşletme Fakültesi İşletme Dergisi*, 35(2), 37-62.
- [26] Arıkan Kargı, V. S. & Aydın, Z. B. (2018). İstatistiksel Kalite Kontrol Teknikleri ile Otomotiv Sektöründe Bir Uygulama. *Yönetim ve Ekonomi Araştırmaları Dergisi*, 16(1), 41-63.
- [27] Surinova, Y. & Skurkova, K.L. (2013). Brief Review Of German Standards For Quality Audits In Automotive Production. *Production Engineering Archives*, 1(1), 35-37.
- [28] Luczak, J. & Wolniak, R. (2015). Problem Solving And Developing Quality Management Methods And Techniques On The Example Of Automotive Industry, *Manager*, 22(1), 237-250.
- [29] Özçalıcı, M. (2017). Veri Madenciliğinde Birliktelik Kuralları ve İkinci El Otomobil Piyasası Üzerine Bir Uygulama. *Ordu Üniversitesi Sosyal Bilimler Araştırmaları Dergisi*, 7(1), 45-58.
- [30] Köksal, G., Batmaz, İ., Testik, M. C., & Güntürkün, F. (2010). İmalat Sektöründe Kalite İyileştirmede Veri Madenciliği Tekniklerinin Kullanımı. *Verimlilik Dergisi*, 47-65.
- [31] Reuter, C., Brambring, F., Weirich, J., & Kleines A. (2016). Improving Data Consistency In Production Control By Adaptation Of Data Mining Algorithms. *Procedia CIRP*, 56, 545-550
- [32] Köksal, G., Batmaz, İ., & Testik, M.C. (2011). A Review Of Data Mining Applications For Quality Improvement In Manufacturing Industry, *Expert Systems with Applications*, 38(10), 13448-13467.
- [33] Çınar, A. & Silahtaroglu, G. (2012). Veri Madenciliği Teknikleri İle Müşteri Memnuniyetine Etki Eden Gizli Nedenlerin Keşfi. *Marmara Üniversitesi İİB Dergisi*, 33(2), 309-330.
- [34] Acar, Ö. E. & Çakırkaya, M. (2016). Bir Üretim Hattında Meydana Gelen Hataların Önem Derecelerinin İstatistiksel Proses Kontrol Tekniklerinden Pareto Analizi ile Belirlenmesi. *Mustafa Kemal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 13(36), 272-288.

- [35] Liu, C. & Sun, Y. (2009). Application Of Data Mining İn Production Quality Management. Third International Symposium on Intelligent Information Technology Application, 21 November, China, 284-287.
- [36] Yalçın Pirinççiler, E. C. & Şen, A. (2012). Süreç İyileştirme Çalışmalarında Veri Madenciliği Yaklaşımının Kullanılması Üzerine Bir Çalışma. Muğla Sıtkı Koçman Üniversitesi Sosyal Bilimler Enstitüsü Dergisi 29, 57-77.
- [37] Wallis, R., Erohin, O., Klinkenberg, R., Deuse, J. & Stromberger, F. (2014). Data Mining-Supported Generation Of Assembly Process Plans. CIRP, Volume 23, 178-183.
- [38] Yılmaz, B. (2000). Hata Türü ve Etki Analizi. Dokuz Eylül Üniversitesi Sosyal Bilimler Enstitüsü Dergisi 2(4), 133-150.
- [39] Asilkan, Ö. (2011). İkinci El Otomobillerin Güncel Pazar Fiyatlarının Veri Madenciliği Yöntemleriyle Modellenmesi. Akademik Bakış Dergisi, 1-19.
- [40] Vazan, P., Janikova, D., Tanuska, P., Kebisek, M., & Cervenaska, Z. (2017, 50 1). Using Data Mining Methods For Manufacturing Process Control. IFAC, 6178-6183.
- [41] Akteke Öztürk, B., Weber, G. W., Kayaligil, S. (2007). Kalite İyilestirmede Veri Kümeleme: Döküm Endüstrisinde Bir Uygulama. Yöneylem Araştırması ve Endüstri Mühendisliği 27. Ulusal Kongresi (YA/EM 2007) Bildiri Kitabı, 1207-1212, Temmuz 2-4, 2007, İzmir.
- [42] Koldere, A. Y. (2008). Veri Madenciliğinde Kümeleme Algoritmaları ve Kümeleme Analizi”, Doktora Tezi. Marmara Üniversitesi Sosyal Bilimler Enstitüsü Ekonometri Anabilim Dalı İstatistik Bilim Dalı İstanbul, 57.
- [43] Gürsoy Şimşek, U. T. (2009). Veri Madenciliği ve Bilgi Keşfi. Pegem Akademi, 1.Baskı, Ankara.
- [44] <http://www.uzaktancrmegitimi.com/> (Erişim Tarihi: 15.01.2017)
- [45] Kıyak, E. (2006). CRISP-DM Yöntembilimi Kullanılarak Deniz Kuvvetleri Verisi Üzerinde Veri Madenciliği Sınıflandırma Tekniklerinin Karşılaştırılması. Yüksek Lisans Tezi, Kocaeli Üniversitesi Fen Bilimleri Enstitüsü, Kocaeli.
- [46] Küçüksille, E. (2009). Veri Madenciliği Süreci Kullanılarak Portföy Performansının Değerlendirilmesi ve MKB Hisse Senetleri Piyasasında Bir Uygulama. Doktora Tezi, Süleyman Demirel Üniversitesi Sosyal Bilimler Enstitüsü, Isparta.

- [47] Oğuzlar, A. (2003). Veri Ön İşleme. Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, Sayı: 21, 67-76.
- [48] Gürbüz, F. (2009). Havacılık Sektöründe Veri Madenciliği ile Farklı Sınıflandırma Tekniklerinin Karşılaştırılmalı Olarak Uygulanması. Doktora Tezi, Erciyes Üniversitesi, Fen Bilimleri Enstitüsü, Kayseri.
- [49] Şentürk, Z. K. (2011). Veri Madenciliği ile Kanser Teşhisi. Yüksek Lisans Tezi, Düzce Üniversitesi, Fen Bilimleri Enstitüsü, Düzce.
- [50] Altıntaş, Y.Y. (2010). Veri Madenciliğinin Tıpta Kullanımı ve Bir Uygulama: Hemodiyaliz Hastaları İçin Risk Seviyelerine Göre Risk Faktörlerinin Etkileşimlerinin İncelenmesi. Yüksek Lisans Tezi, Gazi Üniversitesi, Fen Bilimleri Enstitüsü, Ankara.
- [51] Baloğlu, U.B. (2006). DNA Sırasındaki Tekrarlı Örüntülerin ve Potansiyel Motiflerin Veri Madenciliği Yöntemiyle Çıkarılması. Yüksek Lisans Tezi, Fırat Üniversitesi, Fen Bilimleri Enstitüsü, Elazığ.
- [52] Danacı, M., Çelik, M., & Akkaya, A. E. (2010). Veri Madenciliği Yöntemleri Kullanılarak Meme Kanseri Hücrelerinin Tahmin ve Teşhisi. Akıllı Sistemlerde Yenilikler ve Uygulama Sempozyumu, 9-12, Kayseri.
- [53] Şaylan, Ç. A., (2013). Böbrek Nakli Geçirmiş Hastalarda Akıllı Yöntem Tabanlı Yeni Öznitelik Seçme Algoritması Geliştirilmesi. Yüksek Lisans Tezi, Kadir Has Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.
- [54] Çakır Ö., (2008). Veri Madenciliğinde Sınıflandırma Yöntemlerinin Karşılaştırılması; Bankacılık Müşteri Veri Tabanı Üzerinde Bir Uygulama, Doktora Tezi, Marmara Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul.
- [55] Hand, D., Mannila H. & Smyth P. (2001). Principles of Data Mining, MIT, 546p, USA.
- [56] Piramuthu, S. (2003). Evaluating Feature Selection Methods for Learning in Data Mining Applications. European Journal of Operational Research, Article In Press, 1-11.
- [57] Anne, M.D.B. (2005 Mayıs). Identification of Factors Associated with Postoperative Pneumonia Using a Data Mining Approach, Boston College, p53-54, USA.
- [58] Özmen, Ş. (2003). Veri Madenciliği Süreci, Veri Madenciliği ve Uygulama Alanları, İstanbul: İstanbul Ticaret Üniversitesi, s.37.

- [59] Kiremitçi, B. (2005), Veri Ambarlarında Veri Madenciliği ve Ulaştırma-Lojistik Sektöründe Bir Uygulama, (Yayımlanmamış Yüksek Lisans Tezi), İstanbul Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul.
- [60] Gürgen, G. (2008). *Birliktelik Kuralları ile Sepet analizi ve Uygulaması*, Yüksek lisans tezi, *Marmara üniversitesi Sosyal Bilimler Enstitüsü*, İstanbul.
- [61] Erdoğan, N. K., Gülacan, B. & Karamaşa, Ç. (2017). Birliktelik Kuralları ve Uygulamaları: literatür taraması (2000-2014), http://bildiri.anadolu.edu.tr/papers/bildirimakale/10199_b846w59.docx (Ziyaret Tarihi: 09.04.2018).
- [62] Doğrul, G., Akay, D., & Kurt, M. (2015). Trafik Kazalarının Birliktelik Kuralları ile Analizi. *Gazi Mühendislik Bilimleri Dergisi*, 1(2), 265-284.
- [63] Özçakır, F. C. & Çamurcu, A. Y. (2007). Birliktelik Kuralı Yöntemi İçin Bir Veri Madenciliği Yazılımı Tasarımı ve Uygulaması, *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, 6(12), 21-37.
- [64] <http://yazilimagiris.com/2017/12/apriori-algoritmasi/?print=pdf> (Ziyaret Tarihi: 09.04.2018).
- [65] Han, J. & Kamber, M., (2000), *Data Mining: Concepts and Techniques*, Morgan Kaufmann Publishers, 550p Burnaby. USA.
- [66] Agrawal, R. & Srikant, R. (1994). Fast Algorithms For Mining Association Rules. In *Proc. 20th Int. Conf. Very Large Data Bases, VLDB*, p487-499, San Francisco, USA.
- [67] Yıldırım, H. & Karaca, E. (2013). Üretim Sürecinde İstatistiksel Proses Kontrol (İpk) Uygulamaları ve Elektronik Sektöründe Bir İnceleme. *Öneri Dergisi*, 10(39), 77-87.
- [68] Özcan, S. (2012). İstatistiksel Proses Kontrol Tekniklerinden Pareto Analizi ve Çimento Sanayiinde Bir Uygulama. *C.Ü. İktisadi ve İdari Bilimler Dergisi*, 2(2), 151-174.
- [69] Bircan, H. & Gedik, H. (2003). Tekstil Sektöründe İstatistiksel Proses Kontrol Teknikleri Uygulaması Üzerine Bir Deneme. *C.Ü. İktisadi ve İdari Bilimler Dergisi*, 4(2), 69-79.
- [70] Baykal, A. (2007). Kalite Çalışmalarında İstatistiksel Süreç Kontrol Tekniklerinin İncelenmesi ve Örnek Uygulama. Yüksek Lisans Tezi, İstanbul Ticaret Üniversitesi Fen Bilimleri Enstitüsü, İstanbul.

- [71] Çelik, B. (2019). Mekanik Arızaların Veri Madenciliği Apriori Algoritması ile Analiz Edilmesi. Yüksek Lisans Tezi, Kocaeli Üniversitesi, Fen Bilimleri Enstitüsü, Kocaeli.
- [72] Çakırkaya, M. & Acar, Ö. E. (2016). Bir Üretim Hattında Meydana Gelen Hataların Önem Derecelerinin İstatistiksel Proses Kontrol Tekniklerinden Pareto Analizi ile Belirlenmesi. Mustafa Kemal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 13(36), 272-288.
- [73] Kobu, B. (1987) Endüstriyel Kalite Kontrolü. İstanbul Üniversitesi Yayınları, İstanbul.
- [74] Jafari, A. Y. (2013). Potansiyel Abonelerin Kazanılmasına Yönelik Pazarlama Stratejilerinin Veri Madenciliği ile Geliştirilmesi: Konya Şehri Doğalgaz Örneği. Yüksek Lisans Tezi, Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Konya.
- [75] <http://ninova.itu.edu.tr/tr/dersler/bilisim-enstitusu/195/bbl-606/ekkaynaklar?g8396> (Ziyaret Tarihi: 09.05.2018).
- [76] Ayık, Y. Z., Özdemir, A. & Yavuz, U. (2010). Lise Türü ve Lise Mezuniyet Başarısının, Kazanılan Fakülte İle İlişkisinin Veri Madenciliği Tekniği İle Analizi. Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 10(2), 441-454.
- [77] Zhu, Z., Labbé S., Peña, M. M. O. & Thiele, D. J. (1998). Copper Differentially Regulates the Activity and Degradation of Yeast Mac1 Transcription Factor. Journal of Biological Chemistry, 273(3), 1277-1280.
- [78] Tektaş, N., Çiçek, C. (2019). Türkiye’de Veri Zarflama Konusunda Yazılan Lisansüstü Tezlerin Analizi, II. International Conference on Empirical Economics and Social Science (ICEESS’ 19) June 20-21-22, 2019 / Bandırma, Türkiye
- [79] Şahin, H. (2019). Dumlupınar Üniversitesi Meslek Yüksekokulları’nın Etkinliklerinin Veri Zarflama Analizi ile Karşılaştırılması. Mühendislik Bilimleri ve Araştırmaları Dergisi, 1(1), 49-63.
- [80] Karaibrahimoğlu, A. (2014). Veri Madenciliğinden Birliktelik Kuralı ile Onkoloji Verilerini Analiz Edilmesi: Meram Tıp Fakültesi Onkoloji Örneği, Doktora Tezi, Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Konya.
- [81] Witten I.H., Frank E. (2005). Data Mining: Practical Machine Learning Tools and Techniques, 2nd ed., Elsevier, Amsterdam.
- [82] Çelik, U. (2020). Akıllı Ulaşım Sistemleri ve Büyük Veri, BANU-ITSC’18 1. Uluslararası AUS Konferansı April 19-20-21, 2018 / Bandırma, Türkiye

ÖZGEÇMİŞ

Kişisel Bilgiler		Fotoğraf
Adı Soyadı	Yücel KURTULUŞ	
Doğum Yeri		
Doğum Tarihi		
Uyruğu		
Telefon		
E-Posta Adresi		
Web Adresi	-	

	Eğitim Bilgileri
	Lisans
Üniversite	
Fakülte	
Bölümü	
Mezuniyet Yılı	

	Yüksek Lisans
Üniversite	
Enstitü Adı	
Anabilim Dalı	
Programı	

Makale ve Bildiriler
Şahin, H., Kurtuluş, Y & Tektaş, M. (2021). Kentiçi Otopark Analizi: Çanakkale İli Örneği. Akıllı Ulaşım Sistemleri ve Uygulamaları Dergisi, 4(1), 16-26.