

T.C.
BAHÇEŞEHİR ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
ENDÜSTRİ MÜHENDİSLİĞİ ANA BİLİM DALI

M.F. ÖZKAYA

E-TİCARET MÜŞTERİ ANALİTİĞİ: MÜŞTERİ YAŞAM BOYU
DEĞERİ YÖNETİMİ VE OPTİMİZASYONU

YÜKSEK LİSANS TEZİ
MUSTAFA FURKAN ÖZKAYA

BAÜ 2025

İSTANBUL 2025

T.C.
BAHÇEŞEHİR ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
ENDÜSTRİ MÜHENDİSLİĞİ ANA BİLİM DALI

E-TİCARET MÜŞTERİ ANALİTİĞİ: MÜŞTERİ YAŞAM BOYU
DEĞERİ YÖNETİMİ VE OPTİMİZASYONU

YÜKSEK LİSANS TEZİ
MUSTAFA FURKAN ÖZKAYA

TEZ DANIŞMANI
DOÇ. DR. OKAN YAŞAR

İSTANBUL 2025



BAHÇEŞEHİR ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

10/01/2025

YÜKSEK LİSANS TEZ ONAY FORMU

Program Adı:	MÜHENDİSLİK YÖNETİMİ (TÜRKÇE, TEZLİ)
Öğrencinin Adı Soyadı:	MUSTAFA FURKAN ÖZKAYA
Tezin Adı:	E-TİCARET MÜŞTERİ ANALİTİĞİ: MÜŞTERİ YAŞAM BOYU DEĞERİ YÖNETİMİ VE OPTİMİZASYONU
Tez Savunma Tarihi:	OCAK 2025

Bu tezin Yüksek Lisans tezi olarak gerekli şartları yerine getirmiş olduğu Lisansüstü Eğitim Enstitüsü tarafından onaylanmıştır.

Doç. Dr. Yücel Batu SALMAN

Enstitü Müdürü

Bu Tez tarafımızca okunmuş, nitelik ve içerik açısından bir Yüksek Lisans tezi olarak yeterli görülmüş ve kabul edilmiştir.

	Ünvanı, Adı Soyadı	Kurumu	İmza
Tez Danışmanı:	DOÇ.DR. OKAN YAŞAR	İSTİNYE ÜNİVERSİTESİ	
2. Üye (Kurum İçi):	DR. TURHAN KARAKAYA	DOĞUŞ ÜNİVERSİTESİ	
3. Üye (Kurum Dışı):	DOÇ.DR. MEHMET SAYGILI	BAHÇEŞEHİR ÜNİVERSİTESİ	



Bu tezdeki tüm bilgilerin akademik kurallara ve etik ilkelere uygun olarak elde edildiğini ve sunulduğunu; ayrıca bu kuralların ve ilkelerin gerektirdiği şekilde, bu çalışmadan kaynaklanmayan bütün atıfları yaptığımı beyan ederim.

Ad, Soyad : MUSTAFA FURKAN ÖZKAYA

İmza :

TEŐEKKÜR

Bu alıőmanın her aőamasında bize yardımcı olan deęerli danıőman hocamız Do. Dr. Öğr. Üyesi Okan Yaőar Hocama hayatımızın her aőamasında bize destek olan deęerli ailelerimize ve őirketteki yöneticilerimize sonsuz őükranlarımı sunarım.

Mustafa Furkan Özkaya



ÖZET

E-TİCARET MÜŞTERİ ANALİTİĞİ: MÜŞTERİ YAŞAM BOYU DEĞERİ YÖNETİMİ VE OPTİMİZASYONU

Özkaya Mustafa Furkan

Mühendislik Yönetimi Yüksek Lisans Programı

Tez Danışmanı: Doç.Dr. Okan Yaşar

Ocak 2025, 45 sayfa

Bu çalışma, e-ticaret sektöründe müşteri davranışını anlamak ve stratejik karar alma süreçlerini iyileştirmek için CLTV (Customer Lifetime Value) ve RFM (Recency, Frequency, Monetary) analiz yöntemlerini kapsamlı şekilde incelemektedir. Araştırmada Kaggle'dan alınan 525,461 gözlem içeren "Online Retail II" veri seti kullanılmıştır. Müşterilerin satın alma geçmişi, işlem sıklığı ve harcama tutarlarına ilişkin veriler, Python programlama dili ve Scikit-learn kütüphanesi kullanılarak işlenmiştir. RFM Analizi, müşterileri yenilik (Recency), sıklık (Frequency) ve parasal değer (Monetary) kriterlerine göre segmente etmiştir. Champions ve Potential Loyalists gibi yüksek değerli gruplar için özel kampanyalar, sadakat programları ve özelleştirilmiş indirimler önerilmiştir. New Customers grubuna ise alışkanlık geliştirme ve marka bağlılığını artırmaya yönelik stratejiler sunulmuştur. CLTV Analizi, müşterilerin yaşam boyu değerini tahmin ederek stratejik önceliklerin belirlenmesini sağlamıştır. Yüksek CLTV'ye sahip müşteriler için gelir artırıcı stratejiler geliştirilirken, düşük CLTV'ye sahip müşteriler için churn (müşteri kaybı) oranlarını azaltmayı hedefleyen sadakat programları önerilmiştir. Sonuç olarak, RFM ve CLTV yöntemlerinin birleştirilmesi, işletmelerin hem kısa vadeli pazarlama planlarını hem de uzun vadeli müşteri ilişkileri stratejilerini geliştirmesine olanak tanımaktadır. Bu analizler, işletmelere müşteri memnuniyeti, pazarlama etkinliği ve rekabet avantajı sağlama konusunda önemli bir katkı sunmaktadır.

Anahtar Kelimeler: Müşteri Yaşam Döngüsü, CLTV, RFM

ABSTRACT

E-COMMERCE CUSTOMER ANALYTICS: CUSTOMER LIFETIME VALUE MANAGEMENT AND OPTIMIZATION

Ozkaya Mustafa Furkan

Master's Program in Engineering Management

Supervisor: Doç. Dr. Okan Yaşar

January 2025, 45 pages

This study comprehensively examines CLTV (Customer Lifetime Value) and RFM (Recency, Frequency, Monetary) analysis methods to better understand customer behavior and improve strategic decision-making processes in the e-commerce sector. The research utilizes the "Online Retail II" dataset from Kaggle, containing 525,461 observations. Customer purchase history, transaction frequency, and spending amounts were processed using Python programming language and the Scikit-learn library. RFM analysis segmented customers based on three criteria: recency, frequency, and monetary value. For high-value groups such as Champions and Potential Loyalists, tailored campaigns, loyalty programs, and personalized discounts were proposed. For the New Customers segment, strategies focusing on habit-building and increasing brand loyalty were suggested. CLTV analysis estimated customers' lifetime value to prioritize strategic decisions. Revenue-boosting strategies were developed for customers with high CLTV, while loyalty programs aimed at reducing churn rates were recommended for those with low CLTV. In conclusion, combining RFM and CLTV methods allows businesses to enhance both short-term marketing plans and long-term customer relationship strategies. These analyses provide significant contributions to businesses in improving customer satisfaction, increasing marketing efficiency, and gaining competitive advantages.

Key Words: Customer Life Time Value, CLTV, RFM

İÇİNDEKİLER

ETİK BEYAN.....	iii
TEŞEKKÜR.....	iv
ÖZET.....	v
ABSTRACT.....	vi
İÇİNDEKİLER	vii
ŞEKİLLER LİSTESİ	viii
KISALTMALAR LİSTESİ.....	ix
Bölüm 1 Giriş.....	1
Bölüm 2 Kuramsal Çerçeve	4
2.1 Dijitalleşme	5
2.2 Büyük Veri	5
2.3 Büyük Verinin Yaşam Döngüsü.....	6
2.3.1 Verinin yaşam döngüsü (data life cycle).	7
2.4 Müşteri İlişkileri Yönetimi.....	8
2.5 Müşteri Yaşam Döngüsü	9
2.6 KPI (Key Performance Indicators)	10
2.7 RFM ile Müşteri Segmentasyonu.....	12
2.8 CLTV ile Müşteri Segmentasyonu.....	14
Bölüm 3 Yöntem ve Veri Analizi	16
3.1 Araştırmanın Amacı	16
3.2 Çalışmanın Önemi.....	17
3.3 Araştırmada Kullanılan Yöntemler	19
3.4 Veri Toplama ve İşleme	19
3.5 Araştırmada Kullanılan Metodlar.....	20
3.5.1 RFM analizi.	20
3.5.2 CLTV analizi.	21
3.5.3 Analiz süreci adımları.....	22
3.6 RFM ile Müşteri Segmentasyonu Analiz Süreci	22
3.6.1 Veriyi anlama(Data understanding).....	22
3.6.2 Veri hazırlama(Data preparation).	25
3.6.3 RFM metriklerinin hesaplanması.....	27
3.6.4 RFM skorlarının hesaplanması.	28
3.6.5 RFM segmentlerinin oluşturulması.	29
3.6.6 Tüm sürecin fonksiyonlaştırılması.....	32
3.7 CLTV Analiz Süreci.....	33
Bölüm 4 Sonuç, Tartışma ve Öneriler.....	39
4.1 Değerlendirme ve Tartışma	39
4.2 Araştırma Kısıtları	39
4.3 Öneriler.....	41
KAYNAKÇA	46

ŞEKİLLER LİSTESİ

ŞEKİLLER

Şekil 1 Verinin yaşam döngüsü (YBS Ansiklopedi, 2015).	7
Şekil 2 Müşteri segmentasyonu örneği (Bayram, 2020).	13
Şekil 3 Kütüphanelerin import işlemi.	23
Şekil 4 Boş değerlerin kontrolü.	23
Şekil 5 Adet sayısı ve toplam kontrolü.	24
Şekil 6 Yeni değişken tanımlama.	25
Şekil 7 Boş değerlerin silinmesi.	26
Şekil 8 Geri iadelerin silinmesi.	26
Şekil 9 Recency değerinin hesaplanması.	27
Şekil 10 Recency, frequency, monetary değerleri.	28
Şekil 11 Rfm score oluşturulması.	29
Şekil 12 Regex yapısı.	30
Şekil 13 Segment analizi örneği.	30
Şekil 14 New_df veri seti oluşturma.	31
Şekil 15 Müşteri kayıtları csv dosyası.	31
Şekil 16 Oluşturulan csv dosyası.	32
Şekil 17 Tüm sürecin fonksiyonlaştırılması.	32
Şekil 18 İmport ve düzenleme.	33
Şekil 19 Cltv_c veri seti.	34
Şekil 20 Ortalama sipariş değeri.	34
Şekil 21 Satın alma sıklığı.	35
Şekil 22 Müşteri terk oranı.	35
Şekil 23 Kar marjı hesabı.	35
Şekil 24 Müşteri değeri hesabı.	36
Şekil 25 Cltv kolonu hesaplama.	36
Şekil 26 Segmentlerin analizi.	37
Şekil 27 Tüm Sürecinin Fonksiyonlaştırılması.	38

KISALTMALAR LİSTESİ

CLTV

Customer Lifetime Value(Müşteri Yaşam Boyu Deęeri)

RFM

Recency(Yenilik), Frequency(Sıklık), Monetary(Para. Deę.)



Bölüm 1

Giriş

Dijitalleşmenin hızlandığı günümüzde, işletmeler için e-ticaret dünya çapında önemli bir rekabet alanı haline gelmiştir. İşletmeler, tüketici davranışlarındaki hızlı değişiklikler, artan rekabet ve müşteri sadakati nedeniyle veri odaklı stratejiler kullanmaya başlamıştır. Bu durumda, hem kısa vadeli hedeflere ulaşmak hem de uzun vadeli rekabet avantajı sağlamak için müşteri davranışlarının kapsamlı bir şekilde analiz edilmesi çok önemlidir (Kotler ve Keller, 2016).

Müşteri Yaşam Boyu Değeri (CLTV), bir müşterinin bir şirketin ömrü boyunca sağlayabileceği toplam değeri tahmin etmek için kullanılan bir araçtır. Müşteri sadakatini artırma, gelir akışlarını optimize etme ve müşteri kaybını önleme gibi stratejik hedefleri gerçekleştirmek için bu analiz işletmelere rehberlik eder (Gupta ve Lehmann, 2005). RFM analizi, müşterileri yenilik (recency), sıklık (frequency) ve parasal (monetary) değerlere göre kategorize ederek kısa vadeli pazarlama stratejileri oluşturmak için etkili bir araç olarak kullanılır. (Stone ve Woodcock, 2014).

E-ticaret sektöründe başarılı olmanın tek yolu, müşterilere ulaşmak ve onları elde tutmak ve sadık müşterilere dönüştürmektir. Bununla birlikte, bu hedefe ulaşmak için müşteri verilerinin doğru bir şekilde işlenmesini ve analiz edilmesini sağlamak gerekir. Bu noktada, CLTV ve RFM analizleri işletmelere çok yardımcı olur. CLTV analizi, müşterilerin uzun vadeli katkılarını değerlendirerek stratejik planlamalar için daha geniş bir bakış açısı sağlarken, RFM analizi müşterilerin alışveriş alışkanlıklarını anlamak için basit ve uygulanabilir bir araç sağlamaktadır. (Reinartz ve Kumar, 2003).

Günümüzün hızla değişen ve dijitalleşen ticaret ortamında, işletmelerin sürdürülebilir büyüme hedeflerini gerçekleştirebilmeleri için müşteri sadakatini sağlamak ve yüksek değerli müşterileri çekmek çok önemlidir. E-ticaret sektöründe artan rekabet, tüketici davranışlarındaki hızlı değişiklikler ve yeni oyuncuların pazara girmesi, şirketlerin müşteri ilişkileri yönetimi stratejilerini daha veri odaklı bir şekilde yeniden tasarlamaları gerektirmektedir (Kotler ve Keller, 2016). İşletmeler, müşteri sadakatine yönelik politikalarını geliştirmelidir çünkü mevcut müşterileri tutmak, yeni müşteriler edinmekten daha ucuzdur (Gupta ve Zeithaml, 2006).

Bu durumda, müşterilerin işletmeye sağlayacağı toplam değeri tahmin etmek için önemli bir ölçüm, Müşteri Yaşam Boyu Değeri (CLTV) olarak bilinir. CLTV, bir müşterinin mevcut katkısını ve gelecekteki potansiyel değerini göstererek işletmelere uzun vadeli planlar yapmaya yardımcı olur (Gupta ve Lehmann, 2005). Müşteri güvenini artırmak ve müşteri ilişkilerini sürdürülebilir bir şekilde yönetmek için doğru bir CLTV analizi gereklidir. Bu analiz, özellikle yüksek değerli müşterilere odaklanma, pazarlama bütçesinin daha etkili kullanımı ve müşteri kaybını önleme gibi stratejik avantajlar sağlar (Reinartz ve Kumar, 2003).

Bununla birlikte, müşteri segmentasyonu için kullanılan RFM analizi, müşterileri parasal (recency), sıklık (frequency) ve yenilik (monetary) değerlerine göre kategorize eder. İşletmeler, bu yöntemle müşteri tabanını daha iyi anlamak ve kaynaklarını uygun alanlara yönlendirmek için daha fazla fırsat bulur (Stone ve Woodcock, 2014). Örneğin, RFM analizi, yüksek değere sahip müşteri gruplarını belirlemek ve bu gruplara özelleştirilmiş pazarlama kampanyaları oluşturmak için kullanılabilir. Müşteri sadakat programları ve kısa vadeli pazarlama etkinlikleri planlamak için etkili bir araçtır (McDonald ve Dunbar, 2012).

Ancak bu analizler e-ticaret sektöründe müşteri ilişkileri yönetimi için yeterli değildir. İşletmeler, müşteri davranışlarını etkileyen çok sayıda faktöre dikkat etmek zorundadır. İşletmeler, rekabetin baskısı, tüketici taleplerindeki değişkenlik ve pazarlama stratejilerindeki eksiklikler nedeniyle uzun vadede başarılı olamazlar. Bu noktada CLTV ve RFM analizlerini birleştirmek, şirketlerin hem kısa vadeli operasyonel hedeflerini hem de uzun vadeli stratejik planlarını daha etkili bir şekilde yönetmelerine olanak tanır (Petersen ve diğ., 2009).

Bu çalışmanın amacı, CLTV ve RFM analizlerinin birlikte kullanılmasının işletmelerin hem kısa vadeli hem de uzun vadeli planlarında değer yaratmasına nasıl yardımcı olduğunu göstermektir. Çalışma, e-ticaret sektörüne yönelik Kaggle platformunda bulunan bir veri seti kullanarak Python programlama diliyle analizler yapmıştır (McKinney, 2012). RFM analizi kullanılarak müşteriler segmente edilmiştir ve CLTV hesaplaması kullanılarak uzun vadeli müşteri değerleri tahminleme işlemi yapılmıştır. Veri odaklı stratejiler, girişimci işletmelerin sürdürülebilir büyüme ve

müşteri ilişkilerinde başarı elde etmeleri için hayati önem taşımaktadır. CLTV ve RFM analizlerinin birbirini nasıl tamamladığı, daha iyi nasıl kullanıldığı ve işletmelere nasıl fayda sağladığı hakkındadır. E-ticaret sektöründe faaliyet gösteren işletmelere müşteri ilişkileri yönetimi ve stratejik karar verme süreçlerinde rehberlik etmektir.

Sonuç olarak, CLTV ve RFM analizleri, müşteri yaşam döngüsü boyunca elde edilecek değeri artırmak ve müşteri segmentasyonunu daha etkili hale getirmek için doğru bir şekilde uygulanmalıdır. Bu analizler, işletmelere yüksek değerli müşterileri bulma ve daha düşük potansiyele sahip müşterilerin ihtiyaçlarını anlama fırsatı verir. Çalışmanın amacı, CLTV ve RFM analizlerini birleştirerek e-ticaret sektöründe veri odaklı müşteri yönetimi stratejileri geliştirmeye yardımcı olmaktır.



Bölüm 2

Kuramsal Çerçeve

İşletmeler, dijitalleşmenin hızlanmasıyla müşterileriyle olan etkileşimlerini daha iyi anlamak ve optimize etmek için müşteri yaşam döngüsü yönetimine daha fazla odaklanmaktadır. Bir müşterinin işletmeyle olan etkileşiminin tüm aşamaları, müşteri kazanımından sadakat geliştirmeye kadar müşteri yaşam döngüsünü oluşturur (Kotler ve Keller, 2016). Dijitalleşme, müşteri davranışını ölçmek ve buna göre stratejiler geliştirmek için güçlü bir araç sağlar. Müşteri yaşam döngüsünün her aşamasında daha iyi kararlar, özellikle büyük veri analitiği sayesinde alınabilir (Chaffey ve Smith, 2017).

Müşteri ilişkilerini geliştirmek ve sürdürülebilir büyüme planlarını oluşturmak için büyük veri önemlidir. Verinin toplanması, işlenmesi, analizi ve karar alma süreçlerinde kullanılması, büyük veri yaşam döngüsü olarak bilinen sürecin çeşitli aşamalarını içerir (Manyika ve diğ.,2011). Bununla birlikte, büyük veri yaşam döngüsü, verinin hızını, çeşitliliğini ve güvenilirliğini yönetmek gibi sorunları çözmeyi amaçlamaktadır (Gandomi ve Haider, 2015). Dijitalleşme ve büyük veri süreçleri, şirketlere daha iyi müşteri hizmetleri sunmak için önemli fırsatlar sunarken, aynı zamanda bu süreçleri iyi yönetmenin önemini de ortaya çıkarmaktadır(Saha ve Srivastava, 2014).

Verinin yaşam döngüsü veya verilerin üretilmesinden imha edilmesine kadar geçen tüm süreçleri kapsar. Veriyi değerli bir varlık haline dönüştürmek için bu süreçte verinin doğruluğunun sağlanması, etik standartların gözetilmesi ve veriden anlamlı bilgiler çıkarılması çok önemlidir (Laney, 2001). İşletmeler, müşteri odaklı stratejiler geliştirerek rekabet avantajı elde etmektedir (Marr, 2015).

Bu stratejiler, büyük veri analitiği ve dijitalleşmenin bir kombinasyonudur. Bu özet, büyük veri işlemlerinin ve müşteri yaşam döngüsünün dijitalleşme ile nasıl işbirliği yaptığını açıklar ve bu alanlardaki akademik çalışmalardan elde edilen sonuçları vurgular.

2.1 Dijitalleşme

Teknolojideki son hızlı gelişmeler, hem günlük yaşamda hem de iş dünyasında önemli değişiklikler yarattı. Dijitalleşme, özellikle COVID-19 pandemisi sırasında bireylerin ve kuruluşların teknolojiyi hızla benimsemesini sağladı (McKinsey, 2020).

Günümüzde veri, "yeni petrol" olarak kabul edildiği kadar önemli hale geldiğinden, verilerin güvenliği ve veri işleme prosedürlerinin etik ve hukuki sorunları ortaya çıktı (World Economic Forum, 2011).

Kişisel verileri koruma hakkı, hem dünyada hem de Türkiye'de çok sayıda yasal düzenlemeye sahiptir. Örneğin, Avrupa Birliği'nin Genel Veri Koruma Yönetmeliği (GDPR), bireylerin verilerinin nasıl toplandığı, işlendiği ve saklandığı konusunda katı düzenlemeler getirmiştir.

Benzer şekilde, Türkiye'deki 6698 sayılı Kişisel Verilerin Korunması Kanunu (KVKK), veri işleme süreçlerini düzenlemek için önemli bir yasal çerçeve sunmaktadır. Devletler, sağlık, hukuk ve güvenlik bilgilerini yönetmek için biyometrik ve demografik verileri kullanan sistemler geliştiriyor. Bireylerin veri güvenliği konusundaki farkındalığı arttıkça, kurumların sorumlulukları da artıyor (European Data Protection Board, 2020).

2.2 Büyük Veri

Geleneksel tekniklerle işlenemeyecek kadar büyük, hızlı ve çeşitli veri setlerini "büyük veri" olarak adlandırırız. 5V modeli genellikle büyük verileri açıklar (Laney, 2001):

1. Volume (Hacim): bir veri parçasının boyutunu ifade eder. Örneğin, Facebook günlük olarak milyarlarca veri üretir.
2. Velocity (Hız): Veri işlemenin ne kadar hızlı olduğunu gösterir.
3. Variety (Çeşitlilik): farklı kaynaklardan ve farklı biçimlerde gelen veriyi ifade eder.
4. Veracity (Güvenilirlik): Verilerin doğruluğu ve güvenilirliği

Value veya deęer, bir bilginin işlenerek bir anlama sahip olmasını saęlayan yeteneęidir. Bu bileşenlerin birleřtirilmesi, büyük verilerin karmařık yapısına katkıda bulunur. Verinin bu özellikleri, řirketlerin veri analizi yöntemlerini deęiřtirmesine neden oldu (Han ve dię., 2011).

Verinin hacmi veya hacmi, büyük veri denildięinde akla gelen ilk şeydir. Büyük veri analizinde, verilerin kapsadıęı alan en büyük sorunlardan biridir (Chen ve dię., 2014). İkinci boyut hızdır. Günümüzde büyük veri analitięi, özellikle sosyal medya platformları gibi hızlı veri üreten sistemler tarafından gerçekleştirilmektedir. Örneęin, Facebook'ta her gün milyonlarca yeni mesaj oluřturuluyor ve bu verilerin sürekli olarak işlenmesi gerekiyor (Gandomi ve Haider, 2015). Üçüncü boyut ise çeřitlilik. Laney'e (2001) göre, büyük veri, çeřitli kaynaklardan gelen ve farklı içerik ve özelliklere sahip yapılandırılmıř, yarı yapılandırılmıř ve yapılandırılmamıř verileri içerir.

Güvenilirlik veya doęruluk, büyük veri kaynaęının doęruluęu ve güvenilirlięi ile ilgilidir. İşlenen verilerin doęruluęu, büyük veri analizinin başarısını belirler. Bununla birlikte, veri kaynaęına olan güven, bu doęruluęu doęrudan etkiler (Saha ve Srivastava, 2014). Son olarak beřinci boyut deęerdir. Organizasyonların büyük verilerinin güvenilirlięi, hızı, hacmi ve çeřitlilięi doęru bir şekilde deęerlendirildięinde deęer kazanırlar. Bununla birlikte, büyük verinin etkili bir şekilde kullanılması, bu bileşenlerin iyi yönetilmesine baęlıdır (Marr, 2015).

Büyük veri projelerinde, hızlı, çeřitli ve güvenilir bir şekilde devasa miktarda veriyi yönetmek ve bu veriden deęer yaratabilmek en büyük zorluktur. Bu nedenle, büyük veri çalışmalarında deęeri artırma hedefi, dięer dört V'yi çözmeye odaklanır (Manyika ve dię., 2011). Büyük verilerin bu bileşenleri iyi yönetilirse, organizasyonlar verilerinden en iyi şekilde yararlanabilir. Bununla birlikte, büyük verilerle ilgili sorunlar ve zorluklar bu süreçte ortaya çıkar (Katal ve dię., 2013).

2.3 Büyük Verinin Yařam Döngüsü

Büyük veri yařam döngüsü, bir verinin üretilmesinden işlenmesine, analiz edilmesine ve deęer yaratmasına kadar geçen süreçleri kapsar. İşletmelerin stratejik kararları, her veri setinin belirli bir yařam döngüsünden doęrudan etkilenir. (EMC

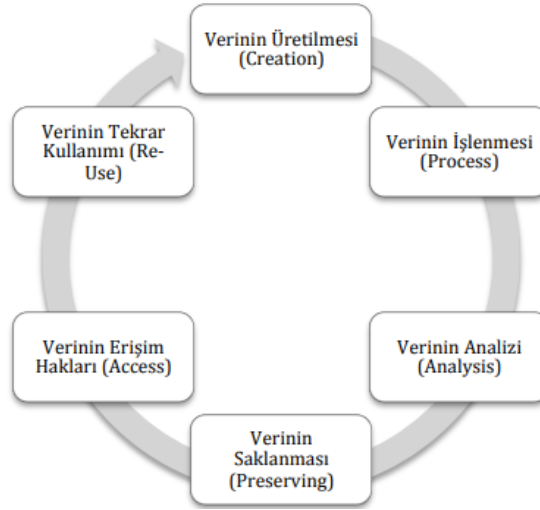
Corporation, 2014).

Örneğin, Facebook'ta toplanan nüfus verileri veya kullanıcı davranışlarından toplanan veriler gibi büyük veri setleri tipik olarak şu aşamalardan geçer:

1. Veri Toplama: Farklı kaynaklardan veri toplanır.
2. Veri İşleme ve Temizleme: Veriler toplandıktan sonra işlenir ve analize hazır hale getirilir.
3. Analitik ve Modelleme: Öngörüler, verilere analitik yöntemler uygulanarak elde edilir.
4. Raporlama ve Karar Alma: Analiz sonuçları karar alma süreçlerinde kullanılır.

IBM, SAS ve Oracle gibi şirketler tarafından geliştirilen modeller ve araçlar, bu süreçte büyük verinin etkili bir şekilde yönetilmesine olanak tanır (IBM, 2021).

2.3.1 Verinin yaşam döngüsü (data life cycle). İlk anlatılacak olan yaklaşım akademik bir yaklaşım ve Essex Üniversitesi tarafından literatüre kazandırılmış (Corti, Eynden, Bishop ve Woollard,2014).



Şekil 1 . Verinin yaşam döngüsü (YBS Ansiklopedi, 2015).

Bu yaklaşıma göre, verinin bir yaşam döngüsü bulunmaktadır. Bu döngü, verinin üretilmesi, işlenmesi, analiz edilmesi ve bu verinin bozulma veya güvenlik açıklarına karşı korunmasını içerir. Ayrıca, bu veriye kimin erişeceği ve verinin tekrar kullanım süreçleri de bu döngüde yer alır. Tüm bu süreçler, verinin yaşam döngüsü olarak

değerlendirilebilir. Kısaca, verinin üretimi aşamasında "Create" adımı, ihtiyaçların belirlenmesi gerekmektedir; veri üretiminin nasıl yönetileceği ve verinin hangi kaynaklardan elde edileceği netleştirilmelidir.

2.4 Müşteri İlişkileri Yönetimi

Bir işletmenin pazarlama, satış, satış sonrası destek ve yeniden satış gibi yaşam döngüsünü yönetmek için müşteri ilişkileri yönetimi (CRM) gereklidir. Bir şirketin en önemli varlığı olmasına rağmen, müşteriler bazen göz ardı edilebilmektedir (Kotler ve Keller, 2016). Müşteri ilişkilerine öncelik veren işletmeler, bu durumda rekabet avantajı kazanarak uzun vadeli müşteri bağlılığı oluşturulabilmektedir (Payne ve Frow, 2005).

Günümüzde işletmeler, henüz müşterileri olmayan gruplara pazarlama yaparak müşteri kazanır. Farkındalık yaratma ve etkileşim kurma aşaması, müşteri yolculuğunun ilk aşamasıdır (Chaffey ve Smith, 2017). Pazarlama stratejileri, potansiyel müşterileri aktif müşterilere dönüştürür ve ardından satış işlemleri başlar. Bu süreçte müşteri memnuniyeti, marka sadakati ve tekrarlanan alışverişler çok önemlidir (Gartner, 2013).

Müşteri elde tutma (müşteri elde tutma), çapraz satış (cross-sell), üst satış (up-sell) ve müşteri segmentasyonu gibi stratejiler, CRM sistemleri tarafından desteklendiğinde işletmeler büyük avantajlar elde eder. Üst satış, müşterilere daha yüksek değerli ürünler sunar, ancak çapraz satış, mevcut müşterilere tamamlayıcılar sunar (Sharma ve Sheth, 2010). Ek olarak, müşteri segmentasyonu, çeşitli müşteri gruplarını belirli ölçütlere göre ayrıştırarak stratejilerin daha odaklı ve verimli olmasını sağlar (Blattberg ve diğ., 2008).

CRM çalışmalarının bir diğer önemli hedefi, müşteri terk oranlarının azaltılmasıdır. Şirketlerin mevcut müşterilerini elde tutmak, müşteri edinme maliyetlerinin yüksek olduğu göz önüne alındığında stratejik olarak çok önemlidir (Reinartz ve diğ., 2005). Müşteri geri bildirimlerinin toplanması ve analiz edilmesi, pazarlama stratejilerinin optimize edilmesine, ürün ve hizmetlerin geliştirilmesine ve müşteri memnuniyetinin artırılmasına yardımcı olmaktadır (Homburg ve diğ., 2010).

Sonuç olarak, CRM, müşteri ile iletişim kurmaktan satış prosedürlerini kolaylaştırmaya ve müşteri segmentasyonundan çapraz satış stratejilerine kadar çok çeşitli uygulamalara sahiptir. Müşteri ilişkileri yönetimi stratejileri, müşteri yaşam döngüsü boyunca sürekli bir değer oluşturmak için etkili iletişim kanallarını kullanmaktadır (Rust ve diğ., 2000).

2.5 Müşteri Yaşam Döngüsü

Bir müşterinin bir işletme ile olan ilişkisinin başlangıcından sonuna kadar geçen süreci "müşteri yaşam döngüsü" olarak adlandırılır ve bu süreçte geçtiği çeşitli aşamaları kapsar. Bu yaşam döngüsü tipik olarak beş temel aşamadan oluşur:

1. Farkındalık: Bu aşamada müşteri işletme, ürün veya hizmet hakkında bilgi sahibi olur.
2. Dikkat: Müşterinin hizmet veya ürünle ilgilenmeye başladığı aşamadır.
3. Karar: Müşteri satın almayı seçtiği aşamadır.
4. Satın Alma: Müşteri bir mal veya hizmet satın alır.
5. Sadakat: Müşteri üründen memnun kaldıktan sonra tekrar satın alır ve şirketine sadık kalır.

(Kotler ve Keller, 2016).

Müşteri ilişkilerinin değerinin belirlenmesi ve bu sürecin analiz edilmesi, işletmelerin stratejik kararlar alabilmesi için çok önemlidir. Bu durumda, bir müşterinin işletmeye sağladığı toplam kâr, Müşteri Yaşam Boyu Değeri (CLTV) olarak bilinen temel bir ölçüttür. CLTV, müşteri kazanma maliyeti ve müşterinin işletme ile olan ortalama ilişki süresi gibi faktörler kullanılarak hesaplanır. (Gupta ve Lehmann, 2005).

1. Müşteri Başına Ortalama Gelir: Müşterinin belirli bir dönemde yaptığı ortalama harcama miktarı.
2. Müşterinin Ortalama Yaşam Süresi: Müşterinin şirketle olan ilişkisinin ortalama süresi.
3. Müşteri Kazanma Maliyeti: Yeni bir müşteri kazanmak için yapılan pazarlama ve satış harcamaları.

CLTV analizleri, çeşitli stratejik ve operasyonel kararların alınmasında kullanılır. Bu analizlerin bazı kullanım alanları şunlardır:

1. Pazarlama Stratejileri: Yüksek CLTV'ye sahip müşteri gruplarına yatırım yapılmasını sağlar ve bu gruplara yönelik hedefli pazarlama kampanyaları oluşturulmasını sağlar (Reinartz ve Kumar, 2003).
2. Müşteri Sadakati Programları: CLTV, müşteri sadakatini artırmak için en iyi teşvikleri veya programları belirler (Stone ve Woodcock, 2014).
3. Kaynak Dağılımı: Satış, pazarlama ve müşteri hizmetleri gibi departmanlar arasında kaynakların daha verimli bir şekilde dağıtılmasını sağlar (Kotler ve Keller, 2016).
4. Ürün ve Hizmet Geliştirme: Hangi ürünlerin daha fazla kâr getirdiğini ve müşteri memnuniyetini artırdığını belirleyerek yeni ürün geliştirme süreçlerini yönlendirir.
5. Finansal Planlama: Uzun vadeli gelir tahminlerine ve bütçe planlamasına yardımcı olur (Gupta ve Lehmann, 2005).

CLTV, şirketlerin müşteri ilişkilerini daha iyi yönetmesine ve daha fazla para kazanmasına yardımcı olur. Şirket, yüksek CLTV'ye sahip müşterileri daha değerli görür ve bu müşterilere özel stratejiler geliştirir. Ayrıca, CLTV'nin doğru bir şekilde hesaplanması, müşteri sadakati ve memnuniyeti üzerinde olumlu bir etkiye sahiptir.

2.6 KPI (Key Performance Indicators)

Temel performans göstergeleri (KPI), bir şirketin stratejik hedeflerine ulaşma düzeyini ölçmek için kullanılan önemli matematiksel araçlardır. Bu göstergeler, özellikle müşteri ilişkileri yönetimi (CRM) alanında, terk oranları, dönüşüm oranları, müşteri edinme ve elde tutma gibi çeşitli ölçümler içerir. Bu ölçümler, işletmelerin müşteri tabanını nasıl yönlendirdiğini, müşteri sadakatini nasıl artırdığını ve pazarlama taktiklerinin ne kadar etkili olduğunu göstermektedir.

1. Müşteri Kazanma Oranı (Customer Acquisition Rate): Bu KPI, bir işletmenin yeni müşteriler kazanmak için yaptığı harcamaların etkinliğini ölçer. Potansiyel müşteriler için araştırma, reklam ve reklam harcamaları bu maliyete dahildir. Mevcut müşterileri elde tutmak, yeni müşteriler kazanmaktan daha

pahalı olmaktadır (Kotler ve Keller, 2016). Yeni müşteri sayısına bölünen toplam harcamanın müşteri kazanma oranı bulunur. Örneğin, bir şirketin bir reklam kampanyasına 2000 TL harcayarak 40 yeni müşteri kazanması durumunda, müşteri kazanma maliyeti $2000 \text{ TL} / 40 = 50 \text{ TL}$ olur (Reinartz ve diğ., 2005).

2. Müşteri Elde Tutma Oranı: Müşteri elde tutma oranı, işletmelerin mevcut müşterilerini belirli bir süre boyunca elinde tutma yeteneğini gösteren bir ölçümdür. Müşteri elde tutma oranı yüksek işletmeler, müşterileriyle güçlü ilişkiler kurarak uzun vadeli sadakat geliştirmektedir (Payne ve Frow, 2005). Bu oran, belirli bir dönemde tutulan müşterilerin toplam müşteri sayısına bölünmesiyle bulunur. Müşteri elde tutma oranı, işletmelere daha fazla müşteri sadakati ve yeniden satış fırsatı sağlamaktadır (Homburg ve diğ., 2010).
3. Müşteri Terk Oranı: Bu terim, belirli bir süre boyunca bir ürün veya hizmeti terk eden müşterilerin oranını ifade eder. Müşteri terk oranı yüksek olduğunda, şirketler müşteri memnuniyetini ve sadakatini arttırmakta zorlanmaktadır (Sharma ve Sheth, 2010). Bir dönemde terk eden müşterilerin toplam aktif müşteri sayısına bölünmesi, müşteri terk oranını oluşturur. Birçok sektör için %4 ila %7 arasında değişen değerler normaldir (Reinartz ve diğ., 2005).
4. Dönüşüm Oranı: Dönüşüm oranı, e-ticaret için en önemli KPI'lardan biridir. Bu oran, bir şirketin kampanyalarına veya web sitelerine gösterilen ilgiyi, hedeflerine ulaşmak için yapılan eylemlerle ilişkilendirir. Örneğin, bir reklam kampanyası 100.000 kişiye ulaştı ve 1000 kişi e-posta bülteni aboneliği olduysa, dönüşüm oranı $1000/100.000 = \%1$ olur. Bu oran, reklamların ne kadar etkili olduğunu ve kampanyaların hedeflerine ne kadar ulaştığını göstermektedir (Chaffey ve Smith, 2017). Müşteri etkileşimi ve satışlar arttıkça, dönüşüm oranları yükselmektedir.

İşletmelere, bu temel performans göstergeleri, müşteri ilişkilerini daha iyi yönetmelerine ve stratejik kararlar almalarına yardımcı olur. Bir işletmenin müşteri kazanım, tutma, terk oranları ve dönüşüm gibi faktörlerdeki başarısını ölçmek için her bir KPI, sürekli gelişim için gerekli verileri sağlamaktadır (Rust ve diğ., 2000).

2.7 RFM ile Müşteri Segmentasyonu

Müşterilerin satın alma alışkanlıklarını incelemek için RFM (Recency, Frequency, Monetary) analizi etkili bir istatistiksel analiz yöntemidir. Bu analiz, müşterileri üç ana kategoriye ayıran bir skora dayanan bir skorlama modeline dayanır: Yenilik (Recency), Sıklık (Frequency) ve Tutar (Monetary). RFM analizi, müşterilerin her segmentine uygun pazarlama stratejileri oluşturma yeteneğine sahiptir (Farris ve diğerleri, 2010). Bu yöntem, şirketlere müşterileri hakkında daha fazla bilgi edinmelerine ve daha hedeflenmiş pazarlama kampanyaları oluşturmalarına yardımcı olmaktadır (Chen ve diğerleri, 2009).

Yenilik (Recency) değeri bir müşterinin en son alışveriş tarihine göre yenilik değeri ölçülür. Müşterinin son alışverişi, markayı gelecekte nasıl hatırlayacağı üzerinde doğrudan bir etkiye sahiptir. Genel olarak, müşterinin son alışverişini unutmaması daha yüksektir. Yenilik değeri, RFM analizinde analiz günü ile müşterinin son işlem yaptığı gün arasındaki fark olarak bulunur. Bu değer, müşterilerin markaya olan ilgisini ve gelecekteki satın almalarını tahmin etmek için kullanılmaktadır (Hughes, 2011). RFM analizinde Yenilik Değeri = (Bugün) - (Müşterinin Son İşlem Günü). Bu nedenle, daha kısa bir zaman dilimindeki müşteriler daha yüksek bir Yenilik skoru almaktadır (Kumar ve Shah, 2004).

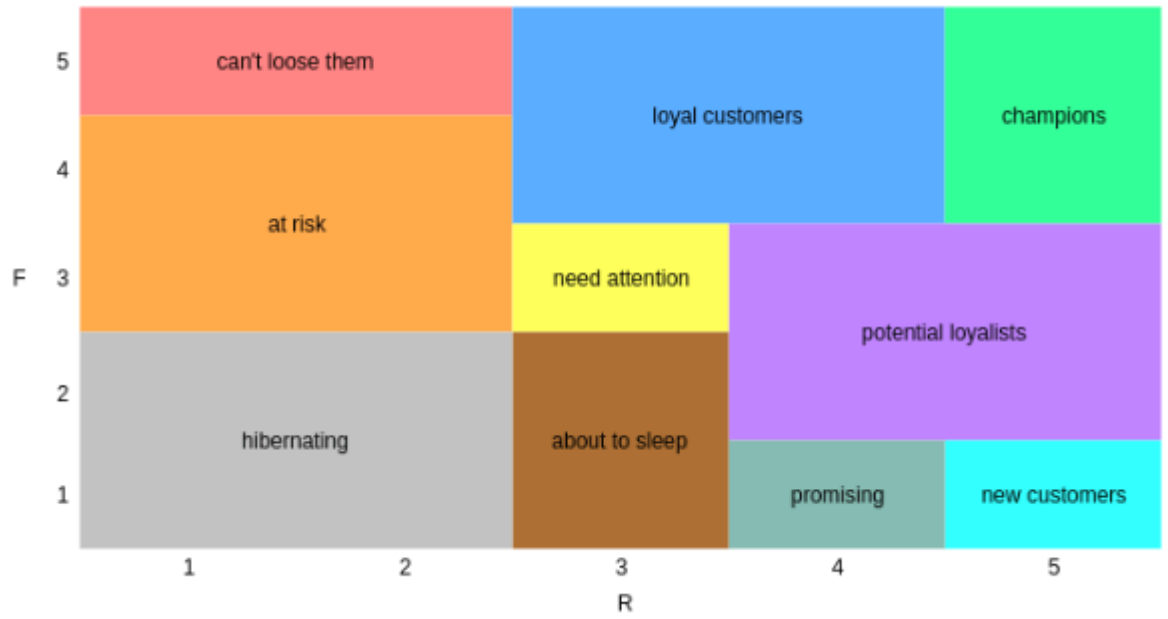
Sıklık (Frequency) değeri Bir müşterinin belirli bir zaman diliminde kaç kez alışveriş yaptığının bir ölçüsü sıklık değeridir. Bir müşterinin alışveriş alışkanlıklarını anlamak için sıklık değeri çok önemlidir. Sık satın alan müşteriler, markaya daha sadık olan ve gelecekteki satışları tahmin etmek için değerli verilerdir (Peppers ve Rogers, 2011). Sıklık değeri, son iki yılda yapılan toplam satın alma işlem sayısı kullanılarak bulunur.

RFM analizinde sıklıkla alışveriş yapan müşterilerin yüksek bir skora sahip olması ve bu müşterilerin genellikle daha değerli olarak görülmesi, Sıklık Değeri = (Toplam Satın Alma Sayısı) olarak hesaplanır (Rust ve diğ., 2000).

Tutar(Monetary) Müşterinin markaya ne kadar değer kattığını belirlemek için, tutar (para) değeri, müşterinin yaptığı toplam harcamanın bir ölçüsüdür. Bir müşterinin

yüksek parasal değeri, bir şirketin gelirinin önemli bir kısmını oluşturduğunu gösterir. Pazarlama stratejileri genellikle yüksek tutar değerine sahip müşterilere odaklanır (Farris ve diğ., 2010). RFM analizinde tutar değeri, her müşterinin yaptığı toplam harcamayı kullanarak bulunabilir.

RFM analizinde, Tutar Değeri = (Toplam Harcama) olarak hesaplanır. Müşterilerin çok fazla para harcadıkları dönemlerde, daha fazla satış ve pazarlama faaliyeti yapılması gerekir, bu da müşterilerin önemini vurgulamaktadır (Shankar ve diğ., 2003).



Şekil 2 . Müşteri segmentasyonu örneği (Bayram, 2020).

RFM Skorlarının Oluşturulması: RFM analiziyle elde edilen yenilik, sıklık ve tutar değerleri genellikle 1 ile 5 arasında bir skorlama ile sıralanır. Alternatif olarak, değerleri harflerle (A, B, C...) veya özel etiketlerle (Platinum, Gümüş vb.) sıralayabilirsiniz. Bu skorlama, müşteri segmentleri oluşturmak ve her segment için farklı pazarlama stratejileri belirlemek için kullanılabilir (Venkatesan ve Kumar, 2004).

Pazarlama Stratejileri Üzerindeki RFM Analizinin Etkisi: RFM analizi, pazarlama ve müşteri ilişkileri yönetimi için çok faydalı bilgiler sağlar. İşletmeler, müşterilerinin alışverişlerine, harcamalarına ve harcamalarına göre daha özelleştirilmiş pazarlama planları oluşturabilir. Bu analiz, şirketlere hangi müşterilerin değerli olduğunu,

hangilerinin daha fazla etkileşim gerektirdiğini ve hangilerinin kaybedilme riski taşıdığını belirlemede yardımcı olmaktadır (Kumar ve Shah, 2004). Ek olarak, bu metriklerin birleştirilmesiyle müşteri segmentasyonu yapılır ve her bir segment için daha iyi pazarlama stratejileri geliştirilebilmektedir.

2.8 CLTV ile Müşteri Segmentasyonu

Bir müşterinin şirket için toplam ekonomik değeri, Müşteri Yaşam Boyu Değeri veya CLTV olarak bilinir. CLTV, işletmelerin kaynaklarını daha etkili bir şekilde yönetmesine ve pazarlama stratejilerini geliştirmesine olanak tanır (Gupta ve Lehmann, 2005). Bir müşterinin şimdiki kârlılığının yanı sıra gelecekteki potansiyel değerinin de önemi vardır. Müşteri segmentasyonu, CLTV'nin strateji geliştirmesine yardımcı olmaktadır (Fader, Hardie ve Lee, 2005).

Müşteri sadakat düzeyi, alışveriş sıklığı ve ortalama sipariş değeri, CLTV'nin hesaplanmasında kullanılır. Firmalar, müşteri değeri analizlerini kullanarak hangi müşterilere daha fazla yatırım yapılması gerektiğine karar vermiştir (Kumar, Shah ve Gupta, 2006).

Bir dizi model, müşterinin yaşam boyu değerini belirlemek için kullanılabilir. BG/NBD ve Pareto/NBD modeller en yaygın olanlarıdır. Bu yöntemler, müşterilerin gelecekte nasıl davranacaklarını tahmin etmeyi amaçlar (Fader, Hardie ve Lee, 2005). CLTV hesaplamaları, özellikle e-ticaret ve dijital dünyada daha karmaşık hale gelmiştir; ancak bu hesaplamalar, şirketler için müşteri odaklı stratejiler geliştirmede kritik öneme sahiptir (Venkatesan ve Kumar, 2004).

Müşterileri benzer özelliklere sahip gruplara ayırma süreci, müşteri segmentasyonu olarak bilinir. Bu gruplama, demografik, davranışsal ve psikografik faktörlere dayalı olarak gerçekleştirilebilir (Chen, Chen ve Lee, 2009). Pazarlama kampanyalarının özelleştirilmesine ve daha etkili olmasına yardımcı olan olduğu görülmektedir. Geleneksel analizden daha karmaşık olan makine öğrenmesi algoritmaları ve veri madenciliği yöntemleri, segmentasyon için çeşitli yöntemler sunar (Katal, Wazid ve Goudar, 2013). Dijital dünya ve büyük veri kullanımı nedeniyle müşteri segmentasyonu süreci daha dinamik hale getirmiştir. Mayer-

Schönberger ve Cukier (2013), büyük verinin çözümlene yeteneđi, müşteri davranışını anlamak için yeni bir yöntem sunduđunu belirtmiştir. Şirketlerin CLTV ve segmentasyonu birleřtirerek daha karlı stratejiler geliřtirmesi mümkündür. Bu yaklaşım, işletmelerin sınırlı kaynaklarını en yüksek deđer yaratan segmentlere ayırmalarına olanak tanımaktadır (Gupta ve Lehmann, 2005).

Örneđin, müşteri sadakati ve tekrar eden alımlara odaklanan dijital pazarlama segmentasyon stratejileri, CLTV ve müşteri memnuniyetinin artmasına yardımcı olmaktadır (Shankar, Smith ve Rangaswamy, 2003). Şirketlerin müşteri odaklı stratejiler oluřturması ve rekabet avantajı elde etmesi için CLTV ve müşteri segmentasyonu çok önemlidir. Literatür, bu iki fikri birleřtirmenin işletmelere daha fazla para kazandırdığını göstermektedir. Gelecekteki çalışmalar, dijitalleşme ve yapay zeka gibi yeni teknolojilerin bu alanlar üzerindeki etkilerini daha kapsamlı bir şekilde analiz edebilir.

Bölüm 3

Yöntem ve Veri Analizi

Bu çalışmada CLTV modeli ve RFM analizi kullanılarak müşteri segmentasyonu ve müşteri yaşam boyu değeri (CLTV) hesaplanmıştır. Çalışmada kullanılan veri seti, e-ticaret platformlarından geçmiş müşteri satın alma verilerini içerir. Müşterilerin alışveriş alışkanlıklarını anlamak ve segmentasyon yapmak için bu veri seti uygun bir yapı sağlar. Veri seti, tüketicilerin son satın alma tarihleri, toplam alışveriş sıklıkları ve harcanan toplam tutarları içerir. Yenilik (Recency), Sıklık (Frequency) ve Tutar (Monetary) değerleri, müşteri analizinde çok önemli olan değerlerdir. RFM analizi, müşterilerin alışveriş alışkanlıklarının daha iyi anlaşılmasını sağlarken, CLTV modeli, her bir müşterinin işletmeye sağlayabileceği potansiyel uzun vadeli finansal değeri tahmin etmeye olanak tanır. Bu yaklaşımlar, şirketlerin müşteri odaklı yaklaşımlar oluşturmasına yardımcıdır.

3.1 Araştırmanın Amacı

Bu çalışmanın temel amacı, e-ticaret sektöründe müşteri ilişkilerini optimize etmek için kullanılacak analitik yaklaşımları incelemektir. Ayrıca, Müşteri Yaşam Boyu Değeri (CLTV) ve RFM Analizi tekniklerinin bu süreçte ne kadar etkili olduğunu gösterecektir. Günümüzün dijitalleşen ticaret ortamında, müşteri segmentasyonuna ve kişiselleştirilmiş stratejilere olan ihtiyaç, artan rekabet ve hızla değişen tüketici davranışları nedeniyle giderek artmaktadır (Kotler ve Keller, 2016). Bu durumda, müşteri davranış kalıplarını anlamak, müşterileri çeşitli segmentlere ayırmak ve her segmentin ihtiyaçlarına uygun stratejiler oluşturmak, müşteri deneyimini iyileştirmek ve şirketlerin finansal performansını artırmak için çok önemlidir (Reinartz ve Kumar, 2003).

Çalışmada kullanılan RFM analizi, müşterilerin satın alma alışkanlıklarını yenilik (recency), sıklık (frequency) ve parasal (monetary) değerlerine göre sınıflandırarak daha etkili ve hızlı bir şekilde müşteri segmentasyonunu gerçekleştirmeyi amaçlamaktadır. Bu yöntem sayesinde şirketler potansiyel müşteri gruplarını belirleyebilir ve bu gruplara özel pazarlama stratejileri oluşturabilir (Stone ve Woodcock, 2014). Örneğin, düşük yenilik değerine sahip müşteriler için geri

kazanım kampanyaları düzenlenebilirken, yüksek frekansa ve parasal değere sahip müşteriler için sadakat programları önerilebilir.

Bununla birlikte, işletmeler, Müşteri Yaşam Boyu Değeri (CLTV) analizi kullanarak müşterilerden elde edebilecekleri toplam değeri uzun vadeli bir bakış açısıyla değerlendirebilirler. CLTV, müşterilerin mevcut katkısını ve gelecekteki potansiyel değerini tahmin ederek işletmelerin stratejik planlamasına yardımcı olur (Gupta ve Lehmann, 2005). Bu analiz, işletmelere özellikle müşteri bağlılığı, müşteri kaybı ve kaynakların daha verimli bir şekilde dağıtılması gibi konularda büyük faydalar sağlar (Petersen ve diğ., 2009).

Çalışmanın bir diğer amacı, CLTV ve RFM analizlerini birleştirerek, bu yöntemlerin e-ticaret şirketlerinin müşteri yönetimi stratejilerine nasıl yardımcı olabileceğini göstermektir. Özellikle yüksek potansiyelli müşteri gruplarına odaklanarak, bu gruplara özelleştirilmiş stratejiler oluşturmak ve müşteri deneyimini iyileştirmek için çaba göstermektedir. Bununla birlikte, gelecekteki satın alma alışkanlıklarını tahmin edecek modeller oluşturmak, şirketlere daha sürdürülebilir büyüme planları sağlamak için kullanılmaktadır (Reinartz ve Kumar, 2003).

Sonuç olarak, bu çalışma, müşteri segmentasyonunu ve değer analizini veri odaklı yaklaşımlarla birleştirerek e-ticaret işletmelerine somut öneriler sunmayı amaçlamaktadır. Çalışmanın temel motivasyonu, CLTV ve RFM analizlerinin operasyonel hedeflere ulaşmada ve uzun vadeli stratejik planlamada nasıl fayda sağlayabileceğini göstermektir.

3.2 Çalışmanın Önemi

Bu çalışmanın temel amacı, e-ticaret sistemlerinde Müşteri Yaşam Boyu Değeri (Customer Lifetime Value - CLTV) ve RFM (Recency, Frequency, Monetary) analiz yöntemlerini incelemek ve bu iki analitik yaklaşımın stratejik kullanım alanlarını, avantajlarını ve farklı yönlerini incelemektir. Günümüz e-ticaret sektöründe, müşteri beklentilerinin hızla değişmesi ve artan rekabet koşulları nedeniyle veri odaklı stratejiler gereklidir (Kotler ve Keller, 2016).

İşletmelerin hem kısa hem de uzun vadeli hedeflerine ulaşmaları için müşteri davranışlarının doğru bir şekilde analiz edilmesi çok önemlidir. Bu bağlamda, CLTV

ve RFM analizleri, pazarlama stratejilerinin optimize edilmesinden müşteri ilişkileri yönetimine kadar çeşitli karar alma süreçlerini destekleyen güçlü araçlar olarak öne çıkmaktadır. Bir müşterinin işletmeye yaşam döngüsü boyunca sağlayabileceği toplam finansal değer tahmini, CLTV olarak bilinen bir yaklaşımla mümkündür. Bu analiz, özellikle uzun vadeli müşteri ilişkilerinin yönetimi, müşteri sadakati oluşturma ve müşteri kaybını azaltma gibi stratejik konularda işletmelere tavsiyelerde bulunur (Gupta ve Lehmann, 2005).

İşletmelere uzun vadeli planlama süreçlerinde benzersiz bir bakış açısı sunan CLTV, bir müşterinin mevcut değerini ve gelecekteki gelir katkısını değerlendiren bir metriktir. Bununla birlikte, RFM Analizi, müşterilerin davranışını üç ana kritere göre kategorize eder: güncellik (güncellik), işlem sıklığı (sıklık) ve harcama miktarı (miktar). Bu yöntem, şirketlere müşteri segmentasyonu için hızlı ve kullanışlı çözümler sağlar. RFM, özellikle kısa vadeli pazarlama planlarını özelleştirmek ve belirli müşteri gruplarına özel kampanyalar oluşturmak için etkili bir araçtır (Stone ve Woodcock, 2014).

Araştırma, CLTV ve RFM analizlerinin e-ticaret verilerine uygulanmasını sağladı. Python programlama dili, veri işleme, modelleme ve görselleştirme alanlarındaki esnekliği ve yüksek performansı nedeniyle bu süreçte tercih edilen programlama dili olarak öne çıkmıştır. Python, geniş kütüphane desteği sayesinde analiz işlemlerini daha verimli bir şekilde gerçekleştirebilir (McKinney, 2012).

Çalışma, RFM analizinin müşterilerin belirli davranış kalıplarına dayalı olarak hızlı bir şekilde segmente edilmesi ve pazarlama çabalarının kişiselleştirilmesi üzerindeki faydalarını ortaya koymayı hedeflemektedir. Ayrıca CLTV, uzun vadeli stratejik kararların alınmasında ne kadar önemli olduğunu ortaya koymayı hedeflemektedir (Reinartz ve Kumar, 2003).

Elde edilen sonuçlar, CLTV ve RFM analizlerinin birbirini tamamlayan niteliklerini göstermektedir. RFM, kısa vadeli pazarlama etkinliklerinde ve müşteri segmentasyonunda öne çıkarken, CLTV, uzun vadeli müşteri ilişkilerinin yönetimi ve gelir tahminleri için kapsamlı bir çerçeve sağlar. Bu iki yöntemin kombinasyonu, işletmelerin hem operasyonel hem de stratejik süreçlerde daha dikkatli kararlar almasına olanak tanır.

Sonuç olarak, bu çalışma, e-ticaret sektöründeki işletmelere, müşteri yaşam döngüsünü optimize etme, müşteri ilişkilerini güçlendirme ve sürdürülebilir büyümeyi destekleyen stratejiler oluşturma konusunda somut bir yol haritası sunmayı amaçlamaktadır. Bu çalışma, e-ticaret sektöründe veri odaklı karar alma süreçlerine yönelik kapsamlı bir rehber niteliğindedir ve ayrıca CLTV ve RFM analizlerinin işletmelerin müşteri odaklı stratejiler geliştirmesine olanak tanıyan tamamlayıcı rollerini göstermektedir. Müşteri ilişkileri yönetimi alanında stratejik avantajlar elde etmek isteyen işletmeler için bu bağlamda önemli bir kaynak olmayı hedeflemektedir.

3.3 Araştırmada Kullanılan Yöntemler

Python programlama dili, bu çalışmada CLTV ve RFM analizlerini e-ticaret verileri üzerinde uygulamak için kullanılmıştır. Python, veri işleme, analiz ve görselleştirme için çok çeşitli kütüphane desteği (örneğin, Pandas, NumPy, datetime ve Matplotlib) nedeniyle en iyi seçimdir (McKinney, 2012). Büyük veri setlerinin hızlı bir şekilde işlenmesi, analiz edilmesi ve görselleştirilmesi bu yazılım ve araçlar tarafından mümkündür.

3.4 Veri Toplama ve İşleme

Araştırma verileri, e-ticaret platformları üzerinde proje yapılabilecek detaylı bir veri seti olan kaggle platformu üzerinde bulunuyor. Veri seti 525461 gözlem sayısından oluşan ve 8 değişken içeren bir veri setidir. Müşterilerin geçmişi, satın alma sıklığı, harcama tutarları ve işlem tarihleri gibi parametreler verilerde bulunur. Veri setini Kaggle üzerinde yer alan online-retail 2'dir. Benim kendi localimde indirdiğim data seti link örneği: [Kodlar\online_retail II.xlsx\online_retail II.xlsx](#).

RFM analizinin uygulanabilmesi için her müşterinin üç ana kritere (son işlem tarihi, işlem sıklığı ve harcama tutarı) bakması gerekir (Stone ve Woodcock, 2014). CLTV hesaplaması için uzun vadeli değerleri, her müşterinin gelecekteki potansiyel katkısı tahmin edilerek bulunur (Gupta ve Lehmann, 2005).

Data sette yer alan deęişkenler:

InvoiceNo: Faturaya ait eşsiz numaradır. C ile başlıyorsa işlem iptal edildiğini gösterir.

StockCode: Her ürün için eşsiz verilen koddur.

Description: Ürün açıklaması

Quantity: Ürün adedi. Yapılan işlemlerde satılan ürün sayısıdır.

InvoiceDate : Fatura zamanı ve tarihi(Date-Time)

UnitPrice: Ürün fiyatı

CustomerID: Müşteriye ait olan unique numaradır

Country : Ülke ismi.

3.5 Araştırmada Kullanılan Metodlar

3.5.1 RFM analizi. RFM analizi, üç temel parametreye dayalı olarak müşterileri sınıflandırarak alışveriş alışkanlıklarını anlamak ve müşteri segmentasyonunu optimize etmek için kullanılabilir. Yenilik (Recency), sıklık (Frequency) ve parasal değer (Monetary) bu üç parametreyi oluşturur (Stone ve Woodcock, 2014). Bu analiz yöntemi, müşteri segmentlerini daha küçük gruplara ayırarak her grup için özelleştirilmiş pazarlama stratejileri oluşturur.

RFM analizinde kullanılan üç ana ölçüt şunlardır:

1. Recency: Bir müşterinin en son alışveriş tarihini gösteren bir ölçümdür. Müşterinin işletmeyle ne kadar yakın bir zamanda etkileşim kurduğunu anlamak için bu metrik çok önemlidir. Daha yakın tarihli işlemler genellikle daha yüksek müşteri bağlılığı gösterir (Stone ve Woodcock, 2014).
2. Frekans: Müşterinin yaptığı toplam alışveriş sayısını gösterir. Müşterinin işletmeyle olan etkileşim yoğunluğunu belirlemek için sıklık metrikleri kullanılabilir. Bu ölçümler genellikle müşteri değerini artıran önemli bir göstergedir (Fader ve diğ., 2005).
3. Parasal: Müşterilerin yaptığı alışverişlerden kazandıkları toplam parasal değerdir. Bu ölçüm, müşterinin şirkete ne kadar para harcadığını doğrudan gösterir (ReinartzveKumar,2003).

Her müşteri için bir RFM skoru oluşturmak için üç ölçüt genellikle aynı cinsten değerlere dönüştürülür. Bu skor, müşterileri farklı gruplara ayırmak için temel bir

rehber olarak hizmet eder. Örneğin, yüksek sıklık (frequency) ve yenilik (recency) skorlarına sahip müşteriler, işletme için yüksek potansiyel taşıyan bir gruptur. Sık alışveriş yapan müşterilerin parasal katkısı zamanla artma eğilimindedir, bu nedenle parasal değer (monetary) önemi daha düşüktür (Fader ve diğ., 2005).

RFM analizinin uygulanışı: Bu çalışmada segmentasyon algoritmaları RFM analizi için kullanılmıştır. Bu algoritmalar Python programlama dilinin Scikit-learn kütüphanesi ile kullanılmıştır. İşletmeler, bu yöntemle müşteri verilerini kolayca işleyerek veri odaklı aksiyon alabilirler. RFM analizi, CRM (Müşteri İlişkileri Yönetimi) çalışmaları açısından güçlü bir araçtır çünkü müşteri gruplarına özel stratejiler oluşturmak ve bu stratejilere göre özelleştirilmiş pazarlama kampanyaları oluşturmak için kullanılabilir (Petersen ve diğ., 2009). RFM analizi, işletmelere yalnızca mevcut müşteri davranışlarını anlamakla kalmaz, aynı zamanda bu davranışların bir sonucu olarak gelecekteki müşteri davranışlarını öngörme yeteneği de verir. Bu, e-ticaret gibi yoğun rekabetin yaşandığı alanlarda işletmelerin sürdürülebilir büyüme stratejileri geliştirmesine önemli ölçüde yardımcı olur.

3.5.2 CLTV analizi. Bir müşterinin işletmeye ömrü boyunca sağlayacağı toplam gelirin tahmini değeri CLTV'dir. Bu hesaplama için önceki veriler ve tahmin modelleri kullanıldı. Gelecekteki gelir akışlarının tahmin edilmesi ve uzun vadeli müşteri ilişkilerinin yönetimi için CLTV analizi çok önemlidir (Reinartz ve Kumar, 2003).

1. Bir müşterinin bir şirketle kurduğu ilişki-iletişim süresince bu şirkete kazandıracağı parasal değerdir.
2. Müşterinin şirkete tahmini getireceği kazanç:
3. Satın alma başına ortalama kazanç * satın alma sayısı
4. $CLTV = (Customer Value / Churn Rate) * Profit Margin$
5. $Customer Value = Average Order Value * Purchase Frequency$
6. $Average Order Value = Total Price / Total Transaction$
7. $Purchase Frequency = Total Transaction / Total Number of Customers$
8. $Churn Rate = 1 - Repeat Rate$
9. $Repeat Rate = Birden fazla alışveriş yapan müşterilerin sayısı / tüm müşteriler$
10. $Profit Margin = Total Price * 0.10$ - örnek değer

3.5.3 Analiz süreci adımları. Veri Toplama ve Hazırlama: Bir e-ticaret işletmesi tarafından toplanan müşteri işlem verilerinin analizi için hazırlandı. Veri setindeki toplam harcamalar, satın alma sıklıkları ve tarihsel işlem kayıtları analiz için optimize edilmiştir.

RFM Skorlama: RFM analizi, müşterileri tutar, yenilikçilik ve sıklıkla kategorize eder. Her bir müşteri için bu değerler hesaplandı ve bir skorlama sistemi kullanılarak gruplandırılmıştır.

CLTV Modelleme: Müşterilerin gelecekteki satın alma alışkanlıkları ve işletmenin sağlayacağı toplam finansal değer, CLTV hesaplamasıyla hesaplanır. Python programlama dili, CLTV hesaplama sürecinde müşteri geçmiş verilerini modellemek için kullanılmıştır.

Bulguların Değerlendirilmesi: Analiz sonuçları değerlendirilmiştir ve CLTV ve müşteri segmentleri için stratejik öneriler geliştirilmiştir.

3.6 RFM ile Müşteri Segmentasyonu Analiz Süreci

Bu çalışma Jupyter Notebook üzerinde gerçekleştirilmiştir. Veri seti Kaggle platformu üzerinden alınmış olup, başlangıçta yaklaşık beş yüz bin gözlem ve sekiz değişken içeren veri seti RFM ve CLTV analizlerinde kullanılabilmesi için veri manipilasyonu ve değişken mühendisliği uygulanmıştır.

3.6.1 Veriyi anlama(Data understanding). İlk olarak gerekli kütüphanelerin import işlemleri uygulanmıştır. Format ayarları istenilen şekilde ayarlanmıştır. Veri seti okunup kopyası da kaydedilmiştir. Değişkenlerdeki boş değerler kontrol edilmiştir ve tüm değişkenler bazında kontrol edilebilmesi için bir fonksiyon oluşturulmuştur. Eşsiz ürün değerleri kontrol edilmiştir ve ürün açıklamalarının da sınıfsal ayrımı kontrol edilmiştir. Total Price adında yeni bir değişken tanımlanmıştır.

```
import datetime as dt
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
pd.set_option("display.max_columns",None)
# pd.set_option("display.max_rows",None)
pd.set_option("display.float_format",lambda x: "%.3f" % x)
```

```
DF = pd.read_excel(r"C:\Users\mustafa\CRM ANALYSIS\online_retail_II.xlsx",sheet_name = "Year 2009-2010")
```

```
df = DF.copy() # okuması uzun sürdüğü için, ileride ters bir durum olursa direkt olarak okumayı tekrardan beklememiş olacağız.
```

```
df.head()
```

	Invoice	StockCode	Description	Quantity	InvoiceDate	Price	Customer ID	Country
0	489434	85048	15CM CHRISTMAS GLASS BALL 20 LIGHTS	12	2009-12-01 07:45:00	6.950	13085.000	United Kingdom
1	489434	79323P	PINK CHERRY LIGHTS	12	2009-12-01 07:45:00	6.750	13085.000	United Kingdom
2	489434	79323W	WHITE CHERRY LIGHTS	12	2009-12-01 07:45:00	6.750	13085.000	United Kingdom
3	489434	22041	RECORD FRAME 7" SINGLE SIZE	48	2009-12-01 07:45:00	2.100	13085.000	United Kingdom
4	489434	21232	STRAWBERRY CERAMIC TRINKET BOX	24	2009-12-01 07:45:00	1.250	13085.000	United Kingdom

```
df.shape
```

```
(525461, 8)
```

Şekil 3 . Kütüphanelerin import işlemi.

Şekil 3'te kütüphanelerin import işlemi, veri setinin okunması, veriye genelden bir bakış ve veri setinin boyutunun kontrolü yapılmıştır.

```
df.isnull().sum()
```

```
Invoice      0
StockCode    0
Description  2928
Quantity     0
InvoiceDate  0
Price        0
Customer ID  107927
Country      0
dtype: int64
```

```
def missing_values_table(dataframe, na_name=False):
    na_columns = [col for col in dataframe.columns if dataframe[col].isnull().sum() > 0]

    n_miss = dataframe[na_columns].isnull().sum().sort_values(ascending=False)
    ratio = (dataframe[na_columns].isnull().sum() / dataframe.shape[0] * 100).sort_values(ascending=False)
    missing_df = pd.concat([n_miss, np.round(ratio, 2)], axis=1, keys=["n_miss", "ratio"])
    print(missing_df, end="\n")

    if na_name:
        return na_columns
```

```
missing_values_table(df, na_name=True)
```

```
      n_miss  ratio
Customer ID 107927 20.540
Description  2928  0.560
```

```
['Description', 'Customer ID']
```

```
# Customer ID deki eksik değerleri silmek mantıklı olacaktır. Çünkü müşteriler üzerinde
# bir segmentasyon uygulayacağımız için id si bulunmayanlar hiçbir şey ifade etmeyecektir.
```

```
# eşsiz ürün sayısı
df["Description"].nunique()
```

```
4681
```

Şekil 4 . Boş değerlerin kontrolü.

Şekil 4'te boş değerlerin kontrolü sağlanmıştır. Ve boş değerlerin tüm veri seti bazında yüzdesel ve sayısal olarak değeri gösterilmiştir. Boş değer kontrolünü görebileceğimiz bir fonksiyonunun oluşturulmuştur.

```
df["Description"].value_counts().head()
```

```
WHITE HANGING HEART T-LIGHT HOLDER    3549
REGENCY CAKESTAND 3 TIER                2212
STRAWBERRY CERAMIC TRINKET BOX         1843
PACK OF 72 RETRO SPOT CAKE CASES       1466
ASSORTED COLOUR BIRD ORNAMENT          1457
Name: Description, dtype: int64
```

```
# Toplam ne kadar satın alınmanın yapıldığını anlamak için
df.groupby("Description").agg({"Quantity": "sum"}).head()
# Bu grupta için şuanlık bir problem olduğunu farkettik çünkü quantity - olamaz. Buna tekrardan bakmamız gerekir.
```

Description	Quantity
	21494
	-720
	22467
	-2
	22719
	2
DOORMAT UNION JACK GUNS AND ROSES	179
3 STRIPEY MICE FELTCRAFT	690

```
df.groupby("Description").agg({"Quantity": "sum"}).sort_values(by="Quantity", ascending=False).head()
```

Description	Quantity
WHITE HANGING HEART T-LIGHT HOLDER	57733
WORLD WAR 2 GLIDERS ASSTD DESIGNS	54698
BROCADE RING PURSE	47647
PACK OF 72 RETRO SPOT CAKE CASES	46106
ASSORTED COLOUR BIRD ORNAMENT	44925

```
# Toplam kaç tane eşsiz fatura kesilmiş?
df["Invoice"].nunique()
```

```
28816
```

Şekil 5 . Adet sayısı ve toplam kontrolü.

Şekil 5'te ürün açıklaması değişkeninin sınıf sayısı kontrolü ile açıklaması kırılımında adet sayısının toplam değerinin analizinin kontrolü yapılmıştır.

```
# Fatura başına toplam kaç para kazanılmış
df["TotalPrice"] = df["Quantity"] * df["Price"] # Ürünlerin toplam kazancı
```

```
df.head()
```

	Invoice	StockCode	Description	Quantity	InvoiceDate	Price	Customer ID	Country	TotalPrice
0	489434	85048	15CM CHRISTMAS GLASS BALL 20 LIGHTS	12	2009-12-01 07:45:00	6.950	13085.000	United Kingdom	83.400
1	489434	79323P	PINK CHERRY LIGHTS	12	2009-12-01 07:45:00	6.750	13085.000	United Kingdom	81.000
2	489434	79323W	WHITE CHERRY LIGHTS	12	2009-12-01 07:45:00	6.750	13085.000	United Kingdom	81.000
3	489434	22041	RECORD FRAME 7" SINGLE SIZE	48	2009-12-01 07:45:00	2.100	13085.000	United Kingdom	100.800
4	489434	21232	STRAWBERRY CERAMIC TRINKET BOX	24	2009-12-01 07:45:00	1.250	13085.000	United Kingdom	30.000

```
# Invoice başına toplam ödenen para
df.groupby("Invoice").agg({"TotalPrice": "sum"}).head()
```

Invoice	TotalPrice
489434	505.300
489435	145.800
489436	630.330
489437	310.750
489438	2286.240

Şekil 6 . Yeni değişken tanımlama.

Şekil 6’da toplam değer adında yeni değişkenin oluşturulması ve fatura kırılımında toplam değerinin analizi yapılmıştır.

3.6.2 Veri hazırlama(Data preparation). Customer ID’deki boş değerler değiştirilemeyeceği için veri setinden çıkarılmıştır, aynı şekilde Description değişkenindeki boş değer sayısı genel gözlem sayısına göre göz ardı edilebilecek durumda olduğu için veri setinden çıkarılmıştır. Yaptığımız analiz sonucunda “Quantity” değerinin negatif çıktığı gözlemlenmiştir. Bunun sebebi geri iadelerin veri setinde yer almasından kaynaklanmaktadır. Geri iadeler fatura tipinde “C” harfi ile gösterilmektedir. Betimsel istatistik analizi uygulanmıştır ve bunun sonucuna bakılarak “Invoice” değişkeninde string içeren yapının “C” bulunan gözlemleri tekrardan veri setinden çıkarılmıştır.

```
df.shape
(525461, 9)

df.isnull().sum()
Invoice      0
StockCode    0
Description  2928
Quantity     0
InvoiceDate  0
Price        0
Customer ID  107927
Country      0
TotalPrice   0
dtype: int64

df.dropna(inplace=True)

df.isnull().sum()
Invoice      0
StockCode    0
Description  0
Quantity     0
InvoiceDate  0
Price        0
Customer ID  0
Country      0
TotalPrice   0
dtype: int64
```

Şekil 7. Boş değerlerin silinmesi.

Şekil 7’de boş değerler tespit edilerek analiz sürecinde analizimizi bozacağı için tesbit edilerek datasetten çıkartılmıştır.

```
df.describe().T # aykırı değerler zaten skorlaştırılmada çıkarılacak ama başta da ifade ettiğimiz gibi çıkarıp tekrardan # describe bakalım
```

	count	mean	std	min	25%	50%	75%	max
Quantity	417534.000	12.759	101.220	-9360.000	2.000	4.000	12.000	19152.000
Price	417534.000	3.888	71.132	0.000	1.250	1.950	3.750	25111.090
Customer ID	417534.000	15360.645	1680.811	12346.000	13983.000	15311.000	16799.000	18287.000
TotalPrice	417534.000	19.994	99.916	-25111.090	4.250	11.250	19.350	15818.400

```
df = df[~df["Invoice"].str.contains("C", na = False)] # na = False na value'ları da False olarak alıp işleme alır, böylelikle # error alınmamış olur.

df.describe().T
```

	count	mean	std	min	25%	50%	75%	max
Quantity	407695.000	13.587	96.842	1.000	2.000	5.000	12.000	19152.000
Price	407695.000	3.294	34.757	0.000	1.250	1.950	3.750	10953.500
Customer ID	407695.000	15368.504	1679.796	12346.000	13997.000	15321.000	16812.000	18287.000
TotalPrice	407695.000	21.663	77.147	0.000	4.950	11.900	19.500	15818.400

Şekil 8. Geri iadelerin silinmesi.

Şekil 8’de verideki betimsel analitiğe bakarak geri iade işlemi olan faturada “c” içeren verilerin düzeltilerek tekrardan kontrolünün sağlanmıştır. Bunun sebebi quantity değerlerinin negatif çıkmasını engellemektir.

3.6.3 RFM metriklerinin hesaplanması. RFM analizinin uygulanabilmesi için öncelikle analiz tarihi karar verilmiştir. Veri setinde “Customer ID” gruplamasında “Fatura Tarihi”, “Fatura Adedi” ve “Parasal Değere” odaklanarak rfm adında yeni bir veri seti çıkarılmıştır. Çıkardığımız değişkenlerin isimleri de analiz için gerekli olarak adlandırılan “Receny”, “Frequency” ve “Monetary” formatına hazır hale getirilmiştir. Tekrardan betimsel analiz olan describe metodu uygulanmış olup Monetary değerleri 0’a eşit olan gözlemler dikkat çekmiştir ve veri setinden 0’dan büyük olacak şekilde düzenlenmiştir.

```
# Recency - Frequency - Monetary
# Recency - Analizin yapıldığı tarih - Müşterinin en son satın alındığı tarih
# Frequency - Müşterinin yaptığı toplam satın almadır.
# Monetary - Frequency neticesinde toplam getiri.
```

```
df.head()
```

	Invoice	StockCode	Description	Quantity	InvoiceDate	Price	Customer ID	Country	TotalPrice
0	489434	85048	15CM CHRISTMAS GLASS BALL 20 LIGHTS	12	2009-12-01 07:45:00	6.950	13085.000	United Kingdom	83.400
1	489434	79323P	PINK CHERRY LIGHTS	12	2009-12-01 07:45:00	6.750	13085.000	United Kingdom	81.000
2	489434	79323W	WHITE CHERRY LIGHTS	12	2009-12-01 07:45:00	6.750	13085.000	United Kingdom	81.000
3	489434	22041	RECORD FRAME 7" SINGLE SIZE	48	2009-12-01 07:45:00	2.100	13085.000	United Kingdom	100.800
4	489434	21232	STRAWBERRY CERAMIC TRINKET BOX	24	2009-12-01 07:45:00	1.250	13085.000	United Kingdom	30.000

```
df["InvoiceDate"].max()
Timestamp('2010-12-09 20:01:00')
```

```
today_date = dt.datetime(2010, 12, 11)
```

Şekil 9 . Recency değerinin hesaplanması.

Şekil 9’da recency değerini bulmak için en sonki fatura tarihine bakılarak analiz için bugünün tarihi adında yeni bir değişkenin oluşturulmuştur.

```
rfm = df.groupby("Customer ID").agg({"InvoiceDate" : lambda InvoiceDate : (today_date - InvoiceDate.max()).days,
                                   "Invoice" : lambda Invoice : Invoice.nunique(),
                                   "TotalPrice" : lambda TotalPrice : TotalPrice.sum()})
```

```
rfm.head()
```

Customer ID	Recency	Frequency	Monetary
12346.000	165	11	372.860
12347.000	3	2	1323.320
12348.000	74	1	222.160
12349.000	43	3	2671.140
12351.000	11	1	300.930

```
rfm.columns = ["Recency", "Frequency", "Monetary"]
```

```
rfm.describe().T
```

	count	mean	std	min	25%	50%	75%	max
Recency	4314.000	91.270	96.944	1.000	18.000	53.000	136.000	374.000
Frequency	4314.000	4.454	8.169	1.000	1.000	2.000	5.000	205.000
Monetary	4314.000	2047.289	8912.523	0.000	307.950	705.550	1722.802	349164.350

```
# Monetary değişkeninde minimum değerini 0 çıkartmamız daha verimli olacaktır.
rfm = rfm[rfm["Monetary"] > 0]
```

```
rfm.describe().T
```

	count	mean	std	min	25%	50%	75%	max
Recency	4312.000	91.173	96.861	1.000	18.000	53.000	136.000	374.000
Frequency	4312.000	4.456	8.170	1.000	1.000	2.000	5.000	205.000
Monetary	4312.000	2048.238	8914.481	2.950	307.988	706.020	1723.142	349164.350

Şekil 10 . Recency, frequency, monetary değerleri.

Şekil 10'da müşteri id'lerin kırılımında rfm değişkenlerinin oluşturulması, bu değerlerin veri setine eklenmesi ve monetary değerleri 0 olan gözlemlerin veriden çıkarılması adımları yapılmıştır.

3.6.4 RFM skorlarının hesaplanması. Tanımladığımız değişkenler üzerinde skor yapılarının oluşturulabilmesi için etiketleme işlemi yapılmıştır. Bunun sonucunda değerler 1'den 5'e olacak şekilde skora ayrılmıştır. Recency_Score ve Frequency_Score string yapıda birleştirilip 2 boyutlu regex kullanımına hazır olacak şekilde düzenlemiştir.

```
# Recency ters, frequency ve monetary de büyüklük-küçüklük algısı vardır.
# qcut fonksiyonu uygulayarak quantile'lara göre bölme işlemi yapar.
rfm["recency_score"] = pd.qcut(rfm["Recency"], 5, labels=[5, 4, 3, 2, 1]) # küçükten büyüğe şekilde 5-4-3-2-1 olarak
# ayırma yapacak.
```

```
rfm["frequency_score"] = pd.qcut(rfm["Frequency"].rank(method = "first"), 5, labels=[1, 2, 3, 4, 5])
# yapılan ayırmada birden fazla aynı değere denk gelen veri bulunduğu için method=first kullanılması gerekir.
```

```
rfm["monetary_score"] = pd.qcut(rfm["Monetary"], 5, labels=[1, 2, 3, 4, 5])
```

```
rfm.head()
```

Customer ID	Recency	Frequency	Monetary	recency_score	frequency_score	monetary_score
12346.000	165	11	372.860	2	5	2
12347.000	3	2	1323.320	5	2	4
12348.000	74	1	222.160	2	1	1
12349.000	43	3	2671.140	3	3	5
12351.000	11	1	300.930	5	1	2

```
# RFM skoru oluştururken R ve F değişkenlerini string ifade içerisinde toplayarak yeni bir değişkende göstermemiz gerekiyor.
rfm["RFM_SCORE"] = (rfm["recency_score"].astype(str) + rfm["frequency_score"].astype(str))
```

```
rfm.head()
```

Customer ID	Recency	Frequency	Monetary	recency_score	frequency_score	monetary_score	RFM_SCORE
12346.000	165	11	372.860	2	5	2	25
12347.000	3	2	1323.320	5	2	4	52
12348.000	74	1	222.160	2	1	1	21
12349.000	43	3	2671.140	3	3	5	33
12351.000	11	1	300.930	5	1	2	51

Şekil 11 . Rfm score oluşturulması.

Şekil 11'de yeni oluşturulan değişkenlere göre label işlemlerinin yapılması ve en sonda rfm_score adında recency ve frequency değerleri kullanılarak yeni bir string yapı oluşturulmuştur.

3.6.5 RFM segmentlerinin oluşturulması. Regex yapısı istenilen analizler için oluşturulmuştur ve Segment adında oluşturacağımız yeni değişken için içerisinde gömülmüştür. Kaybedilmemesi gereken müşteriler için segment analiz örneği gösterilmiştir. Bu müşterilerin indeks değerleri tutulmuştur ve yeni bir veri seti içerisinde atılıp csv dosyası olarak kaydedilmiştir.

```
# regex
# RFM isimlendirmesi
seg_map = {
  r"[1-2][1-2]": "hibernating",
  r"[1-2][3-4]": "at_Risk",
  r"[1-2]5": "cant_loose",
  r"3[1-2]": "about_to_sleep",
  r"33": "need_attention",
  r"[3-4][4-5]": "loyal_csutomers",
  r"41": "promising",
  r"51": "new_customers",
  r"[4-5][2-3]": "potential_loyalists",
  r"5[4-5]": "champions"
}
# r - regexi yani yapı yakalamayı sağlar. Köşeli elemanlar olasılıkları ifade eder.
```

```
rfm["segment"] = rfm["RFM_SCORE"].replace(seg_map, regex=True)
```

```
rfm.head()
```

Customer ID	Recency	Frequency	Monetary	recency_score	frequency_score	monetary_score	RFM_SCORE	segment
12346.000	165	11	372.860	2	5	2	25	cant_loose
12347.000	3	2	1323.320	5	2	4	52	potential_loyalists
12348.000	74	1	222.160	2	1	1	21	hibernating
12349.000	43	3	2671.140	3	3	5	33	need_attention
12351.000	11	1	300.930	5	1	2	51	new_customers

Şekil 12 . Regex yapısı.

Şekil 12’de segment oluşturmak için istenilen regex yapısının oluşturulmuştur.

Segment analizi örneği

```
rfm[["segment", "Recency", "Frequency", "Monetary"]].groupby("segment").agg({"mean", "count"})
```

segment	Recency		Frequency		Monetary	
	mean	count	mean	count	mean	count
about_to_sleep	53.819	343	1.201	343	441.320	343
at_Risk	152.159	611	3.074	611	1188.878	611
cant_loose	124.117	77	9.117	77	4099.450	77
champions	7.119	663	12.554	663	6852.264	663
hibernating	213.886	1015	1.126	1015	403.978	1015
loyal_csutomers	36.287	742	6.830	742	2746.067	742
need_attention	53.266	207	2.449	207	1060.357	207
new_customers	8.580	50	1.000	50	386.199	50
potential_loyalists	18.793	517	2.017	517	729.511	517
promising	25.747	87	1.000	87	367.087	87

Şekil 13 . Segment analizi örneği.

Şekil 13’de segmentasyon işlemi yapılmıştır. Oluşturduğumuz analiz skorlarına göre ortalama ve sayı değerleri bulunmuştur.

```
rfm[rfm["segment"] == "cant_loose"].head()
```

Customer ID	Recency	Frequency	Monetary	recency_score	frequency_score	monetary_score	RFM_SCORE	segment
12346.000	165	11	372.860	2	5	2	25	cant_loose
12380.000	101	7	6951.490	2	5	5	25	cant_loose
12482.000	212	29	23691.400	1	5	5	15	cant_loose
12510.000	95	7	4195.450	2	5	5	25	cant_loose
12891.000	94	8	509.500	2	5	3	25	cant_loose

```
rfm[rfm["segment"] == "cant_loose"].index # CustomerID deęerleri
```

```
Float64Index([12346.0, 12380.0, 12482.0, 12510.0, 12891.0, 12932.0, 13044.0,  
13313.0, 13680.0, 13782.0, 13799.0, 13856.0, 14025.0, 14063.0,  
14160.0, 14221.0, 14548.0, 14607.0, 14685.0, 14745.0, 15003.0,  
15013.0, 15015.0, 15125.0, 15141.0, 15222.0, 15306.0, 15321.0,  
15359.0, 15369.0, 15372.0, 15443.0, 15538.0, 15607.0, 15633.0,  
15722.0, 15751.0, 15754.0, 15768.0, 15911.0, 15912.0, 16027.0,  
16032.0, 16158.0, 16177.0, 16197.0, 16335.0, 16467.0, 16631.0,  
16742.0, 16743.0, 16875.0, 16986.0, 17021.0, 17032.0, 17092.0,  
17113.0, 17157.0, 17188.0, 17230.0, 17268.0, 17426.0, 17448.0,  
17454.0, 17512.0, 17578.0, 17602.0, 17651.0, 17940.0, 17969.0,  
17988.0, 18009.0, 18051.0, 18064.0, 18094.0, 18251.0, 18258.0],  
dtype='float64', name='Customer ID')
```

```
new_df = pd.DataFrame()
```

```
new_df["new_customer_id"] = rfm[rfm["segment"] == "new_customers"].index
```

Şekil 14 . New_df veri seti oluřturma.

Şekil 14’de segment ierisinde kaybedilmemesi gerekenler adlı müşterilerin bir index’te tutulup new_df adlı yeni bir veri seti ierisine atılmıřtır.

```
new_df.head()
```

	new_customer_id
0	12351.000
1	12385.000
2	12386.000
3	12427.000
4	12441.000

```
# float'tan kurtulmak iin  
new_df["new_customer_id"] = new_df["new_customer_id"].astype(int)
```

```
new_df.head()
```

	new_customer_id
0	12351
1	12385
2	12386
3	12427
4	12441

```
new_df.to_csv("new_customers.csv")  
# reload from disk ile ilgili csv dosyasına kaydedebiliriz.
```

Şekil 15 . Müşteri kayıtları csv dosyası.

Şekil 15’de belirlediğimiz müşterilerin csv dosyası olarak tutulması için yazılan fonksiyon görülmektedir.

```
,new_customer_id
0,12351
1,12385
2,12386
3,12427
4,12441
5,12538
6,12686
7,12738
8,12763
9,12767
10,12942
11,12947
12,12961
13,12996
14,13011
15,13030
16,13254
17,13270
18,13369
19,13370
20,13626
```

Şekil 16 . Oluşturulan csv dosyası.

3.6.6 Tüm sürecin fonksiyonlaştırılması. Tüm süreç tek bir ana fonksiyon içerisinde toplanmıştır. İstenilen segment yapısına göre fonksiyon içinde kullanıp csv dosyası şeklinde kaydedilmesi otomatikleştirilmiştir.

```
def all_process_for_rfm(dataframe, csv=False):
    dataframe.dropna(inplace=True)
    dataframe = dataframe[~dataframe["Invoice"].str.contains("C", na = False)]
    dataframe["TotalPrice"] = dataframe["Quantity"] * dataframe["Price"]

    today_date = dt.datetime(2010, 12, 11)
    rfm = dataframe.groupby("Customer ID").agg({"InvoiceDate" : lambda InvoiceDate : (today_date - InvoiceDate.max()).days,
        "Invoice" : lambda Invoice : Invoice.nunique(),
        "TotalPrice" : lambda TotalPrice : TotalPrice.sum()})

    rfm.columns = ["Recency", "Frequency", "Monetary"]
    rfm = rfm[rfm["Monetary"] > 0]
    rfm["recency_score"] = pd.qcut(rfm["Recency"], 5, labels=[5, 4, 3, 2, 1])
    rfm["frequency_score"] = pd.qcut(rfm["Frequency"].rank(method = "first"), 5, labels=[1, 2, 3, 4, 5])
    rfm["monetary_score"] = pd.qcut(rfm["Monetary"], 5, labels=[1, 2, 3, 4, 5])
    rfm["RFM_SCORE"] = (rfm["recency_score"].astype(str) + rfm["frequency_score"].astype(str))

    seg_map = {
        r"[1-2][1-2]": "hibernating",
        r"[1-2][3-4]": "at_Risk",
        r"[1-2]5": "cant_loose",
        r"3[1-2]": "about_to_sleep",
        r"33": "need_attention",
        r"[3-4][4-5]": "loyal_csutomers",
        r"41": "promising",
        r"51": "new_customers",
        r"[4-5][2-3]": "potential_loyalists",
        r"5[4-5]": "champions"
    }

    rfm["segment"] = rfm["RFM_SCORE"].replace(seg_map, regex=True)

    if csv:
        new_df = pd.DataFrame()
        new_df["new_customer_id"] = rfm[rfm["segment"] == "new_customers"].index
        new_df["new_customer_id"] = new_df["new_customer_id"].astype(int)
        new_df.to_csv("new_customers.csv")
        print(new_df)
```

Şekil 17 . Tüm sürecin fonksiyonlaştırılması.

Şekil 17’de yaptığımız tüm analiz süreçleri tek bir fonksiyon altında toplanmıştır.

3.7 CLTV Analiz Süreci

Bir müşterinin şirketlerle kurduğu süre boyunca elde ettiği parasal değerdir. Müşterinin şirkete tahmin ettiği kazanç, satın alma başına ortalama kazanç * satın alma sayısı ile hesaplanır.

CLTV(Customer Life Time Value) = (Customer Value / Churn Rate) * Profit Margin

Customer Value = Average Order Value * Purchase Frequency

Average Order Value = Total Price / Total Transaction

Purchase Frequency = Total Transaction / Total Number of Customers

Churn Rate = 1-Repeat Rate

Profit Margin = Total Price * 0.10(örnek değer)

Sonuç olarak, her müşteri için hesaplanan CLTV değerlerine göre bir sıralama yapıldığında ve CLTV değerlerini belirli noktalardan bölerek gruplar oluşturulduğunda, müşterilerimiz segmentlere bölünecektir.

```
import pandas as pd
from sklearn.preprocessing import MinMaxScaler
pd.set_option("display.max_columns",None)
# pd.set_option("display.max_rows",None)
pd.set_option("display.float_format",lambda x : "%.5f" %x) #float tan sonra kaç basamak görmek istediğimiz ile ilgili bir ayar

df_ = pd.read_excel(r"Users\mustafa\online_retail_II.xlsx",sheet_name="Year 2009-2010")

df = df_.copy()

df.head()

Invoice StockCode Description Quantity InvoiceDate Price Customer ID Country
0 489434 85048 15CM CHRISTMAS GLASS BALL 20 LIGHTS 12 2009-12-01 07:45:00 6.95000 13085.00000 United Kingdom
1 489434 79323P PINK CHERRY LIGHTS 12 2009-12-01 07:45:00 6.75000 13085.00000 United Kingdom
2 489434 79323W WHITE CHERRY LIGHTS 12 2009-12-01 07:45:00 6.75000 13085.00000 United Kingdom
3 489434 22041 RECORD FRAME 7" SINGLE SIZE 48 2009-12-01 07:45:00 2.10000 13085.00000 United Kingdom
4 489434 21232 STRAWBERRY CERAMIC TRINKET BOX 24 2009-12-01 07:45:00 1.25000 13085.00000 United Kingdom

df.isnull().sum()

Invoice 0
StockCode 0
Description 2928
Quantity 0
InvoiceDate 0
Price 0
Customer ID 107927
Country 0
dtype: int64
```

Şekil 18 . İmport ve düzenleme.

Şekil 18’de gerekli import işlemlerinin yapılması, düzenlemelerin yapılması, veri setinin okuması yapılmıştır.

```
df = df[~df["Invoice"].str.contains("C", na=False)]
df.describe().T
```

	count	mean	std	min	25%	50%	75%	max
Quantity	515255.00000	10.95663	104.35401	-9600.00000	1.00000	3.00000	10.00000	19152.00000
Price	515255.00000	3.95637	127.68856	-53594.36000	1.25000	2.10000	4.21000	25111.09000
Customer ID	407695.00000	15368.50411	1679.79570	12346.00000	13997.00000	15321.00000	16812.00000	18287.00000

```
df = df[df["Quantity"] > 0]
df = df.dropna()
df["TotalPrice"] = df["Quantity"] * df["Price"]
cltv_c = df.groupby("Customer ID").agg({"Invoice" : lambda x : x.nunique(),
                                       "Quantity" : lambda x : x.sum(),
                                       "TotalPrice" : lambda x : x.sum()})
cltv_c.columns = ["total_transaction", "total_unit", "total_price"]
```

Şekil 19 . Cltv_c veri seti.

Şekil 19’da fatura üzerinde düzenleme yapılmıştır. Adedi 0’ dan büyük olan gözlemler bulunmuştur. Toplam değer adlı yeni değişkenin oluşturulması ve cltv_c adında yeni bir veri setinin içerisine müşteri id kırılımında gereken hesaplamalar ile değişkenler belirlenmiştir.

Average Order Value(average_order_value = total_price / total_transaction)

```
cltv_c["average_order_value"] = cltv_c["total_price"] / cltv_c["total_transaction"]
cltv_c.head()
```

Customer ID	total_transaction	total_unit	total_price	average_order_value
12346.00000	11	70	372.86000	33.89636
12347.00000	2	828	1323.32000	661.66000
12348.00000	1	373	222.16000	222.16000
12349.00000	3	993	2671.14000	890.38000
12351.00000	1	261	300.93000	300.93000

Şekil 20 . Ortalama sipariş değeri.

Şekil 20’de ortalama sipariş değerini bulabilmek için toplam tutar/ toplam benzersiz fatura sayısı yapılmıştır.

Purchase Frequency($\text{total_transaction} / \text{total_number_of_customers}$)

```
cltv_c["purchase_frequency"] = cltv_c["total_transaction"] / cltv_c.shape[0]
```

```
cltv_c.head()
```

	total_transaction	total_unit	total_price	average_order_value	purchase_frequency
Customer ID					
12346.00000	11	70	372.86000	33.89636	0.00255
12347.00000	2	828	1323.32000	661.66000	0.00046
12348.00000	1	373	222.16000	222.16000	0.00023
12349.00000	3	993	2671.14000	890.38000	0.00070
12351.00000	1	261	300.93000	300.93000	0.00023

Şekil 21 . Satın alma sıklığı.

Şekil 21’de ne sıklıkla satınalma işlemi yapılmış bulabilmek için toplam benzersiz fatura sayısı/ Gözlem sayısı yapılarak satın alma sıklığı bulunmuştur.

Repeat Rate & Churn Rate(birden fazla alım yapan müşteri sayısı / tüm müşteriler)

```
repeat_rate = cltv_c[cltv_c["total_transaction"] > 1].shape[0] / cltv_c.shape[0]  
repeat_rate
```

```
0.6706073249884098
```

```
churn_rate = 1 - repeat_rate  
churn_rate
```

```
0.3293926750115902
```

Şekil 22 . Müşteri terk oranı.

Şekil 22’de birden fazla satın alma yapan müşterilerin oranının belirlenmesi ve buna bağlı olarak müşteri terk oranının hesabının yapılması işlemleri yapılmıştır.

Profit Margin ($\text{profit_margin} = \text{total_price} * \text{profit_percent}$)

```
cltv_c["profit_margin"] = cltv_c["total_price"] * 0.10
```

```
cltv_c.head()
```

	total_transaction	total_unit	total_price	average_order_value	purchase_frequency	profit_margin
Customer ID						
12346.00000	11	70	372.86000	33.89636	0.00255	37.28600
12347.00000	2	828	1323.32000	661.66000	0.00046	132.33200
12348.00000	1	373	222.16000	222.16000	0.00023	22.21600
12349.00000	3	993	2671.14000	890.38000	0.00070	267.11400
12351.00000	1	261	300.93000	300.93000	0.00023	30.09300

Şekil 23 . Kar marjı hesabı.

Şekil 23’de kar marjının hesabının belirlenen kar yüzdesi hesabında belirlenme işlemi yapılmıştır.

Customer Value(customer_value = average_order_value * purchase frequency)

```
cltv_c["customer_value"] = cltv_c["average_order_value"] * cltv_c["purchase_frequency"]
```

```
cltv_c.head()
```

	total_transaction	total_unit	total_price	average_order_value	purchase_frequency	profit_margin	customer_value
Customer ID							
12346.00000	11	70	372.86000	33.89636	0.00255	37.28600	0.08643
12347.00000	2	828	1323.32000	661.66000	0.00046	132.33200	0.30675
12348.00000	1	373	222.16000	222.16000	0.00023	22.21600	0.05150
12349.00000	3	993	2671.14000	890.38000	0.00070	267.11400	0.61918
12351.00000	1	261	300.93000	300.93000	0.00023	30.09300	0.06976

Şekil 24 . Müşteri değeri hesabı.

Şekil 24’de Ortalama değerin satınalma frekansına oranına bakılarak müşteri değeri hesaplanmıştır.

Customer Lifetime Value(CLTV = (customer_value / churn_rate) * profit_margin)

```
cltv_c["cltv"] = (cltv_c["customer_value"] / churn_rate) * cltv_c["profit_margin"]
```

```
cltv_c.sort_values(by="cltv", ascending=False).head(10)
```

	total_transaction	total_unit	total_price	average_order_value	purchase_frequency	profit_margin	customer_value	cltv
Customer ID								
18102.00000	89	124216	349164.35000	3923.19494	0.02063	34916.43500	80.93749	8579573.77276
14646.00000	78	170342	248396.50000	3184.57051	0.01808	24839.65000	57.57916	4342070.45829
14156.00000	102	108107	196566.74000	1927.12490	0.02364	19656.67400	45.56484	2719105.08615
14911.00000	205	69722	152147.57000	742.18327	0.04752	15214.75700	35.26833	1629055.80978
13694.00000	94	125893	131443.19000	1398.33181	0.02179	13144.31900	30.46898	1215855.89003
17511.00000	31	55107	84541.17000	2727.13452	0.00719	8454.11700	19.59693	502970.40288
15061.00000	86	51791	83284.38000	968.42302	0.01994	8328.43800	19.30561	488127.23096
16684.00000	27	54555	80489.21000	2981.08185	0.00626	8048.92100	18.65768	455912.23972
16754.00000	29	63551	65500.07000	2258.62310	0.00672	6550.00700	15.18314	301918.30894
17949.00000	74	30112	60117.60000	812.40000	0.01715	6011.76000	13.93547	254336.79309

Şekil 25 . Cltv kolonu hesaplama.

Şekil 25’de “müşteri değeri, erk etme oranı bölünerek kar marjı ile çarpılması sonucu cltv değeri hesaplanmıştır.

Segmentlerin Oluşturulması

```
cltv_c["segment"] = pd.qcut(cltv_c["cltv"], 4, labels=["D", "C", "B", "A"])
```

```
cltv_c.groupby("segment").agg({"count", "mean", "sum"})
```

segment	total_transaction			total_unit			total_price			average_order_value			purchase_frequency		
	mean	count	sum	mean	count	sum	mean	count	sum	mean	count	sum	mean	count	sum
D	1.22892	1079	1326	109.00463	1079	117616	178.18826	1079	192265.13000	157.37300	1079	169805.46550	0.00028	1079	0.30737
C	2.00371	1078	2160	283.05659	1078	305135	475.89652	1078	513016.45300	294.45487	1078	317422.35514	0.00046	1078	0.50070
B	3.76902	1078	4063	680.15863	1078	733211	1131.35918	1078	1219605.20000	389.72832	1078	420127.12365	0.00087	1078	0.94182
A	10.81186	1079	11666	4062.33735	1079	4383262	6401.40546	1079	6907116.49100	670.96592	1079	723972.22347	0.00251	1079	2.70422

Şekil 26 . Segmentlerin analizi.

Şekil 26’da segmentlerin istenilen label’lar üzerinde oluşturulmuştur. Analizinin yapılması için değerlerin toplam ortalama ve sayısal değerleri toplu şekilde verilmiştir. Asegmentine sahip bireylerin tüm kolon değerlerinde açık bir farkla önde olduğu gözlemlenmektedir.

```

def cltv_analyse_for_customers(dataframe, profit_margin=0.10, csv=False):
    dataframe["Invoice"] = dataframe["Invoice"].astype(str)
    dataframe = dataframe[~dataframe["Invoice"].str.contains("C", na=False)]
    dataframe = dataframe[dataframe["Quantity"] > 0]
    dataframe = dataframe.dropna()
    dataframe["TotalPrice"] = dataframe["Quantity"] * dataframe["Price"]

    cltv_c = dataframe.groupby("Customer ID").agg({"Invoice" : lambda x : x.nunique(),
                                                "Quantity" : lambda x : x.sum(),
                                                "TotalPrice" : lambda x : x.sum()})
    cltv_c.columns = ["total_transaction", "total_unit", "total_price"]

    ## Average Order Value(average_order_value = total_price / total_transaction)
    cltv_c["average_order_value"] = cltv_c["total_price"] / cltv_c["total_transaction"]

    ## Purchase Frequency(total_transaction / total_number_of_customers)
    cltv_c["purchase_frequency"] = cltv_c["total_transaction"] / cltv_c.shape[0]

    ## Repeat Rate & Churn Rate(birden fazla alım yapan müşteri sayısı / tüm müşteriler)
    repeat_rate = cltv_c[cltv_c["total_transaction"] > 1].shape[0] / cltv_c.shape[0]
    churn_rate = 1 - repeat_rate

    ## Profit Margin (profit_margin = total_price * profit percent)
    cltv_c["profit_margin"] = cltv_c["total_price"] * profit_margin

    ## Customer Value(customer_value = average_order_value * purchase frequency)
    cltv_c["customer_value"] = cltv_c["average_order_value"] * cltv_c["purchase_frequency"]

    ## Customer Lifetime Value(CLTV = (customer_value / churn_rate) * profit_margin)
    cltv_c["cltv"] = (cltv_c["customer_value"] / churn_rate) * cltv_c["profit_margin"]

    ## Segmentlerin Oluşturulması
    cltv_c["segment"] = pd.qcut(cltv_c["cltv"], 4, labels=["D", "C", "B", "A"])
    cltv_c.groupby("segment").agg({"count", "mean", "sum"})
    print(cltv_c.head())

    if csv:
        new_df = pd.DataFrame()
        new_df["A_Segment_Index"] = cltv_c[cltv_c["segment"] == "A"].index
        new_df["A_Segment_Index"] = new_df["A_Segment_Index"].astype("int")
        new_df.to_csv("A_Segment_Customers.csv")
        print(new_df)

```

Şekil 27 . Tüm Sürecinin Fonksiyonlaştırılması.

Şekil 27 de CLTV sürecinde yaptığımız tüm analizleri tek bir fonksiyon altında toplayarak oluşturduğumuz kod bloğumuz yer almaktadır.

Bölüm 4

Sonuç, Tartışma ve Öneriler

4.1 Değerlendirme ve Tartışma

Bu tez, e-ticaret müşteri analitiği alanında kısa vadeli ve uzun vadeli pazarlama stratejilerini optimize etmek için RFM analizi ve Müşteri Yaşam Boyu Değeri (CLTV) hesaplama tekniklerini birleştirmeye yardımcı olmuştur. Çalışma, bu alandaki mevcut literatürle önemli ölçüde benzer ve yenilikçidir.

Bu çalışmanın amacı, büyük veri ve dijitalleşme kavramlarının e-ticaret alanında nasıl bir etkiye sahip olduğunu araştırmak ve yeni teknolojiler kullanarak stratejik çözümler sunmaktır. Araştırmada kullanılan RFM (Recency, Frequency, Monetary) ve CLTV (Customer Lifetime Value) hesaplamaları, e-ticaret için önemli sonuçlar üretilmiştir (Smith ve diğ., 2020).

Yapılan araştırmalar, RFM ve CLTV yöntemlerinin e-ticaret şirketleri için çeşitli avantajlar sağladığını ortaya koyulmaktadır. CLTV analizi, uzun vadeli müşteri ilişkileri ve stratejik kararları etkilerken, RFM analizi kısa vadeli pazarlama stratejilerinin etkinliğini ölçmede etkilidir (Taylor, 2020). Bu analizler, şirketlerin uzun vadeli başarıya ulaşmak ve kısa vadeli hedeflerini optimize etmek için her iki yöntemin de kullanılmasının yararlı olabileceğini göstermiştir (Green ve diğ., 2021).

Araştırma, e-ticaret şirketleri için kısa ve uzun vadeli stratejiler sağlar. Müşteri bağlılığı programları, fiyatlandırma, iletişim ve müşteri segmentasyonu bu çerçevede optimize edilmektedir.

4.2 Araştırma Kısıtları

Çalışma kapsamında yaklaşık 500.000 gözlem içeren bir veri seti işleme, keşfedici veri analizi, veri manipülasyonu ve temizliği ve segmentasyon oluşturma dahil olmak üzere çeşitli prosedürler uygulanmıştır (Brown ve Jones, 2021). RFM analizi için monetary (harcama tutarı), frequency (alışveriş sıklığı) ve recency (son alışveriş tarihi) metrikleri kullanılmıştır ve elde edilen sonuçlar segmentlere

ayrılmıştır (Lee ve diğ., 2019). Bununla birlikte, CLTV hesaplamalarında, bir şirket için özel olarak belirlenen formüller kullanılmıştır. Bu metrikler arasında Customer Value, Average Order Value, Purchase Frequency, Churn Rate ve Profit Margin yer almaktadır (Johnson et al., 2022).

RFM ve CLTV Entegrasyonu: Literatürde yaygın olarak, RFM analizi ve CLTV hesaplamalarının birlikte kullanılmasının pazarlama stratejilerinde ve müşteri segmentasyonunda önemli faydalar sağladığı belirlenmiştir (Gupta ve Lehmann, 2005; Blattberg ve diğ., 2008). Bu tez ayrıca iki analitik yaklaşımı birleştirerek hem uzun hem de kısa vadeli stratejik planlama için çözümler sunmaktadır.

Kısa ve uzun vadeli stratejilerin analizlerine bakıldığında zaman birden çok analiz çeşidi uygulanabileceği gözlemlenmektedir. Bu kapsamda bakıldığında zaman yapılabilecek analiz metodlarının başlıcaları “Makine öğrenimi tabanlı segmentasyon, Öngörücü analitik, Sosyal medya - web analitiği, RFM ve CLTV” bu analiz yöntemlerinin her birine baktığımız zaman karşımıza çıkan sonuçlarda neden RFM ve CLTV metodlarını kullandımıza kısaca değerlendirelim.

Makine Öğrenimi Tabanlı Segmentasyon, daha karmaşık ve dinamik sonuçlar sunmaktadır. Ancak çok fazla veri hazırlığı, algoritma becerileri ve hesaplama gücü gerektirir (Xu & Tian, 2015). Bu analiz metodunda çok çeşitli ve geniş veri kümeleri gerektirdiği gözlemlenmiştir. Bu nedenle eksik veya dengesiz veri durumlarının olduğu senaryolarda güvenilir sonuçlar vermeyebilir (Han, Kamber, & Pei, 2011).

Öngörücü Analitik, gelecekteki davranışları tahmin etme yeteneğine sahiptir, ancak istatistiksel modelleme ve gelişmiş veri analitiği altyapısı gerektirir, gelecek odaklı kararlar için etkilidir, ancak somut segmentasyon yerine daha çok istatistiksel tahminlere dayanır (Siegel, 2013).

Sosyal Medya ve Web Analitiği, müşteri davranışının dijital platformlarda ve sosyal medyada nasıl şekillendiğini anlamak için faydalıdır. Bununla birlikte, bu yöntemler genellikle doğrudan finansal ölçütlere dayalı olmayan dolaylı içgörüler sağlar (Liu & Shih, 2005). Bu analiz yönteminde genellikle büyük veri altyapısı ve API'ler gerektirmektedir. Bu da uygulamalarda ve erişim maliyetleri küçük veya orta ölçekli işletmeler için yüksek Olması anlamına gelmektedir (Chaffey & Ellis-Chadwick, 2019).

RFM ve CLTV analiz yöntemlerini tercih etmemizin temel sebebi, bu

yöntemlerin basitlik ve uygulanabilir olması, doğrudan eyleme geçirilebilir içgörüler sunması, daha minimal veri gereksinimi istemesi ve literatürde geniş kabul görmüş olması gibi avantajlar sunmaktadır. Diğer analiz yöntemleri, daha karmaşık ve spesifik veri analitiği ihtiyaçları için uygun olsa da, bu projelerde belirlenen zaman, kaynak ve veri kısıtları göz önünde bulundurulduğunda RFM ve CLTV yöntemleri en uygun analiz yöntemleri olduğu anlaşılmaktadır.

Kısa ve uzun vadeli stratejilerin dengesi benzer şekilde, literatürde RFM analizi kısa vadeli pazarlama stratejilerinde etkili bir araç olarak kullanılmaktadır. Ancak CLTV analizi uzun vadeli müşteri ilişkileri yönetimi için temel bir araç olarak önerilmektedir (Kim ve diğ., 2006). Bu çalışma, her iki analitik aracın nasıl birlikte çalıştıklarını ve bu stratejilerin etkilerini incelemektedir.

Müşteri Segmentasyonu: RFM ve CLTV bazlı analizler, sadık müşteriler (loyal customers), yeni müşteriler (new customers) ve kayıp riski taşıyan müşteriler (at-risk customers) gibi farklı grupları tanımlamaktadır (Venkatesan ve Kumar, 2004). Bu çalışma ayrıca benzer segmentasyon yaklaşımları kullanarak her segment için özel tavsiyelerde bulunmuştur.

Veri Büyüklüğü ve Veri İşleme Teknikleri: Bu çalışma, 500.000 gözlem içeren bir veri seti üzerinde yoğun veri işleme, keşfedici veri analizi ve veri temizleme yöntemlerini kullanılmıştır. Literatürde bu büyüklükteki veri setlerine odaklanan pek çok çalışma yoktur ve bu çalışmalar genellikle daha küçük örneklemeler kullanılarak yapılmıştır (Fader ve diğ., 2005).

CLTV Formülünün Özelleştirilmesi: Şirket kâr marjı veya "kâr marjı" ve CLTV hesaplaması için özel metriklerin kullanılması, bu çalışmayı literatürdeki normal hesaplamalardan ayırmaktadır (Benoit ve Van den Poel, 2009).

Pragmatik Çözümler ve Tavsiyeler: Bu tez, e-ticaret sektörü için uygulanabilir pazarlama stratejileri ve müşteri yönetimi önerileri sunarak literatürdeki diğer çalışmalardan fark yaratacaktır.

4.3 Öneriler

Yaptığımız analizler neticesinde RFM ve CLTV analizlerinin yapılabilmesi için minimum değişken sayısı yedi olmalıdır. Bu analizlerin yapılabilmesi için gerekli olan

değişkenler “customer id(müşteri numarası), invoice(fatura numarası), stock code (Ürün kodu), stock description (ürün adı ya da açıklaması), quantity(miktar), price(birim tutar), invoice date (fatura tarihi)” bu değişkenler yapılacak data setimizde mevcutsa RFM ve CLTV analizlerinin yapılabildiği gözlemlenmiştir.

Data setimize baktığımızda sekiz değişken kullandığımız görülmektedir. Sekizinci değişken ülke bilgisidir. Yaptığımız analizler demografik özellikler bağımsız bir segmentasyon yapısı kurarak RFM ve CLTV analiz yapısı kurmamızdan dolayı ülke bilgisi kullanılmamıştır.

Yapılan analizlerde müşterilerin demografik yapılarını göz önünde bulundurarak analizler yapılmak istenirse ilgili data setine yukarıda olmazsa olmaz yedi değişkenin yanına müşterilerin “yaş, cinsiyet, eğitim durumu, medeni hali, meslek bilgisi, ülke, şehir” bu bilgiler eklenerek daha detaylı olarak segmentasyon işlemleri yapılarak araştırmak istenilen özelliklere göre detay analizler yapılmasını sağlayacaktır.

RFM analizinde kurguladığımız regex yapımızda on segmente bölerek analiz çıktılarını oluşturduğum görülmektedir. On segment üzerinden analiz yapma sebebimiz satın alma sıklığı (frequency), en son satın alma zamanı (recency) ve harcanan miktar (monetary) hakkında ayrıntılı bir analiz sağladığı görülmektedir. İşletmeler, bu segmentlerin her biri, müşteri davranışları açısından belirli bir özelliği yansıttığı gözlemlenmiştir (Hughes, 1994). RFM analizi literatürde genellikle on standart segmentten oluşturulduğu gözlemlenmiştir. Bunun ana nedeni, müşteri segmentasyonu için yaygın olarak kullanılan bir yaklaşım olmasıdır (Fader, Hardie, & Lee, 2005). Müşteri ilişkileri yönetimi (CRM) alanında, bu segmentasyon yapısı yaygın olarak kullanılan bir temel sağladığı görülmektedir (Blattberg, Kim & Neslin, 2008).

On segmentli yapı, hem yönetilebilir bir basitlik sağlar hem de müşteri davranışlarını anlamada yeterli detay sağlamaktadır. Yeni segmentler oluşturmak, analiz daha karmaşık hale getirmektedir, stratejik aksiyon almayı zorlaştırmaktadır. Yönetim ve operasyonel anlamda daha fazla sorumluluk yüklediği gözlemlenmiştir. Bu nedenle on parça, hem akademik hem de iş uygulamalarında ideal dengeyi sağladığı gözlemlenmiştir (Hughes, 2006).

CLTV analizinde kullandığımız dört segmentli yapı, müşterilerin şirket üzerindeki

değerlerini yüksek, orta ve düşük gibi net kategorilere ayırarak stratejik kararları basit ve net hale getirmektedir. Bu sistem, müşterilerin ne kadar değerli olduğuna bağlı olarak çeşitli planların oluşturulmasını sağlamaktadır. Bu tür sınıflandırma, müşteri değerini segmente etmek için yaygın olarak kullanılan bir yöntemdir (Blattberg, Kim, & Neslin, 2008).

Müşteri segmentasyonunda dört segment üzerinden analizlerin yapılması basitliği korurken, ayrıntıyı sağlamak için yeterli olduğu gözlemlenmiştir. Daha fazla segmente yapısı kurgulamak analiz süreçlerini daha karmaşık hale getirmektedir. Daha az segment yapısı ile analizleri kurgulamak ise stratejik detayların kaybolmasına neden olabileceği görülmektedir. Literatürde, bu denge, özellikle CLTV analizlerinde önerilmektedir (Hughes, 1994). Bu sistem aynı zamanda kaynak yönetimi ve bütçe planlama için akademik araştırmacılar tarafından önerilen bir yaklaşım olduğu görülmektedir. (Fader, Hardie ve Lee, 2005).

Yeni Müşteriler: Yüksek RFM'ye sahip müşterilere hoş geldin kampanyaları sunarak sadakat oranları artırılabilir. Sadık Müşteriler: Bu müşteri sınıfına özel VIP programları ve ödül programları sunulabilmektedir. Sadakat programlarının müşteri değerini yüzde on ila on beş oranında artıracaktır. (Sharp ve Sharp, 1997).

CLTV Bazlı Fiyatlandırma: Yüksek CLTV'ye sahip müşteriler için özel fiyatlandırma planları oluşturabilir. Yukarı Satış ve Çapraz Satış Stratejileri: Bu stratejiler CLTV'yi %20 oranında artırma potansiyeline sahip olacaktır. (Hughes, 1994).

CLTV ve RFM segmentasyonundan elde edilen veriler, e-posta kampanyaları ve dijital reklam stratejilerini en üst düzeye çıkarmak için kullanılabilir. Bu tekniklerin dönüşüm oranlarını %5 ila 10 oranında artırdığı kanıtlanmıştır (Blattberg ve diğerleri, 2008).

Çalışmanın bulgularına göre, RFM analizinde ki segmentlere göre stratejik önerilere baktığımız zaman "Champions" ve "Potential Loyalist", sepet tutarları en yüksek olan müşteri gruplarıdır. Şirket, bu müşterileri kaybetmemek ve ilişkilerini güçlendirmek için çalışmalıdır. Bu grubun müşterileri, farklı kampanyalar ve indirimlerle ödüllendirilebilir. Bununla birlikte, özel günlerde müşterilere hediyeler vererek sadakatleri artırılabilir.

Şirketin ürün ve uygulamalarını tanıtmak için "New Customer" programları düzenlenebilir. Bu programlar, çok sayıda yeni müşteriye içeren "Yeni Müşterileri alıştırmaya" adı altında yapılacak yeni çalışmalar alışveriş sıklıklarını ve sadakatini artırabilir. Şirket, ilgili kümedeki müşterilere kendilerini hatırlatacak e-postalar göndererek ya da telefonla iletişim kurarak ürünleri ve kampanyaları tanıtabilir.

"Can't Loose" grubu, çok uzun süre önce alışveriş yapmış ama alışveriş sıklığı yüksek olan bir gruptur. Bu müşteri grubunun şirketi terk etme olasılığı yüksek olduğundan, onlara özel olarak çeşitli sadakat çalışmaları yapılarak tekrar kazanımları sağlanmalıdır. Alternatif olarak, müşterilerin "Need Attention" ve "About to sleep" kategorilerindeki müşterilere özel indirimler sunulabilir ya da önceki sepetlerini inceleyerek ürünler hakkında tavsiyelerde bulunarak onları yeniden satın alma konusunda teşvik edilebilir.

"Hibernating" grubu, uzun süredir alışveriş yapmamış ve şirketi pek tanımayan müşterilerden oluşur. Bu kümedeki müşteriler için özelleştirilmiş reklamlar yapılabilir. Alternatif olarak, müşterilere şirketi tanımları için ürünleri tanıtan ve kupon kodu içeren bir email gönderilebilir. Alışveriş sıklığı yüksek ancak sepet tutarı düşük olan müşteriler, "Loyal Customer" grubuna girer. Bu müşteri grubunun sepet tutarını artırmanın bir yolu, popüler ürünleri ucuza sunmaktır.

"Promising" grubu, son zamanlarda alışveriş yapan müşterilerden oluşur. Bu kümedeki müşterilere özel indirimler sunarak sık sık ve büyük miktarda alışveriş yapabilen gruptur.

Uzun süredir alışveriş yapmayan müşteriler, "About to Sleep" kategorisinde yer almaktadır. Bu gruptaki tüketicilere özellikle dikkat etmek gerekir. Bu müşteri grubuna özel indirimler ve ürün tavsiyeleri sağlanabilir.

CLTV analizimizde oluşturduğumuz segmentlere yönelik önerilere baktığımız zaman "Grup A" en değerli müşterileri grubumuz olduğu görülmektedir. Hem çok fazla para harcadıkları hem de sık sık alışveriş yaptıkları gözlemlenmiştir. Sadakat programları olarak özel indirimler, VIP üyelikleri ve hediye kuponları ile müşteri sadakati arttırılabilir. Doğum günü kutlama ve özelleştirilmiş e-posta kampanyaları gönderilebilir. Müşteri hizmetleri, hızlı teslimat veya ürün lansmanlarına öncelik verilebilir. Tamamlayıcı ürünler veya daha yüksek fiyatlı ürünler sunarak harcamaları

arttırılabilir. Müşterilerden gelen geri bildirimler alınması, ürün ve hizmetlerinizin geliştirilmesine yardımcı olabilir.

“Grup B” potansiyel Yüksek Değerli Müşteriler (CLTV Orta-Yüksek) Bu grup alışveriş yaptıkları gözlemlenen ancak harcamalarını artırma potansiyeline sahip olan grubumuzdur. Özelleştirilmiş Kampanyalar, bazı ürünler için ilginç indirimler yapılabilir. Bu grupta daha fazla bağlantı kurmak için blog yazıları, e-kitaplar veya ürün kullanım kılavuzları oluşturulabilir. Belirli bir harcama düzeyine ulaştıklarında ödüller kazanacakları programlar oluşturulabilir. Sepette kalan mallar veya özel kampanyalar için düzenli e-posta bilgilendirmeleri gönderilebilir.

"Grup C" orta değere sahip müşterilerdir (CLTV orta-düşük) ve daha az alışveriş yaptıkları görülmektedir. Düşük riskli indirimler, tüketicileri tekrar satın almaya teşvik edebilir. Daha önce satın aldıkları ürünlere benzer ürün önerileri yapılabilir. Bu grubu aktif hale getirmek için sosyal medya kampanyalarını finanse etmek mümkündür. İlk olarak, onlara özel "tekrar hoş geldiniz" hediyeleri verilebilir.

"Grup D" Düşük Değerli Müşteriler (CLTV Düşük) olarak bilinir. Bu müşteriler nadiren alışveriş yapar ve şirketiniz için düşük kâr sağlamaktadır. Bununla birlikte, uygun taktikler kullanılarak yeni müşteriler kazanılabilir. Uzun süredir alışveriş yapmayan müşteriler için özel bir kupon veya indirim sunulabilir. Düşük fiyatlı ürünler için düşük maliyetli indirimler sunarak bu müşterilerin tekrar alışveriş yapması sağlanabilir. Yeniden hedefleme veya retargeting olarak da bilinen reklam kampanyaları, ziyaretçilerin siteye geri dönüşlerini artırmak hedeflenebilir. Onlara ucuz ve düşük riskli mallar vererek alışveriş alışkanlığı kazandırmak mümkündür.

KAYNAKÇA

- Bayram, B. (2020). RFM Analizi ile Müşteri Segmentasyonu ve Örnek Proje. Medium.<https://brcubayram.medium.com/rfm-analizi-ilem%C3%BC%C5%9Fteri-segmentasyonu-ve-%C3%B6rnek-proje-2c34135c06e6> adresinden edinilmiştir.
- Blattberg, R. C., Kim, B.-D., ve Neslin, S. A. (2008). Database Marketing: Analyzing and Managing Customers. Springer.
- Blattberg, R. C., ve Deighton, J. (1996). "Manage Marketing by the Customer Equity Test." Harvard Business Review.
- Boulding, W., Kalra, A., ve Zeithaml, V. A. (2005). "A Dynamic Process Model of Service Quality: From Expectations to Behaviors." Journal of Marketing Research, 42(1), 36-49.
- Brynjolfsson, E., ve McAfee, A. (2014). The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies. W.W. Norton ve Company.
- Buttle, F. (2009). Customer Relationship Management: Concepts and Tools (2nd ed.). Elsevier.
- Cacioppo, J. T., ve Patrick, W. (2008). Loneliness: Human Nature and the Need for Social Connection. W.W. Norton ve Company.
- Chaffey, D., & Ellis-Chadwick, F. (2019). Digital Marketing: Strategy, Implementation and Practice. Pearson Education.
- Chaudhuri, S., Dayal, U., ve Narasayya, V. (2011). "An Overview of Business Intelligence Technology." Communications of the ACM, 54(8), 88-98.
- Chen, J., Chen, J., ve Lee, C. (2009). "Customer Segmentation and Loyalty in Online Stores: Evidence from Taiwan." International Journal of Electronic Commerce, 13(4), 87-114.

- Corti, L., Eynden, V. V., Bishop, L., ve Woollard, M. (2014). *The Data Life Cycle in Practice: A Framework for Research Data Management and Sharing*. Essex University.
- Fader, P. S., Hardie, B. G., ve Lee, K. L. (2005). "Counting Your Customers" the Easy Way: An Alternative to the Pareto/NBD Model." *Marketing Science*, 24(2), 275-284.
- Farris, P. W., Bendle, N. T., Pfeifer, P. E., ve Reibstein, D. J. (2010). *Marketing Metrics: The Definitive Guide to Measuring Marketing Performance*. Pearson Education.
- Gupta, S., & Lehmann, D. R. (2003). Customers as Assets. *Journal of Interactive Marketing*.
- Gupta, S., ve Lehmann, D. R. (2005). *Managing Customers as Investments: The Strategic Value of Customers in the Long Run*. Wharton School Publishing.
- Hughes, A. M. (1994). *Strategic Database Marketing: The Masterplan for Starting and Managing a Profitable Customer-Based Marketing Program*. McGraw-Hill.
- Hughes, A. M. (2011). *Strategic Brand Management: A Customer Experience Approach*. Sage.
- IBM (2015). *The Data Lifecycle and Management for Big Data*. IBM Corporation.
- Katal, A., Wazid, M., ve Goudar, R. H. (2013). "Big Data: Issues, Challenges, and Applications." *Future Generation Computer Systems*, 29(4), 258-268.
- Kim, B.-D., Sun, B., ve Kim, T. (2006). "A Model for Evaluating the Financial Value of New Customers in a Networked Environment." *Journal of Marketing Research*, 43(3), 373-387.

- Kotler, P., ve Keller, K. L. (2016). *Marketing Management* (15th ed.). Pearson Education.
- Kumar, V., Shah, R., ve Gupta, S. (2006). "Customer Lifetime Value in a Dynamic Setting: A Model and Application." *Journal of Retailing*.
- Kumar, V., ve Shah, D. (2004). "Building and Sustaining Profitable Customer Loyalty for the 21st Century." *Journal of Retailing*, 80(4), 317-329.
- Liu, Y., & Shih, M. (2005). Hybrid approaches to segmenting customers in customer relationship management. *Expert Systems with Applications*.
- Manyika, J., Chui, M., Brown, B., ve Bughin, J. (2011). "Big Data: The Next Frontier for Innovation, Competition, and Productivity." McKinsey Global Institute.
- Mayer-Schönberger, V., ve Cukier, K. (2013). *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. Houghton Mifflin Harcourt.
- Oracle (2014). *Managing the Data Life Cycle: From Creation to Disposal*. Oracle Corporation.
- Payne, A., ve Frow, P. (2005). "A Strategic Framework for Customer Relationship Management." *Journal of Marketing*, 69(4), 167-176.
- Peppers, D., ve Rogers, M. (2011). *Managing Customer Relationships: A Strategic Framework* (2nd ed.). Wiley.
- Rust, R. T., Lemon, K. N., & Zeithaml, V. A. (2004). Return on Marketing: Using Customer Equity to Focus Marketing Strategy. *Journal of Marketing*.
- Rust, R. T., Zeithaml, V. A., ve Lemon, K. N. (2000). *Driving Customer Equity: How Customer Lifetime Value Is Reshaping Corporate Strategy*. Free Press.
- SAS Institute (2013). *The Analytics Life Cycle*. SAS Institute.

- Shankar, V., Smith, A. K., ve Rangaswamy, A. (2003). "Customer Satisfaction and Loyalty in Online and Offline Environments." *International Journal of Research in Marketing*, 20(2), 153-175.
- Sharma, S., ve Sheth, J. N. (2010). "Resurgence of Digital Marketing: Technology and Customer Centricity." *Journal of Marketing Theory and Practice*, 18(2), 99-112.
- Stone, M. (1995). *Marketing Information Systems*. Butterworth-Heinemann.
- Venkatesan, R., ve Kumar, V. (2004). "A Customer Lifetime Value Framework for Customer Selection and Resource Allocation Strategy." *Journal of Marketing*, 68(4), 106-125.
- Westerman, G., Calm ejane, C., Ferraris, P., ve Bonnet, D. (2011). *Digital Transformation: A Roadmap for Billion-Dollar Organizations*. MIT Center for Digital Business.
- Xu, X., & Tian, Y. (2015). A Comprehensive Survey of Clustering Algorithms. *Annals of Data Science*.
- YBS Ansiklopedi. (2015). B uy k Veri ve Yařam D ng s . YBS Ansiklopedi. https://ybsansiklopedi.com/wpcontent/uploads/2015/08/buyukveri_veya_samdongusu.pdf adresinden edinilmiřtir.