

T.C.
DOKUZ EYLÜL ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
DENİZCİLİK İŞLETMELERİ YÖNETİMİ ANABİLİM DALI
LOJİSTİK YÖNETİMİ PROGRAMI
YÜKSEK LİSANS TEZİ

TÜRK KARGO HİZMET SAĞLAYICILAR İLE İLGİLİ
TWİTTER' DA ÜRETİLEN İÇERİĞİN ANALİZİ

Artun TOMATIR

Danışman
Prof. Dr. Ömür Yaşar SAATÇIOĞLU

İZMİR – 2018

TEZ ONAY SAYFASI



YEMİN METNİ

Yüksek Lisans Tezi olarak sunduğum“Türk Kargo Hizmet Sağlayıcılar İle İlgili Twitter’ Da Üretilen İçeriğin Analizi” adlı çalışmanın, tarafımdan, akademik kurallara ve etik değerlere uygun olarak yazıldığını ve yararlandığım eserlerin kaynakçada gösterilenlerden oluştuğunu, bunlara atıf yapılarak yararlanılmış olduğunu belirtir ve bunu onurumla doğrularım.



05/07/2018

Artun TOMATIR

İmza

ÖZET

Yüksek Lisans Tezi

Türk Kargo Hizmet Sağlayıcılar İle İlgili Twitter’ da Üretilen İçeriğin Analizi

Artun TOMATIR

Dokuz Eylül Üniversitesi

Sosyal Bilimler Enstitüsü

Denizcilik İşletmeleri Yönetimi Anabilim Dalı

Lojistik Yönetimi Programı

Twitter’ in işlemek ve saklamak durumunda olduğu veri miktarı pek çok araştırmacı tarafından büyük veri olarak nitelendirilmektedir. Büyük veri yeni bir kavram olmasına rağmen günümüzde pek çok disiplinde tartışılmaktadır. Artan veri büyüklüğü ve karmaşıklığı yeni teknoloji ve yaklaşımları gerektirmekte ve yeni tekniklerle işletmelere pratik faydalar sağlayabilmektedir. Büyük veri üzerine yapılan araştırmaların temel olarak iki problem üzerine odaklandığı görülmektedir. İlk problem büyük veriyi işleme ve analiz etmek için metot, teknoloji ve algoritmaların geliştirilmesidir. İkinci problem büyük veri analizinin farklı disiplinlere ve problemlere uyarlanması ve kullanılmasıdır. Bu çalışmada *Twitter* tarafından sağlanan kullanıcı tarafından üretilen içerik değerlendirilmiş ve büyük veri analizinin lojistik yönetimde kullanımına ilişkin öneriler geliştirilmeye çalışılmıştır.

Kimi araştırmacılar büyük veriyi değişkenlerin kapsamı ve verinin güvenilirliği konusunda eleştirmektedir. İfadelerin kapsamı bu çalışmanın ilk problemi. Bu nedenle daha önce yapılmış araştırmalardan birinin değişkenlerine ilişkin ifadeler *tweet*lerin içinde aranmıştır. İkinci problem kullanıcı tarafından üretilen içeriğin lojistik yönetimi için nasıl kullanılabilirliği. Farklı *Twitter* kullanıcılarının *Twitter’* ı kullanma amaçları ve çıkarılabilecek kullanışlı bilgiler keşfedilmeye çalışılmıştır. Çalışmanın güçlü yönü büyük veri analizi tekniklerinin lojistik yönetimde kullanımına ilişkin Türkiye’ de yapılan ilk yüksek lisans tezi olmasıdır. Aynı zamanda Türkiye’ de kargo firmaları üzerine yapılan bir kaç çalışma bulunmaktadır.

Araştırmanın zayıf yönü ise sınırlı sayıda *tweet* üzerinde yapılmış olması ve sadece kullanıcı tarafından oluşturulan içeriğin değerlendirilmiş olmasıdır. Pek çok tanıma göre ilişkisel veri tabanlarında saklanabilen ve elle ya da konvansiyonel metotlarla değerlendirilebilecek miktarda veri, örneklem büyük veri kaynağından alınmış olsa bile, büyük veri olarak tanımlanamamaktadır. Bu nedenle bu çalışma büyük veri analizinin lojistik yönetimde kullanımına ilişkin bir ön çalışmadır. Daha önce yapılan çalışmada 31 değişken çalışılmış ve 29' una ilişkin ifadeye örneklemde rastlanılmıştır. Müşterilerin kullanım amacına ilişkin en yüksek frekans şikâyet, konu ise teslimattır. Firmaların kullanım amacına ilişkin en yüksek frekans yardımıdır ve ilk yaptıkları takip numarası, iletişim bilgilerini istemektir. Araştırmacılara göre büyük veri analizinin lojistikte kullanımının üç boyutu bulunmaktadır. Bunlar operasyonel verimlilik, müşteri deneyimi ve yeni iş modelleri geliştirilmesidir. Kimi araştırmacılara göre müşteri hizmetleri lojistik firmalarının fonksiyonlarından biri olarak görülmektedir. Bu çalışma sonunda büyük veri tekniklerinin sosyal medyayı şikâyetleri bulma amacıyla tarama, iletişim bilgileri ve takip numaralarını isteyip ilgili taraflara konu ile birlikte gönderilmesi amacıyla kullanılabilceği düşünülmektedir. Araştırmacılar tarafından *Twitter* konum bilgisi sağlayabilen geo sosyal bir ağ olarak tanımlanmaktadır. Kimi kullanıcılar da rakip analizi ve tedarikçi seçiminde kullanılabilcek konum bilgisi sağlamakta ve kargo firmalarını birbiri ile karşılaştırmaktadır. Bu önerilerin uygulanabilirliği ve işe yararlığı başka araştırmaların konusunu oluşturmaktadır.

Anahtar Kelimeler: Twitter, Sosyal Medya, Büyük Veri, Kargo Firmaları, Müşteri, Teslimat, Şikâyet, Lojistik Yönetimi

ABSTRACT

Master's Thesis

An Analysis of the Content Produced on Twitter about Courier Express

Companies

Artun TOMATIR

Dokuz Eylül University

Graduate School of Social Sciences

Department of Maritime Business Administration

Logistics Management Program

The amount of data which *Twitter* operates on and stores is accepted as a big data by many researchers. The big data is a recent concept; however, it is discussed in many disciplines. The increasing amount of data and complexity both require new technology and approaches and can provide the companies with practical benefits through new analysis methods. The researches on big data are mainly focused on two basic problems. The first problem is developing method, technology and algorithms in order to operate and analyze big data. The second problem is using and adapting the big data analysis to different disciplines and problems. Having analyzed the content that has been produced by *Twitter* and generated by the provided users, this study aims at coming up with proposals related to the uses of big data analysis for logistics management.

Some researchers criticize the big data in terms of the extent of the variables and reliability of the data. The first problem of this study is the scope of the expressions. Thus, the expressions about one of the former researches' variables have been looked up among the *tweets*. The second problem is how the user generated content can be used for the logistics management. It has been attempted to find out the aims of different *Twitter* users and the useful information that can be gained. The strength of the study is that it has been the first master thesis in Turkey about the use of big data analysis techniques in logistics management. There are also few researches on the cargo firms in Turkey. The weakness of the study is that it has been conducted with a limited

number of the *tweets* and only the content of the users has been examined. According to many definitions, although the amount of data that can be stored in relational databases and evaluated through manual or conventional methods has been withdrawn from the samples of big data source, it is not defined as big data. Therefore this study is a pre-research about using big data in logistics management. 31 variables have been studied in the former study and the expressions about 29 of them have been found in the sample. The highest frequency about the purposes of the clients is the complaints and the topic is delivery. The highest frequency about the purposes of the firms is support and the first tracking they do is to ask for the contact information of them. According to the researchers, the use of big data analysis in logistics has three dimensions; these are operational efficiency, customer experience and the development of new business models. Some researchers refer the customer service as one of the functions of logistics firms. As a result of this study, it is thought that big data techniques can be used for scanning social media in order to find complaints, to ask for the contact details and tracking numbers and send them to the related ones together with the subjects. *Twitter* is described as a geo social network which is able to provide the information of location by the researchers. Some users also provide location information that can be used for rival analysis and the choice of suppliers and they compare the delivery firms with one another. Applicability and utility of these proposals compiles the subject for further researches.

Keywords: Twitter, Social Media, Big Data, Cargo Firms, Customer, Delivery, Complaint, Logistics Management

TÜRK KARGO HİZMET SAĞLAYICILAR İLE İLGİLİ TWİTTER' DA ÜRETİLEN İÇERİĞİN ANALİZİ

İÇİNDEKİLER

TEZ ONAY SAYFASI	ii
YEMİN METNİ	iii
ÖZET	iv
ABSTRACT	vi
İÇİNDEKİLER	viii
KISALTMALAR	x
TABLolar LİSTESİ	xi
ŞEKİLLER LİSTESİ	xiii
GİRİŞ	1

BİRİNCİ BÖLÜM BÜYÜK VERİ KAVRAMI

1.1. BÜYÜK VERİ TANIMI	4
1.2. BÜYÜK VERİ ANALİZİ ÜZERİNE GÖRÜŞLER	7
1.2.1. Duygu Analizi	13
1.2.1.1 Veri Sözlüğü Tabanlı Yaklaşım ve Sorunları	15
1.3. BÜYÜK VERİ HAKKINDA TARTIŞMALAR	16
1.4. BÜYÜK VERİ ÜZERİNE YAPILAN ÇALIŞMALAR	21

İKİNCİ BÖLÜM
LOJİSTİK VE KARGO FİRMALARI ÜZERİNE ARAŞTIRMALAR

2.1. LOJİSTİK YÖNETİMİNDE BÜYÜK VERİ ÜZERİNE YAPILAN ÇALIŞMALAR	24
2.2. KARGO FİRMALARI ÜZERİNE YAPILAN ÇALIŞMALAR	30

ÜÇÜNCÜ BÖLÜM
TÜRK KARGO HİZMET SAĞLAYICILAR İLE İLGİLİ TWİTTER' DA ÜRETİLEN İÇERİĞİN ANALİZİ

3.1. ARAŞTIRMANIN EVRENİ VE ÖRNEKLEMİ	33
3.2. ARAŞTIRMA PROBLEMİNİN BELİRLENMESİ	33
3.3. ARAŞTIRMANIN AMACI	33
3.4. ARAŞTIRMANIN KISITLARI	34
3.5. ARAŞTIRMANIN METODOLOJİSİ	34
3.6. VERİLERİN TOPLANMASI VE ANALİZİ	35
3.7. ARAŞTIRMANIN DEĞİŞKENLERİNİN BELİRLENMESİ	37
3.8. BULGULAR	38
3.8.1. Alıcı Profilinden Atılan <i>Tweetler</i>	47
3.8.2. Gönderici Profilinden Atılan <i>Tweetler</i>	50
3.8.3. Firma Profilinden Atılan <i>Tweetler</i>	52
SONUÇ VE ÖNERİLER	59
KAYNAKÇA	63

KISALTMALAR

API	Uygulama Programı Arabirimi(<i>Application Program Interface</i>)
NLP	Dođal dil iřleme (<i>Natural language processing</i>)
NLTK	Dođal Dil Aracı (Natural Language Tool Kit)
NoSQL	SQL ve daha fazlası (Not only SQL)
RFID	Radyo Frekans ile Kimlik Belirleme (Radio Frequency Identification)
SDD	Aynı Gn Teslimat (<i>Same Day Delivery</i>)
SPSS	Sosyal Bilimler İstatistik Programı (Statistical packages for social sciences)
SQL	Yapısal Sorgulama Dili (Structured Query Language)
Tm	Metin Madenciliđi (Text mining)

TABLULAR LİSTESİ

Tablo 1: Makalelerde Var Olan Büyük Veri Tanımları ve Büyük Veride Olması Gereken dört tema:*I - Bilgi, *T - Teknoloji, *M - Metot, *P - Etki.	s. 6
Tablo 2: Büyük veri Zorluklar ve Çözümleri	s. 9
Tablo 3: <i>Tweeter</i> Search API Parametreler ve Açıklamaları	s. 36
Tablo 4: Atılan Profillere Göre <i>Tweetler</i> (%)	s. 38
Tablo 5: <i>Retweet</i> Oranları (%)	s. 38
Tablo 6: Konularına Göre <i>Tweetler</i> (%)	s. 39
Tablo 7: Amacına Göre <i>Tweetler</i> (%)	s. 40
Tablo 8: Konum Bilgisi Sağlama Durumuna göre <i>Tweetler</i> (%)	s. 40
Tablo 9: Memnuniyet ifadesi oranları(%)	s. 41
Tablo 10: Memnuniyetsizlik Ifadesi Oranları(%)	s. 41
Tablo 11: Karşılaştırma Oranları(%)	s. 41
Tablo 12: Çakmak ve Özkan' ın Çalışmasında Ölçülen İfadelerin Örneklem İçindeki Sıklıkları(%)	s. 42
Tablo 13: Çalışma Gün ve Saatleri Hakkında Atılan <i>Tweetler</i> (firma isimleri gizlenmiştir)	s. 46
Tablo 14: Sorun Halinde Tutanak Tutulması Konusunda Rastlanan <i>Tweetler</i> (firma isimleri gizlenmiştir)	s. 47
Tablo 15: Alıcı Profilinden Atılan <i>Tweetler</i> de Takip No Bildirme Oranları(%)	s. 47
Tablo 16: Alıcı Profili Konularına Göre <i>Tweetler</i> (%)	s. 48
Tablo 17: Alıcı Profili Amaçlarına Göre <i>Tweetler</i> (%)	s. 48
Tablo 18: Alıcı Profili Konum Bilgisi Sağlanma Oranına Göre <i>Tweetler</i> (%)	s. 49
Tablo 19: Alıcı Profili Memnuniyet Ifadesi Bulunma Durumunda Göre <i>Tweetler</i> (%)	s. 49
Tablo 20: Alıcı Profili Memnuniyetsizlik Ifadesine Göre <i>Tweetler</i> (%)	s. 49
Tablo 21: Alıcı Profili Karşılaştırma Yapılmasına Göre <i>Tweetler</i> (%)	s. 50
Tablo 22: Gönderici Profili Konusuna Göre <i>Tweetler</i> (%)	s. 50
Tablo 23: Gönderici Profili Amacına Göre <i>Tweetler</i> (%)	s. 50
Tablo 24: Gönderici Profili Konum Bilgisi Sağlama Oranına Göre <i>Tweetler</i> (%)	s. 51

Tablo 25: Gönderici Profili Memnuniyet İfadesi Bulunma Oranına Göre <i>Tweetler</i> (%)	s. 51
Tablo 26: Gönderici Profili Memnuniyetsizlik İfadesi Bulunma Oranına Göre <i>Tweetler</i> (%)	s. 51
Tablo 27: Gönderici Profili Karşılaştırma Oranına Göre <i>Tweetler</i> (%)	s. 52
Tablo 28: Firma Profili Takip Numarası Talebine Göre <i>Tweetler</i> (%)	s. 52
Tablo 29: Firma Profili Konularına Göre <i>Tweetler</i>	s. 53
Tablo 30: Firma Profile Amacına Göre <i>Tweetler</i>	s. 53
Tablo 31: Müşteri destek hesabından paylaşılan <i>tweetler</i>	s. 54



ŞEKİLLER LİSTESİ

Şekil 1: Duygu Sınıflandırma Teknikleri	s. 14
Şekil 2: Sosyal medya madenciliği veya analizinin 3 basamağı	s. 15
Şekil 3: Profillere Göre Atılan <i>Tweet</i> lerin Özeti	s. 54
Şekil 4: Genel Olarak Atılan <i>Tweet</i> lerin Özeti	s. 55
Şekil 5: Lojistiğin Fonksiyonlarına Göre Atılan <i>Tweet</i> ler	s. 55



GİRİŞ

Twitter işlemek ve saklamak zorunda olduğu veri miktarı bakımından kimi araştırmacılar tarafından büyük veri olarak tanımlanmaktadır. Buna rağmen, literatürden anlaşıldığı kadarıyla, büyük veri analizinden bahsedebilmek için büyük veri analizinde kullanılan teknoloji ve yöntemlerin kullanılması, analiz edilen verinin analizi bakımından bazı zorluklar taşınması veya çok büyük hacme, hıza ve çeşitliliğe sahip olması beklenmektedir. Bu bakımdan *Twitter*' dan analizi için alınan sınırlı sayıdaki örneğin büyük veri olarak tanımlanabilmesi için büyük verinin niteliği olarak geçen hacim, hız ve çeşitlilik gibi özelliklerin tamamını sağlamayabileceği yapılan araştırmalarda göz önünde bulundurulmadır. Tedarik zinciri kapsamında *tweet* analizi için yapılmış çalışmalara literatürde rastlanmaktadır, analiz edilen *tweet* sayısı düşük olmasına rağmen söz konusu çalışma büyük veri analizi bağlamında *tweet*lerin analizi için yöntem önermektedir. Büyük veri analizinin lojistik yönetiminde üç boyutta kullanılabileceği belirtilmekte bunlar operasyonel verimlilik, müşteri deneyimi ve yeni iş modelleri olarak geçmektedir. Başka bir araştırma sosyal medyayı, müşteri deneyimi başlığı altında, müşteriler hakkında bilgi toplanabilecek bir kaynak olarak ifade etmektedir. Bu çalışma bir ön çalışma olarak büyük veri analizinin lojistik yönetiminde kullanımını tartışmayı amaçlamaktadır.

Büyük verinin terim olarak kullanımı 1970'lere uzanmaktadır ve yapılan çalışmalar gün geçtikçe artmakta, büyük verinin işletme ve araştırmalara sağlayabileceği fayda potansiyelinden söz edilmektedir. Geçmişte araştırmacılar tarafından dikkate alınmayan kimi kaynakların araştırma konusu olabildiği görülmekte ve büyük veri analizi tekniklerinin araştırmacılara ve işletmelere yeni imkânlar sağladığından bahsedilmektedir. Söz konusu imkânlar daha çok değişken arasında muhtemel ilişkilerin bulunması, popülasyon verisini elde ederek daha isabetli çıkarımların yapılması, olguların daha iyi açıklanması, tahminlerin daha isabetli olması gibi pek çok başlıkta değerlendirilmektedir.

Söz konusu yöntemlerin farklı alanlarda kullanımına ilişkin çalışmalar olduğu görülmekte, kimi çalışmalarda gemi hızlarının optimizasyonlarında karşılaşılan tahmini yakıt tüketimi ile gerçek yakıt tüketimi arasındaki sapmanın hava

durumunun göz önünde bulundurularak azaltılabileceği belirtilmekte; kimi çalışmalarda sosyal ağ üzerinden seçim tahminleri yapıldığı görülmektedir.

Büyük veri analizi tekniklerinin lojistik yönetimde kullanılmasına ilişkin araştırmalar da bulunmakta ve büyük veri analizi tekniklerinin lojistik yönetimde kullanılabileceği muhtemel alanlar konusunda araştırmacıların öneriler sunduğu görülmektedir. Literatüre bakıldığında verinin artan hızda üretilmesinin yeni yöntem ve teknolojileri zorunlu kıldığı, söz konusu yeni yöntem ve teknolojilerin ise büyük miktardaki veriden kullanışlı bilgiler elde etme imkânı sağlayabildiği ifade edilmektedir. Bu bağlamda literatür incelendiğinde büyük verinin işleme, saklama, analiz etme gibi konularda kimi zorluklar barındıracak ve bu problemlerin çözümü için yeni teknoloji ve yaklaşımlar gerektirecek büyüklükte veri setleri olduğu anlaşılmaktadır. Bu nedenle araştırmacılar genellikle, büyük veriyi tanımlamak için, büyük verinin kimi özelliklerine atıf yapmakta ve bunu 3V ile açıklamaktadır. Söz konusu karakteristikler hacim (*volume*), hız (*velocity*), çeşitlilik (*variety*) olarak geçmektedir. Yine de literatüre bakıldığında büyük verinin üzerinde uzlaşılan bir tanımının bulunmadığı görülmektedir. Literatürden anlaşıldığı kadarıyla büyük veri analizinin ise söz konusu çok büyük; işleme, analiz etme bakımından kimi zorluklar içeren ve çeşitliliği bulunan veri setlerinden kullanışlı bilgiler çıkarmayı amaçlayan yöntem ve teknolojiler olduğu görülmektedir. Büyük veri analizi için daha çok yöntemsel çalışmanın yapılması gerektiği pek çok araştırmacı tarafından dile getirilmektedir. Büyük veri analizinin alışılan yöntemlerin dışında bir yaklaşım olduğuna dair çalışmalara da literatürde rastlanmaktadır. Söz konusu çalışmalar tümevarım, nedensellik, kuram konusunda geleneksel yöntemle büyük veri analizi arasında bir takım farklılar bulunduğunu belirtmektedir.

Literatürde büyük veri analizi konusunda rastlanan unsurlardan biri de uzman sayısının yetersizliği, kullanılan teknolojilerin farklılığı ve doktora seviyesinde verilen araştırma metot ve tekniklerinin büyük veri araştırmaları için yeterli olmamasıdır. Ayrıca metot üzerine daha çok çalışma yapılması gerektiği belirtilmektedir. Kimi araştırmacıların söz konusu teknikleri disiplinlerine uyarlama konusunda çalışmalar yaptığı görülmektedir. Bu nedenle araştırmacılar tarafından derlenmiş olan büyük veri analizine hızlı başlangıç kılavuzu ve analiz tekniklerine, lojistik yönetimde büyük veri analizi teknikleri kullanılması durumunda, hangi

araştırma yöntemlerine ihtiyaç duyulabileceğine dair bilgi sağlaması bakımından bu çalışmada yer verilmiştir.

Farklı disiplinlerden araştırmacılar konuyu ihtiyaçları bağlamında farklı yönleri ile ele almaktadır, ancak pek çok disiplinde büyük veri analizinin tartışıldığı görülmüştür. Kimi araştırmacılar hızın yönetim ve organizasyon çalışmaları için önemli olmadığını belirtmekte, kimi araştırmacılar ise söz konusu verinin geçerliliği, güvenilirliği, toplanma şekline ilişkin ve araştırma değişkenlerinin tamamını içerip içermediğine ilişkin sorular sormaktadır. Büyük veri analizi tekniklerinin farklı disiplinlerde uygulanması kapsamında metoda ilişkin çalışmalarda literatür de bulunmaktadır. Buna rağmen büyük verinin olguların açıklanmasında potansiyel taşıdığı; şehirlerin iyileştirilmesinden, salgın hastalıkların izlenmesine pek çok alanda kullanılabileceği görülmektedir.

Büyük veri analizinin kullanımına ilişkin tartışmaların bulunduğu bir alan da lojistik yönetimidir. Bu konuda yapılan çalışmalarda büyük veri analizinin üç boyutta kullanılabileceği, bunların operasyonların verimliliği, müşteri deneyimi ve yeni iş modelleri olduğunun altı çizilmektedir. Sosyal medya söz konusu başlıklarda müşteri deneyimi konusunda geçmektedir. Bu nedenle kargo firmalarının etiketlendiği *tweet*ler bu çalışmada incelenerek, kullanıcıların *Twitter*' i kullanım amaçları, sağladığı yararlı olabilecek bilgiler tespit edilmeye ve bu konuda öneriler geliştirilmeye çalışılmıştır. Ülkemizde kargo firmaları konusunda memnuniyet analizi ve önem başarımları analizi yapıldığı görülmüş, literatürde yöneltilen araştırma değişkenlerinin hepsini taşıyıp taşımadığı eleştirisi göz önünde bulundurularak, bütün değişkenleri kapsayıp kapsamadığına bakılmıştır.

BİRİNCİ BÖLÜM

BÜYÜK VERİ KAVRAMI

1.1. BÜYÜK VERİ TANIMI

Ularu ve diğerleri (2012) büyük veri kavramının Roger Magoulas tarafından 2005 yılında geleneksel veri yönetimi tekniklerinin yönetemeyeceği ve işleyemeyeceği büyüklük ve karmaşıklığa sahip verileri tanımlamak amacıyla kullanıldığını belirtmektedir. *The IBM Big Data Flood Infographic*'e göre dijital evrende 2,7 zettabayt verinin bulunduğu, 2020 yılına kadar yıllık veri üretiminin sosyal ağlardaki aktiviteyle 35 zettabayta ulaşacağı ifade edilmektedir. Büyük verinin taşıdığı zorluklar büyük verinin anlaşılması, yeni geliştirilen teknolojiler ve bilişim uzmanı ihtiyacı olarak geçmektedir. Makalede büyük veri analizinin kullanılabileceği yerlerden biri, müşteri tatminini yükseltmek amacıyla, müşteri hizmetleri olarak verilmiştir. Ularu ve diğerleri (2012) büyük veriyi karakterize eden 4 boyuttan bahsetmektedir. Bunlar hacim (*volume*), hız (*velocity*), çeşitlilik (*variety*), gerçeklik (*veracity*) olarak geçmektedir.

Halevi ve Moed (2012) büyük veri kavramına literatürde ilk olarak 1970 yılında Barbados Oşinografik ve Meteorolojik Deney isimli çalışmada rastlanıldığını belirtmekle beraber, yapılan araştırmaların 2008 yılından sonra artış gösterdiğini ifade etmektedirler. Literatürü büyük veri konusunda taramışlar ve 306 araştırma bulmuşlardır. Büyük veri konusunda en çok bilgisayar bilimlerinde araştırma yapılırken; bilgisayar bilimlerini sırasıyla mühendislik, matematik, işletme, fizik ve astronomi, biyokimya-genetik-mikrobiyoloji, sosyal bilimler, malzeme bilimi, tıp, karar bilimleri, disiplinler arası çalışmalar, sanat ve beşeri bilimler takip etmektedir. Bu günlerde büyük verinin farklı bakış açılarıyla farklı disiplinlere uygulanması konusunda çalışıldığını ifade etmektedirler. Bu tezin konusunu da büyük veri konusunda lojistik yönetimindeki tartışmalar ve muhtemel uygulamaları oluşturmaktadır.

Büyük veri pek çok çalışmada 3V ile açıklanmaktadır (Dutta ve Bose, 2015; Romero ve diğerleri, 2015; George ve diğerleri, 2016) Gandomi ve Haider (2015) çalışmalarında 3V' yi şu şekilde açıklamışlardır: Hacim (*Volume*), *terabyte*, *petabyte*

gibi verinin büyüklüğünü ifade etmektedir. Çeşitlilik (*Variety*), video, fotoğraf, ses kaydı gibi ya da yapılandırılmış, yarı yapılandırılmış, yapılandırılmamış gibi çeşitliliğini ifade eder. Hız (*Velocity*) ise verinin üretilme ve analiz hızı olarak alınmıştır.

George ve diğerleri (2014) büyük verinin internetteki tıklamalar, mobil işlemler, kullanıcı tarafından oluşturulan içerik, sosyal medya ile birlikte sensör ağları ve satış sorguları, alış işlemleri gibi ticaret işlemleri tarafından üretildiğini belirtmektedir.

Baro ve diğerleri (2015) sağlık bilimlerinde büyük veri üzerine yapılan çalışmaları araştırmışlar ve yapılan çalışmalarda büyük veri karakteristiklerinin çok büyük hacim ve karmaşıklık ile karakterize olduğunu bulmuşlardır. Büyük verinin sadece çok geniş kapsamı olan yeni veri setleri olmadığını, söz konusu veri setlerinin artan karmaşıklığının bulunduğunu söylemektedirler.

Borgi ve diğerleri (2017) büyük verinin pek çok tanımı olduğunu belirtmekte ve büyük verinin genellikle klasik yöntemlerle seçme, depolama, işleme yapılamayan veri setlerini ifade ettiğini söylemektedirler. Büyük veri analizinin ise matematiksel ve istatistiksel yöntemler kullanılarak söz konusu büyük veri setlerinden yeni ve kullanışlı bilgilerin çıkarılması, verilerin anlaşılması, senaryoların simüle edilmesi, hipotezlerin test edilmesi ve tahminleme çalışmalarının yapılması olarak ifade etmektedirler.

De Mauro ve diğerleri (2016) *co-word* ve Cohen's *K coefficient* ile literatürdeki bulunan büyük veri terimleri üzerinden konsensüs sağlanabilecek bir büyük veri tanımı üzerinde çalışmışlardır. Büyük veri tanımlarını listelemiş, büyük veriyi değere çevirmek için özel teknoloji ve analitik yöntemler gerektiren ve yüksek hacim, hız ve çeşitlilik ile karakterize olan bilgiye ilişkin varlıklar olarak tanımlamışlardır (Tablo 1). Wong ve diğerleri de (2016) büyük verinin farklı disiplinlerden uzmanların üzerinde anlaştığı bir tanımının bulunmadığından bahsetmektedir.

Tablo 1: Makalelerde Var Olan Büyük Veri Tanımları Ve Büyük Veride Olması Gereken Dört Tema: *I - Bilgi, *T - Teknoloji, *M - Metot, *P - Etki.

Kaynak	Tanım	*I	*T	*M	*P
(Beyer & Laney 2012)	Kavrama ve karar vermeyi iyileştirmek için maliyet etkin ve yenilikçi bilgi işleme formlarına ihtiyaç duyulan yüksek hacim, hız ve çeşitliliğe olan bilgi varlıkları.	X		X	X
(Dijcks 2012)	Büyük veriyi tanımlayan 4 nitelik hacim, hız çeşitlilik ve değerdir.	X			X
(Intel 2012)	Karmaşık, yapılandırılmamış veya büyük miktarda veri	X			
(Suthaharan 2013)	Üç veri niteliği kullanılarak tanımlanabilir: nicelik, süreklilik ve karmaşıklık	X			
(Schroeck ve diğerleri. 2012)	Büyük veri günümüzün dijitalleşen piyasasında kurumlara rekabet avantajı kazanmak için fırsat yaratan hacim, çeşitlilik, hız ve gerçeklik birleşimidir.	X			X
(NIST Big Data Public Working Group 2014)	Depolama, işleme ve analiz için ölçeklendirilebilir mimariyi gerektiren, öncelikli nitelikleri hacim, hız ve çeşitlilik olan, çok geniş veri setleridir.	X	X		
(Ward & Barker 2013)	Geniş ve karmaşık veri setlerinin NoSql, MapReduce ve machine learning teknikleri içeren ancak bunlarla sınırlı olmayan bir dizi teknikle depolanma ve işlenmesidir.	X	X	X	
(Microsoft 2013)	Çok büyük ve genellikle karmaşık veri setlerine ciddi miktarda programlama gücü uygulanması sürecidir.	X	X	X	
(Dumbill 2013)	Konvansiyonel veri tabanı sistemlerinin işleme kapasitesini aşan veridir.	X	X		
(Fisher ve diğerleri. 2012)	Kolay bir yöntemle elde edilemeyen ve işlenemeyen veridir.	X		X	
(Shneiderman 2008)	Ekrana sığdırmak için çok büyük olan bir veri seti	X			
(Manyika ve diğerleri. 2011)	Normal veri tabanlarının yakalama, depolama, yönetme ve analiz yeterliliklerinin ötesinde büyüklüğü olan veri setleridir.	X	X	X	
(Chen ve diğerleri. 2012)	Gelişmiş ve özgün veri depolama, yönetme, analiz ve görselleştirme teknolojileri gerektiren çok geniş ve karmaşık veri setleri ve analitik tekniklerdir.	X	X	X	
(Boyd & Crawford 2012)	Teknoloji, analiz ve mitolojinin karşılıklı etkileşimine dayanan kültürel, teknolojik ve akademik bir olgudur		X	X	X
(Mayer-Schönberger & Cukier 2013)	Toplum nasıl anlayacağımız ve düzenleyeceğimiz konusunda veri analizi yöntemine üç temel değişim getiren olgu 1. Daha çok veri 2. Karmaşık (tamamlanmamış) 3. Nedenselliğe erişen korelasyon	X		X	X

Kaynak: De Mauro ve diğerleri, 2016

White ve Breckendridge' in (2014) çalışmalarında aktardıklarına göre National Science Foundation büyük veriyi geniş, dağınık, karmaşık, boylamsal ve/veya sensörler, internet işlemleri, tıklamalar, e-mailler, videolar tarafından üretilen; bu gün veya gelecekte erişilebilecek bütün ulaşılabilen dijital kaynaklar olarak tanımlamıştır.

Rathore ve diğeri (2017) *Twitter, Facebook, Foursquare ve Flickr* gibi geo sosyal ağların konum bilgisi ile birlikte pek çok veri sağladığını, söz konusu ağların milyonlarca kullanıcısı bulunduğunu, kullanıcılar tarafından terabaytlarca heterojen verinin her gün sağlandığını belirtmekte ve bu veriye büyük veri demektedir.

George ve diğeri (2016) özellikle yönetim ve organizasyon araştırmalarında, büyük verinin kullanımının pek çok kurum ve kişi için yeni bakış açıları kazandırabileceğini düşünmektedir. Bu konuda aynı pazarda, benzer ürünlerle rekabet eden iki firmanın araştırma geliştirme faaliyetleri ya da üretimde işbirliğine gidebileceği yönünde örnek verilmiştir. Büyük veriden elde edilecek yeni kavrayışların (*insight*) istatistikleri, anketleri ve arşiv verisini tamamlayabileceği söylenmektedir.

1.2. BÜYÜK VERİ ANALİZİ ÜZERİNE GÖRÜŞLER

George ve diğeri (2016) tarafından büyük veri analizine ilişkin olarak 5 alan belirlenmiştir. Bunlar (1) veri toplama (2) veri saklama (3) veri işleme (4) veri analizi (5) raporlama ve görselleştirme. Veri bilimine ilişkin bir hızlı başlangıç kılavuzu oluşturulmaya çalışılmıştır. Yapılan çalışmada literatürde kullanılan yöntem ve araçlar tablo ile listelenmiştir. Bilgisayar bilimlerinde kimi araştırmacılar büyük veriyi konvansiyonel yöntemlerle işlenmesi mümkün olmayan büyüklük ve çeşitlilikte verinin anlık olarak işlenebilmesi ile ilişkilendirmektedir. George ve diğeri (2016) tarafından yapılan çalışmanın büyük veri tanımı ve analizi konusunda, hem önerilen araçlar hem de 3V' ye yapılan atıfla, farklı disiplinlerden araştırmacıların büyük veri konusundaki ayrışmalarına örnek temsil ettiği düşünülmektedir. George ve diğeri (2016) büyük veri analizi hakkında görüşleri aşağıdaki gibidir:

Veri İşleme: Metinsel verinin kuram geliştirme ve hipotez test etmede kullanılabileceği söylenmiş örnek olarak hakkında iyi haberlerin sık olarak yayımlandığı firmalardan çalışanların ücret beklentilerinin daha yüksek olacağı yönünde bir hipotez verilmiştir. Verinin analizi için işlenerek uygun hale getirilmesi gerekmektedir. Bunun için tipik uygulamaların noktalama işaretlerinin kaldırılması, kelimelerin karakter olarak ayrılması ve bütün karakterlerin aynı karakterleri

yakalayabilmek için küçük harf olarak düzenlenmesi olduđu söylenmektedir. Metin işleme için kullanılabilir araçlar NLTK (Natural Language Processing Tool Kit) ve tm (Text Mining) olarak belirtilmiştir.

Veri Analizi: Veri işleme ile ilgili karşılaşılmaması muhtemel iki problemden bahsedilmektedir. Bunlar çok büyük miktarda açıklayıcı potansiyel değişkenin bulunması ve verinin kişisel bilgisayarlarla işlenemeyecek kadar büyük olmasıdır. Değişken seçiminde açıklayıcı çalışmalarda istatistiksel metotlara ihtiyaç bulunduđu belirtilmiştir. Büyük verinin tek bir bilgisayarda tutulması halinde bile analiz için çok büyük olabileceği belirtilerek, analiz için üç yöntem önerilmiştir. Bunlar *parallelization, bags of little bootstrap, sequential updating* olarak belirtilmektedir.

Raporlama ve Görselleştirme: Görselleştirme için *Bime, Qlik Sense, ve Tableau* önerilmiştir. Raporlama konusunda etkisi az olan değişkenlerle bile veri seti büyüklüğü nedeniyle anlamlı bir ilişki gösterebileceğinden hareketle Bayes istatistiği önerilmiştir.

Veri Saklama: Büyük verinin saklama ihtiyacının sıklıkla masa üstü ve diz üstü bilgisayarları geçtiği ifade edilmektedir. Görece olarak küçük veri setlerinin Excel, SAS, SPSS ve STATA ile saklanabileceği söylenmiştir. Daha büyük veri setleri için SQL, NoSQL veri tabanlarının kullanılabilmesi; söz konusu veri tabanlarının kullanımı için temel seviyede SQL bilmenin gerekeceği iletilmiştir. NoSQL' in daha hızlı veri işlemenin gerektiği durumlarda kullanılabilmesi söylenmiştir.

Literatürden büyük veri analizine ilişkin zorluklar ve çözümleri toplayarak aşağıdaki şekilde tabloştürmüştür.

Tablo 2:Büyük Veri Zorluklar ve Çözümleri

Süreç	Zorluklar	Çözümler	Referanslar
Veri ulaşımı ve toplama	<ul style="list-style-type: none">-Veriye önerilen standartta ulaşım.Bu verilerin pratik bir limitinin bulunmaması limitsiz ölçekte önerilmesi-Detaylı veriyi büyük miktarda temsilciden verimli biçimde temin etmek-Güvenlik, kişisel gizlilik ve veri hakları konusunda iletişim kuralları	<ul style="list-style-type: none">-Sensörler-Web kazıma-Web trafiği ve iletişimini dinleme	<ul style="list-style-type: none">-Chaffin et al. (2015)-Sismeiro and Bucklin (2004)
Veri Saklama	<ul style="list-style-type: none">-Farklı büyük veri setlerinin bütünleştirilmesi için veri saklama araçları-Veri güvenilirliği-Depolama	<ul style="list-style-type: none">-SQL, NoSQL, Apache Hadoop-Sadece ana bilgiyi gerçek zamanlı olarak kaydetme ve güncelleme	<ul style="list-style-type: none">-Varian (2014)-Prajapati (2013)
Veri işleme	<ul style="list-style-type: none">-Kantitatif analiz için numerik olmayan veri kullanımı	<ul style="list-style-type: none">-Metinleri numaralara çevirmek için metin madenciliği araçları-Duygu tanıma	<ul style="list-style-type: none">-Manning, Raghavan, and Schütze (2009)-Teixeira, Wedel, and Pieters (2012)
Veri Analizi	<ul style="list-style-type: none">-Büyük miktarda değişken-Nedensellik-Örtülü konuları bulma ve anlam ile bağlama-Verinin işleme için çok büyük olması	<ul style="list-style-type: none">-Ridge, lasso, principal components regression, kısmi en küçük kareler, regression trees-Topic modeling, latent Dirichlet allocation, entropy-based measures, and deep learning-Cross-validation and holdout samples-Alan deneyi-Parallelization, bags of little bootstrap, sequential analysis	<ul style="list-style-type: none">-Hastie, Tibshirani, and Friedman (2009)-George and McCulloch (1993)-Archak, Ghose, and Ipeirotis (2011)-Tirunillai and Tellis (2012)-Blei, Ng, and Jordan (2003)-LeCun, Bengio, and Hinton (2015)-Lambrech and Tucker (2013)-Wang, Chen, Schifano, Wu, and Yan (2015)-Wedel and Kannan (2016)

Raporlama ve gör-selleştirme	-Açıklamayı kolaylaştırma, dış partnerlerle ve knowledge userlar ile sunum -Kompleks örüntülerin anlaşılmasının zorluğu	-Veri kaynaklarını tanımla -Yöntem ve özelliklerini tanımla -Bayes analizi -Görselleştirme ve grafiklerle açıklama	-Loughran and McDonald (2011) -Simonsohn, Simmons, and Nelson (2015)
------------------------------	--	---	---

Kaynak: George ve diğerleri, 2016

Manyika ve diğerleri (2011) büyük veri analizinde kullanılabilecek yöntemleri aşağıdaki şekilde açıklamış ve listelemiştir.

- A/B testi (*A/B testing*): Bir kontrol grubu birkaç çeşit test grubu ile karşılaştırılarak hangi değişimlerin hedef değişkeni arttırdığına bakılır. Renkler, anahat, resimlerde yapılan değişikliklerin geri dönüş oranı üzerine etkisinin araştırılması örnek olarak verilmektedir. Büyük verinin istatistiksel olarak anlamlı olabilecek büyüklükte veri sağlayabildiğinden bahsedilmektedir.
- İlişkilendirme Kuralı Öğrenme (*Association Rule Learning*): Geniş veri setlerinde ilginç ilişkilerin keşfedilmesi amacıyla çeşitli algoritmaları içeren bir dizi teknik olarak belirtilmektedir. Örnek olarak kullanıcıların hangi ürünleri beraber aldığı araştırılması verilmiştir.
- Sınıflandırma (*Classification*): Daha önceden kategorilere ayrılmış bir veri setine göre yeni verilerin hangi kategoriye ait olduğunun araştırıldığı bir dizi teknik olarak geçmektedir. Örnek olarak segment bazında müşteri davranışı araştırmaları verilmiştir.
- Küme Analizi (*Cluster Analysis*): Benzer karakteristiklere göre nesnelerin daha küçük gruplara bölüdüğü istatistiksel metot olarak geçmektedir.
- Kitle Kaynak Kullanımı (*Crowdsourcing*): Açık bir çağrı yaparak geniş bir grup insandan veri toplanması olarak geçmektedir.
- Veri Füzyonu Veri Entegrasyonu (*Data fusion and data integration*): Birden çok kaynaktan gelen verinin analizi ve entegrasyonu için kullanılan teknikler olarak geçmektedir. Pazarlama kampanyasının etkilerini görmek üzere

natural language processing (NLP) ile işlenen sosyal medya verisinin, gerçek zamanlı satış verisi ile kombine edilebileceği belirtilmiştir.

- Veri madenciliği (*Data mining*): Veri tabanı yönetimi, yapay zekâ ve istatistiksel metotların kombinasyonu ile çok geniş veri setlerinden ilişkilerin çıkarılmasını amaçlayan bir dizi teknik olarak aktarılmaktadır.
- Topluluk Öğrenimi (*Ensemble learning*): Daha iyi tahminleme için birden fazla tahminleme yönteminin kullanıldığı teknik olarak geçmektedir.
- Genetik Algoritmalar (*Genetic algorithms*): En iyinin hayatta kaldığı doğal evrimden esinlenilmiş, her bir bireyin bir model ortamında test edildiği bir teknik olarak belirtilmiştir. Üretim için iş sıralama ve yatırım portföyünün performansının iyileştirilmesi örnek olarak verilmektedir.
- Makine Öğrenimi (*Machine Learning*): Ampirik bir veri setine dayanarak bilgisayarların davranışlarını geliştirdiği teknik olarak açıklanmıştır.
- Doğal Dil İşleme (*Natural language processing*): NLP 'nin bir çeşit *machine learning* tekniği olduğu belirtilmektedir. Sosyal medyada duygu analizi örnek olarak verilmiştir. İnsan dilini analiz etmek için kullanılan algoritmalar olduğu belirtilmektedir.
- Sinir Ağları (*Neural networks*): Sinir ağlarının çalışmasından esinlenen teknik olarak açıklanmaktadır. Doğrusal olmayan örüntülerin bulunmasına uygun bir metot olarak belirtilmektedir. Şirket için önemli olan ve kaybedilme riski taşıyan müşterilerin belirlenmesi ve sigorta dolandırıcılığının tespiti örnek olarak verilmiştir.
- Ağ Analizi (*Network analysis*): Birey ve kurumlar arasındaki bağlantıların incelendiği bilginin nasıl aktarıldığı, kimin başkaları üzerinde daha çok etkisinin bulunduğu ilişkin analiz olarak geçmektedir. Bilgi akışındaki dar boğazların belirlenmesi uygulamalara örnek olarak verilmiştir.
- Optimizasyon (*Optimization*): Karmaşık sistem ve süreçlerin yeniden dizaynını amaçlayan bir dizi teknik olarak belirtilmektedir. Uygulamalara örnek olarak zaman çizelgeleri, rotalama gibi alanlar verilmektedir.
- Örüntü Tanıma (*Pattern recognition*): Bir dizi yapay zekâ tekniği olarak belirtilmektedir. Sınıflandırma teknikleri örnek olarak verilmiştir.

- Tahminli Modelleme (*Predictive modeling*): Matematiksel modeller ile hangi tahminin en olası olduğunu bulmayı amaçlayan tekniklere verilen isimdir. Uygulamalara örnek olarak tedarikçi değiştirmesi muhtemel olan müşterilerin belirlenmesi verilmiştir. Regresyon analizi pek çok *predictive modelling* tekniğinden biri olarak belirtilmektedir.
- Regresyon (*Regression*): Bir veya daha fazla bağımsız değişkenin değiştiği durumda bağımlı değişkenin nasıl değiştiğinin analiz edildiği istatistiksel bir teknik olarak geçmektedir. Genellikle tahminleme amacıyla kullanıldığı belirtilmektedir. Uygulamalara örnek olarak satış hacimlerinin pazar ve ekonomik verileri ilişkin olarak tahminlenmesi verilmektedir.
- Duygu Analizi (*Sentiment analysis*): NLP gibi tekniklerin metinler üzerinde pozitif, negatif, nötr gibi kutuplarının bulunması ve ifadelerin gücünün belirlenmesi amacıyla uygulandığı bir teknik olarak geçmektedir. Sosyal medya ve bloglar üzerinde yapılan araştırmalarla tüketicilerin ve yatırımcıların bir firmanın ürünlerine verdikleri tepkilerin ölçülmesi uygulamalara örnek olarak verilmektedir.
- Sinyal işleme (*Signal processing*): Uygulamalı matematik ve elektronik mühendisliğini içeren bir dizi teknik olarak belirtilmektedir.
- Mekansal Analiz (*Spatial analysis*): Veri setinde bulunan topolojik, geometrik ve coğrafik verilerin analizini amaçlayan bir dizi teknik olarak belirtilmektedir. Konuma göre müşterilerin bir ürünü alma isteğine ilişkin araştırmalar, farklı bölgelerde tedarik zincirinin nasıl performans göstereceğine ilişkin araştırmalar uygulamalara örnek olarak verilmektedir.
- İstatistik (*Statistics*): Özellikle değişkenler arasında ilişki bulmaya dönük analizlerde kullanıldığı, değişkenler arasında bulunan hangi ilişkilerin rastlantısal, hangilerinin istatistiksel olarak anlamlı olduğunu belirlemeyi amaçladığı teknik olarak belirtilmektedir.
- Öğreticili Öğrenme (*Supervised learning*): Bir dizi *machine learning* tekniğidir. İlişkiler veya fonksiyonların bir veri seti üzerinden çıkarıldığı belirtilmektedir.

- Simülasyon (*Simulation*): Karmaşık sistemlerin davranışlarının modellendiği teknik olarak geçmektedir. Belirsizlik durumunda finansal hedeflerin karşılanmasına yönelik analizler uygulamalara örnek olarak verilmektedir.
- Zaman serisi Analizi (*Time series analysis*): *Signal processing* ve istatistiğin kullanıldığı veriden anlamlı parçalar çıkarmaya yarayan bir dizi teknik olarak geçmektedir. Uygulama alanı olarak satış rakamlarının tahmin edilmesi ve bir salgından kaç kişinin etkileneceğine ilişkin çalışmalar verilmiştir.
- Öğreticisiz Öğrenme (*Unsupervised learning*): Veri içerisinde gizli yapıları bulmayı amaçlayan bir dizi teknik olarak geçmektedir.
- Görselleştirme (*Visualization*): Büyük veri analizi sonuçlarının görselleştirilmesini amaçlayan tekniklerdir.

1.2.1.Duygu Analizi

Duygu analizi metinlerin, hizmet ve ürünler konusunda bilgi alma, kampanyaların başarısını ölçme, daha iyi kararlar verme amacıyla olumlu, olumsuz ya da nötr şeklinde kategorilere ayrıldığı bir analiz çeşididir. (Augustyniak ve diğerleri; 2016; Kabir ve diğerleri, 2018; Pak and Paroubek, 2010; Çoban ve diğerleri; 2014; Ziora, 2016). Sosyal medya platformları, farklı insanların farklı fikirleri ifade ettiği, muazzam miktarda ve her gün büyüyen mesajların bulunduğu, ünlülerden politikacılara pek çok farklı kişi tarafından kullanılan, pek çok ülkeden insanların kullandığı platformlardır (Pak ve Paroubek, 2014). Söz konusu platformların pazarlama, iş zekâsı ve sosyal psikologlar için zengin bir kaynak oluşturduğu belirtilmektedir (Çoban ve diğerleri, 2015). Literatüre bakıldığında duygu analizinin farklı seviyelerde ve farklı araçlarla yapılabildiği görülmektedir. Seviyeler döküman, cümle, öz ve yön bakımından olarak belirtilmektedir (Liu, 2012). İki araca rastlanmıştır. Augustyniak ve diğerleri ilk yöntemi geçmişte yapıldığı belirtilen yorumcuların bilgi ve görüşlerine dayanan insanla yapılan analiz olarak belirtmektedirler. Bir grup yorumcu tarafından metnin değerlendirildiğini ve kesinliği sağlamak için aynı metne kaç yorumcunun aynı değeri verdiğinin ölçüldüğünü söylemektedirler. Bu yöntemin pahalı, zaman maliyeti yüksek ve yorumcuların seçimi konusunda kimi zorluklar taşıdığını ifade etmektedirler. Ayrıca

Augustyniak ve diğerlerine göre internet ve sosyal medyanın gelişimiyle ölçek, sıklık ve hacim bakımından büyüklük, duygu analizinin otomatik olarak yapılma zorunluluğunu getirmiştir. Augustyniak ve diğerlerine göre bilgisayarla duygu analizinde temelde üç yöntem görülmüştür, bunlar kelime tabanlı yöntem (lexicon based), öğretici ile öğrenme (supervised learning) ve öğreticisiz öğrenmedir (unsupervised learning). Çoban ve diğerleri Türkçe için duygu analizi yönünde çalışmalar yapmışlardır. Genel olarak verinin toplanıp ön işlemeden geçirildiği, analiz edildiği ve görselleştirildiği görülmektedir (Kabir ve diğerleri, 2018). Çoban ve diğerleri ön işleme aşamasında kelime köklerini bulmuşlardır. Türkçenin sondan eklemeli bir dil olduğu, Çoban ve diğerlerinin ifade ettiği gibi duygu analizi konusunda Türkçe’ de henüz çok çalışma yapılmadığı göz önünde bulundurulduğunda, duygu analizinin Türkçe’ de kullanımına yönelik çalışmaların önemli olduğu düşünülmektedir. Duygu analizinde kullanılan yöntemler aşağıdaki gibidir.

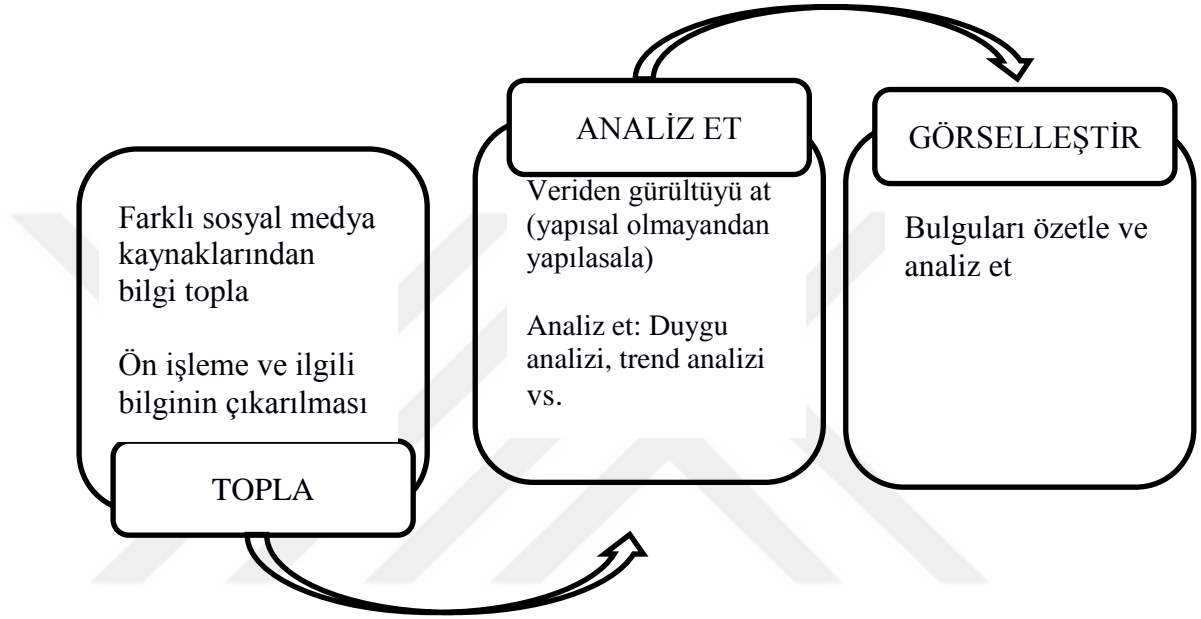
Şekil 1: Duygu Sınıflandırma Teknikleri



Kaynak: Can ve Alataş, 2016

Duygu analizinde Kabir ve diğerleri (2018) GitHub’ dan aldıkları kodlarla R üzerinde 700 *tweet* ile duygu analizi yapmışlardır. Pozitif ve negatif kelimeleri toplayarak kelime tabanlı bir analiz yaptıkları görülmektedir. Duygu analizine yönelik bir örnek araştırmanın akışı aşağıdaki gibidir.

Şekil 2: Sosyal medya madenciliği veya analizinin 3 basamağı



Kaynak: Kabir ve diğerleri, 2018

1.2.1.1 Veri Sözlüğü Tabanlı Yaklaşım ve Sorunları

Bing Liu (2012) duygu göstergelerinin duygu kelimeleri olduğunu belirtmektedir. İyi, harika, muhteşem gibi kelimelerin pozitif duygu kelimeleri olduğu; kötü, düşük kaliteli, berbat gibi kelimelerin negatif duygu kelimeleri olduğunu; bu kelimelerden oluşan listelere ise duygu sözlükleri denildiğini söylemektedir.

Bing Liu (2012) bir duygu kelimesinin farklı anlamları olabileceğini belirtmektedir. Örneğin “sucks” kelimesi cümle içerisinde berbat anlamında kullanılabileceği gibi, elektrik süpürgesinin emiş gücünü ifade etmek için de kullanılmış olabilir. Bing Liu’ ya (2012) göre bir diğer problem duygu kelimelerinin

“bu kameralardan hangisi iyidir” gibi aslında bir duygu içermeyen cümlelerin içinde geçmesidir. Aynı zamanda kinaye de veri sözlüğü tabanlı analizde bir problem olarak bahsedilmektedir. Ne güzel araba! İki günde bozuldu örneğinde olduğu gibi. Duygu kelimelerinin aslında fikirleri kastedebileceği veri sözlüğü tabanlı yaklaşımın sorunlarından biri olarak geçmektedir. Kabir ve diğerleri (2018) de çalışmalarında ironi ve kelimelerin birbiri ile olan ilişkisi göz önünde bulundurulduğunda sosyal medya analizinin zor olduğunu, ancak gelişen yeni analitik araçlar ile anlamlı yorumlar ve analizler sağlanabileceğini belirtmektedirler.

1.3. BÜYÜK VERİ HAKKINDA TARTIŞMALAR

Literatürde büyük veri üzerinde uzlaşma sağlanmış bir tanımın bulunmadığı, metot konusunda çalışmalar yapılması gerektiği belirtilmektedir (De Mauro ve diğerleri, 2016; George ve diğerleri, 2014; White and Breckenridge, 2014). George ve diğerlerine (2014) göre büyük verinin sağladığı çok büyük veri setleri, hem verinin toplanması hem de depolanıp analiz edilmesi anlamında, daha fazla beceri ve çaba gerektirmektedir. De Mauro ve diğerlerinin (2016) çalışmasına bakıldığında büyük verinin kavramsallaştırılması ile ilgili bir problem bulunduğu ve büyük veriden anlaşılan şeyin farklı yönlerinin farklı tanımlarda öne çıkarıldığı görülmektedir. George ve diğerlerinin (2016) büyük veri tanımlarında kullanılan özelliklerden biri olan hız (*velocity*) kavramının mevcut paradigma ile aslında yönetim ve organizasyon çalışmaları için önemli olmadığını belirtmektedir. Araştırmalarda alınan *tweet*lerin örneklem olarak nitelenerek veri tabanlarında saklanabildiği göz önünde bulundurulmalıdır. Bu durumda büyük veri tanımlarına göre, söz konusu çalışmaların büyük veri analizi olarak değerlendirilip değerlendirilemeyeceği sorusu akla gelmektedir. Bu nedenle literatüre bakıldığında De Mauro ve diğerlerinin (2016) büyük veri tanımları üzerine yaptıkları çalışmaların önemli olduğu düşünülmektedir.

George ve diğerleri (2016) istatistiksel analizin anlamlılığa odaklandığını, ancak büyük veri setleri ile çalışırken anlamlılığın daha az önemli olduğunu da ifade etmektedirler. Bir değişken üzerinde, çok küçük etkisi olan herhangi bir değişkenin veri setinin büyüklüğü nedeniyle anlamlı çıkabileceğini, ayrıca klasik yaklaşımın

örneklem üzerinde çalışmayı hedef aldığını ve büyük veride popülasyon verisine ulaşılabilirliğini belirtmektedirler. Söz konusu görüşleri Işıklı' nın (2014) örneklem almanın terki, popülasyon verisini analiz etmenin bir yolunun bulunduğu konusundaki görüşleri ile paralellik göstermektedir. Bununla beraber istatistiksel anlamlılığın öneminin azalması, Işıklı' nın (2014) belirtmiş olduğu araştırmalarda tümevarımın zayıflaması nedeniyle, büyük veri analizinde kullanılacak istatistiksel metotların seçiminin de, işleme-saklama alanında bulunan zorluklar gibi, bir takım zorlukları barındırabileceğini düşündürmektedir. Bu nedenle sınırlı vakayı içeren bir örneklem üzerinde kullanılan istatistiksel metotların, milyonlarca vaka üzerinde kullanılamamasının söz konusu olabileceği, hem yeni teknolojiler hem de analizdeki zorluklar göz önünde bulundurulduğunda araştırmalar için disiplinler arası ekiplerin kurulmasının gerekebileceği izlenimi oluşmaktadır. Literatürdeki çalışmalara bakıldığında örneğin Manyika' nın (2011) önerdiği yöntemlerde ve bu yöntemlere ilişkin örnek uygulama alanlarında büyük verinin kullanılabilirlik yönünün ön plana çıktığı, George ve diğerlerinin ise (2016) araştırmalarında kullandıkları “bilim insanları” teriminden de anlaşılacağı üzere büyük verinin akademide kullanımına yoğunlaştığı görülmektedir. Işıklı (2014) büyük veride kullanılabilirliğin öne çıktığını belirtirken, sosyal bilimlerin istatistiksel metotları kullanmaya başlamaları ile birlikte 20. yüzyılda yarı-bilim özelliği kazandıklarını ifade etmektedir. Hem Işıklı (2014) hem de George ve diğerleri (2016) mevcut paradigmadan söz etmekte, George ve diğerleri (2016) mevcut paradigma ile yönetim ve organizasyon araştırmalarında hızın önemli olmadığını belirtmekte, Işıklı ise (2014) büyük verinin başka bir paradigma ile açıklama getirmeye çalıştığını ifade etmektedir. Söz konusu ayrımın White ve Breckenridge' in (2014) sormakta oldukları sorularda da bulunduğu düşünülmektedir. White ve Breckenridge (2014) verinin toplanma şekli, analizi, araştırma değişkenleri, geçerliliği ve güvenilirliği konusunda eleştiriler getirirken, büyük veri analizi konusuna bakış açılarının geleneksel yöntemle yapılan çalışmalardaki yaklaşımımızla benzerlik gösterdiği düşünülmektedir. Büyük veri analizi konusunun çok yeni olması nedeniyle, bu eleştirilerinde haklı olabilecekleri göz önünde bulundurulmakla beraber, literatüre bakıldığında kimi veri setlerinin eldeki yaklaşım ve metotlarla zaten analiz edilemeyeceği kanaati oluşmuştur ve anlaşıldığı kadarıyla aslında büyük veri analizi eldeki yöntemlerle zaten analizi

mümkün olmayan verilerin analizine yoğunlaşmaktadır. Işıklı (2014) büyük veride yaklaşım farkından bahsetmekte ve getirilen açıklama biçimlerinin geleneksel yaklaşımdan kuram, örnekleme gibi konularda farklılaştığını belirtmektedir. Bu nedenle söz konusu veri setlerinin analizi konusunda metotlara güvenilip güvenilemeyeceği tartışılabilir olmakla birlikte, eldeki metotlarla analiz edilemeyen veri setlerinin analizi için yapılan çalışmaların değerli olduğu düşünülmektedir. George ve diğerleri (2016) yönetim alanında yapılan çalışmalarda büyük veri analizi için doktora seviyesinde öğretilen araştırma metot ve tekniklerinin yetersiz kaldığını da belirtilmektedir. Bu nedenle büyük veri analizinin akademik anlamda kullanımı için araştırma metot ve tekniklerine ilişkin derslerin kapsamının ileride genişlemesinin söz konusu olabileceği kanaati doğmuştur.

Büyük verinin George ve diğerlerinin (2016), ayrıca Işıklı' nın (2014) söylediği gibi popülasyona ilişkin veri sağlayabileceği belirtilmekte, ancak Visb ve diğerleri(2013) gibi kimi araştırmacılar analiz ettikleri veriyi örneklem verisi olarak tanımlamaktadır. Benzer şekilde yapılan araştırmalarda örnekleme yöntemi ve örneklemin popülasyonu temsil yeteneği yönünden herhangi bir bilgiye rastlanamamıştır. Işıklı (2014) bu konuda pozitivist tümevarımcı bir yaklaşım benimsenerek, evrenin bir prototipi sayılabilecek örneklem üzerinden genellemeler yapıldığını söylemekte, ancak büyük veriyle yapılan araştırmaların evrenin tümüne bakmanın bir yolu olduğunu belirtmektedir. Bu anlamda bir yöntem çerçevesinde bir sonuç elde etmenin mümkün olabileceği, ancak sonuçların Fujimoto ve diğerleri'nin (2017) çalışmasındaki gibi gerçeğe uygun olabileceği gibi, tutarsız olmasının da mümkün olabileceği kanaati oluşmuştur.

Büyük veriye ilişkin pek çok araştırmacı tarafından yöntem konusunda daha çok çalışma yapılması gerektiği belirtilmiştir (George ve diğerleri, 2014; White and Breckenridge, 2014). Bongsug (2015) tedarik zincirine ilişkin sosyal medya verisinin analizi için bilinen pratik bir yöntem olmadığını belirtmektedir. Bu nedenle Visb ve diğerlerinin (2013) büyük verinin sosyal bilimlerde uygulanması alanında yaptıkları çalışmanın önemli olduğu, benzer çalışmalara lojistik yöntemi için de ihtiyaç duyulduğu düşünülmektedir. Farklı disiplinlere ilişkin yöntemlerin geliştirilmesi ve uygulama alanında çalışma yapılmasının, araştırmacı ve araştırma sayısını arttırabileceği kanaati oluşmuştur.

Büyük veri analizinin yöntem olarak bazı üstünlükler taşıdığı ve kimi zayıf yönlerinin bulunduğu literatüre bakıldığında anlaşılmaktadır. Weiß ve diğerlerinin (2013) çalışması incelendiğinde tahminlerde yapılan sapmaların azaltılabileceği, müşteri bazında satışların azalması ve müşterinin yaşadığı gecikmeler arasında ilişki bulunabileceği görülmektedir. Büyük verinin daha çok değişken arasında muhtemel ilişkileri ortaya çıkarabileceği George ve diğerleri (2016) tarafından da belirtilmekte, ancak değişkenler üzerindeki kontrolün nispeten az olacağı söylenmektedir. Değişkenler üzerindeki kontrolün zayıflamasının ve detaylı bilgi alma konusunda George ve diğerlerinin (2016) literatürden derlemiş olduğu zorluk ve analiz için yeni teknolojiler ile daha fazla çaba ve beceri gerektirdiği göz önünde bulundurulduğunda, söz konusu noktaların büyük verinin zayıf yönü olabileceği kanaati uyanmıştır. Deniz taşımacılığı alanında Habin ve diğerleri (2017) geleneksel fonksiyonların hava koşullarının geminin yakıt tüketimi üzerindeki etkisini göz ardı ettiğini belirtmekte ve tahmin ile gerçekleşen tüketim arasında fark bulunduğunu söylemektedirler. Anlaşıldığı kadarıyla büyük veri, araştırmacılara hız optimizasyonunda tarihsel hava durumu verisini kullanma imkânı sağlamakta ve bu sayede hızın optimizasyonuna ilişkin bir başka değişkenin etkisinin ölçülmesine imkân tanımaktadır. Literatürden anlaşıldığı kadarıyla büyük veri analizi geleneksel yöntemlerle yapılan araştırmalara kıyasla daha çok değişkenin gözlenmesine olanak sağlamakta ve tahminleri iyileştirebilmektedir. Bu noktanın yöntemin güçlü yanı olabileceği düşünülmektedir.

Weiß ve diğerlerinin (2013) çalışmasına bakıldığında büyük veri analizinin lojistik alanında stok dışı kalmayı engelleme, rotaların optimizasyonu gibi alanlarda kullanılabileceği belirtilmektedir. DHL tarafından yayınlanmış olan rapor büyük veri analizinin lojistik sektöründe kullanımına ilişkin çalışmaların olduğunu göstermektedir. Bongsug (2015) tedarik zinciri araştırmaları için sosyal medyanın kullanımına ilişkin çalışma yapmıştır. Uzman sayısının ve çalışmaların artması ile birlikte büyük veri analizinin ülkemizde de yoğun olarak lojistik yönetimde kullanılması söz konusu olabilir. Kullanılan istatistiksel yöntemlerin de veri setine göre değişebileceğinin ifade edildiği de göz önünde bulundurulduğunda (George ve diğerleri, 2016) araştırmacıların daha farklı bir araştırma yöntemleri dersi içeriğine ihtiyaç duyabileceği düşünülmektedir.

Borgi ve diğeri (2017) sosyal medyanın lojistik firmalarına her bir müşteri için veri sağlayabileceğini belirtmektedir. Bongsug' un (2015) *tweet* analizi için önerdiği yöntem ve Weiß ve diğeri (2013) tarafından DHL için hazırlanan raporda, büyük verinin müşteri değerinin hesaplanmasında kullanımına ilişkin uygulamalar göz önünde bulundurulduğunda, sosyal medyada kullanıcı tarafından üretilen içeriğin lojistik firmaları tarafından kullanımının ileride yaygınlaşabileceği kanaati oluşmuştur.

Büyük veri analizinin Pakistan seçimlerinde seçim sonuçlarının tahmini için kullanımında seçim sonuçları ile tahminin örtüşmediği bu nedenle *Twitter'* in Pakistan nüfusunu temsil etme bakımından uygun olmadığı sonucuna varılmıştır (Mahmood ve diğeri, 2013). Diğer yandan Weiß ve diğeri (2013) katalog ürünlerinin geleneksel yöntemlerle tahmin edilmesi halinde tahminlerin %63 oranında tutmadığını, ancak geliştirilen yeni yöntem ile bu oranın %11'e indirildiğini belirtmekte; başka bir çalışmada Japonya nüfusunun konum bilgilerinin atılan *tweet*lerin konum bilgileri ile örtüştüğü belirtilmektedir (Fujimoto ve diğeri; 2017). Bu durum ülkelere, problemin konusuna göre sosyal medyanın bir araştırma için kullanışlı olabileceği veya olmayabileceğini düşündürmüştür. Sosyal medyanın bu nedenle geleneksel yöntemlerle yapılan çalışmaların problemlerine uygun kullanılabilir veri sağlayıp sağlamadığı konusunun göz önünde bulundurulmasının, kullanılabilir sonuçlara ulaşmak için önemli olabileceği düşünülmektedir.

Özet olarak büyük veri analizinin Işıklı' nın (2014) belirttiği üzere geleneksel yaklaşımdan farklı bir yaklaşımla olguları ele almakta, tanım, örneklem-popülasyon konusunda ayrışmalar bulunmaktadır. Ayrıca kapsamı ve yeni teknoloji-metotlara duyulan ihtiyaç nedeniyle, araştırmalar uygulama ve sonuçları yorumlama konusunda da kimi zorluklar içermektedir. Yine de söz konusu veri setlerinin başka yöntemlerle analizi mümkün olmayabileceği için bu konuda yapılan araştırmaların yeni bakış açıları sağlaması bakımından değerli olduğu düşünülmektedir. Bununla birlikte yöntemin araştırmalarda hali hazırda kabul edilen yaklaşım ve uygulanan bilimsel metottan farklılık göstermesi nedeniyle, sonuçların akademik araştırmalarda kabul edilebilmesi için daha çok yöntemsel çalışma yapılması gerekebileceği göz önünde bulundurulmalıdır.

1.4.BÜYÜK VERİ ÜZERİNE YAPILAN ÇALIŞMALAR

Baro ve diğerleri (2015) büyük veri üzerine yapılan çalışmalarını incelemişlerdir. Literatür üzerinde yaptıkları araştırma ile büyük veri tanımı üzerinde çalışmış, sağlık bilimlerinde yapılan çalışmaların 2003 yılından 2013 yılına 1' den 79' a yükseldiğini bulmuştur.

Gupta ve Hewett (2017) gündelik olarak üretilen *tweet*lerin hızı ve büyüklüğü nedeniyle *tweet* analizinin büyük verinin esas problemlerinden biri olduğunu belirtmektedir. Pek çok çalışmanın *tweet*leri *hashtag* ve sözlüklerle kategorize etmeye dayandığını belirterek, çalışmalarında *tweet*leri konularına göre gruplandırmayı amaçladıklarını ifade etmektedirler.

Yeruva ve diğerleri (2017) Amerika Birleşik Devletlerinde nüfusun 2/3' ünün fazla kilolu olduğunu belirterek *Twitter* üzerinden sağlıklı beslenme sosyal etki bağlamında bir büyük veri çalışması yaptıklarını belirtmektedir. Mesaj/resim sınıflaması ve duygu analizi ile araştırma yapılmıştır. Çalışmalarında yiyecek fotoğraflarını sağlıklı ya da sağlıksız olarak gruplamış ve yiyeceklere ilişkin ifadeleri olumlu ya da olumsuz olarak sınıflandırmışlardır. Büyük veri analizi tekniklerinin sağlıklı/sağlıksız yiyecekler anlamında atılan *tweet*leri etkin olarak sınıflayıp, duygu analizi yapılabileceği sonucuna varmışlardır. Çalışmalarının hastalık kontrol ve önleme merkezleri için faydalı olabileceğini düşünmektedirler.

Visb ve diğerleri (2013) *Twitter* üzerinden ayaklanmalara ilişkin bir araştırma yapmışlardır. Sosyal bilimlerde büyük verinin kullanımına ilişkin metodolojik bir çalışma yapmayı amaçlamışlardır. Çalışmalarında içerik analizi ve bilgisayarların kullanımı ile büyük veri setleri üzerinde çalışmayı amaçlamışlardır. 700.000 kişiden gelen 2,6 milyon *tweet*i analiz etmelerine rağmen, ellerindeki veriyi örneklem olarak tanımlamaktadırlar. *Tweet*leri bir ilişkisel veri tabanına aktarmışlar ve analiz için NLP kullanmışlardır. İnsanların *Twitter*' i hangi amaçla kullandıklarını incelemiş ve *tweet* atanların profillerini belirlemişlerdir. Çalışmalarının gelecekte toplamayı düşündükleri veri göz önünde bulundurulduğunda mütevazı kaldığını, NLP konusunda daha çok çalışma yapma konusunda cesaretlendiklerini ve sosyal bilim araştırmacılarının söz konusu bilgisayar teknolojileri ve büyük veri teknikleri konusunda eğitilmesinin gerektiğini öne sürmektedirler. Çalışmalarında polis, acil

durum hizmetleri ve devlet kurumlarının kriz anında sosyal medyayı kullanmada çeşitli güçlüklerle karşılaştıklarını bulmuşlardır. Visb ve diğerlerinin yaptığı araştırmanın büyük veri tanımları konusundaki farklılıklar açısından örnek teşkil ettiği düşünülmektedir. Pek çok çalışmada görülmediği halde örnekleme yöntemlerini metodolojilerinde anlatmışlardır. Büyük verinin popülasyona ilişkin bilgi verebileceği kimi araştırmacılar tarafından söylene de, çalışmada elde edilen verinin örnekleme ait olduğu belirtilmektedir.

Rathore ve diğerleri (2017) sosyal medyanın insanlara, kendilerine ilişkin bilgileri, konum bilgisi ile birlikte paylaşma imkânı tanıdığını belirtmektedir. Geo sosyal veri sağlayan hizmetlerin hızla arttığından bahsedilmekte ve *Twitter* bu hizmetlerden biri olarak belirtilmektedir. Sosyal medya aracılığıyla davranışların, yaşam tarzının, olayların, doğal afetlerin, salgınların ve pek çok şeyin sosyal medya aracılığıyla paylaşıldığı belirtilmektedir. Milyonlarca ağ kullanıcısının terabaytlarca, heterojen ve anlık veri üretmekte olduğunu belirtmektedirler. Bu bilgiye büyük veri denildiğini söylemektedirler. Rathore ve diğerleri (2017) çalışmalarında *Twitter'* in bir büyük veri kaynağı olduğunu, sosyal medyanın konum bazında bilgi sağlayabildiğini, büyük veriden karar verme ve planlama amacıyla yararlanılabildiğini belirtmektedirler. Konum bazında verilen bilgilerin kargo şirketleri ve kargo kullanıcılarının planlama faaliyetlerinde kullanılabileceği düşünülmektedir.

Mahmood ve diğerleri (2013) *Twitter* üzerinden Pakistan 2013 seçim sonuçlarını tahmin etmeye çalışmışlardır. Seçimi kazanmasını tahmin ettikleri aday ülkede başarı göstermesine rağmen, seçim tahmin edilen adayın dışında bir aday tarafından kazanılmıştır. Bu nedenle *Twitter'* in seçim sonuçları ile ilişkili olabileceğini düşünmekle birlikte, tüm nüfusu temsil edemeyeceği sonucuna varmışlardır.

Sotiriadis ve Zyl (2013) hangi kullanıcıların *Twitter* üzerinden paylaşılan bilgilere göre turizm hizmetlerini satın aldığını ve kararlarının hangi faktörlerden etkilendiğini çalışmışlardır. *Twitter'* in seçilme nedenini gerçek zamanlı olması, içeriğin sürekli genişlemesi ve *Twitter* üzerinden deneyimlerini paylaşan kullanıcıların turistlere yardımcı olabileceği düşüncesi olarak belirtmektedirler. Araştırma hipotezlerini güvenilirlik, ilişki, iletişim sıklığı ve kaynakların bilgi ve

uzmanlığı üzerine kurmuşlardır. Araştırmayı internet üzerinden oluşturdukları bir anket aracılığıyla yürütmüşlerdir ve söz konusu anket 5 ay içerisinde 500 kullanıcı tarafından cevaplanmıştır. Cronbach's metodu ile anketin güvenilirliğini ölçmüşlerdir. Araştırmalarında Cronbach's Alfa 0,74 ve 0,90 arasında değişmektedir, bu nedenle alınan sonuçların lineer regresyon için uygun olduğunu belirterek, regresyon analizi uygulamışlardır. Güvenilirlik, ilişki ve uzmanlığın karar vermede etkili olduğu sonucuna ulaşmışlardır. Sotiriadis ve Zyl'in (2013) araştırmaları büyük veri araştırması olmamasına rağmen, tüketicilerin karar verme süreçlerinde sosyal medya üzerinden yapılan paylaşımların etkinliğini ortaya koyması bakımından önemli olduğu düşünülmektedir. Sotiriadis ve Zyl (2013) araştırmalarında sosyal medya üzerinde kullanıcı tarafından üretilen içeriğin turistlere yardımcı olabileceğini belirtmektedir. Benzer şekilde kargo firması seçiminde, kullanıcılar tarafından verilen geri bildirimlerin kargo hizmeti satın alanlara yardımcı olup olamayacağı başka araştırmaların konusudur.

Gandomi ve Haider (2015) çalışmalarında büyük verinin tanımını yapmışlardır. Metin, ses, video, sosyal medya verisi için analitik teknikleri incelemişlerdir. Büyük veri için yeni istatistiksel teknikler ortaya koymuş ve büyük veride gelecekte olması beklenen gelişmelerin altını çizmişlerdir.

İKİNCİ BÖLÜM

LOJİSTİK VE KARGO FİRMALARI ÜZERİNE ARAŞTIRMALAR

2.1. LOJİSTİK YÖNETİMİNDE BÜYÜK VERİ ÜZERİNE YAPILAN ÇALIŞMALAR

Waller ve Fawket (2013) büyük verinin tedarik zincirine uygulanması ile ilgili yaptıkları çalışmada; büyük veri analizinin taşıyanlar için teslimat zamanı tahmininde, hava durumunun etkileri, sürücü karakteristikleri tahmininde; gerçek zamanlı kapasite yönetiminde; optimal rotaların belirlenmesi, hava ve trafik durumunun, sürücü karakteristiklerinin hesaba katılmasında ve insan kaynaklarında sürücü devir oranını, sürücü tahsisini düşürmede, ayrıca duygu analizinde kullanılabileceğini belirtmiştir. Sonuç olarak büyük verinin, kullanımı son zamanlarda yaygınlaşmış yeni bir kavram olmasına rağmen, tedarik zincirinde önemli uygulamaları olduğunu ve araştırma ve eğitim için fırsatlar sunduğunu belirtmektedirler.

Tan ve diğerleri (2015) günümüzde firmaların *tweet*, video, tıklama gibi yapılandırılmamış veri kaynaklarının müşteri, pazar ve ürünlerini anlamak veya yeni fikirler geliştirmek için kullanılabileceğini belirtmekte; ancak veri analizi tekniklerinde eksiklik bulunduğundan bahsetmektedir. Bu nedenle tedarik zincirinde yeniliklerin geliştirilmesi amacıyla, veri analizine yönelik bir çalışma yapmışlardır. Büyük verinin yeni bir kavram olmamasına rağmen tedarik zincirine uygulanmasının yeni bir uygulama olduğunu belirtmektedirler. Çalışmalarının firmalara yeni ürün geliştirme fikirleri, alt yükleniciler ya da departmanların yeni ürün fikirleri geliştirmede maliyet etkin bir biçimde nasıl bir arada çalışabileceği konusunda imkân sağladığını belirtmektedirler.

Mirović ve diğerleri (2018) yaptıkları çalışmada denizcilik sektörünün gemi ve kıyı arasındaki iletişimin sınırlı olması nedeniyle geleneksel olarak bilgiye duyarlı olmadığını söylemektedirler. Yine de navigasyon sistemlerindeki gelişmelerle, sensör ve takip sistemlerinden üretilen büyük miktarda veriden kullanışlı bilgiler elde edilebileceğini, veri miktarı, çeşitliliği ve hızının problemi bir büyük veri problemi haline getireceğini belirtmektedirler. Çalışma büyük verinin denizcilik sektörüne

uygulanması konusunda genel bir bakış sağlamayı amaçlamaktadır. Çalışmalarında armatörlerin çok büyük bir sistem içerisinde ve belirsizlik durumunda faaliyetlerine devam ettiğini belirterek, büyük veri analizi tekniklerinin *network design*, enerji verimliliği, maliyetlerin düşürülmesi, denizde güvenliğin iyileştirilmesi ve çevresel faydalar anlamında denizcilik sektöründe faydalı olabileceğini söylemektedirler.

Zhong ve diğerleri (2016) nesnelerin internetinin üretimde kullanılmaya başlaması ile RFID teknolojilerinin çok büyük miktarda veri üretebileceğini belirtmektedir. RFID' den günlük olarak toplanabilecek veri miktarının terabaytları bulabileceği ifade edilmektedir. Bu verinin yenilikçi bir yaklaşım ile analiz edilmesinin üretim ve karar verme konusunda faydalı olabileceği belirtilmiştir. Çalışmada örnek olay incelemesi yapılmış ve söz konusu yenilikçi yaklaşım kullanılmıştır. RFID-Cuboids' in uygulanmasından sonra lojistik güzergâhın (*logistics trajectory*) görselleştirilebileceği ve kullanışlı bilgilerin, lojistik trendlerin karar verme süreçlerini destekleyebileceğini bulmuşlardır. Yönetimsel anlamda zamanı kesin olarak ölçebilmenin üretim planlamaya katkı sağlayabileceği, operatörlerin performanslarının ölçümünün iş bölümünün sağlanması için faydalı olabileceği önerilmektedir.

Bodell (2014) büyük veriyi standart veritabanı yazılımları ile analiz edilmesi mümkün olmayan çok büyük miktarda ve sıklıkla yapılandırılmamış biçimde olan veri olarak tanımlamaktadır. Enerji sektörünün yıllardır büyük veri ile çalışmakta olduğunu belirtmektedirler. Büyük verinin müşteri deneyimini geliştirmede, tedarik zincirinin iyileştirilmesinde, risk ve belirsizliklerin anlaşılmasında kullanılabileceğini belirtmektedir. Amazon' un müşterilerine ilişkin ürün önerilerini kişiselleştirmeye 20 yıl önce başladığı, bu durumun başka firmalar tarafından da izlendiği belirtilmektedir. Walmart gibi diğer perakendecilerin büyük veriyi talep tahmininde, tedarik zinciri ve envanter seviyelerini iyileştirme, stok dışı kalmaktan kaçınma amacıyla kullandıkları belirtilmektedir.

Kaur ve Prakash (2017) çalışmalarında büyük verinin tedarik zinciri ve operasyon yönetimine entegrasyonuna ilişkin literatürü incelemişlerdir. Çalışmaları üç alana odaklanmaktadır bunlar tedarik zinciri ve operasyon yönetimi, üretim, satın alma ve lojistik olarak belirtilmektedir. Geçmiş çalışmaların gözden geçirilmesine ek

olarak büyük verinin tedarik zincirine uygulanmasına ilişkin önerilerde bulunmaktadır.

Habin ve diğerleri (2017) gemi hızlarının optimizasyonu için hava durumu arşivlerini kullanarak büyük veri analizi ile bir karar destek sistemi üzerinde çalışmışlardır. Literatürde hız optimizasyonu problemi üzerine yapılan çalışmaların bir teorik yakıt tüketimi fonksiyonuna dayandığı ifade edilmektedir. Hava koşullarındaki değişkenlik nedeniyle söz konusu fonksiyonun hava koşullarından etkilenen gemilerin gerçek yakıt tüketimini yansıtmadığı belirtilmektedir. Bir Türk armatörüne ait tarihsel yakıt tüketimi verisi söz konusu teorik yakıt tüketimi fonksiyonu ile karşılaştırılmıştır. Geminin yakıt tüketiminin temel olarak gemi hızına bağlı olduğu, ancak hava koşullarının yakıt tüketimini etkileyebildiği belirtilerek, karşılaştırma sonucunda tahmin edilen yakıt tüketimi ile gerçek yakıt tüketimi arasındaki farkın seyirde geçen süre arttıkça arttığı bulunmuştur. Çalışmaları bu nedenle hava koşulları göz önünde bulundurularak gemi hızının optimize edilmesine dayanmaktadır ve bu probleme ilişkin bir karar destek sisteminin geliştirilmesini amaçlamaktadır. George ve diğerleri (2016) deneylerde değişken sayısının sınırlı olduğunu belirtmekte ve büyük veri analizi tekniklerinin çok daha fazla değişken arasında potansiyel ilişkilerin bulunmasına yardımcı olabileceğini söylemektedirler. Söz konusu çalışmada bahsedilen teorik fonksiyonun yakıt tüketimi ve gemi hızı ilişkisini temel aldığı ancak hava durumu gibi diğer bazı faktörlerin söz konusu değişken üzerindeki etkisinin ölçülebilmesi ile probleme ilişkin çözümün iyileştirilebileceği yönündeki önerinin George ve diğerleri'nin (2016) büyük veri analizinin açabileceği yeni ufuklar anlamında örnek oluşturabileceği düşünülmektedir.

Bongsug (2015) işletmeler ve araştırmacıların sosyal medyayı daha çok önemsemekte olduğunu belirtmekte; ancak tedarik zinciri yönetimi konusunda uygulama ve araştırmaların görece az olduğunu söylemektedir. Araştırmasında 22399 supplychain *hashtag*lı *tweet* üzerinde çalışarak; tedarik zinciri üzerine atılan *tweet*lerin çeşitli tedarik zinciri profesyonelleri ve kurumlar tarafından bilgi paylaşımı, işe alma, hissedarlarla iletişim amacıyla atıldığını; tartışılan konuların lojistik, sosyal sorumluluk, risk, üretim, tedarik zinciri yönetimi, bilişim teknolojileri ve hatta insan haklarını kapsadığını; bazı *tweet*lerin teslimat, satış, çevre standartları,

risk ve tedarik zincirindeki bozulmalar konusunda güçlü duygular içerdiğini bulmuştur.

Zhong ve diğerleri (2015) RFID' nin lojistik yönetiminde büyük oranda kullanıldığından bahsetmektedir. RFID üzerinden büyük miktarda veri toplanabileceğini belirtmektedirler. RFID üzerinden toplanan verinin lojistik yönetiminde kullanımı üzerine çalışma yapmışlardır.

Karlskind (2014) firmaların mobil uygulamaları büyük veri yönetimi için nasıl kullandığını anlatmıştır. Karlskind'e (2014) göre bu uygulamalar envanter yönetiminde, iş yükünün optimize edilmesinde, rotaların analiz ve uyarlanmasında, teslim kanıtı üretilmesinde, sürücü güvenliği amacıyla kullanıldığını belirtmektedir.

Chang ve diğerlerinin (2017) yaptığı araştırmada büyük veriye ilişkin tahminsel yöntemlerin kullanılması ile askeri lojistikte ilaç ve tıbbi malzemelere ilişkin karar destek sistemleri geliştirilmesi konusunda çalışılmıştır.

Borgi ve diğerleri (2017) gelişen imkânlarla birlikte, dijitalleşme ve teknolojiye ayak uydurmanın taşımacılık için bir zorunluluk haline geldiğini belirttikleri çalışmalarında, büyük verinin taşımacılık ve lojistik için kullanımına ilişkin bir inceleme yapmışlardır. Taşımacılıkta teknolojinin daha yoğun kullanımının çok büyük miktarda, çeşitli ve anlık veri sağlayabileceği, söz konusu verinin operasyonların iyileştirilmesi ve iş geliştirme için faydalı olabileceği belirtilmektedir.

Wang ve diğerleri (2016) internet üzerinde üretilen veri büyüklüğünün artmasından ve bu veriden yararlanmak isteyen kurumlar için zorluklar bulunmasından hareketle literatür üzerinde bir çalışma yapmış ve iş zekâsı anlamında lojistik yönetiminde kullanılabilecek büyük veri analizi tekniklerini araştırmışlardır. Büyük veri analizinin müşteri satın alma süreçleri, bakım periyotları, maliyetlerin düşürülmesi ve daha iyi hedeflenmiş kararlar verilmesi konusunda faydalı olabileceği düşünülmektedir.

Mehmood ve Graham (2013) sağlık hizmetlerine ulaşım konusunu Markov modeli ile ele almışlardır. Büyük veri lojistiğinin şehir içi taşımacılığın ve dağıtım sistemlerinin GPS, cep telefonu, şirketlerin işlem verilerinin insan tarafından üretilen aktivitelerle kombine edilmesi olarak tanımlandığı söylenmektedir.

Borgi ve diğeri (2017) yeni bilgi teknolojilerinin uygulanmasının lojistik ve taşımacılık firmaları için bir zorunluluk haline geldiğini belirttikleri araştırmalarında, büyük verinin lojistik alanında kullanımına ilişkin çalışmaları ve yapılabilecek araştırmaları konu almışlardır. Verinin çok büyük miktarlarda üretilmeye başlamasının fırsatlarla birlikte zorluklar da getirdiğini belirtmektedirler. Dünya çapında her gün milyonlarca sevkiyat yapıldığını; söz konusu sevkiyatlardan konum, içerik, büyüklük, ağırlık, yükleme ve tahliye noktaları konusunda pek çok veri sağlandığını, söz konusu verilerin büyük veri setlerini oluşturduğunu ve bu setler üzerinde yapılacak analizlerin operasyonel verimliliği artırma, müşteri deneyimini iyileştirme ve yeni iş modellerinin oluşturulmasına imkân sağladığını belirtmektedirler. Taşıma ve lojistiğin önemli bir veri kaynağı olduğu ve büyük veri teknolojilerinin bu alanda önemli işletme değerleri içerdiği belirtilmektedir. Tedarik zinciri yönetimi, kargo, e-ticaret, ters lojistik, nakliye, deniz, hava ve kara lojistiği yeni bilgi teknolojilerinin uygulanabileceği alanlar olarak belirtilmektedir. Büyük verinin lojistiğe muhtemel katkıları üç başlık altında incelenmektedir. Bunlar operasyonel verimlilik, müşteri deneyimi ve yeni iş modelleridir. Büyük veri analizinin operasyonları etkinleştirilmesi konusunda daha iyi karar verme sağlaması, süreçlerin kalitesini iyileştirmesi ve kaynak kullanımının optimize etmesinden bahsetmektedirler. Aynı zamanda uzun vadeli alt yapı yatırımlarına ilişkin tahminleri iyileştirebileceği ve riski düşürebileceğinden bahsedilmektedir. Sevkiyatların takibi ile rotaların iyileştirilebileceği ve trafik koşulları konusunda bilgi toplanabileceği ifade edilmektedir. Müşteri deneyimine ilişkin başlıkta, müşteri sadakatinin sağlanması için internetin kendisi bir kaynak olarak gösterilmekte ve sosyal medyanın lojistik firmalarına her bir müşteri konusunda bilgi verebileceği belirtilmektedir. Büyük veri analizi tekniklerinin kullanılması ile lojistik yöneticilerine proaktif önlemler almaları için imkânlar sağlanabileceği ifade edilmektedir. Yeni iş modellerinin kurulması konusunda araçların kamera ve sensörlerle donatılması halinde emlak, şehir planlama, çevre planlama konusunda; trafik yoğunluğu, gürültü, kirlilik, sıcaklık, nem gibi değerli olabilecek bilgilerin sağlanması örnek olarak verilmektedir. Büyük verinin tahminleme ile lojistik servis sağlayıcılarının hizmet servis kalitelerini ve süreç verimliliklerini yükseltebileceği, daha kısa teslimat zamanları sağlayabileceği belirtilmektedir.

Kaur ve Singh (2017) büyük veri ile tedarik zinciri için çevresel olarak sürdürülebilir satın alma ve lojistik modeli kurmayı amaçlamaktadır. Modelleri gerçek zamanlı veriye dayanarak optimal sürdürülebilir satın alma ve nakliye kararları sunmaktadır. Büyük veriden çıkarılacak kullanışlı bilgilerin tedarik zinciri operasyonları için önemli olduğunu belirtmektedirler. ERP' nin kullanılmaya başlamasıyla büyük miktarda verinin üretilmekte olduğu ifade edilmektedir. Büyük verinin tedarik zinciri yönetiminde başlıca satın alma, envanter yönetimi, lojistik ve dağıtım planlama konularında kullanılabileceği belirtilmektedir.

Weiß ve diğerleri (2013) büyük verinin lojistikte kullanımına ilişkin yaptıkları çalışmada, büyük verinin operasyonel verimlilik, müşteri deneyimi ve yeni iş modelleri konusunda kullanılabileceğini belirtmişlerdir. Anlık gelişmelere göre rotaların güncellenebileceği ve yolda olan araçların şoförlerinin anlık olarak bir sonraki en iyi teslimat noktasını alabileceğinden bahsedilmektedir. Aynı zamanda alıcılar tarafından sağlanan uygun olma ve konum bilgileri göz önünde bulundurularak gerçekleştirilecek teslimat girişimlerinden kaçınılabileceği belirtilmektedir.

Wei ve Lin (2018) bir örümcek yardımıyla topladıkları veriyi hibrit bir analiz metoduyla değerlendirmişlerdir. Çalışmada veri bir örümcek yardımı ile toplanmış, NLP ve *topic mining* için SAS kullanılmıştır. Web sitelerinin botlarını engellemesi nedeni ile araştırma görevlileri tarafından verinin elle toplandığı ifade edilmektedir. Çalışmalarının e-ticaret lojistiği ile ilgilenen firmaların strateji geliştirmesinde faydalı olabileceğini düşünmektedirler.

Büyük veri analizinin kargo firmaları tarafından da talep tahmini konusunda kullanılabileceği düşünülmektedir. Perakendecilerin talep tahmini ve envanter seviyelerinin iyileştirilmesi konusundaki uygulamalarının, konteynır pozisyonlama ve talep tahmini konusunda denizcilikte kullanılıp kullanılmayacağına başka araştırmalara konu oluşturabileceği düşünülmektedir. Crainic ve diğerleri (1993) boş konteynır pozisyonlama konusunda 1992 yılında yaptıkları bir çalışmada Avrupa'nın başlıca hat işletmeciliği yapan firmalarından birinin 1986 yılında 300,000 konteynırının karada taşındığını, maliyetinin \$50.000.000 olduğunu ve söz konusu konteynırların %40' nın boş olduğunu belirtmektedir. Söz konusu hattın Fransa, Almanya, Belçika ve Hollanda' da bahsi geçen yıllarda yirmi liman arasında hizmet

verdiği belirtilmektedir. Shen ve Khoong (1995) ise gemilerde boş konteynırların taşınması için atanabilecek slotların sınırlı olduğunu ifade etmektedirler. Habin ve diğerleri (2017) tahmini hava koşulları, demirde, limanda geçirilen sürelerin farklılığı nedeniyle bazı belirsizlik bulunduğunu çalışmalarında ifade etmektedir. Büyük veri analizinin kullanımıyla farklı limanlarda konteynır hareketlerinin analiz edilerek, belirsizliklerin azaltılması, ekipman seviyelerinin iyileştirilmesi ve gemilerde boş konteynırlar tarafından işgal edilen yeri düşürülmesinin, büyük veri analizinin George ve diğerleri (2014) tarafından ifade edilen potansiyeli ile söz konusu olabileceği düşünülmektedir.

Büyük veri analizinin rota optimizasyonu konusunda tarihsel trafik yoğunluğu verisini kullanarak (Waller ve Fawket, 2014) dağıtım için kullanılacak rotaları günün saatlerini ve trafik yoğunluğunu göz önünde bulundurarak araçların etkinliğini artırması söz konusu olabilir. Weiß ve diğerleri de (2013) büyük veri analizinin rotaların optimizasyonu ve başarısız teslimatların önüne geçmekte kullanılabileceğinden bahsetmektedirler.

2.2. KARGO FİRMALARI ÜZERİNE YAPILAN ÇALIŞMALAR

Deniz ve Gökmerdan (2011) Erzurum ilinde, kolayda örnekleme ile 400 kişilik bir örneklem üzerinde anket çalışması uygulayarak, kargo firması kullanıcılarının kargo firmalarının sunduğu hizmete yönelik tutum ve davranışları ile karşılaştıkları problemleri araştırmışlardır. Araştırmada personelin sorunları zamanında çözümlendiğini, müşteriyi iyi bir biçimde karşıladığını, personelin yeterli deneyime sahip olduğunu ayrıca firmanın malları hasarsız olarak teslim ettiğini düşündükleri belirlenmiştir. Araştırmaya göre temel problem gecikmedir. Bu problemi kargonun iade edilmesi ve ekstra ödemeler takip etmektedir. En az karşılaşılan problem ise kargonun başkasına teslimi takip etmektedir. Deniz ve Gökmerdan (2011) lojistiğin 15 fonksiyon ve faaliyetini belirtmiştir. Bunlar müşteri hizmetleri, ulaştırma ve trafik yönetimi, envanter yönetimi, sipariş yönetimi, depolama, talep tahmini, dökümantasyon, tedarik, koruyucu ambalajlama, malzeme aktarılması, bilgi yönetimi, yedek parça ve servis desteği, fabrika ve depo yeri seçimi, geri dönen malların yönetimi, hurda döküntü ve artıkların elden çıkarılmasıdır.

Çakmak ve Özkan (2017) kargo kullanıcılarının önem verdiği hizmetleri ve kargo firmalarının bu hizmetleri ne ölçüde yerine getirebildiğini belirlemeye yönelik olarak bir araştırma yürütmüştür. Kullanıcıların en çok önem verdiği unsurlar gönderinin eksiksiz teslimi, kurcalanmaması, belirtilen adrese teslimidir. En az önem verilen unsurlar müşteriden taşıma için yardım istenmesi, firmanın fiziksel koşulları ve kampanyalardır. 302 kişilik bir örneklem üzerinde anket uygulanmıştır.

Büyükkeklik ve diğerleri (2014) kargo firmalarının hizmet kalitesinin müşteri tatmini ve tekrar satın alma üzerine etkisini incelemiştir. Yüz yüze anket yoluyla uygulama yapılmıştır. Hizmet kalitesinin görünüm, güvenilirlik, yanıt verebilirlik, güvence ve empati şeklinde beş boyutunun bulunduğu belirtilmiştir. Hizmet kalitesinin müşteri tatmini ve tekrar satın alma üzerinde doğrudan etkisi olduğu sonucuna ulaşılmıştır.

Ghat ve Pal (2016) hizmet kusurları üzerinde çalışma yürütmüştür. Kargo hizmeti pazarının 3 segmenti bulunduğu bunların işletmeler arası, işletmeden tüketiciye, tüketiciden tüketiciye olduğu belirtilmiştir.

Hafız ve Alam (2016) e-ticaret firmalarından alışveriş yapan müşteriler servis kalitesi ve kargo hizmeti arasında korelasyon bulmayı, SERVQUAL kullanarak, amaçlamıştır. Çalışmada empati, güvenilirlik ve taahhüt boyutları ile kargo hizmeti arasında pozitif etki bulunmuştur.

Kortick ve O' Brien (2008) Servis kalitesinin yükseltilmesine yönelik örnek olay incelemesi yapmıştır. Beyzbola dayanan bir oyun geliştirilerek, personelin performansını, firmanın performans ölçüm kriterlerine göre, yükseltmeyi amaçlamaktadırlar. Bunun için ekip, takımlara bölünerek beyzbol ligine benzer bir yapıda oyunlar oynamaktadır. Firmanın servis kalitesini değerlendirmek için üç kriteri bulunmaktadır bunlar yanlış arabaya yüklenen paket sayısının doğru yüklenen paket sayısına oranı, yükleme kalitesinin (istiflemenin 1-10 arasında değerlendirildiği) oranı, hasarları minimize etmek için yükün kayışlarla tespit edilme oranlarıdır. Bir maçı kazanmak için bu üç kriterden ikisinde rakipten iyi olmak gerekmektedir. Takımlar arasına 12 haftalık bir program yapılmıştır. Beyzbol kalite kontrol sisteminin bu oyunlardan sonra bütün kuzeydoğu bölgesinde uygulandığı ifade edilmektedir. Beyzbol analog formatının sevkiyat performansını arttırabileceği ifade edilmektedir.

Lin ve diğeri (2017) aynı gün teslimat (*Same day delivery*;SDD) paradigmasını değerlendirerek topla ve dağıt, ticari filo ile SDD, kitle kaynaklı SDD paradigmaları ile karşılaştırmıştır. Topla ve dağıtın en uygun maliyetli yaklaşım olduğu ve aynı zamanda aynı gün teslim için uygun olduğu bulunmuştur.

Stout (2013) muhasebe eğitimi üzerine yazdığı makalesinde öğrencilerin bir işletme danışmanının rolü; maliyet, hacim, kar analizi ve kazanç planlamasını Excel ile geliştirmeyi, müşteri tarafından değerlendirilen bir kargo firması üzerinden anlatmıştır.

Chonhenchob ve diğeri (2012) dağıtımda kullanılan küçük ve orta büyüklükteki araçların titreşim seviyelerini ölçmüş ve büyük araçlarla karşılaştırmıştır. Hem küçük kamyonlar hem de minibüsler için dikey eksenli titreşim yanal ve boylamsal eksenli titreşimden anlamlı olarak yüksek bulunmuştur. Küçük kamyonlarda dikey titreşimin, minibüslere oranla, anlamlı olarak yüksek olduğu sonucuna varılmıştır. Çalışmalarının küçük ve orta büyüklükteki kamyonlar ve minibüsler için testler geliştirilmesinde kullanılabileceğini ifade etmektedirler.

Pacciarelli ve diğeri (2011) İtalya' daki bir kargo firmasının deposunun değişim mühendisliği ve tasarım optimizasyonu ile ilgili RFID uygulaması yönünde çalışma yapmıştır. RFID' nin uygulama maliyetinin ölçülmesine ilişkin yeni bir model önerilmiştir. Maliyet minimizasyonu için yeni bir model geliştirilmiştir. RFID sistemlerinin daha kesin, sık ve hızlı bilgi sağladığı ifade edilmektedir.

ÜÇÜNCÜ BÖLÜM

TÜRK KARGO HİZMET SAĞLAYICILAR İLE İLGİLİ TWITTER' DA ÜRETİLEN İÇERİĞİN ANALİZİ

3.1. ARAŞTIRMANIN EVRENİ VE ÖRNEKLEMİ

Araştırma evrenini *Twitter* kullanıcılarının kargo firmalarına attıkları *tweet*ler oluşturmaktadır. @yurtiçikargo, @mngkargo, @araskargo olarak etiketlenen *tweet*ler taranmıştır. Bu konudaki *tweet*ler '*Twitter Search API*' ile alınmıştır. *Twitter Search API* 7 gün içerisinde *public* olarak atılan *tweet*lerin bir bölümünü vermektedir. 2107 *tweet* üzerinden kullanıcının oluşturduğu içerik değerlendirilmiştir.

3.2. ARAŞTIRMA PROBLEMİNİN BELİRLENMESİ

Türkiye' de yapılan çalışmalara bakıldığında kargo firmalarına ilişkin müşteri memnuniyeti ve önem başarımları analizi konusunda araştırmalar yapıldığı görülmektedir. Literatürde büyük verinin araştırmanın bütün değişkenlerini içerip içermediği yönünde sorulara rastlanmıştır. Literatürdeki araştırmalar doğrultusunda, seçilen örneklemin daha önce yapılmış bir araştırmanın değişkenlerine ilişkin ifadeleri bulundurup bulundurmadığı bu araştırmanın sorularından ilkinin oluşturmaktadır. Büyük veriden ticari ve akademik amaçla nasıl yararlanılabileceği konusunda da araştırmalar yapılmıştır. Araştırmanın ikinci problemi ise kullanıcı tarafından üretilen içerik üzerinden yapılan çalışma ile büyük veri analizinin lojistik yönetiminde kullanımına ilişkin öneriler geliştirmektir.

3.3. ARAŞTIRMANIN AMACI

Twitter da kullanıcı tarafından üretilen içeriğin, seçilen bir araştırmanın, bütün değişkenlerine ilişkin ifade içerip içermediğinin test edilmesi; alıcı, gönderici ve firmaların *Twitter'* ı kargo hizmeti bağlamında kullanım amaçlarının ve paylaşımlarının konularının belirlenmesi, yapılan paylaşımlar üzerinden elde edilen

bilgilerin firmaların faaliyetlerinin iyileştirilmesi ve maliyetlerinin düşürülmesi konusunda nasıl kullanılabileceğine ilişkin önerilerin geliştirilmesidir.

3.4. ARAŞTIRMANIN KISITLARI

Büyük verinin uzlaşılabilir bir tanımı olmamakla birlikte; pek çok tanımda kişisel bilgisayarlar tarafından depolanamayacak büyüklükte, sıklıkla kullanılan istatistik paket programları, hesap tabloları gibi geleneksel yöntemlerle işlenemeyecek, karmaşık yapısı ve çeşitliliği nedeni ile işlenmesi zor olan çok büyük veri setleri olduğu ifade edilmektedir. Bu nedenle araştırmanın kısıtlarından birini büyük veri analizi için gereken teknolojilerin kullanımının mümkün olmaması ve büyük miktarda verinin nasıl işleneceğinin bu tezin konusunun dışında kalması oluşturmuştur.

Araştırma sırasında sadece *Twitter* üzerinde kargo firmalarının *mention*landığı *tweet*ler dikkate alınmış, çağrı merkezi kayıtları, e-postalar, müşteriler ve firmalar tarafından üretilen videolar, fotoğraflar, web siteleri üzerinde kargo firmalarına ilişkin görüşler araştırmanın dışında bırakılmıştır.

Büyük veri analizi için doktora seviyesinde öğretilen araştırma metot ve tekniklerinin yeterli olmadığı literatürde belirtilmektedir. Aynı zamanda büyük veri analizi için yeni metot ve tekniklerin geliştirilmesi gerektiğinden de bahsedilmektedir.

3.5. ARAŞTIRMANIN METODOLOJİSİ

Araştırmada 2107 *tweet* içeriği değerlendirilmiş ve değişkenlere ilişkin ifade içerip içermedikleri, konuları, amaçları bakımından gruplanmıştır. Araştırma incelenen içeriğin örnekleme oluşturması, değerlendirmenin bilgisayarlarla ve büyük veri analizi teknikleri ile yapılmaması, veri büyüklüğünün büyük veriye ilişkin tanımlarda geçen konvansiyonel yöntemlerle analiz edilemeyecek büyüklükte olmaması bakımından, pek çok büyük veri tanımına göre, bir büyük veri analizi çalışması değildir; fakat gelecekte büyük veri analizinin lojistikte kullanımı açısından bir bakış açısı oluşturmaktadır.

3.6. VERİLERİN TOPLANMASI VE ANALİZİ

13 Kasım 2017 - 23 Kasım 2017 arasında 3 kargo şirketine dair atılmış olan 2107 *tweet* ulaşılmıştır. Söz konusu *tweet*ler veri tabanına alınmış ardından hesap tablosunda toplanarak profil, konu, amaç, konum bilgisi, memnuniyet ya da memnuniyetsizlik ifadesi içermesi, kargo firmaları arasında karşılaştırma yapıp yapılmadığı yönünden içerik değerlendirmesi yapılarak; Çakmak ve Özkan' ın (2017) araştırmalarında kullandıkları değişkenlere ilişkin ifade içerip içermediğine bakılmıştır.

Literatüre bakıldığında *Twitter Search API*' nin yoğun olarak kullanıldığı görülmektedir. *Twitter*' in sitesi üzerinde *Search API*' in *tweet*lerin bir bölümünü döndürdüğü söylene de örneklemin nasıl sağlandığına ilişkin herhangi bir bilgiye rastlanamamıştır. Literatüre bakıldığında *Twitter Search API* ile duygu ve içerik analizi için veri çekmenin mümkün olduğu görülmektedir. *Search API* ile ilgili açıklamalara tezde bu nedenle yer verilmiştir, ancak *Search API*' in *Twitter* tarafından da ifade edilmekte olduğu üzere sorgulanan parametrelerin tamamını vermediği ve versiyonlarının, dolayısı ile imkân ve sınırlarının değişebileceği göz önünde bulundurulmalıdır.

Twitter dünyada ne olup bittiği ve insanların ne hakkında konuştuğu konusunda bilgi sağlamaktadır. *Twitter* tarafından ifade edildiği kadarıyla *Twitter* üzerindeki bilginin mümkün olduğunca geniş bir biçimde paylaşılması için firmalara, geliştiricilere ve kullanıcılara API vasıtası ile erişim sağlanmaktadır. *Twitter* verisine bu şekilde erişmek için kayıt yapılması gerekmektedir. API hesap yönetimine, *tweet* atma, cevaplama, arama, mesaj göndermeye imkân tanımaktadır help.twitter.com/en/rules-and-policies/twitter-api, (29.07.2018).

Standard *Search API* basit sorgulamalara izin verir ve son 7 gün içerisinde atılmış olan *tweet*lerden bir örneklem sunar. Belirtilen bir sorgu için eşleşen *tweet*lerin bir bölümünü döndürmektedir. *Search API*' in tüm *tweet*leri teferruatlı bir şekilde döndürmediği göz önünde bulundurulmalıdır. *Search API*' ı kullanmak için yetkilendirme gerekmektedir. User auth ile 15 dakikada 180 arama, app auth ile 450 arama yapılabileceği ifade edilmektedir <https://developer.twitter.com/en/docs/tweets/search/overview/standard>, (29.07.2018).

Tablo 3: *Tweeter* Search API Parametreler ve Açıklamaları

Parametre	Gerekli/İsteğe bağlı	Tanım
q	Gerekli	Maksimum 500 karakterlik arama sorgusu için
geocode	İsteğe Bağlı	Belirli konumdan atılan tweetlerin alınması için
lang	İsteğe Bağlı	Tweetleri dile göre sınırlar
Locale	İsteğe Bağlı	Gönderilen sorgu dilini belirtir
result_type	İsteğe Bağlı	Sonuç tercihi için kullanılır
count	İsteğe Bağlı	Bir sayfada gelecek tweetleri belirlemek için kullanılır. Varsayılan 15'tir, maksimum 100'dür.
until	İsteğe Bağlı	Verilen tarihten önceki tweetleri döndürür. Search API' da 7 günden eski tweetler döndürülemez.
since_id	İsteğe Bağlı	Verilen nodan büyük olan noya sahip tweetleri döndürür.
max_id	İsteğe Bağlı	Verilen noya eşit veya küçük olan noya sahip tweetleri döndürür.
include_entities	İsteğe Bağlı	False olarak ayarlandığında entity nodu dahil edilmez

Kaynak: <https://developer.twitter.com/en/docs/tweets/search/api-reference/get-search-tweets>, (29.07.2018).

Araştırmada verilerin toplanması için aşağıdaki gibi bir sorgu kullanılmıştır.

'[https://api.twitter.com/1.1/search/tweets.json?q=%40YurticiKargo&max_id='.\\$max_id.'&src=typd&count=100](https://api.twitter.com/1.1/search/tweets.json?q=%40YurticiKargo&max_id='.$max_id.'&src=typd&count=100)'

3.7. ARAŞTIRMANIN DEĞİŞKENLERİNİN BELİRLENMESİ

Çakmak ve Özkan'ın (2017) araştırmalarında önemleri ölçülen değişkenler, atılan *tweet*lerin içerisinde aranmıştır. Bu değişkenler aşağıdaki gibi verilmiştir:

- Firmanın müşteri bilgilerini güncellemesi
- Çağrı merkezi ile iletişim
- Kargo şubesinin çalışma gün ve saatleri
- Mevcut kampanyaların uygulanması
- Gönderinin teslim alınırken kontrol edilmesi
- Gönderinin belirtilen adrese teslimi
- Gönderinin belirtilen kişiye teslimi
- Kargo şubesiyle telefon iletişimi
- Kargo çalışanlarının üslup ve davranışları
- Gönderinin internetten takibi
- Kargo çalışanlarının yaşanan problemleri çözmesi
- Gönderiyi belirtilen yerden istenildiğinde almaları
- Kargo firmasının gönderi ücretleri
- Ödeme imkânları çeşitliliği
- Gönderinin eksiksiz teslimi
- Kargo firmasına ait sitenin güncel olması
- Firma ve çalışanlar tarafından verilen bilgilerin doğruluğu
- Kargo çalışanlarının işini özenle yapması
- Kargo şubesinin fiziksel koşulları
- Gönderinin açılmaması kurcalanmaması
- Gönderinin teslim alındığı şekliyle teslimi
- Kargo şirketinin teknolojik olarak gelişmişliği
- Kargo takip sistemindeki bilgilerin doğruluğu
- Gönderinin kapıya getirilmesi
- Gönderinin zamanında teslimi
- Kargo taşımada müşteriden yardım istenmesi
- Kargo firmasının fatura ve benzeri evrakları her zaman düzenlemesi
- Müşteriye mesajla bilgi verilmesi

- Kargo firması şoförlerinin trafik kurallarına uyması
- Gönderinin teslim süresi
- Gönderi ile ilgili sorunlar yaşandığında tutanak tutulması

Kullanıcıların *Twitter*' ı kullanma amaçları, profilleri, atılan *tweet*lerin konuları, konum bilgisi, memnuniyet ya da memnuniyetsizlik ifadesi olup olmadığı, kargo firmaları arasında karşılaştırma yapılıp yapılmadığı ise diğer değişkenleri oluşturmaktadır.

3.8. BULGULAR

Profillerine göre *tweet*ler incelendiğinde; 1578 *tweet* alıcılar tarafından, 136 *tweet* firmalar tarafından, 80 *tweet* göndericiler tarafından atılmış; 313 *tweet*in hangi profilden geldiği belirlenememiştir, Alıcılar tarafından atılan *tweet*ler %74,89' a, firmalar tarafından atılan *tweet*ler, %6,45' e, göndericiler tarafından atılan *tweet*ler yüzde 3,80' e tekabül etmektedir, *Tweet*i atan profilin belirlenemediği *tweet*ler örneklemin %14,86' sını oluşturmaktadır (Tablo 4).

Tablo 4:Atılan Profillere Göre *Tweet*ler (%)

Profil	n	Yüzde	Kümülatif Yüzde
Alıcı	1578	74,89	74,89
Belirsiz	313	14,86	89,75
Firma	136	6,45	96,20
Gönderici	80	3,80	100,00
Toplam	2107	100,00	

Tablo 5: *Retweet* Oranları (%)

<i>Retweet</i>	n	Yüzde	Kümülatif Yüzde
Var	198	9,40	9,40
Yok	1909	90,60	100,00
Toplam	2107	100,00	

Retweete 2107 *tweet*in 198' inde rastlanmıştır, 1909 *tweette* ise bulunamamıştır. Atılan *tweet*lerin %9,40' ını *retweeter*ler oluşturmaktadır (Tablo 5).

Tablo 6:Konularına Göre *Tweet*ler (%)

Konu	n	Yüzde	Kümülatif Yüzde
Performans	76	3,61	3,61
Belirsiz	656	31,13	34,74
İletişim	275	13,05	47,79
Altyapı	53	2,52	50,31
Gönderici İle İlgili	57	2,71	53,01
Müşteri Hizmetleri	73	3,46	56,48
Teslimat	833	39,53	96,01
Personel İle İlgili	84	3,99	100,00
Toplam	2107	100,00	

Konularına göre *tweet*lere bakıldığında; 833 *tweet*, 39,53% oranla en çok teslimat konusunda *tweet* atılmıştır, 656 *tweet*in konusu belirlenemezken, 275 *tweet* iletişim konusunda atılmıştır (Tablo 6).

Tablo 7: Amacına Göre *Tweetler* (%)

Değer	n	Yüzde	Kümülatif Yüzde
Anma	1	0,05	0,05
Belirsiz	389	18,46	18,51
İletişim	25	1,19	19,7
Karşılaştırma	1	0,05	19,75
Memnuniyet İfadesi	24	1,14	20,89
Öneri	6	0,28	21,17
Sorun Çözümü	1	0,05	21,22
Sponsor Arayışı	2	0,09	21,31
Şikâyet	1539	73,04	94,35
Tanıtım	4	0,19	94,54
Tebrik	1	0,05	94,59
Tehdit	1	0,05	94,64
Yardım	113	5,36	100
Toplam	2107	100	

Amaçlarına göre *tweet*lerde en yüksek oran %73,04, 1539 *tweet* ile şikâyet olarak bulunmuştur, 389 *tweet*in amacı belirlenemezken, yardım amacıyla atılan 113 *tweet* örneklemin %5,36' ını oluşturmaktadır (Tablo 7).

Tablo 8: Konum Bilgisi Sağlama Durumuna göre *Tweetler* (%)

Değer	n	Yüzde	Kümülatif Yüzde
Var	169	8,02	8,02
Yok	1938	91,98	100,00
Toplam	2107	100,00	

Konuma ilişkin ifadeye, 2107 *tweet*in 169' unda rastlanırken, 1938 *tweette* rastlanmamıştır. Konuma ilişkin ifadeye rastlanan *tweetler* örneklemin %8,02' sini, rastlanamayan *tweetler* %91,98' ini oluşturmaktadır (Tablo 8).

Tablo 9: Memnuniyet İfadesi Oranları (%)

Değer	n	Yüzde	Kümülatif Yüzde
Var	26	1,23	1,23
Yok	2081	98,77	100,00
Toplam	2107	100,00	

Memnuniyet ifadesine 26 *tweette* rastlanırken, 2081 *tweette* rastlanmamıştır. Memnuniyet ifadesinin bulunduğu *tweetler* örneklemin %1,23' ünü oluştururken, bulunmadığı *tweetler* %98,77' sini oluşturmaktadır (Tablo 9).

Tablo 10: Memnuniyetsizlik İfadesi Oranları (%)

Değer	n	Yüzde	Kümülatif Yüzde
Var	495	23,49	23,49
Yok	1612	76,51	100,00
Toplam	2107	100,00	

Memnuniyetsizlik ifadesine rastlanan *tweetler* örneklemin %23,49' unu oluştururken, rastlanmayan *tweetler* %76,51' ini oluşturmaktadır. Memnuniyetsizlik ifadesi bulunan *tweet* sayısı 495, bulunmayan *tweet* sayısı 1612' dir (Tablo 10).

Tablo 11: Karşılaştırma Oranları(%)

Değer	n	Yüzde	Kümülatif Yüzde
Var	65	3,08	3,08
Yok	2042	96,92	100,00
Toplam	2107	100,0	

2107 *tweete* bakıldığında kullanıcılar 65 *tweette* kargo firmalarını birbiri ile kıyaslamış, 2042 *tweette* herhangi bir kıyaslamaya rastlanamamıştır. Kıyaslama yapılan *tweetler* örneklemin %3,08' ini, yapılmayan *tweetler* %96,92' sini oluşturmaktadır (Tablo 11).

Tablo 12:Çakmak ve Özkan’ ın Çalışmasında Ölçülen İfadelerin Örneklem İçindeki Sıklıkları(%)

Değişken	Var		Yok		Toplam (%)
	n	%	n	%	
Müşteri Bilgileri Güncelleme	3	0,14	2104	99,86	100
Çağrı Merkezi İletişim	60	2,85	2047	97,15	100
Çalışma gün ve saatleri	---	---	2017	100	100
Kampanyaların Uygulanması	2	0,09	2105	99,91	100
Gönderinin Teslim Alınırken Kontrolü	2	0,09	2105	99,91	100
Gönderinin Belirtilen Adrese Teslimi	230	10,92	1877	89,08	100
Gönderinin Belirtilen Kişiye Teslimi	64	3,04	2043	96,96	100
Kargo Şubesi ile Telefon İletişimi	120	5,70	1987	94,30	100
Kargo Çalışanlarının Uslup ve Davranışları	191	9,07	1916	90,93	100
Gönderinin İnternette Takibi	20	0,95	2087	99,05	100
Çalışanların Yaşanan Problemi Çözmesi	61	2,90	2046	97,10	100
Gönderinin Belirtilen Adresten İstenildiğinde Alınması	28	1,33	2079	98,67	100
Kargo Firması Gönderi Ücretleri	22	1,04	2085	98,96	100
Ödeme İmkanları Çeşitliliği	5	0,24	2102	99,76	100
Gönderinin Eksiksiz Teslimi	60	2,85	2047	97,15	100
Kargo Firmasına Ait Sitenin Güncel Olması	3	0,14	2104	99,86	100
Firma ve Çalışanlar Tarafından Verilen Bilgilerin Doğruluğu	188	8,92	1919	91,08	100
Kargo Çalışanlarının İşini Özenle Yapması	88	4,18	2019	95,82	100
Kargo Şubesi Fiziksel Koşulları	1	0,05	2106	99,95	100
Gönderinin Açılmaması Kurcalanmaması	3	0,14	2104	99,86	100
Gönderinin Teslim Alındığı Şekliyle Teslimi	15	0,71	2092	99,29	100
Kargo Şirketinin Teknolojik Olarak Gelişmişliği	4	0,19	2103	99,81	100
Kargo Takip Sistemindeki Bilgilerin Doğruluğu	18	0,85	2089	99,15	100
Gönderinin Kapıya Getirilmesi	135	6,41	1972	93,59	100
Gönderinin Zamanında Teslimi	414	19,65	1693	80,35	100
Kargo Taşımada Müşteriden Yardım İstenmesi	2	0,09	2105	99,91	100
Kargo Firmalarının Evrakları Her Zaman Düzenlemesi	7	0,33	2100	99,67	100
Müşteriye Mesajla Bilgi Verilmesi	27	1,28	2080	98,72	100
Trafik Kurallarına Uyulması	2	0,09	2105	99,91	100
Gönderinin Teslim Süresi	71	3,37	2036	96,63	100
Sorun Halinde Tutanak Tutulması	---	---	2107	100	100

Müşteri bilgilerinin güncellenmesine ilişkin ifadeye rastlanan *tweet*ler örneklemin %0,14’ ünü, rastlanmayan *tweet*ler %99,86’ sını oluşturmaktadır(Tablo 12).

Çağrı merkeziyle iletişime ilişkin 60 *tweet*e rastlanmıştır, Söz konusu *tweet*ler örneklemin %2,85’ini oluşturmaktadır, 2047 *tweet*te çağrı merkezi ile iletişime ilişkin ifadeye rastlanmamıştır. Çağrı merkezi ile iletişime ilişkin ifadeye rastlanmayan *tweet*ler örneklemin %97,15’ ini oluşturmaktadır (Tablo 12).

Kampanyaların uygulanmasına ilişkin ifadeye rastlanan *tweetler* örneklemin %0,09' unu, rastlanmayan *tweetler* %99,91'ini oluşturmaktadır. Kampanyaların uygulanmasına ilişkin ifadeye rastlanan *tweet* sayısı 2, rastlanmayan *tweet* sayısı 2015' tir (Tablo 12).

Gönderinin teslim alınırken kontrolüne ilişkin 2 *tweete* rastlanmıştır, Gönderinin teslim alınırken kontrolüne ilişkin ifadeye rastlanan *tweetler* örneklemin %0,09' unu, rastlanmayan *tweet* sayısı %99,91'ini oluşturmaktadır (Tablo 12).

Gönderinin belirtilen adrese teslimine ilişkin 230 *tweete* rastlanmıştır, Söz konusu *tweetler* örneklemin %10,92' sini oluşturmaktadır. Gönderinin belirtilen adrese teslimine ilişkin ifadeye rastlanmayan *tweet* sayısı 1877, oranı %89,08' dir (Tablo 12).

Gönderinin belirtilen kişiye teslimine ilişkin 64 *tweete* rastlanmıştır. Söz konusu *tweetler* örneklemin %3,04' ünü oluşturmaktadır. Gönderinin belirtilen kişiye teslimine ilişkin ifadeye rastlanmayan *tweet* sayısı 2043, oranı %96,96' dır (Tablo 12).

Kargo şubesi ile telefon iletişimine ilişkin 120 *tweette* ifadeye rastlanmıştır. 1987 *tweette* kargo şubesi ile telefon iletişimine ilişkin ifadeye rastlanmamıştır. Kargo şubesi ile telefon iletişimine ilişkin ifadeye rastlanan *tweetler* örneklemin %5,7' sini, rastlanmayan *tweetler* %94,3' ünü oluşturmaktadır (Tablo 12).

Kargo çalışanlarının uslu ve davranışları konusunda 191 *tweette* ifadeye rastlanmıştır, Söz konusu *tweetler* örneklemin %9,07' sini oluşturmaktadır. Kargo çalışanlarının uslu ve davranışları konusunda 1916 *tweette* ifadeye rastlanmamıştır. Uslu ve davranışları konusunda ifadeye rastlanmayan *tweetlerin* oranı %90,93' tür (Tablo 12).

Gönderinin internetten takibi konusunda 20 *tweette* ifadeye rastlanmıştır. Gönderinin internetten takibi konusunda ifadeye rastlanan *tweetlerin* örnekleminde oranı %0,95' tir. 2087 *tweette* gönderinin internetten takibi konusunda ifadeye rastlanmamıştır. Bu konuda ifadeye rastlanmayan *tweetler* örneklemin %99,05' ini oluşturmaktadır (Tablo 12).

Çalışanların yaşanan problemi çözmesi konusunda 61 *tweette* ifadeye rastlanmıştır. Söz konusu *tweetler* örneklemin %2,9' unu oluşturmaktadır.

Çalışanların yaşanan problemi çözmesi konusunda ifadeye rastlanmayan *tweet* sayısı 2046 iken, oranı %97,1' dir (Tablo 12).

Gönderinin belirtilen adresten istenildiğinde alınması konusunda 28 *tweette* ifadeye rastlanmıştır. Bu konuda ifadeye rastlanan *tweet* sayısı %1,33' tür. Gönderinin belirtilen adresten istenildiğinde alınması konusunda ifadeye rastlanmayan *tweet* sayısı 2079 iken, oranı %98,67' dir (Tablo 12).

Kargo firması gönderi ücretleri konusunda 22 *tweette* ifadeye rastlanmıştır. Kargo firması gönderi ücretleri konusunda ifadeye rastlanan *tweetlerin* örnekleme oranı %1,04' tür. Kargo firması gönderi ücretleri konusunda ifadeye rastlanmayan *tweetler* %98,96 oranla 2085 adettir (Tablo 12).

Ödeme imkânları çeşitliliği konusunda 5 *tweete* rastlanmıştır. Söz konusu *tweetler* örneklemin %0,24' ünü oluşturmaktadır. Ödeme imkânları çeşitliliği konusunda ifadeye rastlanmayan *tweet* sayısı 2102, oranı %99,76' dır (Tablo 12).

Gönderinin eksiksiz teslimi konusunda 60 *tweete* rastlanmıştır. Oranı %2,85' tir. Gönderinin eksiksiz teslimi konusunda ifadeye rastlanmayan *tweet* sayısı 2047, oranı %97,15' tir (Tablo 12).

Kargo firmasına ait sitenin güncel olması konusunda 3 *tweete* rastlanmıştır. Oranı %0,14' tür. Kargo firmasına ait sitenin güncel olması konusunda ifadeye rastlanmayan *tweet* sayısı 2104, oranı %99,86' dır (Tablo 12).

Firma ve çalışanlar tarafından verilen bilgilerin doğruluğu konusunda 188 *tweette* ifadeye rastlanmıştır. Oranı %8,92' dir. Firma ve çalışanlar tarafından verilen bilgilerin doğruluğu konusunda ifadeye rastlanmayan *tweet* sayısı 1919 adetken oranı %91,08' dir (Tablo 12).

Kargo çalışanlarının işini özenle yapması konusunda 88 *tweette* ifadeye rastlanmıştır. İfadeye rastlanan *tweetlerin* örnekleme oranı %4,18' dir. Kargo çalışanlarının işini özenle yapması konusunda ifadeye rastlanmayan *tweetlerin* sayısı 2019, oranı %95,82' dir (Tablo 12).

Kargo şubesi fiziksel koşulları konusunda 1 *tweette* ifadeye rastlanmıştır. Oranı %0,05' tir. Bu konuda ifadeye rastlanmayan *tweet* sayısı 2016, oranı %99,95' tir (Tablo 12).

Gönderinin açılmaması kurcalanmaması konusunda 3 *tweette* ifadeye rastlanmıştır. Bu konuda ifadeye rastlanan *tweetler* örneklemin %0,14' ünü

oluşturmaktadır. Gönderinin açılmaması kurcalanmaması konusunda ifadeye rastlanmayan *tweet*lerin sayısı 2104, oranı %99,86' dır(Tablo 12).

Gönderinin teslim alındığı şekliyle teslimi konusunda 15 *tweette* ifadeye rastlanmıştır. Bu konuda ifadeye rastlanan *tweet*lerin oranı %0,71' dir. Gönderinin teslim alındığı şekliyle teslimi konusunda ifadeye rastlanmayan *tweet* sayısı 2092, oranı %99,29' dur (Tablo 12).

Kargo şirketlerinin teknolojik gelişmişliği konusunda ifadeye rastlanan *tweet* sayısı 4, oranı %0,19' dur. Kargo şirketlerinin teknolojik gelişmişliği konusunda ifadeye rastlanmayan *tweet* sayısı 2103, oranı %99,81' dir (Tablo 12).

Kargo takip sistemindeki bilgilerin doğruluğu konusunda 18 *tweette* ifadeye rastlanmıştır, oranı %0,85' tir. Kargo takip sistemindeki bilgilerin doğruluğu konusunda ifadeye rastlanmayan *tweet* sayısı 2089, oranı %99,15' tir (Tablo 12).

Gönderinin kapıya getirilmesi konusunda 135 *tweette* ifadeye rastlanmıştır. Gönderinin kapıya getirilmesi konusunda ifadeye rastlanan *tweet* sayısı %6,41' dir. Gönderinin kapıya getirilmesi konusunda ifadeye rastlanamayan *tweet* sayısı 1972, oranı %93,59' dur (Tablo 12).

Gönderinin zamanında teslimi konusunda 414 *tweette* ifadeye rastlanmıştır. Söz konusu *tweet*ler örneklemin %19,65' ini oluşturmaktadır. Gönderinin zamanında teslimi konusunda ifadeye rastlanmayan *tweet*ler 1693 adetle örneklemin %80,35' ini oluşturmaktadır (Tablo 12).

Kargo taşımada müşteriden yardım istenmesi konusunda 2 *tweette* ifadeye rastlanmıştır. Kargo taşımada müşteriden yardım istenmesi konusunda ifadeye rastlanan *tweet*ler örneklemin %0,09' unu oluşturmaktadır. Kargo taşımada müşteriden yardım istenmesi konusunda ifadeye rastlanmayan *tweet*lerin sayısı 2105, oranı %99,91' dir (Tablo 12).

Kargo firmalarının evrakları her zaman düzenlemesi konusunda 7 *tweette* ifadeye rastlanmıştır. Kargo firmalarının evrakları her zaman düzenlemesi konusunda ifadeye rastlanan *tweet*ler örneklemin %0,33' ünü oluşturmaktadır. Kargo firmalarının evrakları her zaman düzenlemesi konusunda 2100 *tweette* ifadeye rastlanmamıştır. Söz konusu ifadeye rastlanmayan *tweet*ler örneklemin %99,67' sini oluşturmaktadır (Tablo 12).

Müşteriye mesajla bilgi verilmesi konusunda 27 *tweette* ifadeye rastlanmıştır. Müşteriye mesajla bilgi verilmesi konusunda ifadeye rastlanan *tweetler* örneklemin %1,28' ini oluşturmaktadır. Müşteriye mesajla bilgi verilmesi konusunda ifadeye rastlanmayan *tweetlerin* sayısı 2080, oranı %98,72' dir (Tablo 12).

Trafik kurallarına uyulması konusunda 2 *tweette* ifadeye rastlanmıştır. Trafik kurallarına uyulması konusunda ifadeye rastlanan *tweetler* örneklemin %0,09' unu oluşturmaktadır. Trafik kurallarına uyulması konusunda ifadeye rastlanmayan *tweetlerin* sayısı 2105, oranı %99,91' dir (Tablo 12).

Gönderinin teslim süresi konusunda 71 *tweette* ifadeye rastlanmıştır. Söz konusu *tweetlerin* örneklem içerisinde oranı %3,37' dir. Gönderinin teslim süresi konusunda ifadeye rastlanmayan *tweetlerin* sayısı 2036, örneklem içindeki oranları %96,63' tür (Tablo 12).

Örneklem içerisinde çalışma gün ve saatleri ile sorun halinde tutanak tutulması konusunda ifadeye rastlanamayınca, bu konuda gerçekten hiç *tweet* atılmadığını kontrol etmek üzere sorgu kelimeleri “@*****kargo”+”tutanak”, “@*****kargo”+”çalışma”+”gün” (firma adı gizlenmiştir) olarak değiştirilerek atılan *tweetler* *Twitter* search üzerinden kontrol amaçlı tekrar kontrol edilmiştir, Bu konularda atılan *tweetler* olduğu görülmüştür. Çalışma gün ve saatleri hakkında atılan *tweetlere* ait örnekler Tablo 13 verilmiştir.Sorun halinde tutanak tutulmasına ilişkin *tweet* örnekleri tablo 14' teki gibidir.

Tablo 13:Çalışma Gün veSaatleri Hakkında Atılan *Tweetler* (firma isimleri gizlenmiştir)

05:09 - 14 Oca 2017	Çalışma saatlerinizi kontrolden geçirin akşam 5'e kadar çalıştığınızı yazdığınız Beykent şubeniz saat 15,30da kapattık diyor
08:01 - 3 Ağu 2016	Çalışma saatleri 09:00-18:00 ama saat 17:00'den sonra gelen kargolara fiş verilmiyor ne kadar tuhaf bir çalışma prensibi
00:18 - 16 Ara 2015	Hayir o zaman calisma saati neden 9 da basliyo yaziyosun ?? 1 saat 18 dakikadir ariyorum bekle ki biri acsin?! hafta ici bu!!

Tablo 14:Sorun Halinde Tutanak Tutulması Konusunda Rastlanan *Tweetler* (firma isimleri gizlenmiştir)

06:42 - 2 Şub 2018	Teslim aldığım paketin içinde eksik ürün olduğunu kargoyu açınca fark ettim, baktım kutu kenardan yırtılmış, paketi teslim aldığım için tutanak tutmuyorlar, şube haberi olmasına rağmen bakıp dönücez, yan yattık, çamura battık diyor, nasıl yapcaz
02:29 - 30 Ara 2017	Arkadaslar aradılar,Paketi kapıda teslim alamadığımdan(aynı gün gelen ***** kargo teslimat yapmisti)on yuz bin milyon paketi asarak subelerine girdim ve oradan aldım,Tutanak tutturmadığım için bir sey yapamazlarmis,Bilgilendirmek istedim.
05:19 - 13 Mar 2017	Cuma günü Pendik şubenizin tarafıma teslim ettiği tutanak tutulmasını istediğimde de bana siz açtınız ben böyle teslim **
14:05 - 15 Kas 2016	Giyimkent şb, hasarlı koliye tutanak tutmuyor, Subeye gitmiş olmama rağmen adreste bulamadık muamelesi yapıyor.

3.8.1.Alıcı Profilinden Atılan *Tweetler*

Alıcı profili için *retweetler* dikkate alınmamıştır: Alıcı profilinden atılan *tweetlerin* sayısı 1578'dir.

Tablo 15:Alıcı Profilinden Atılan *Tweetlerde* Takip No Bildirme Oranları(%)

Değer	n	Yüzde	Kümülatif Yüzde
Var	17	1,08	1,08
Yok	1561	98,92	100,00
Toplam	1578	100,00	

Alıcıların sadece 17' si atılan *tweetlerde* kargo numarasını bildirmiştir. Kargo numarasını bildiren alıcıların, tüm alıcılar tarafından atılan *tweetlere* oranı %1,08' dir (Tablo 15).

Tablo 16: Alıcı Profili Konularına Göre *Tweetler*(%)

Değer	n	Yüzde	Kümülatif Yüzde
Altyapı	12	0,76	0,76
Belirsiz	458	29,02	29,78
Gönderici ile ilgili	47	2,98	32,76
İletişim	174	11,03	43,79
Müşteri hizmetleri	22	1,39	45,18
Performans	45	2,85	48,04
Personel ile ilgili	72	4,56	52,60
Teslimat	748	47,40	100,00
Toplam	1578	100,0	

Alıcıların 748 adet, %47,40 oranla en çok teslimat konusunda *tweet* attıkları görülmüştür (Tablo 16).

Tablo 17: Alıcı Profili Amaçlarına Göre *Tweetler*(%)

Değer	n	Yüzde	Kümülatif Yüzde
Belirsiz	253	16,03	16,03
İletişim	3	0,19	16,22
Karşılaştırma	1	0,06	16,29
Memnuniyet İfadesi	18	1,14	17,43
Öneri	5	0,32	17,74
Şikâyet	1297	82,19	99,94
Tehdit	1	0,06	100,00
Toplam	1578	100,00	

Alıcılar 1297 adet %82,19 oranla en çok şikâyet amacıyla *tweet* atmaktadır (Tablo 17).

Tablo 18: Alıcı Profili Konum Bilgisi Sağlanma Oranına Göre *Tweetler*(%)

Değer	n	Yüzde	Kümülatif Yüzde
Var	146	9,25	9,25
Yok	1432	90,75	100,00
Toplam	1578	100,00	

1578 *tweet*in 146' sında konuma ilişkin bilgilere yer verilmiştir. Konuma ilişkin bilgilere yer verilen *tweet*lerin oranı %9,25; yer verilmeyen *tweet*lerin sayısı 1432 oranı %90,75' tir (Tablo 18).

Tablo 19: Alıcı Profili Memnuniyet İfadesi Bulunma Durumunda Göre *Tweetler*(%)

Değer	n	Yüzde	Kümülatif Yüzde
Var	22	1,39	1,39
Yok	1556	98,61	100,00
Toplam	1578	100,00	

Memnuniyet ifadesi sadece 22 *tweette* geçmektedir, alıcılar tarafından atılan *tweet*lere oranı %1,39' dur (Tablo 19).

Tablo 20: Alıcı Profili Memnuniyetsizlik İfadesine Göre *Tweetler*(%)

Değer	n	Yüzde	Kümülatif Yüzde
Var	424	26,87	26,87
Yok	1154	73,13	100,00
Toplam	1578	100,0	

424 *tweette* memnuniyetsizlik ifadesine rastlanmıştır. Memnuniyetsizlik ifadesine rastlanan *tweet*ler örneklemin %26,87' sini oluşturmaktadır (Tablo 20).

Tablo 21: Alıcı Profili Karşılaştırma Yapılmasına Göre *Tweetler*(%)

Değer	n	Yüzde	Kümülatif Yüzde
Var	53	3,36	3,36
Yok	1525	96,64	100,00
Toplam	1578	100,00	

Alıcıların %3,36' sını attıkları *tweet*lerde kargo firmalarını birbiri ile kıyaslayarak karşılaştırma yapmıştır (Tablo 21).

3.8.2.Gönderici Profilinden Atılan *Tweetler*

Gönderici profilinden atılan *tweetler* için *retweetler* dikkate alınmamıştır. Göndericiler tarafından atılan *tweet*lerde kargo takip numarası bilgisi sağlamamaktadır. Toplam 80 *tweet*in gönderici profilinden atıldığı tespit edilmiştir.

Tablo 22: Gönderici Profili Konusuna Göre *Tweetler*(%)

Değer	n	Yüzde	Kümülatif Yüzde
Altyapı	26	32,50	32,50
Belirsiz	11	13,75	46,25
Gönderici ile ilgili	3	3,75	50,00
İletişim	5	6,25	56,25
Müşteri Hizmetleri	14	17,50	73,75
Performans	4	5,00	78,75
Personel ile ilgili	2	2,50	81,25
Teslimat	15	18,75	100,00
Toplam	80	100,00	

Göndericilerin en çok altyapı konusunda *tweet* attıkları, altyapı konusunu teslimatın takip ettiği görülmektedir. Altyapı konusunda atılan *tweet*lerin sayısı 26, oranını %32,50' dir. Teslimat konusunda atılan *tweet*lerin sayısı 15 oranını %18,75' tir (Tablo 22).

Tablo 23: Gönderici Profili Amacına Göre *Tweetler*(%)

Değer	n	Yüzde	Kümülatif Yüzde
Belirsiz	5	6,25	6,25
Sponsor arayışı	2	2,50	8,75
Şikâyet	72	90,00	98,75
Yardım	1	1,25	100,00
Toplam	80	100,00	

Göndericilerin attığı *tweet*lerin %90' ının şikâyet amacıyla atıldığı görülmektedir (Tablo 23).

Tablo 24:Gönderici Profili Konum Bilgisi Sağlama Oranına Göre *Tweet*ler (%)

Değer	n	Yüzde	Kümülatif Yüzde
Var	7	8,75	8,75
Yok	73	91,25	100,00
Toplam	80	100,00	

Göndericilerin 7 *tweette* konuma ilişkin bilgi verdiği görülmektedir. Konuma ilişkin bilgi verilen *tweet*ler örneklemin %8,75' ini oluşturmaktadır. Konuma ilişkin bilgi verilmeyen *tweet*lerin sayısı 73, oranı %91,25' tir(Tablo 24).

Tablo 25:Gönderici Profili Memnuniyet İfadesi Bulunma Oranına Göre *Tweet*ler(%)

Değer	n	Yüzde	Kümülatif Yüzde
Var	3	3,75	3,75
Yok	77	96,25	100,00
Toplam	80	100,00	

3 *tweette* göndericilerin memnuniyetlerini ifade ettiği görülmüştür. Memnuniyet ifadesi görülen *tweet*lerin oranı %3,75' tir.77 *tweette* memnuniyet ifadesine rastlanmamıştır. Memnuniyet ifadesine rastlanmayan *tweet*lerin sayısı 77, oranı %96,26' tir (Tablo 25).

Tablo 26:Gönderici Profili Memnuniyetsizlik İfadesi Bulunma Oranına Göre *Tweet*ler(%)

Değer	n	Yüzde	Kümülatif Yüzde
Var	22	27,50	27,50
Yok	58	72,50	100,00
Toplam	80	100,00	

22 *tweette* memnuniyetsizlik ifadesine rastlanmıştır. Memnuniyetsizlik ifadesine rastlanan *tweet*lerin oranı %27,50' dir. Memnuniyetsizlik ifadesine rastlanmayan *tweet*lerin sayısı 58, oranı %72,50' dir(Tablo 26).

Tablo 27:Gönderici Profili Karşılaştırma Oranına Göre *Tweetler* (%)

Değer	n	Yüzde	Kümülatif Yüzde
Var	4	5,00	5,00
Yok	76	95,00	100,00
Toplam	80	100,00	

Göndericilerin 4 *tweette* kargo firmalarını birbiri ile kıyasladığı görülmüştür. Kargo firmalarının birbiri ile kıyaslandığı *tweetlerin* oranı %5' tir. Kargo firmalarının birbiri ile kıyaslanmadığı *tweetlerin* sayısı 76, oranı %95' tir (Tablo27).

3.8.3 Firma Profilinden Atılan *Tweetler*

Tablo 28: Firma Profili Takip Numarası Talebine Göre *Tweetler* (%)

Değer	n	Yüzde	Kümülatif Yüzde
Var	114	83,82	83,82
Yok	22	16,18	100,00
Toplam	136	100,00	

Firma profili için atılan *retweetler* dikkate alınmamıştır, Toplam 136 *tweetin* firma profilinden atıldığı görülmektedir. Firmalar attıkları *tweetlerin* 114' ünda kargo numarası talep etmişlerdir. Kargo numarası talep edilen *tweetlerin* oranı %83,82'dir. Kargo numarası talep edilmeyen *tweetlerin* sayısı 22, oranı %16,18' dir (Tablo 28).

Tablo 29:Firma Profili Konularına Göre *Tweetler*(%)

Değer	n	Yüzde	Kümülatif Yüzde
Altyapı	2	1,47	1,47
Belirsiz	8	5,88	7,35
İletişim	87	63,97	71,32
Müşteri Hizmetleri	31	22,79	94,12
Performans	5	3,68	97,79
Personel ile ilgili	2	1,47	99,26
Teslimat	1	0,74	100,00
Toplam	136	100,00	

Firmaların %63,97 ile en çok iletişim konusunda *tweet* attığı görülmektedir. Firmalar tarafından atılan 136 *tweet*in 87'sini iletişim konusu oluşturmaktadır. İletişim konusunu müşteri hizmetleri konusunda atılan *tweetler* takip etmektedir. Müşteri hizmetleri konusuna ilişkin atılan *tweetlerin* sayısı 31, oranı %22,79' dur (Tablo 29).

Tablo 30:Firma Profili Amacına Göre *Tweetler* (%)

Değer	n	Yüzde	Kümülatif Yüzde
Belirsiz	8	5,88	5,88
İletişim	17	12,50	18,38
Sorun Çözümü	1	0,74	19,12
Tanıtım	3	2,21	21,32
Yardım	107	78,68	100,00
Toplam	136	100	

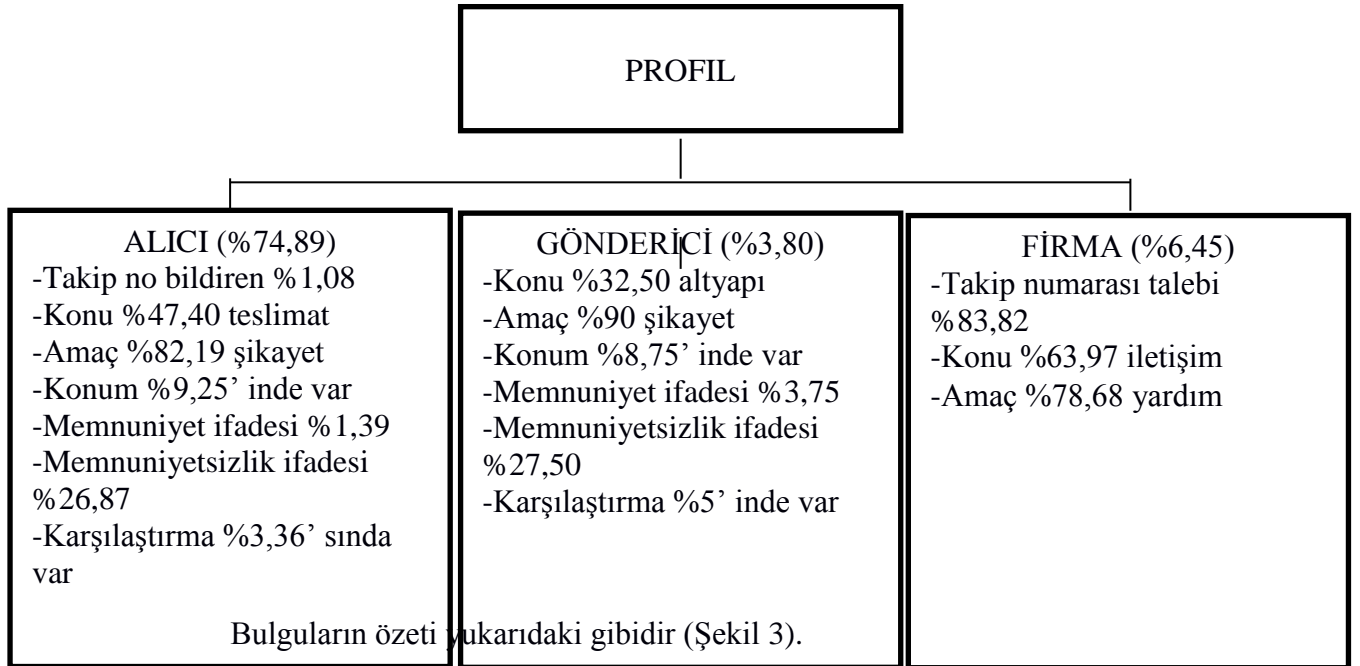
Firmalar 107 *tweet* ile en çok yardım amacıyla *tweet* atmaktadır. Yardım amacıyla atılan *tweetlerin* oranı %78,68' dir. Yardım amacıyla atılan *tweetleri* iletişim izlemektedir. İletişim amacıyla atılan *tweetlerin* oranı %12,50' dir (Tablo 30).

Tablo 31:Müşteri destek hesabından paylaşılan *tweetler*

Mon Nov 20 05:48:28 +0000 2017	Merhaba ,Kargo takip numaranızı ve iletişim bilgilerinizi DM'den paylaşabilir misiniz?
Mon Nov 20 05:44:45 +0000 2017	Merhaba, kargo işleminizle ilgili size yardımcı olmak isteriz, bilgilerinizi dm olarak bizimle paylaşabilir misiniz?
Mon Nov 20 05:35:23 +0000 2017	Merhaba, iletişim bilgilerinizi dm üzerinden ilettiğinizde size memnuniyetle yardımcı olabiliriz, teşekkürler

Cümle yapılarındaki farklılık sosyal medya üzerinden iletilen şikâyetlerin çalışanlar tarafından taranıp, cevaplandırıldığını düşündürmektedir. Destek hesaplarından atılan *tweetlere* bakıldığında ilk olarak müşterinin iletişim detayları ve kargo takip numaralarının istendiği görülmektedir. Bu nedenle müşteri şikâyetlerinin büyük veri analizi teknikleri ile taranıp, konularına göre gruplanıp; müşterilerden iletişim bilgilerinin alınabileceği, şikâyet konusu ve iletişim bilgileri ile birlikte ilgililere çalışanların müdahalesi olmadan iletebileceği düşünülmektedir (Tablo 31).

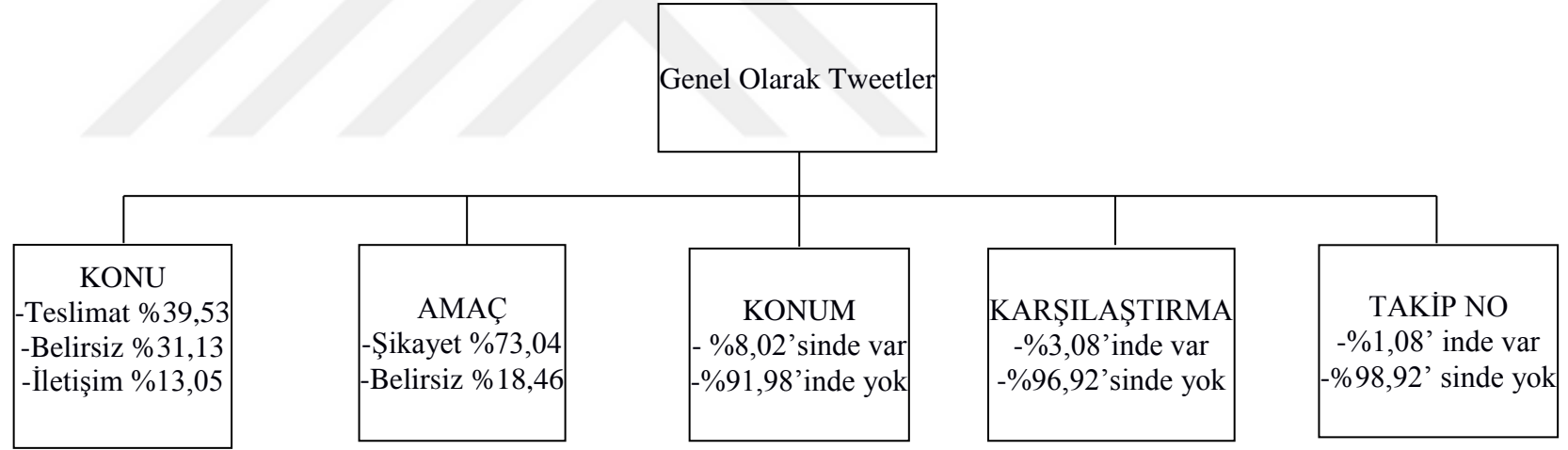
Şekil 3:Profillere Göre Atılan *Tweetlerin* Özeti



Büyük veri analizi tekniklerinin müşteri desteęi kapsamında kullanımının alıřanlar üzerindeki iř yükünü azaltabileceęi ve müşterilerin sorunlarının özümü için harcanan zamanı düşürebileceęi düşünölmektedir. Önerinin uygulanabilirlięi başka arařtırmaların konusudur (Tablo30).

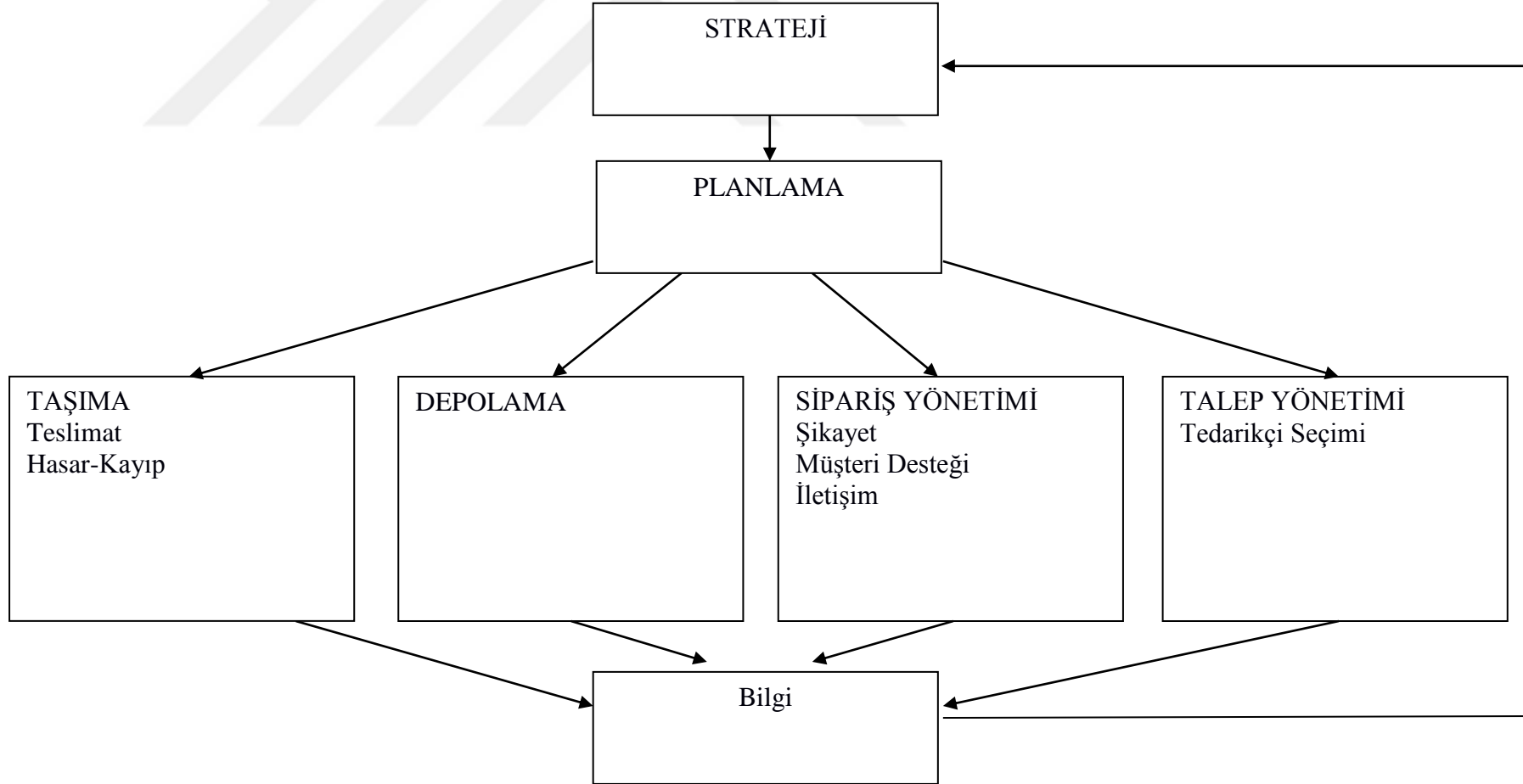


Şekil 4: Genel Olarak Atılan *Tweet*lerin Özeti



Genel olarak tweetlerin örnekleri yukarıdaki şekilde verilmiştir (Şekil 3).

Şekil 5: Lojistiğin Fonksiyonlarına Göre Atılan *Tweetler*



Lojistik fonksiyonlarına göre örneklemeden elde edilebilen bilgiler yukarıdaki şekilde gösterilmiştir. Depolama konusunda herhangi bir bilgiye erişilememiştir. Bunun dışında taşıma, sipariş yönetimi ve talep yönetimine ilişkin kimi ifadeler rastlanmıştır. Buradan çıkarılacak olan bilgilerin strateji geliştirme, planlama ve uygulamada kullanılabileceği düşünülmektedir (Şekil 5).



SONUÇ VE ÖNERİLER

Araştırmada kargo firması müşterilerinin *Twitter*' ı ağırlıklı olarak şikâyet amaçlı kullandığı, müşterilerin büyük kısmının şikâyetlerine konu olan gönderi numaralarını ise *tweet*lerinde paylaşmadıkları görülmüştür. Deniz ve Gökmerdan (2011) lojistiğin 15 fonksiyonundan biri olarak müşteri hizmetlerini vermiştir. Firmaların attığı *tweet*lere bakıldığında müşteri hizmetleri desteğinin bir bölümünün sosyal medya üzerinden verildiği görülmektedir. Firmaların attığı *tweet*lerdeki cümle yapılarının farklılıkları söz konusu desteğin çalışanlar tarafından sağlandığını, çalışanların ise problem çözümü için öncelikli olarak iletişim bilgileri ve gönderi numarasına ihtiyaç duyduğunu göstermektedir. Büyük veri analizinde veri toplamak için geçen web kazıma teknikleri ile sosyal medyayı şikâyetler için taramanın ve müşterilerden iletişim bilgileri ve takip numaralarını bilgisayar ile almanın, Bongsug'un çalışmasında yaptığı üzere konularına göre gruplamanın ve ilgililere şikâyetleri konuları ile iletmenin mümkün olabileceği düşünülmektedir. Bu durumun da iş yükünü azaltarak verimliliği arttırabileceği beklenmektedir. Bu sayede cevap sürelerinin ve müşteri desteği için gereken iş gücünün düşebileceği umulmaktadır. Söz konusu önerinin uygulanabilirliği ve sektörde böyle bir ihtiyacın olup olmadığı başka araştırmaların konusunu oluşturmaktadır.

Kullanıcıların *tweet*lerde firmaları birbiri ile kıyasladığı ve *tweette* konuma ilişkin bilgi sağlayabildiği görülmektedir. Kullanıcı tarafından üretilen içeriğin bu anlamda e-ticaret firmalarının tedarikçi seçimi için karar destek sistemleri geliştirilmesinde kullanılabilmesi, konum bazında kullanıcılar tarafından üretilen verinin değerlendirilmesi ile kargonun gideceği muhitte hangi kargo firmasının performansının karşılaştırmalı olarak daha iyi olduğuna ilişkin değerlendirmeler yapılabileceği düşünülmektedir. Benzer analizler kargo firmalarının şube performanslarına ilişkin konum bazında rakip analizinde de kullanılabilir. Yine söz konusu önerilerin uygulanabilirliği ve sektörde ihtiyaç hissedilip hissedilmediği başka araştırmaların konusunu oluşturmaktadır.

Örnekleme bakıldığında pek çok müşterinin evde olduğu halde kargosunun kendisine teslim edilmediğini iletmediği, takip sistemine teslimata ilişkin girilen ve kargo firmaları tarafından alıcıya adresinde ulaşılamadığı yönünde sağlanan

bilgilerin doğru olmayabildiğini söylediği görülmektedir. Literatürde mobil uygulamaların büyük veri bağlamında lojistik yönetimde kullanımına ilişkin çalışmalar bulunmaktadır. Mobil uygulamalar ile alıcıdan teslimata ilişkin doğrulama yapmasının istenmesi halinde büyük veri analizi ile kargosu adreste olduğu halde teslim edilmeyen kullanıcılara ilişkin teslimat hizmetinin iyileştirilebileceği, hatta kargo firmalarının alıcılara ilişkin uygulamalardan konum bilgisi alması halinde alıcının teslimat adresinde olup olmadığının kontrol edilebileceği ve rotaların bu bilgilere göre anlık olarak iyileştirilebileceği düşünülmektedir. Literatürden anlaşıldığı kadarıyla büyük veri analizi teknikleri trafik durumu, alıcının konumu gibi pek çok değişkeni anlık olarak takip ederek rotaları iyileştirme ve başarısız teslimatlardan kaçınma imkânı sağlayabilecek potansiyele sahiptir. Söz konusu öneriler başka araştırmaların konusunu oluşturmaktadır.

Türkiye’de kargo firmaları üzerine iki araştırma müşteri memnuniyeti ve kargo firmalarına ilişkin önem başarımları analizi üzerine yapılmıştır. Müşterilerin *Twitter*’ı ağırlıklı olarak şikâyet amaçlı kullandıkları göz önünde bulundurulduğunda, sadece *Twitter* bağlamında ve araştırmada kullanılan tarama kelimeleri ile, kullanıcı tarafından üretilen içeriğin müşteri memnuniyetine ilişkin kullanılıp kullanılmayacağı konusunda soru işaretleri olduğu düşünülmektedir. Elbette büyük veri sadece *Twitter*’da bulunan kullanıcı tarafından üretilen içerikten oluşmamaktadır. Kullanıcıların servisten memnun olup olmadığına ilişkin görüşlerini paylaştıkları herhangi bir başka kaynak olup olmadığı, kullanıcı tarafından üretilen içeriğin değerlendirildiği çalışmalarla müşteri memnuniyeti analizi yapıp yapılamayacağı başka araştırmaların konusudur. Araştırmada sadece kargo firmalarının mentionlandığı *tweet*ler dikkate alınmıştır. Almış oldukları hizmetten memnun kalan kullanıcıların örneğin gönderilerin fotoğraflarını çekerek de paylaşımları söz konusu olabilir. Literatürden de anlaşıldığı kadarıyla büyük veri analizi zaten çeşitli verilerin bir bütün olarak değerlendirildiği bir yaklaşım olarak geçmektedir. Bu nedenle memnun olan müşterileri bulmak için de yöntemlerin geliştirilebilmesinin söz konusu olabileceği düşünülmektedir. Turizm ve otelcilik üzerine yapılan çalışmalarda kullanıcıların oteller hakkındaki yorumlarının toplandığı ve hizmetin puanlandığı *web* siteleri üzerinde çalışılmıştır. Müşterilerin hizmeti puanladığı *web* siteleri oluşturulması halinde ordinal verilerin alınabileceği

ve arařtırmalarda kullanılabileceđi dūřünülmektedir. Mūřterilerin *Twitter*' ı ađırlıklı olarak Őikāyet amaçlı kullanması, içeriđi Őikāyet olan *Tweet*lerin sayısının, hizmetten duyulan memnuniyetin ifade edildiđi *Tweet*lerden dūřük olması Tūrkiye' de kargo firmalarına iliřkin memnuniyetin ölçüldüğü çalıřmaların bulguları ile çeliřmektedir. Bu nedenle aldıđı hizmetten memnun olan insanların, bařka bir Őekilde herhangi bir geri bildirimde bulunup bulunmadıđı bařka arařtırmaların konusudur.

Bu arařtırmada, örneklem içerisinde çalıřma gün ve saatleri ile sorun halinde tutanak tutulması konusunda ifadeye rastlanamamıřtır, bu nedenle bu konuda gerçekten hiç *tweet* atılıp atılmadıđı kontrol edilmiřtir. Bu konularda atılan *tweet*ler olduđu görülmüřtür ve bu durumun popülāsyon verisi, daha büyük veri setlerinin gözden kaçabilecek kimi deđiřkenlere iliřkin bilgi sađlayabileceđi görüřünü güçlendirdiđi ve büyük veri analizi tekniklerinin geleneksel yöntemlere kıyasla bu konudaki üstünlüđünü gösterdiđi dūřünülmektedir.

Literatürde büyük veri analizinin popülāsyon verisini analiz edebildiđinden bahsedilmektedir, ancak aldıđı hizmetten memnun olduđu halde elektronik ortamda yakalanabilecek herhangi bir veri sađlamayan mūřterilerin olması durumunda popülāsyona iliřkin verinin elde edilmesi mümkün olmayacaktır. Büyük veri konusunda metoda iliřkin daha çok arařtırma yapılması gerekliliđinin literatürde altı çizilmektedir. Tanımlardaki ayırmda dikkate alındıđında, büyük veriye iliřkin kuramsal çalıřmalara da ihtiyaç duyulduđu dūřünülmektedir. Arařtırma sınıřımızı sadece *Twitter* üzerinden kargo firmaları hakkında atılan public *tweet*ler olarak belirlediđimizde popülāsyon verisine kısmen ulařmamız mümkün olabilecekken, Tūrkiye'de kargo kullanıcılarının memnuniyet seviyesini ölçmek istediđimizde memnuniyete iliřkin e-posta, çağrı merkezi kaydı, *tweet* gibi herhangi bir yolla geri bildirimde bulunmayan mūřterilerin veri setimizin dıřında kalması olasıdır. Literatürde arařtırmaların iki genel amaçla yapıldıđı görülmektedir. Göze çarpan birinci amaç çok büyük miktar ve çeřitlilikte verinin anlık olarak nasıl iřlenebileceđi iken, ikinci amaç herhangi bir yapılandırılmıř ölçek, deney ya da gözlemlerle toplanmamıř, ancak hali hazırda bulunan verinin ticari ve akademik anlamda kullanılabilecek Őekilde nasıl analiz edilebileceđi Őeklinde karřımıza çıkmaktadır. Literatürde büyük veri analizinin pek çok deđiřken arasında iliřki bulmak için kullanılabileceđi ancak deđiřkenler üzerindeki kontrolün deney ya da gözlemlerle

olduđu gibi yüksek olmadıđı da belirtilmektedir. Bu bakımdan büyük veri analizi ile yapılacak alıřmaların büyük verinin řahsına münhasır kendi problemlerinin bulunması gerektiđi, büyük veri tanımlarına bakıldıđında da büyük veri setlerinin analizi, iřlenmesi ve saklanması bakımından zaten problemlerinin bulunduđu ve analizlerin geleneksel yöntemlerle yapılamadıđı kanaati oluřmuřtur.

Büyük veri ile ilgili yazılan makalelere baktıđımızda geleneksel yöntemde gözlemden kurama gidildiđi ve gözlemi mümkün, veriyi anlamlı kılan řeyin kuram olduđu görülmektedir (Iřıklı; 2014). Iřıklı (2014) büyük veri analizinde ise kuramın yerini modelin aldıđı, nedensel görülmeyen iliřkiler kurulduđu; öngörünün yerini tahminin aldıđını, yasanın yerinin ise trendler tarafından alındıđını belirtmektedir. White ve diđerlerinin (2014) verinin toplanma, analiz yönünde getirdikleri eleřtiriler, George ve diđerlerinin (2016) doktora seviyesinde verilen arařtırma yöntemlerinin büyük veri analizi için yeterli olmadıđı eleřtirileri de göz önünde bulundurulduđunda. Büyük veri analizinin hem metod olarak, hem alınan sonuçlar anlamında geleneksel yöntemden farklı olduđu düşünölmektedir. Buna rađmen örnekleme üzerinde yapılan alıřmamızda, örnekleminde bulunmayan, sorun halinde tutanak tutulması ve alıřma gün ve saatlerine iliřkin ifade ieren *tweet*lere rastlanmıřtır. Bu nedenle büyük veri setleri üzerinde alıřabilme imkânının arařtırmacılara, literatürde de ifade edildiđi üzere, daha çok deđiřkeni göz önünde bulundurabilme imkânı tanıdıđı sonucuna ulařılmıřtır. Iřıklı (2014) büyük veri analizinin pratiklik yönüne dikkat çekerek, büyük veri paradigmasında ıvır zıvır denilen bilgilerden pratik faydalar elde edilebileceđini söylemektedir. Bu nedenle metod ve sonuçların geleneksel yöntemden farklı olacađı düşünölmüřtür. Bir diđer zorluk yeni teknoloji ve yöntemleri kullanabilen uzmanların eksikliđi olarak belirtilmektedir. Bu problemlerin Iřıklı' nın da belirttiđi üzere disiplinler arası alıřmalar, lojistik yönetiminde büyük veri analizinin kullanımına iliřkin yöntemsel alıřmalar ve büyük veri analistleri yetiřtirmek için çift anadal programları gibi yöntemlerle özölebileceđi düşünölmektedir.

KAYNAKÇA

Augustyniak, L. Szymanski, P. Kajdanowicz, T. Tuligłowicz, W. (2016). Comprehensive Study on Lexicon-based Ensemble Classification Sentiment Analysis. *Entropy*. 18(4):2-29

Archak, N. Ghose, A.& Ipeirotis, P. G. (2011). Deriving The Pricing Power Of Product Features By Mining Consumer Reviews. *Management Science*. 57: 1485–1509

Baro, E. Degoul, S. Beuscart, R. and Chazard, E. (2015). Toward a Literature-Driven Definition of Big Data in Healthcare. *BiomedResearch International*. 2015:1-9

Beyer, M.A. and Laney, D. (2012). The Importance of “Big Data”: A Definition. Gartner. Stamford, CT.

Blei, D. M. Ng, A. Y.& Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*. 3: 993–1022

Bodell, T. (2012) How Big Data is Becoming a Bigger Deal for the Power Sector. *Economic Inquiry*. 92 (3):9-9

Bongsug, K.C.(2015). Insights from Hashtag #supplychain and *Twitter* Analytics: Considering *Twitter* and *Twitter* Data for Supply Chain Practice and Research. *Int. J. Production Economics*. (165):247–259

Borgi, T. Zoghalmi, N. Abed, M. (2017). Big Data for Transport and Logistics: A Review International Conference on Advanced Systems and Electric Technologies (IC_ASET)

Boyd, D. and Crawford, K. (2012), Critical Questions For Big Data: Provocations For A Cultural, Technological, And Scholarly Phenomenon. *Information, Communication & Society*. 15 (5): 662-679, available at: <http://doi.org/10.1080/1369118X.2012.678878>

Büyükkelik, A. Özoğlu, B. Bülbül, H. (2014). Kargo Hizmet Sağlayıcılarında Kalitenin Tüketici Davranışına Etkisi: Bireysel Tüketici Araştırması. Selçuk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi. 32:33-43

Can, Ü. Alataş, B. (2017). Review of Sentiment Analysis and Opinion Mining Algorithms. International Journal of Pure and Applied Sciences. 3(1):75-111. Retrieved from <http://dergipark.gov.tr/ijpas/issue/29969/304149>

Chang, F. K. Dente, C. Elster, E. (2017). The Impact of a Big Data Decision Support Tool on Military Logistics: Medical Analytics meets the Mission. Defense Acquisition Research Journal: A Publication of the Defense Acquisition University. 24(3):466-487

Chaffin, D. Heidl, R. Hollenbeck, J. R. Howe, M. Yu, A. Voorhees, C. & Calantone, R. (2015). The Promise And Perils Of Wearable Sensors In Organizational Research. Organizational Research Methods. Published online ahead of print. doi: 10.1177/1094428115617004

Chen, H. Chiang, R. And Storey, V. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. MIS Quarterly. 36(4):1165–1188

Chonhenchob, V. Singh, S. P. Stallings, J. Singh, J. J. Grewal, G. (2012). Measurement and Analysis of Vehicle Vibration for Delivering Packages in Small-Sized and Medium-Sized Trucks and Automobiles. Packaging Technology and Science. 25(1):31-38

Çakmak, A. Ç. Özkan, B. (2017). Kargo Kullanıcılarının Önem Verdikleri Faktörlerin, Kargo Firmaları Tarafından Başarım Düzeylerinin İncelenmesi. Tarih Kültür ve Sanat Araştırmaları Dergisi. 6(4):1010-1028

De Mauro, A. Greco, M. Grimaldi, M. (2016). A Formal Definition Of Big Data Based On Its Essential Features. Library Review. 65(3):122-135

Deniz, A. ve Gökmerdan L. (2011). Müşterilerin Kargo Firmalarının Sunduğu Hizmetlere Yönelik Tutum ve Düşünceleri Üzerine Bir Araştırma. Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi. 15(2):379-396

Dijcks, J. (2013). Oracle: Big Data For The Enterprise. Oracle White Paper. Oracle Corporation. Redwood Shores. CA.

Dumbill, E. (2013). Making sense of big data. *Big Data*, 1(1):1-2

Dutta D. Bose, I. (2015). Data Science, Predictive Analytics, and Big Data: A Revolution Managing a Big Data project: The case of RamcoCements Limited, *Int. J. Production Economics*. 165, pp. 293–306.

Crainic, T.G. Gendreau, Dejax, P. (1993) Dynamic and Stochastic Models for the Allocation of Empty Containers. *Operations Research Society of America*. 41(1): 102

Fisher, D. DeLine, R. Czerwinski, M. and Drucker, S. (2012). Interactions with big data Analytics. *Interactions*, available at: <http://dx.doi.org/10.1145/2168931.2168943>

Fujimoto, S. Atushi I. , Takayuki M, Comparison Between Spatial Distributions Of *Tweet* Base And Population In Japan, 2017 IEEE International Conference on Big Data (Big Data)

Gandomi, A. Haider, M. (2015). Beyond TheHype: Big Data Concepts, Methods, And Analytics. *International Journal of Information Management*. 35(2):137-144

George, G. O. Lavie E. C. Scott, B. A. (2016). Big Data and Data Science Methods For Management Research. *Academy of Management Journal*. 59.(5):1493-1507

Gerard, G. Haas, M.R.Pentland, A. (2014). BigData And Management: From The Editors. *Academy of Management Journal*. 57 (2):321–326

Ghat, R.R. Pal, S.S. (2016). Prioritizing the Service Failures of Package Delivery Business: An Analytical Hierarchy Process Analysis. *IUP Journal of Supply Chain Management*. 13:27-52

George, E. I., & McCulloch, R. E. (1993). Variable selection via Gibbs sampling. *Journal of the American Statistical Association*, 88: 881–889.

Gupta, V. Hewett, R. (2017) Harnessing The Power Of Hashtags In *Tweet* Analytics. IEEE International Conference on Big Data (Big Data)

Habin, L. Aydin, N. Choic, Y. Lekhavat, S. Iranid, Z. (2016). A Decision Support System for Vessel Speed Decision in Maritime Logistics Using Weather Archive Big Data. *Computers and Operations Research*. 1-13

Hafiz, N. Alam F. (2016). Applying Servqual Model to Measure Online Customer Satisfaction in Package Delivery Services. 1st International Research Conference on Economics Business and Social Sciences

Halevi G. And Moed H.F. (2012). The Evolution of Big Data as a Research and Scientific Topic: Overview of the Literature. *ResearchTrends*, <http://www.researchtrends.com>

Intel, (2012). Big Data Analytics. Intel's IT Manager Survey on How Organizations Are Using Big Data

Hastie, T., Tibshirani, R., & Friedman, J. 2009. *The Elements Of Statistical Learning: Data Mining, Inference And Prediction* (2nd ed.). Berlin, Germany: Springer.

help.twitter.com/en/rules-and-policies/twitter-api

<https://developer.twitter.com/en/docs/tweets/search/api-reference/get-search-tweets>

<https://developer.twitter.com/en/docs/tweets/search/overview/standard>

Intel IT Center (2012). Big Data Analytics: Intel's IT Manager Survey on How Organizations Are Using Big Data. Intel IT Center, Intel Corporation, Santa Clara, CA.

Işıklı, Ş. (2014) Büyük Veri, Epistemoloji ve Etik Tartışmalar. *Online Academic Journal of Information Technology*. 5(17):89-122

Weiß, F. Jeske, M. Grüner, M. (2013) Big Data in Logistics a Dhl Perspective on How to Move Beyond the Hype. Technical report, DHL Customer Solutions And Innovation.

http://www.dhl.com/content/dam/downloads/g0/about_us/innovation/CSI_Study_e_BIG_DATA.pdf

- Kabir, A.I. Karim R. Newaz S. Hossain M.I. The Power of Social Media Analytics: Text Analytics Based on Sentiment Analysis and Word Clouds on R. *Informatica Economică*. 22(1):25-38
- Karlskind, M. (2014). How 5 Companies Use Mobile Apps to Manage Big Data. *Material Handling&Logistics*. 69:27-31.
- Kaur ,H. Singh, S.P. (2017). Heuristic Modeling for Sustainable Procurement and Logistics in a Supply Chain Using Big Data. *Computers and Operations Research*.
- Kortick, S. A. O'Brien, R. M. (2008). The World Series of Quality Control: A Case Study in the Package Delivery Industry. *Journal of Organizational Behavior Management*. Vol: 16:77-93
- Lambrecht, A.,& Tucker, C. 2013. When Does Retargeting Work? Information Specificity In Online Advertising. *Journal of Marketing Research*, 50:561–576.
- LeCun, Y., Bengio, Y., & Hinton, G. 2015. Deep learning. *Nature*, 521: 436–444.
- Lin, J., Zhou, W., & Du, L. (2017). Is On-demand Same Day Package Delivery Service Green? *Transportation Research Part D: Transport and Environment*. DOI: 10.1016/j.trd.2017.06.016
- Liu B. (2012) Sentiment analysis and opinion mining in Synthesis Lectures on Human Language Technologies. Morgan & Claypool Publisher
- Loughran, T.,& McDonald, B. 2011. When Is A Liability Not A Liability? Textual Analysis, Dictionaries, And 10-Ks. *The Journal Of Finance*. 66: 35–65
- Mahmood, T. Iqbal, T. Amin, F. Lohanna, W. Mustafa, (2013). Mining *Twitter* Big Data To Predict 2013 Pakistan Election Winner. *INMIC, Lahore*, 49-54
- Mehmood, R. Meriton, R. Graham, G. Hennelly, P. & Kumar, M. (2017). Exploring The Influence Of Big Data On City Transport Operations: A Markovian Approach. *International Journal of Operations and Production Management*. 37(1):75-104. <https://doi.org/10.1108/IJOPM-03-2015-0179>

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C. And Byers, A. H. (2011). *Bigdata: The Next Frontier For Innovation, Competition, and Productivity*. McKinsey Global Institute, McKinsey Global Institute

Mayer-Schönberger, V. and Cukier, K. (2013). *Big Data: A Revolution That Will Transform How We Live, Work and Think*, John Murray, London.

Manning, C. D., Raghavan, P., & Schütze, H. (2009). *Introduction to information retrieval*. Cambridge, England: Cambridge University Press.

Microsoft Research (2013). *The big bang: how the big data explosion is changing the world* available at: www.microsoft.com/enterprise/it-trends/big-data/articles/the-big-bang-how-the-big-data-explosion-is-changing-the-world.aspx#fbid!mH4SjMUrH9Y

Mirović, M. Miličević, M. Obradović, I. (2018). *Big Data in the Maritime Industry*. *Nase More*. 65 (1):56-62.

NBD-PWG NIST (2014). *NIST big data public working group. Draft of Big Data Definition*.

Ozyer, Gulsah & Ozyer, Baris. (2015). *Sentiment Analysis for Turkish Twitter Feeds*. 2015 23rd Signal Processing and Communications Applications Conference, SIU 2015 - Proceedings. 10.1109/SIU.2015.7130362.

Pacciarelli, D. D'Ariano, A. Scotto, M. (2011). *Applying RFID in Warehouse Operations of an Italian Courier Express Company*. *NETNOMICS: Economic Research and Electronic Networking*. 12(3):209–222.

Pak, Alexander & Paroubek, Patrick. (2010). *Twitter as a Corpus for Sentiment Analysis and Opinion Mining*. *Proceedings of LREC*. 10.

Prajapati, V. (2013). *Big data analytics with R and Hadoop*. Birmingham, England: Packt Publishing.

Rathore, M. Ahmad, A. Paul, A. Hong, W. Seo, H. (2017) *Advanced Computing Model for Geosocial Media Using Big Data Analytics*. *Multimedia Tools & Applications*. 76(3):24767-24787.

- Romero, O. Herrero, V. Abelló, A. Ferrarons, J. (2015). Tuning Small Analytics on Big Data: Data Partitioning and Secondary Indexes in the Hadoop Ecosystem. *Information Systems*. 54:336–356
- Schroeck M, R. Shockley, J. Smart, D. Romero-Morales, and P. Tufano, (2012). Analytics: The Real-World Use of Big Data. IBM report: 1–20.
- Shen, W.S., Khoong, C.M., (1995). A DSS for Empty Container Distribution Planning. *Decision Support System*. Vol.15(1):75-82
- Shneiderman, B. (2008). Extreme visualization: squeezing a billion records into a million pixels. Proceedings of the 2008 ACM SIGMOD International Conference on Management of data. 3-12, available at: <http://doi.org/10.1145/1376616.1376618>
- Simonsohn, U., Simmons, J. P., & Nelson, L. D. (2015). Specification curve: Descriptive and inferential statistics on all reasonable specifications. Available online at SSRN. doi: 10.2139/ssrn.2694998
- Sismeiro, C., & Bucklin, R. E. (2004). Modeling purchase behavior at an e-commerce web site: A task-completion approach. *Journal of Marketing Research*, 41: 306–323.
- Sotiriadis, M.D. Zyl, C.V. (2013). Electronic Word-of-Mouth and Online Reviews in Tourism Services: the Use of *Twitter* by Tourists, *Electronic Commerce Research*. 13:103–124
- Stout, D. E. (2013). Pack-and-Go Delivery Service: A Multi-Component Cost-Volume-Profit (CVP) Learning Resource. *Accounting Education*. 23(1): 75-94
- Suthaharan, S. 2013. Big Data Classification: Problems and Challenges in Network Intrusion Prediction with Machine Learning. *ACM SIGMETRICS Performance Evaluation Review*. 41(4):70-73.
- Tan, K.H. Zhan, Y.Z. Ji, G. Ye, F. Chang, C. (2015). Harvesting Big Data to Enhance Supply Chain Innovation Capabilities: An Analytic Infrastructure Based on Deduction Graph. *Int. J. Production Economics*. (165):223–233.

- Teixeira, T., Wedel, M., & Pieters, R. 2012. Emotion induced engagement in internet video advertisements. *Journal of Marketing Research*. 49:144–159.
- Ularu, E.G. Puican, F. C. Apostu, A. Velicanu, M. (2012). Perspectives on Big Data and Big Data Analytics. *Database Systems Journal*. 3(4):3-14
- Visb, F. Voss, A. Proctera, R. (2013) Reading the Riots on Twitter: Methodological Innovation for the Analysis of Big Data, *International Journal of Social Research Methodology*. 16(3):197–214
- Waller, M.A. ve Fawket, S.E. (2013). Data Science, Predictive Analytics, and Big Data: A Revolution That Will Transform Supply Chain Design and Management. *Journal of Business Logistics*. 34(2):77-84
- Wang, G. Gunasekaran, A. Ngai, E.W.T. Papadopoulos, T. (2016). Review: Big Data Analytics in Logistics and Supply Chain Management: Certain Investigations for Research and Applications. In *International Journal of Production Economics*. 176:98-110
- Ward, J.S. and Barker, A. (2013). Undefined by data: a survey of big data definitions. available at: [arXiv:1309.5821\[cs.DB\]](https://arxiv.org/abs/1309.5821)
- Wei, P. J., Lin, K.C. (2018). Unstructured Big Data Analytics for Retrieving e-commerce Logistics Knowledge. *Telematics and Informatics*. 35(1):237-244.
- White, P. Breckenridge, R.S. (2014). Trade-Offs, Limitations, and Promises of Big Data in Social Science Research. *Review of Policy Research*. 31: 331-337
- Wang, C., Chen, M. H., Schifano, E., Wu, J., & Yan, J. 2015. *Statistical Methods And Computing For Big Data*. arXiv. org [Website]. Ithaca, NY: Cornell University. E-print available at <http://arxiv.org/pdf/1502.07989v2.pdf>.
- Wong, H.T. Chiang, V.C. Choi K.S. Loke A.Y. (2016) The Need for a Definition of Big Data for Nursing Science: A Case Study of Disaster Preparedness. *International Journal Of Environmental Research And Public Health*. 13 (10):1-13

Tirunillai, S., & Tellis, G. J. 2012. Does chatter really matter? Dynamics Of User-Generated Content And Stock Performance. *Marketing Science*. 31: 198–215.

Yeruva, VK, Junaid, S. Lee, Y. (2017) Exploring Social Contextual Influences on Healthy Eating Using Big Data Analytics. *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*

Varian, H. R. (2014). Big Data: New Tricks For Econometrics. *The Journal of Economic Perspectives*. 28: 3–27.

Zhong, R.Y. Huang, G.Q. Lan, S. Dai, Q. Y. Xud, C. Zhang, T. (2015). A Big Data Approach for Logistics Trajectory Discovery from RFID-Enabled Production Data. *Int. J. Production Economics*. (165):260–272.

Zhong, R.Y. Lan, S. Xu, C. Dai, Q. Huang, G. Q. (2016). Visualization of RFID-enabled shop floor logistics Big Data in Cloud Manufacturing. *International Journal of Advanced Manufacturing Technology*. 84(1-4):5-16.

Ziora, L. (2016). The Sentiment Analysis As A Tool Of Business Analytics In Contemporary Organizations. *Studia Ekonomiczne*. 281:234-241