

**T.C.
SAKARYA ÜNİVERSİTESİ
İŞLETME ENSTİTÜSÜ**

**DEVLET ANALİTİĞİ BAĞLAMINDA İŞ
İLANLARININ METİN MADENCİLİĞİ
YÖNTEMİYLE ANALİZİ VE COĞRAFİ BİLGİ
SİSTEMLERİNE ENTEGRASYONU**

**YÜKSEK LİSANS TEZİ
Burak BULDU**

Enstitü Anabilim Dalı : Yönetim Bilişim Sistemleri

Tez Danışmanı: Doç. Dr. Adem AKBIYIK

Ağustos 2020

Burak Buldu tarafından hazırlanan “Devlet Analitiđi Bađlamında İş İlanlarının Metin Madenciliđi Yöntemiyle Analizi Ve Cođrafi Bilgi Sistemlerine Entegrasyonu” bařlıklı bu tez, 24/08/2020 tarihinde Sakarya Üniversitesi Lisansüstü Eđilim ve Öğretim Yönetmeliđi'nin ilgili maddeleri uyarınca yapılan Tez Savunma Sınavı sonucunda bařarılı bulunarak, jürimiz tarafından Yüksek Lisans olarak kabul edilmiřtir.

Danıřman: Doç. Dr. Adem Akbıyık
Sakarya Üniversitesi

Jüri Üyeleri: Prof. Dr. Aykut Hamit Turan
Sakarya Üniversitesi

Prof. Dr. Ayhan Demiriz
Gebze Teknik Üniversitesi



SAKARYA
ÜNİVERSİTESİ

T.C.
SAKARYA ÜNİVERSİTESİ
İŞLETME ENSTİTÜSÜ
TEZ SAVUNULABİLİRLİK VE ORJİNALLİK BEYAN FORMU

Sayfa : 1/1

Öğrencinin

Adı Soyadı	:	BURAK BULDU
Öğrenci Numarası	:	Y189054013
Enstitü Anabilim Dalı	:	YÖNETİM BİLİŞİM SİSTEMLERİ
Enstitü Bilim Dalı	:	
Programı	:	<input checked="" type="checkbox"/> YÜKSEK LİSANS <input type="checkbox"/> DOKTORA
Tezin Başlığı	:	DEVLET ANALİTİĞİ BAĞLAMINDA İŞ İLANLARININ METİN MADENCİLİĞİ YÖNTEMİYLE ANALİZİ VE COĞRAFİ BİLGİ SİSTEMLERİNE ENTEGRASYONU
Benzerlik Oranı	:	% 2

İŞLETME ENSTİTÜSÜ MÜDÜRLÜĞÜNE,

Sakarya Üniversitesi İşletme Enstitüsü Lisansüstü Tez Çalışması Benzerlik Raporu Uygulama Esaslarını inceledim. Enstitünüz tarafından Uygulama Esasları çerçevesinde alınan Benzerlik Raporuna göre yukarıda bilgileri verilen tez çalışmasının benzerlik oranının herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi beyan ederim.

11/08/2020

B.Buldu

Sakarya Üniversitesi İşletme Lisansüstü Tez Çalışması Benzerlik Raporu Uygulama Esaslarını inceledim. Enstitünüz tarafından Uygulama Esasları çerçevesinde alınan Benzerlik Raporuna göre yukarıda bilgileri verilen öğrenciye ait tez çalışması ile ilgili gerekli düzenleme tarafımda yapılmış olup, yeniden değerlendirilmek üzere@sakarya.edu.tr adresine yüklenmiştir.

Bilgilerinize arz ederim.

...../...../20.....

İmza

Uygundur

Danışman
Unvanı / Adı-Soyadı: Doç. Dr. Adem AKBIYIK

Tarih: 11.08.2020

İmza: *Adem Akbiyik*

KABUL EDİLMİŞTİR

REDDEDİLMİŞTİR

EYK Tarih ve No:

Enstitü Birim Sorumlusu Onayı

ÖNSÖZ

Bu tez çalışmasının başından itibaren tüm aşamalarında, çalışmamı titizlikle takip eden, emeğini ve değerli vaktini esirgemeyen katkılarından dolayı danışmanım Doç. Dr. Adem AKBIYIK'a en içten teşekkürlerimi ve saygılarımı sunarım. Bu tezde desteği ve sağladığı katkılardan dolayı Yönetim Bilişim Sistemleri'nde görev yapan Arş. Gör. Bahadır AKTAŞ'a teşekkür ederim. Lisans ve yüksek lisansım boyunca desteklerini her zaman hissettiren ve değerli vakitlerini paylaşan Sakarya Üniversitesi Yönetim Bilişim Sistemleri bölümünde görevi devam eden ve geçmişte görev yapmış tüm hocalarıma ve asistanlarına teşekkür ederim. Yüksek lisans ders dönemimde Sakarya Üniversitesi Teknokent'te birlikte çalışma fırsatı bulduğum, maddi ve manevi desteğini paylaşan değerli hocam Prof. Dr. Ayhan DEMİRİZ'e ve eski çalışma arkadaşlarıma teşekkürlerimi sunarım. Yüksek öğrenim hayatımın her aşamasında desteğini fazlasıyla gösteren aileme ve bu süreçlerde yanımda olan sevgili arkadaşım Feyza Yıldız YURTAL ve Zonguldak Bülent Ecevit Üniversitesi'ndeki görevli değerli dostum Arş. Gör. Murat VARUL'a çok teşekkür ediyorum. Son olarak tüm bu süreçlerde manevi desteğini esirgemeyen ve hep yanımda olduklarını hissettiren adını saymadığım tüm arkadaşlarıma şükranlarımı sunarım.

Burak BULDU

11.08.2020

İÇİNDEKİLER

KISALTMALAR	iii
TABLO LİSTESİ	iv
ŞEKİL LİSTESİ	v
ÖZET	viii
ABSTRACT	ix

GİRİŞ	1
BÖLÜM 1: DEVLET ANALİTİĞİ	5

1.1. Devlet Analitiği Kavramı	5
1.1.1. Devlet Analitiğinin Tanımlanması	6
1.2. Devlet Analitiği Alanındaki Çalışmalar	8
1.2.1. ABD'deki Çalışmalar	8
1.2.2. Avustralya'daki Çalışmalar	8
1.2.3. BAE'deki Çalışmalar	9
1.2.4. Alandaki Çalışmalarının Değerlendirilmesi	11
1.3. Devlet Analitiğinin Unsurları ve Çerçevesi	12
1.3.1. Ulusal Beşeri Sermaye Analitiği	14
1.3.1.1. İK Analitiklerinin Tarihçesi	15
1.3.1.2. İK Analitiklerinde Büyük Verinin Kullanımı	18
1.3.1.3. Ulusal Beşeri Sermaye Analitiğinin Kapsamı	19

BÖLÜM 2: ARAŞTIRMANIN YÖNTEMİ	24
--	-----------

2.1. Araştırma Yöntemindeki Yaklaşımlar	24
2.1.1. Metin Madenciliği	24
2.1.1.1. Metin Madenciliğinin Tarihçesi	24
2.1.1.2. Metin Madenciliği Kavramları	25
2.1.2. Coğrafi Bilgi Sistemleri	30
2.2. Araştırma Modelinin Ve Veri Setinin Hazırlanması	31
2.2.1. İş İlanlarının İnternet Üzerinden Toplanması	31
2.2.2. Araştırmanın Modeli	35
2.2.3. Veri Setinin Analizlere Hazırlanması	35

BÖLÜM 3: ARAŞTIRMANIN ANALİZ VE BULGULARI.....	41
3.1. Frekans Analizleri	41
3.2. Sık Kullanılan Kelime Grupları.....	58
3.3. Çıkarılmış Varlık İsimleri.....	70
3.4. Çapraz Tablo Analizleri	73
3.5. Coğrafi Analizler	78
SONUÇ VE ÖNERİLER.....	88
KAYNAKÇA	92
ÖZGEÇMİŞ.....	96



KISALTMALAR

ABD	: Amerika Birleşik Devletleri
ACT	: Avustralya Başkent Bölgesi
ACIT	: İleri Teknolojiler Araştırma Şirketi, Dubai, BAE
AR-GE	: Araştırma ve Geliştirme
BAE	: Birleşik Arap Emirlikleri
CAGS	: İleri Devlet Araştırmaları Merkezi, Maryland, ABD
CBS	: Coğrafi Bilgi Sistemleri
EFQM	: Avrupa Kalite Yönetimi Vakfı
İK	: İnsan Kaynakları
KDS	: Karar Destek Sistemleri
MIC	: Mubadala Yatırım Şirketi, Abu Dabi, BAE
TANMIA	: Ulusal İnsan Kaynakları ve İstihdam Kurumu, Dubai, BAE
TMS	: Yetenek Yönetim Sistemi
YBS	: Yönetim Bilişim Sistemleri
YY	: Yetenek Yönetimi

TABLO LİSTESİ

Tablo 1: Literatürdeki İnsan Kaynakları Analitiği İsimlendirmeleri	16
Tablo 2: Popüler Metin Madenciliği Algoritmaları (Intellspot, 2020).....	29
Tablo 3: Çıkarılmış Varlık İsimleri	71
Tablo 4: Çapraz Tablo Analizi	74
Tablo 5: En Çok Ve En Az Geçen Nitelikler	77



ŞEKİL LİSTESİ

Şekil 1: Beşeri Sermaye Analitiği Çerçevesi.....	20
Şekil 2: Yetenek Keşfetme Analizlerinin Devlet Analitiği Hiyerarşisi.....	23
Şekil 3: DIKW Piramidi	26
Şekil 4: Metin Madenciliği Sürecinin Aşamaları	29
Şekil 5: Veri Madenciliği Sürecinin Aşamaları.....	30
Şekil 6: Veri Kazıma Algoritmasının Akış Şeması	33
Şekil 7: Kariyer.Net veri kazıma işlemleri akış şeması.....	34
Şekil 8: Araştırmanın KDS modeli.....	35
Şekil 9: Enlem Ve Boylam Bilgileri İndirme Betiği	37
Şekil 10: Firma Bilgisi Olmayan İlan içeriklerini İndirme Betiği	38
Şekil 11: İllere Göre Yatırım Teşvik Bölgeleri Haritası.....	39
Şekil 12: Frekans Tablosu Kelime Bulutu.....	41
Şekil 13: Analitik – Yatırım Teşvik Bölgesi Dağılımı	42
Şekil 14: Analitik – Aylık Dağılımı	42
Şekil 15: Bilgisayar – Yatırım Teşvik Bölgesi Dağılımı.....	43
Şekil 16: Bilgisayar – Aylık Dağılımı	43
Şekil 17: Elektronik – Yatırım Teşvik Bölgesi Dağılımı	44
Şekil 18: Elektronik – Aylık Dağılımı.....	44
Şekil 19: Gayrimenkul – Yatırım Teşvik Bölgesi Dağılımı	45
Şekil 20: Gayrimenkul – Aylık Dağılımı.....	45
Şekil 21: Gıda – Yatırım Teşvik Bölgesi Dağılımı	46
Şekil 22: Gıda – Aylık Dağılımı.....	46
Şekil 23: Kalite– Yatırım Teşvik Bölgesi Dağılımı	47
Şekil 24: Kalite – Aylık Dağılımı.....	47
Şekil 25: Lojistik – Yatırım Teşvik Bölgesi Dağılımı.....	48
Şekil 26: Lojistik – Aylık Dağılımı	48
Şekil 27: Mağazacılık – Yatırım Teşvik Bölgesi Dağılımı	49
Şekil 28: Mağazacılık – Aylık Dağılımı.....	49
Şekil 29: Muhasebe – Yatırım Teşvik Bölgesi Dağılımı.....	50
Şekil 30: Muhasebe – Aylık Dağılımı	50
Şekil 31: Mühendislik– Yatırım Teşvik Bölgesi Dağılımı.....	51

Şekil 32: Mühendislik – Aylık Dağılımı	51
Şekil 33: Otomotiv – Yatırım Teşvik Bölgesi Dağılımı	52
Şekil 34: Otomotiv – Aylık Dağılımı	52
Şekil 35: Perakende – Yatırım Teşvik Bölgesi Dağılımı.....	53
Şekil 36: Perakende – Aylık Dağılımı	53
Şekil 37: Sağlık – Yatırım Teşvik Bölgesi Dağılımı	54
Şekil 38: Sağlık – Aylık Dağılımı	54
Şekil 39: Üretim – Yatırım Teşvik Bölgesi Dağılımı.....	55
Şekil 40: Üretim – Aylık Dağılımı	55
Şekil 41: Yazılım – Yatırım Teşvik Bölgesi Dağılımı	56
Şekil 42: Yazılım – Aylık Dağılımı.....	56
Şekil 43: Yenilikçi – Yatırım Teşvik Bölgesi Dağılımı	57
Şekil 44: Yenilikçi – Aylık Dağılımı.....	57
Şekil 45: Kelime Grupları Kelime Bulutu	58
Şekil 46: AR-GE – Yatırım Teşvik Bölgesi Dağılımı	59
Şekil 47: AR-GE – Aylık Dağılımı	60
Şekil 48: Dış Ticaret – Yatırım Teşvik Bölgesi Dağılımı	60
Şekil 49: Dış Ticaret – Aylık Dağılımı.....	61
Şekil 50: Elektrik-Elektronik – Yatırım Teşvik Bölgesi Dağılımı	61
Şekil 51: Elektrik-Elektronik – Aylık Dağılımı.....	62
Şekil 52: E-Ticaret – Yatırım Teşvik Bölgesi Dağılımı	62
Şekil 53: E-Ticaret – Aylık Dağılımı.....	63
Şekil 54: Gayrimenkul Değerleme – Yatırım Teşvik Bölgesi Dağılımı.....	63
Şekil 55: Gayrimenkul Değerleme – Aylık Dağılımı	64
Şekil 56: İngilizce Bilen – Yatırım Teşvik Bölgesi Dağılımı	64
Şekil 57: İngilizce Bilen – Aylık Dağılımı	65
Şekil 58: İnsan Kaynakları – Yatırım Teşvik Bölgesi Dağılımı.....	65
Şekil 59: İnsan Kaynakları – Aylık Dağılımı	66
Şekil 60: İş Sağlığı ve Güvenliği – Yatırım Teşvik Bölgesi Dağılımı	66
Şekil 61: İş Sağlığı ve Güvenliği – Aylık Dağılımı.....	67
Şekil 62: Part-Time – Yatırım Teşvik Bölgesi Dağılımı.....	67
Şekil 63: Part-Time – Aylık Dağılımı	68
Şekil 64: Satın Alma – Yatırım Teşvik Bölgesi Dağılımı	68

Şekil 65: Satın Alma – Aylık Dağılımı	69
Şekil 66: Sosyal Medya – Yatırım Teşvik Bölgesi Dağılımı	69
Şekil 67: Sosyal Medya – Aylık Dağılımı	70
Şekil 68: Çapraz Tablo Analizi Kısıtları	73
Şekil 69: Benzeşme Alan Grafiği	76
Şekil 70: AR-GE Sıcaklık Haritası	79
Şekil 71: Gayrimenkul Değerleme Sıcaklık Haritası.....	79
Şekil 72: İş Sağlığı ve Güvenliği Sıcaklık Haritası	80
Şekil 73: Lojistik Sıcaklık Haritası.....	81
Şekil 74: Mağazacılık Sıcaklık Haritası	81
Şekil 75: Otomasyon Sıcaklık Haritası.....	82
Şekil 76: Otomotiv Sıcaklık Haritası.....	82
Şekil 77: Part-Time Sıcaklık Haritası	83
Şekil 78: Satın Alma Sıcaklık Haritası	84
Şekil 79: Savunma Sıcaklık Haritası	84
Şekil 80: Sosyal Medya Sıcaklık Haritası	85
Şekil 81: Tekstil Sıcaklık Haritası	86
Şekil 82: Turizm Sıcaklık Haritası	86
Şekil 83: Yazılım ve Yazılım Geliştirme Sıcaklık Haritası.....	87

Tezin Başlığı: Devlet Analitiği Bağlamında İş İlanlarının Metin Madenciliği Yöntemiyle Analizi Ve Coğrafi Bilgi Sistemlerine Entegrasyonu	
Tezin Yazarı: Burak BULDU	Danışman: Doç. Dr. Adem AKBIYIK
Kabul Tarihi: 11.08.2020	Sayfa Sayısı: vii (ön kısım)+ 95(tez)
Anabilim Dalı: Yönetim Bilişim Sistemleri Bilim Dalı:	
<p>Bilişim sistemlerinde, her geçen gün veri trafiği artmaktadır. Teknolojiyle beraber, veri analizi gereksinimleri ile zorlukları değişmektedir ve büyük veri, veri analitiği gibi alanlar doğmuştur. Analizlerin başarısı, karar vermeyi kolaylaştırmaktadır. Bu açıdan, büyük verinin kullanım alanı yeni analitik yaklaşımlarla artmaya devam etmekte ve karar destek sistemleri bu gelişmelerden etkilenmektedir.</p> <p>Sermaye kuramlarında geçen, işgücü analitikleri mikro ölçekte insan kaynakları alanında çalışılmaktadır. Bu konuda, büyük veri olarak iş ilanları insan kaynakları analitiklerinde kullanılabilir. Dolayısıyla devlet kurumlarına kamu politikalarında fayda üretmek için karar desteği açısından makro bir yaklaşımla bir devlet analitiği çerçevesi oluşturularak, beşeri sermayeye yönelik analitikler gerçekleştirilebilir.</p> <p>Devlet analitiklerinde, sosyal politikalar üretmek üzere yurtdışında birçok faaliyet yürütülmektedir. Günümüzde araştırmacıların sosyal medya analizleri ile çalıştığı bu veri analizi yaklaşımı, siyaset bilimi ve uluslararası ilişkiler alanlarında yardımcı olmaktadır. Ancak makro bakış açısıyla geliştirilebilecek yeni analitik modeller farklı kamu alanlarında istihdam, göç, nüfus, şehirleşme, sanayileşme, eğitim v.b. konularda karar desteği sağlayabilir.</p> <p>İş ilanları, özellikle işgücü piyasalarını anlamada ve istihdam politikalarında başta olmak üzere birçok değerli bilgiler barındırabilir. Bu bağlamda, devlet analitiği için insan kaynakları analitiklerinde kullanılan modeller uyarlanabilir. Tez çalışmasının yönteminde kullanılan metin madenciliği ve coğrafi bilgi sistemleri teknikleri ile iş ilanları analiz edilmiştir ve bu yöntemde oluşturulan makro politikalara uygun karar destek sistemi modeli için deneysel çalışma yapılmıştır.</p> <p>Nihayetinde, devlet analitiği ve insan kaynakları analitiği gibi kavramlara yeni bir bakış açısı getirilmiş ve kullanılan modellerde yeni teknikler uyarlanarak farklı bir boyut kazandırılmıştır. Metin madenciliği ve coğrafi bilgi sistemi yaklaşımları ile kamusal alanda büyük veri araştırmalarına rehber oluşturabilecek karma bir karar destek sistemi ortaya koyulmuştur. Ayrıca bu araştırma, ortaya koyduğu yeni analitik kavramlar ile yönetim bilişim sistemleri alanının çok disiplinli araştırmalardaki konumunu farklı bir yere taşımak adına, makro seviyedeki analizleriyle yeni bir bakış açısı geliştirmiştir.</p>	
Anahtar Kelimeler: Devlet analitiği, İnsan kaynakları analitiği, Metin madenciliği, Coğrafi bilgi sistemleri, Karar destek sistemleri, Karar verme, Büyük veri, İş ilanları	

Title of Thesis: Analysis of Job Postings by Text Mining Method and Integration to Geographic Information Systems In The Context Of Government Analytics	
Author of Thesis: Burak Buldu	Supervisor: Assoc. Prof. Dr. Adem AKBIYIK
Accepted Date: 11.08.2020	Np: vii (pre text) + 95(main body)
Department: Management Information Systems	Subfield:
<p>In information systems, data traffic increases day by day. With technology, data analysis needs and challenges vary, and areas such as big data and data analytics were born. The success of the analysis facilitates decision making. In this respect, the usage area of big data continues to increase with new analytical approaches and decision support systems are affected by these developments.</p> <p>Workforce analytics in the capital theories are researched in the field of human resources on a micro scale. In this topic, job advertisements be used in human resources analytics as big data. Accordingly, a government analytics framework can be created with a macro approach in terms of decision support in order to produce benefits in public policies, and analytics for human capital can be applied.</p> <p>In government analytics, many practices are performed abroad to produce social policies. Today, these data analytics approach that researchers are working with by social media analysis, is helping in the fields of political science and international relations. But new analytical models that can be developed from a macro point of view in public areas include employment, migration, population, urbanization, industrialization, education, etc. can provide decision support on issues.</p> <p>Job postings can contain valuable information, especially in understanding labor markets and employment policies. In this context, models used in human resource analytics for government analytics can be adapted. Job postings were analyzed with text mining and geographic information systems techniques used in the thesis study method and an experimental study has been conducted for the decision support system model in accordance with the macro policies created in this method.</p> <p>In the end, a new perspective has been brought to concepts such as government analytics and human resources analytics, and a different dimension has been gained by adapting new techniques in the models used. A hybrid decision support system has been exhibited that can guide big data research in the public spaces with text mining and geographic information system approaches. In addition, this research has developed a new perspective with its macro-level analysis in order to move its position to somewhere different in the multidisciplinary research of management information systems with the new analytical concepts it has revealed.</p>	
Keywords: Government analytics, Human resources analytics, Text mining, Geographic information systems, Decision support systems, Decision making, Big data, Job postings	

GİRİŞ

İnternetin yaygınlaşmasıyla birlikte, iletişim trafiği ve bilgisayarlarda işlenen veri miktarı artmıştır. Bu trend veri biliminin gelişmesine ve kurumsal bilişim sistemlerinin yayılmasına, böylece yeni veri tabanı teknolojilerinin, veri ambarlarının ve bunlarla birlikte büyük veri kavramının doğuşuna neden olmuştur (Demirool ve diğerleri, 2019). Böylece günümüzde araştırmacılar, kamu kuruluşları ve özel sektör için büyük veri analizleri popüler bir uzmanlık konusu olmuştur.

Günümüzde özel sektörden kamu kurumlarına kadar her türlü organizasyon teknolojik yatırımlar yapmaktadır. İş analitiği ve büyük veri, bu örgütler açısından karar verme problemlerinde kritik çözümlerin geliştirildiği uzmanlıklardır. Bu çözümlere yatırım yapmak için sermaye kavramları planlama süreçlerine dahil olmaktadır. Bu nedenle, karar verme, teknik olduğu kadar sermaye yönetimini de ilgilendirir.

Özel sektörün analitik ihtiyaçları mikro ölçekte gerçekleşirken, devlet için makro seviye analizler ve bu kapsama uygun geniş bir perspektiften bakışa ihtiyaç vardır. Sermaye kuramı ve büyük veri analitiği bu bakış açısını verebilecek büyüklükteki yaklaşımlardır. Devletlerin politikalarını geliştirmek için analitiklere ihtiyacı arttığını literatürde ifade edilmiştir. Böylece, *Devlet Analitiği* kavramı ortaya çıkmıştır (Davenport ve Jarvenpaa 2008; Richards, 2016).

Her örgüt yapısı gibi, özel sektördeki firmalara benzer devletin de sermaye türleri vardır. Bu sermaye türleri, Bourdieu (1986) tarafından açıklanmaktadır. İK analitiği literatür taramasında karşılaşılan mikro çalışmalarda, işgücü yatırımı ve stratejik planlama konularıyla son yıllarda *beşeri sermaye* kavramı öne çıkmaktadır. İşbu beşeri sermaye analitiklerindeki uygulamalar incelendiğinde analitiklerin isimlendirilmesiyle ilgili *literatür karmaşası* tespit edilmiştir.

Bu araştırmada, büyük veri, sermaye türleri ve mikro ölçekte İK analitiği yaklaşımıyla birlikte değerlendirilen devlet analitiği çerçevesi kurulurken ilgili 3 önemli nokta vardır:

- 1) Devlet analitiği kavramını inşa etmek ve kapsamını vermek
- 2) İK analitiklerini makro ölçekte *ulusal beşeri sermaye* çerçevesinde tasarlamak
- 3) İK analitiğindeki literatüründeki karmaşık yapıyı açıklamaktır.

Günümüzde, beşeri sermaye (İK) analitiği büyük veri araştırmalarının çalışılmaya başladığı uzmanlıklardan biridir. Beşeri sermaye analitiği çerçevesinde, İK alanındaki yetenek yönetimi yaklaşımının bir konusu olarak işletmelerde yetenek kazandırma süreci işlenmektedir. Dolayısıyla yetenek yönetimi devlet analitiği çerçevesinde incelenebilir. Literatüre göre, *bu yetenek kazandırma faaliyeti iş ilanlarıyla doğrudan ilişkilidir ve iş ilanları büyük veri olarak kullanılabilir*. Dolayısıyla iş ilanlarının analizi, bu araştırmada ulusal beşeri sermayeyi analiz etmek için devlet analitiğiyle ilişkilendirilmiştir.

İş ilanları metin içerikleri ve konum bilgileri içeren dokümanlardır. Bilimsel çalışmalara başvuru olarak, metin analizleri ve bölgesel analizler için metin madenciliği ve coğrafi bilgi sistemleri araştırmanın yöntemi olarak belirlenmiştir. Bu teknikler, yöneticilere tahmin ve performans değerlendirme yetkinliği kazandıran, karar desteği veren yöntemlerdir. Nihayetinde bu ampirik çalışma, devlet analitiği kavramıyla farklı disiplin ve yöntemleri bir araya getiren fütüristik bir araştırmadır.

Sonuç olarak, devlet analitiği alanında yöneticilerin gereksinimlerine göre karar destek sistemi geliştirilebilir. Bu çalışmaya göre, beşeri sermaye analizleri ölçeğinde iş ilanlarından metin madenciliği ve CBS ile elde edilecek bulgular ile ülkenin yetenek haritası çıkarılabilir. Araştırmanın analizler ile ekonomik kalkınma, eğitim hizmetleri, göç hareketleri, işgücü planlaması, şehirleşme gibi daha birçok farklı konularda kamu yönetiminde değerlendirilebilecek çıktılar üretilebilir.

Araştırmanın Amacı: Bu çalışmanın temel amacı, büyük veri analizinin devlet düzeyindeki karar alım süreçlerinde kullanılabilirliğini, beşeri sermaye analitiği örneği üzerinden ortaya koymaktır. Bu kapsamda çalışmanın 4 hedefi söz konusudur:

- 1) İş ilanlarında talep edilen niteliklerin (bilgi/beceri/yetkinlik) metin madenciliği ile analizleri yapılarak ulusal bir nitelik gereksinim havuzu ortaya koymak.
- 2) Elde edilen niteliklerin Coğrafi Bilgi Sistemleri desteğiyle bölgesel dağılımlarını ortaya koymak.
- 3) Ulusal ve bölgesel düzeyde elde edilen nitelik gereksinim dağılımları dikkate alınarak devlet analitiği bağlamında olası kazanımları belirlemek
- 4) İlgili alanda Türkçe literatüre kavram, konu bütünlüğü ve örnek uygulama kazandırmaktır.

Araştırmanın Problemi ve Soruları: Bu çalışmanın ana problemi, çevrimiçi Türkçe iş ilanlarının araştırma modelindeki yetenek keşfetme analitiği örneği üzerinden makro ölçekte yetenek ve niteliklerini ortaya çıkarmaktır. Bu ana probleme göre, çalışmanın 5 temel araştırma sorusu bulunmaktadır:

- 1) Çevrimiçi Türkçe iş ilanlarında sık kullanılan kelimeler ve kelime grupları nelerdir?
- 2) Çevrimiçi Türkçe iş ilanlarında hangi varlıklar bulunmaktadır?
- 3) Keşfedilen niteliklerin 12 aylık dönemdeki dağılımları nasıl gerçekleşmiştir?
- 4) Keşfedilen niteliklerin yatırım teşvik bölgelerindeki dağılımları nasıl gerçekleşmiştir?
- 5) Keşfedilen nitelikler Türkiye'nin hangi bölgelerinde yoğunluk veya seyreklik göstermektedir?

Araştırmanın Önemi: Bu çalışma, devlet analitiği kavramının ele alındığı literatürdeki az sayıdaki bilimsel araştırmalardan ve ilk Türkçe literatürdeki çalışmalardan biridir. Bu yüksek lisans tezi, devlet analitiği alanında, literatürdeki en kapsamlı çalışmalardan biridir. Bu tez konusu, devlet analitiği kapsamında insan kaynakları analizlerini ve beşeri sermayeyi değerlendiren öncü bir araştırmadır. Bu bağlamda, beşeri sermaye analitiği araştırmalarının makro ölçekte ve ulusal düzeyde gerçekleştirildiği Türkçe literatürdeki ilk akademik tezlerden biridir. Uluslararası literatürde iş ilanlarını büyük veri olarak yalnızca mikro ölçekte insan kaynakları veya firma düzeyinde incelenmesine karşın, bu çalışma iş ilanlarını makro düzeyde inceleyen ilk araştırmalardan biridir. Araştırmanın yöntemi ele alındığında; metin madenciliği ve coğrafi bilgi sistemleri entegrasyonunun işlendiği Türkçe literatürdeki ilk çalışmalardan biridir. Buna ek olarak, iş ilanlarının karar desteği için coğrafi bilgi sistemlerinde kullanıldığı ilk çalışmalardan biridir.

Araştırmanın Yöntemi: Bu tez çalışmasında, iş ilanlarının metin madenciliğiyle analiz edilip, bir coğrafi bilgi sistemiyle birleştiren bir model ortaya çıkarılmak istenmiştir. Veriler, Python programlama dili ile internet kazıma tekniğiyle toplanarak Microsoft Excel dosya formatlarında saklanmıştır. Veri temizleme ve yapılandırma süreçleri Python dili ve Microsoft Excel üzerinde gerçekleştirilmiştir. Excel dosyasındaki veriler, metin madenciliği ve CBS uygulaması için Wordstat paket yazılımına aktarılmıştır. Wordstat yazılımı metin analizleri için kullanılmaktadır. Wordstat bir başka paket uygulamayla bir CBS entegrasyonu içermektedir. GISViewer, Wordstat ile entegrasyon

yeteneğine sahip, kendi içerisinde de çeşitli analiz özellikleri bulunan bir CBS uygulamasıdır (Akbiyık, 2019). Wordstat yazılımına aktarılan veriler, analiz edilip, araştırma bulgularına ve bu kayıtlar GISWiever'e aktarılarak CBS üzerinde Wordstat'tan elde edilen çıktılara ek yeni araştırma bulgularına ulaşılmıştır. Araştırma raporları, Wordstat ve GISWiever projeleri üzerinden çıktılar halinde alınmıştır.

Araştırmanın Kısıtları: Bu çalışmada tez, Türkiye'deki internette yayınlanan iş ilanlarının metin madenciliğiyle nasıl analiz edileceği ve bu analizlerin bir coğrafi bilgi sistemiyle nasıl entegre olarak çalışılabileceği ile kısıtlanmıştır. Metin madenciliği analizi ve coğrafi bilgi sisteminin sınırları Wordstat yazılımının kapasitesi çerçevesinde değerlendirilir. İş ilanları Kariyer.Net internet sitesi üzerinden 11 ve 12 Ocak 2020 ile 11, 12 ve 13 Şubat 2020 tarihlerinde toplanmıştır. Veri setinde, içeriği görsel veri içeren ilanlar ve tekrar eden ilanlar bulunmamaktadır. Analizler, elde edilen verilerden yalnızca 2019 ve 2020 yıllarında içerikleri yayınlanmış ve güncellenmiş ilanlar ile sınırlıdır. CBS için kullanılan konum bilgileri Türkiye sınırları içerisinde yer alan ilanlarla kısıtlanmış olup, vilayet isimleriyle enlem ve boylam bilgileri eşlenerek hazırlanmıştır. Yatırım teşvik bölgeleri de koordinat bilgileri eşlenmesine benzer olarak veri setine dahil edilmiştir. Veriler analizlerden önce toplama işleminde; ilan numarası, iş ilanı başlığı, firma adı, genel nitelikler, şehir ve aranan bireyin iş tecrübesi miktarı, askerlik durumu, eğitim seviyesi, okuduğu bölüm, yabancı dil, sektör, departman, çalışma şekli, pozisyon seviyesi, alınacak personel sayısı, başvuru sayısı, yayınlanma tarihi, son güncelleme tarihi ve kapanış tarihi alanlarına göre yapılandırılmıştır. Analizler için içeriklerin olduğu genel nitelikler sütunu ve bunlarla ilişkilendirilmiş şehir, koordinat bilgileri, yatırım teşvik bölgeleri, çalışma şekli, pozisyon seviyesi ve yayınlanma tarihi alanları kullanılmış ve tekrar eden içerikler veri setinden temizlenmiştir. Temizlenmiş veri setinden elde edilen frekans analizi, varlık isimleri ve kelime öbeklerine dair bulgular; oluşturulan model ile Wordstat yazılımından alınmış ve çapraz tablo analizi için yeni kısıtlar oluşturularak yeni bulgular ile GISWiever tarafından coğrafi analizler gerçekleştirilmiştir. Bu nedenle, bu analizlerde deney sürecinde doğan kısıtlar araştırma tekniği açısından Wordstat ve GISViewer yazılımları çerçevesinde sınırlıdır.

BÖLÜM 1: DEVLET ANALİTİĞİ

Bu bölümde ilk olarak devlet analitiği kavramı dört başlık altında özetlenmiştir. İlk kısımda, devlet analitiğinin literatür içerisindeki gelişimi ve bu kavramın hedefleri aktarılmıştır. İkinci olarak devlet analitiğinin literatürde yer alan örnekleri incelenmiştir. Üçüncü kısımda devlet analitiği kavramının tanımlanması gerçekleştirilmiştir. Son kısımda ise, devlet analitiğinin unsurları ve çerçevesi çizilmiş ve bu tezin özelinde devlet analitiği unsurlarından ulusal beşeri sermayenin tarihçesi, bu analitiklerdeki büyük verinin yeri ve ulusal beşeri sermaye analitiğinin kapsamı ortaya koyulmuştur.

1.1. Devlet Analitiği Kavramı

Bilişim teknolojileri sayesinde veri bilimi günümüzde ivmelenerek araştırmacılar arasında popüler bir bilim dalı haline gelmiştir. Veri analizlerinde kullanılan teknikler her geçen gün yeni yöntemlerle geliştirilmektedir. Buna paralel olarak, veri analitiği alanı yalnızca bilişim dünyası ve mühendislik disipliniyle birlikte değil, sosyal bilimler alanında da araştırmalara konu olmaktadır. Bu bağlamda, veri analitiği çok disiplinli ve hibrit araştırmalarda genellikle uygulanan yöntemlerden biri olmuştur.

Günümüzde veri analitiğinin çalışmaları tıp, uzay araştırmaları, mühendislik ve savunma sanayi başta olmak üzere tüm üretim endüstrisi sektörlerinden sosyoloji, psikoloji, siyaset, politika gibi sosyal hayatın girdiği alanlara kadar geniş bir yelpazesi vardır. Bu bağlamda, uluslararası kuruluşlar, devletler, özel sektör firmaları, vakıf ve dernekler veri analizine ve verinin gücüne gün geçtikçe daha çok önem vermektedir. Bu konuda, *büyük verinin analitiği* araştırmaların içerisinde önemli bir başlık olarak ele alınmaktadır.

Uzun yıllardır kamu kuruluşlarının temel yeteneklerinden biri veri analizi çalışmalarıdır. Bilim insanları ve kamu kuruluşları demografik bilgi toplamayı ve istatistik çalışmalarına devam etmektedirler. Bunlarla birlikte politikaların etkileri, politik tahminler için karar vermeyi kolaylaştırmak üzere veri analizlerinde iyileştirmeler yapılmaktadır. Richards'a (2017) göre, analitik tekniklerin uygulama alanlarının genişlemesi, büyük verinin kuruluşlar tarafından kullanılabilirliğinin artması ve katı bilimsel veya büyük veri için finansal kaynaklara ihtiyacın azalmasının bu alanın kamu

kurumlarınca ilgilenilmesinde etkili olmuştur. Bu bağlamda, Richard 2017’de yayınladığı kitabında *Devlet Analitiği* ile ilgili iki önemli soruya cevap vermiştir.

Birincisi, “*Neden devlet analitiği?*”: Devlet yönetiminde politik belirsizlikler ve dış değişkenlerin çokluğundan ötürü artan bir karmaşık durum söz konusudur ve devlet özel sektöre göre yapısal olarak farklılıklara sahiptir. Dolayısıyla, yapısı gereği de bir karmaşıklık sahibidir ve politikaların ve dış etmenlerin getirdiği belirsizlik bu karmaşıklık düzeyini arttırmaktadır. Bu bağlamda, büyük veri analitiği bu karmaşıklığı ve belirsizliği gidermek için önemli bir araç olarak kullanılabilir.

İkincisi, “*Neden şimdi?*”: Büyük veri analitiğinde kullanılan araçlar artık kolaylıkla erişilebilir hale gelmiştir. Böylece verimliliği arttırmak ve iç görü edinmek için analitik çalışmalar yapılmalıdır. Bu konuda, Richard 3 faktöre dikkat çekmektedir ve bunlardan birincisi, herkes tarafından rahatça *verilere erişebilmek*, ikincisi analistler ve yöneticilerin gözüyle kolaylıkla *veri analizlerini anlamak* ve son olarak üçüncüsü yönetsel kararlarının analizlere ve kanıtlara dayandırıldığı *veriye dayalı bir kültür inşa etmektir*.

1.1.1. Devlet Analitiğinin Tanımlanması

Devlet Analitiği kavramının “devlet” ve “analitik” kavramlarının birleştirilerek oluşturulması kadar basit bir tanım ve küçük bir kapsamı yoktur. Devlet kurum, kuruluş ve organizasyonlarının gerçekleştirilen her analitik devlet analitiğinde bağlamında ele alınamaz. Bu alandaki çalışmalar incelendiğinde, *devlet analitiği araştırmaları ulusal seviyede hedeflere sahip olduğu*, makro bir bakış açısıyla uygulandığı görülmüştür.

Dikkat çeken bir diğer husus ise, makro düzeyde gerçekleştirilecek analitikte *kamu yararı* gözetilmesidir. Diğer bir deyişle, devlet analitiği araştırmaları özel sektör ve vatandaşlara devletin de ortak edildiği ulusal çapta analitik katkılar sağlayan araştırmalardır. Dolayısıyla, devlet analitiğinin tek müşterisi devlet değildir. Nihayetinde, devlet analitiğiyle devletin yanı sıra özel sektör ve vatandaşlara da fayda üretilir.

Devlet analitiđi kapsamında, başka bir önemli nokta daha vardır. Devlet analitiđinin hedeflerinden biri, *veri analitiđi kültürünü* devlet, özel sektör ve vatandaşlara *aşılacaktır*. Dolayısıyla geniş bir perspektife sahip devlet analitiđinin liderliğini genellikle devlet kurumları yaptığı görülmüştür. Ancak devlet analitiđi araştırmalarını yalnızca devlet yöneticileri değil, özel sektör veya üniversiteler de imkanlarınca daha küçük ölçekte uygulanabildiđi ve devlet kurumlarına bu araştırmalarını paylaşılabil-diđi tespit edilmiştir. Buna dayanarak devlet analitiđi çalışmalarında, devlet yönetimi çok daha güçlü bir yapı olması sebebiyle devlet yöneticilerinin liderliğinde yürütmek başarıyı yakalamak için olanak ve hız kazandıracaktır denilebilir.

Nihayetinde, devlet analitiđinin kapsamı belirlenmiş ve hedefleri ortaya koyulmuştur. Literatür taramalarında devlet analitiđinin bir tanımı bulunamamıştır ve devlet analitiđi kavramı sınırlı bilimsel kaynaklarda açıklanmaya çalışılmıştır. Bu nedenle devlet analitiđinin daha iyi anlaşılabilmesi için bu kavramın tanımlanması ve kapsamının özetlenmesi gereksinimi duyulmuştur.

Sonuç olarak, devlet analitiđinin farklı üç tanımı yapılabileceđi öngörüsünde bulunulmuştur. Bu tanımlar aşağıdaki gibidir:

- Devlet analitiđi; devlet dairelerine, sivil toplum kuruluşlarına, vakıflara ve derneklere, özel sektör işletmelerine ve tüm vatandaşlara sosyal yetenek, beceri, uzmanlık alanları vb. konularda ve ekonomik, sosyal, beşeri, kültürel problemlere dair bilgiler ve gelecek tahminler elde etmek için yapılan büyük veri analitiđi çalışmalarıdır.
- Devlet analitiđi; bilgi teknolojilerin kullanılarak kamu yararına bulgular elde edilmesi, bu teknolojilere kurumların ve özel sektörün erişiminin kolaylaştırılması ve bilişim kültürünün devletin tüm vatandaşlarına benimsetmesi amacıyla gerçekleştirdiđi makro düzey analitiklerdir.
- Devlet analitiđi; devletin stratejik ve sosyal politikaların performanslarının artırılması ve ölçülmesi için veri bilimi disiplinlerinden ulusal düzeydeki araştırmalarda faydalanmak amacıyla yürüttüğü analitiklerdir.

1.2. Devlet Analitiđi Alanındaki alıřmalar

Önceki bölümde devlet analitiđi kavramının ortaya ıkıřı ve kapsamı verilerek, bu kavram tanımlanmıřtır. Bu bölümde, devlet analitiđi alanındaki farklı alanlardaki alıřmalara dair literatürden derlenen özet bilgiler verilmiřtir. Bu bölümde, devlet analitiđi üzerine kamu kuruluşları, üniversiteler, özel sektör ve arařtırmacıların alıřmalarına yer verilmiřtir.

1.2.1. ABD’deki alıřmalar

Devlet analitiđi alanında bazı akademik kuruluşlarda önemli arařtırmalar yapılmaktadır. John Hopkins Üniversitesi’nde, bu alanda İleri Devlet Arařtırmaları Merkezi (CAGS) atısı altında online kurslar (yüksek lisans, post-lisans/sertifika programları ve kombine programlar) açılmıřtır. Kamu ve özel sektör için bireyleri *liderlik* konusunda yetiřtirmek amacıyla kurulmuř bu programlarda büyükelilikler, devlet kurumları, düşünce kuruluşları ve diđer çeřitli organizasyonların ortaklıđıyla etkinlikler yapılmaktadır. Bu bağlamda, CAGS Amerikan toplumunun tepkisini řekillendiren arařtırmalar ve analizler ile ilgilenmektedir (Johns Hopkins University, 2020).

1.2.2. Avustralya’daki alıřmalar

Stratejik düzeyde deđerlendirilebilecek, devlet analitiđi konusundaki bir diđer alıřma Avustralya devleti tarafından gerekleřtirilmiřtir. Sađlık hizmetleri için hükümet politikalarına katkı vermek amacıyla ve bu konuda analitik alıřmalar yürüten *ACT Devlet Analitiđi Laboratuvarı* bulunmaktadır. Toksikoloji, adli kimya ve mikrobiyoloji gibi fen bilimleri bařta olmak üzere biliřim ve halk sađlıđı gibi birok disiplini bir araya getiren bilimsel testler yapmak amacıyla kurulmuřtur. ACT’nin liderlik ettiđi, özellikle *salgın, bulařıcı hastalıklar, hava kirliliđi* gibi toplum sađlıđını ve psikolojisini ilgilendiren konularla ilgili alıřmalara büyük veri analitiđi arařtırmaları bulunmaktadır (Tran ve diđerleri, 2018; Kelly, 2017; Australian Capital Territory Government 2020).

1.2.3. BAE'deki Çalışmalar

Devlet analitiği çalışmalarına bir başka örnek olarak, kurumsal altyapı ise, Birleşik Arap Emirlikleri'nde bulunmaktadır. Richard'ın (2017) kitabının 3. bölümünde bu konuda *Dubai Yönetimi Mükemmellik Programı* adı verilen bir araştırmaya yer verilmiştir. Avrupa Kalite Yönetim Vakfı (EFQM) referanslı bu çalışmanın bir *BAE devlet stratejisi* bağlamında İK odaklı bir program olduğu anlaşılmaktadır (Khattab ve Tyagi, 2017).

Literatüre göre BAE devlet politikasında devlet analitiği kritik bir rol oynamaktadır. Bu politika *Emirlikleştirme* denilmektedir. Arap dünyasında ve benzer devlet politikaları *Katarlaştırma*, *Ummanlaştırma* ve *Suudileştirme* (Nitaqat olarak da bilinir) bu örneğe benzemektedir. Bu politikaların genel anlamda hedefi, *ulusal beşeri sermayeyi geliştirmek* ve *ekonomiyi kalkındırmak* diye özetlenebilir. Ancak olumsuz bir etki bağlamında bir karşıt görüş olarak, bu politikaların *navitizm* akımından beslendiği, göç karşıtlığını ve yabancı düşmanlığını büyüttüğü düşünülmektedir (Emiratisation, 2020; Navitism, 2020).

Bu bağlamda, devlet analitiği çalışmaları üzerine BAE'deki örneklerin emirlikleştirme politikaları ve İK kapsamında ele alınması gerektiği anlaşılmıştır. Khattab ve Tyagi, *Dubai Yönetimi Mükemmellik Programı* için dikkat edilmesi gereken bazı hususları belirlemişlerdir. Bu konuda, EFQM kriterlerinden yararlanarak Dubai'deki iş ortamı ve genel BAE yerel kültürüne uyacak şekilde bir çerçeve oluşturmuşlardır.

Bu çerçevede şu başlıklar belirlenmiştir (Khattab ve Tyagi, 2017):

- İnsan yönetimi.
- Çevre.
- Emirlikleştirme (tüm devlet kurumlarında, özellikle yerel halkın az çalıştığı işlerde yerli BAE vatandaşlarının sayısının artırılması).
- Yerel halk için özel liderlik geliştirme programları.
- Tüm organizasyonlarda ikinci / yedek liderlik ekibinin geliştirilmesi.
- Kalite, sağlık, güvenlik ve çevre standartlarının ve yönetim sisteminin güçlendirilmesi.

BAE özelinde akademik arařtırmaların yanı sıra, devlet analitiđine kurumsal bir bakıř aısı olarak ABD ve Avustralya'daki gibi devlet kurumlarının oluřturulduđu veya ortak olduđu yapılar bulunmaktadır. Khattab ve Tyagi'nin alıřmasında ne ıktıđı zere İK alanında BAE'nin bu konuda kurumsallařtıđı, *byk veri ve yapay zeka* gibi biliřim teknolojileriyle entegre olabilen analitik bir kltr inřası misyonuna sahip bir devlet yaklařımı olduđu anlařılmaktadır. Bu konuda, literatrdeki *byk veri, yetenek analizi ve devlet analitiđi* geninde gerekleřen arařtırmaların  BAE kurumu tarafından icra edildiđi gzlemlenmiřtir.

Devlet analitiđiyle ilgili literatr taramasında BAE'de ne ıkan kurumlardan biri *İleri Teknolojiler Arařtırma řirketi* (ACIT) olarak hkmetin halka aık holding řirketi olan *Mubadala* (MIC) bnyesinde 2009 yılında faaliyete gemiřtir. ncelikli misyonu yarı iletken sektr ve mikro iřlemciler zerine AR-GE faaliyetleri olmak zere, devlet politikalarına gre alıřma gcn ynlendirmek, insan sermayesini glendirmek iin kamu kurumları ve zel sektrn byk veri, yapay zeka, web servisleri gibi gereksinimlerini karřılaması amalanmıřtır (Mubadala Investment Company, 2020; Mubadala, 2020; Advanced Technology Investment Company, 2020).

BAE, beřeri sermaye iin teknoloji odaklı yatırımlarının yanı sıra analitik kltrn inřası iin kurumlar oluřturmuřtur. Bu řirketler, BAE'nin yalnızca İnsan Kaynakları ve Emirlikleřtirme Bakanlıđı'na ait mdrlkleri, merkezleri veya daireleri gibi deđildir. Eski adı alıřma Bakanlıđı olan İnsan Kaynakları ve Emirlikleřtirme Bakanlıđı, emirlikleřtirme politikalarına zel sektr teřvik etmektedir. Dolayısıyla sıradan bir devlet dairesinin aksine, řirket gibi ynetilmesi ve faaliyette bulunması beklenen bu organizasyonlara ve bu kltr inřasına bakanlık liderlik etmektedir. Bu bađlamda, 1999 yılında ıkarılan bir yasa ile 2000 yılında *Ulusal İnsan Kaynakları ve İstihdam Kurumu* (TANMIA) kurulmuřtur. BAE'de TANMIA, stratejik olarak iřgc piyasalarını desteklemek, glendirmek ve devletin bu konuda karar almasını kolaylařtırmak amacıyla liderlik ve mkemmellik vizyonu sunmaktadır (The National Human Resources and Employment Authority, 2020). Dolayısıyla, TANMIA devlet analitiđinin beřeri sermayenin geliřtirilmesi iin aktif rol aldıđı nemli kuruluř rneklerinden biridir.

TANMIA'nın ve alt kuruluşlarının çalışma faaliyetlerini özetleyen bir anlatım, Gonzalez ve diğerleri (2008) beşeri sermaye üzerine yaptıkları bir çalışmada geçmektedir. Burada *İşgücü Piyasası Araştırma ve Enformasyon Merkezi (CLMRI)* adı verilen organizasyon, TANMIA'daki verileri işgücü piyasaları araştırmalarında kullanan bir iç varlık olarak tanımlanmıştır. Sonuç olarak, BAE'nin devlet analitiği alanındaki uygulamaları özetle, ACIT ile teknolojiye erişim kolaylığı ve TANMIA ve CLMRI v.b. kurumlarla işgücü piyasaları ve beşeri sermaye odaklı araştırmalar olarak özetlenebilir.

1.2.4. Alandaki Çalışmalarının Değerlendirilmesi

Önceki bölümlerde devlet analitiği tanımlanmış, sonra bu alandaki örnek çalışmalarla ilgili altyapı geliştiren devletler ve kurumları, şirketler ve araştırma merkezlerine dair geniş bir bölüm ayrılmıştır. Bu bağlamda, devlet analitiği araştırmalarının gereksinimleri kurumsal zemin ve kültür açısından ülke bazlı aktarılmıştır. Kimi ülkelerin yönetim şekilleri, kültürleri ve politikaları sebebiyle literatüre siyasi bakış açıları daha yoğun yansıdığı gözlemlenmiştir. Bununla birlikte bu tezin uygulama tarafına daha yakın olmasından dolayı, BAE'nin beşeri sermaye odaklı çalışmalarına daha geniş bir bölüm ayrılmıştır.

Sonuç olarak, devletlerin veri analizleri gereksinimlerinin ve bu alanda devlet analitiği kavramının ortaya çıkışında bilişim teknolojilerinin ilerlemesi, büyük verinin kullanım alanları ve stratejik politikalar etkili olmaktadır (Davenport ve Jarvenpaa 2008). Bu bağlamda, devlet analitiği uygulamalarının ABD, Avustralya ve BAE gibi çeşitli coğrafyalarda ve ülkelerde sosyolojik ve psikolojik olarak toplum tepkisinin ölçülmesinden, sağlık alanında salgın kontrol planlarına, afet yönetimine, kriminal ve terör gibi suç faaliyetlerin engellenmesine ve ekonomik kalkınma, göç idaresi gibi idari konularda işgücü piyasalarına dair öngörü yeteneği kazanılmasına ve tahminler yapılmasına, ulusal sermayenin verimliliğine ve bu özelde ulusal beşeri sermayenin iş becerilerinin keşfedilmesine ve genel işsizlik, göçmenlerin uyum sorunları gibi sosyal problemlere çözüm geliştirilmesine kadar çeşitli imkanlar sunduğu anlaşılmıştır.

1.3. Devlet Analitiğinin Unsurları ve Çerçevesi

Devlet analitiğinin amaçlarının üç kategoride ele alınması uygun bulunmuştur. Bu kategoriler, devlet analitiği tanımlanırken öne çıkan kavramlardır. Bu amaçların sınıflandırılması, devletin asli görevi olan *kamu yararı*, oluşturulacak *bilişim* alt yapısı ve bu analitiklerin etkilediği *çevre* olarak değerlendirilmiştir. Kamu yararına yönelik amaçlar özellikle devletin sahip olduğu sermayeleri (örneğin işgücü, üretim, doğal kaynaklar v.b.) tüm vatandaşların analitik ile doğrudan birebir ilişki kurduğu noktalara dokunmaktadır. Bununla birlikte, özel sektördeki tüzel kişilikler de vatandaşlar gibi bu kapsamda değerlendirilmelidir. Kısacası kamu yararına yönelik hedeflerin *kimlik sahipliğiyle ilgili* bir bağlantısı bulunmaktadır. Bilişime yönelik amaçlar, analitiklerle birlikte ortaya koyulan teknolojinin kullanımına yönelik metriklerden ileri gelmektedir. Son olarak çevreye yönelik amaçlar ise, devlet analitiğini uygulamasını yürüten aktörlerin (devletler, uluslararası kuruluşlar, özel sektör, yöneticiler vb.) dış çevresiyle ilgili konumlanmasıyla ilgilidir. Bu bağlamda, devlet analitiğinin hedefleri şu şekilde özetlenebilir:

- **Kamu yararına yönelik hedefler:**
 - Yeni uzmanlıklar oluşturmak ve ihtiyaç olan uzmanlıkları keşfetmek
 - İhtiyaç duyulan uzmanlara erişimi kolaylaştırmak
 - Vatandaşların sosyal ve kariyer yeteneklerini ortaya çıkarmak
 - Vatandaşların sahip olduğu becerilerinden faydalı bilgiler üretmek
 - Devlet, özel sektör ve vatandaşları bu alanda bir araya getirmek
- **Bilişime yönelik hedefler:**
 - Bilişim teknolojilerine erişimi kolaylaştırıcı zemin hazırlamak
 - Bilişim kültürünü örgütlere, yöneticilere ve vatandaşlara kazandırmak
 - Stratejik ve sosyal politikalarda teknolojiden faydalanmak
 - Bilişim teknolojileriyle faaliyetleri kolaylaştırmak
 - Yöneticilerin aldığı kararların performanslarını değerlendirmek
 - Karar vericilerin performansını yükseltecek bilgiler keşfetmek

- **Çevreye yönelik hedefler**

- Ulusal çapta, makro düzeydeki stratejilere katkıda bulunmak
- Uluslararası alandaki iş ortaklarıyla bilişimin etkinliğini arttırmak
- Bilişim teknolojisiyle rekabet gücü elde etmek
- Veri analitiğini makro düzeyde rekabet unsuru haline getirmek
- Makro düzeyde veri analitiği kullanımının ana omurgasını oluşturmak

Bu hedefler içerisinde *makro politikaların rolü* kritiktir. Bununla birlikte mikro ölçekte işletmeler için veri analitiğinin bir çerçevesi mevcuttur. Mikro ölçekteki bu çerçeve, işletmelerin temel fonksiyonlarına göre şekillendirilmiştir. Günümüz bilgi çağında ise, işletmeler temel fonksiyonlarının yanı sıra kendi misyonlarına göre sahip oldukları görevlerle kendi fonksiyonlarını ve bu analitiklerin çerçevesini genişletebilmektedirler. Böylece iş analitiği bağlamında; e-ticaret analitiği, finans analitiği, insan kaynakları analitiği, muhasebe analitiği, müşteri analitiği, pazarlama analitiği, stok analitiği, tedarik zinciri analitiği üretim analitiği, ürün analitiği v.b. çok farklı türde ve bazılarının iç içe geçtiği analitikler türetilmiştir. Bu konuda, özellikle İK alanındaki literatürde analitiklerle ilgili bir kavram karmaşası olduğu görülmüştür (Huselid, 2018).

Günümüzde ekonominin etkisiyle, devletlerin yönetim şekilleri, kurumsal yapılanmaları zamanla değişim halindedir (Öztürk, 2006). Dolayısıyla, çok çeşitli devlet yönetimleri ve şirketlerden daha karmaşık olan devlet yapısı, makro anlamda devlet analitiği çerçevesi oluşturmak için bir zorluk barındırır. Ayrıca analitik alanındaki kavram karmaşıklığının (İK alanındaki gibi) doğurduğu öz bilgiye ulaşmada bir başka zorluk vardır.

Bu bağlamda, makro anlamda devlet analitiğinin unsurları belirlenirken devletlerin varlık biçimi ve işletmelere (sistem olarak) benzeyen organizasyon yapısından yola çıkılmıştır ve bu devlet analitiği çerçevesi oluşturulurken Bourdieu'nin (1986) *sermaye türleri kuramı* kritik bir rol almaktadır. Bourdieu'nun sermaye kuramına göre, devletin sahip olduğu sermaye türleri; *ulusal ekonomik sermaye, ulusal fiziki sermaye, ulusal beşeri sermaye, ulusal kültürel sermaye ve ulusal sosyal sermaye* olarak geçer.

Analitiğinin çerçevesi oluşturulurken, devlet kavramına olan yaklaşım; devletin bir varlık olarak belirlenmiş ve devletin temel sorumluluğu, sahip olduğu varlığı korumak, geliştirmek ve büyütmek olarak ele alınmıştır. Dolayısıyla, Bourdieu'nun sermaye türleri kuramına göre bu yorumla devletin sorumluluğu olan 5 farklı sermaye türü olduğu yorumu getirilmiştir. Bu bağlamda, devlet analitiği, bu sermayelerin her birinin tek tek analitiğinden oluşan bir bütün, genel bir makro analitik olarak açıklanabilir. Bu bakış açısıyla devlet analitiğinin unsurları aşağıdaki belirlenmiştir:

- Ulusal Ekonomik Sermaye Analitiği
- Ulusal Fiziki Sermaye Analitiği
- Ulusal Sosyal Sermaye Analitiği
- Ulusal Kültürel Sermaye Analitiği
- Ulusal Beşeri Sermaye Analitiği

Sonuç olarak, devlet analitiğinin unsurları ve çerçevesi Bourdieu'nin sermaye türleri kuramından beslenmektedir. Aynı zamanda devlet analitiğinin unsurları oluşturulurken, mikro ölçekteki insan kaynakları analitiği konusundaki kavram karmaşası (Huselid, 2018) ve bu düzeyde beşeri sermaye analitiğinin bilindik bir kavram olmasının araştırmada etkisi olmuştur. Başka bir ifadeyle, makro bir çerçeve oluştururken sermaye kuramından esinlenmenin bir nedeni de literatürde mikro ölçekteki kabul gören kavramlardır.

1.3.1. Ulusal Beşeri Sermaye Analitiği

Bu bölümde ilk olarak beşeri sermaye kavramı açıklanacaktır. Daha sonraki alt bölümlerde, İK analitiğinin tarihi, İK analitiği ile büyük verinin kullanımına dair özet bir bilgi ve son kısımda başta İK analitiği ve büyük veri örnekleriyle oluşturulan Ulusal Beşeri Sermaye Analitiğinin Kapsamı verilecektir.

Beşeri sermaye (insan sermayesi olarak da bilinir), İK uzmanlarının başta olmak üzere işletmecilerin ve iktisatçıların aşına olduğu bir kavramdır. Seyidoğlu (2002: 59) ve Özcan'ın (2011: 107) tanımlarına göre; *Beşeri sermaye, bireyin bilgi ve yeteneklerini geliştirmek üzere deneyim ve kültürel eylemlere katılması sonucu biriktirdiği sermayedir.*

Diğer bir beşeri sermaye tanımı ise; *Beşeri sermaye, sosyal ve ekonomik refaha erişebilmek için bireyin bilgisini, becerisini, yeteneklerini ve niteliklerini ortaya çıkarmasıdır* (Healy, T. ve Côté, 2001).

Bu bağlamda, Albayrak ve Abdioğlu'na (2020) göre ulusal beşeri sermaye, *ekonomik büyümeyle doğru orantılıdır ve bir ülkenin nüfusunun sahip olduğu eğitim, sağlık ve mesleki becerilerinin bütünü* olarak ifade edilir. Bu yorumlardan yola çıkarak; *“Ulusal beşeri sermaye, bir ülkedeki vatandaşların bireysel olarak sahip olduğu eğitim, kültürel ve mesleki deneyim, beceri ve yeteneklerin birleşerek sosyo-ekonomik kalkınma amaçlı işgücü kaynağıdır.”* şeklinde tanımlanabilir.

Sonuç olarak ulusal beşeri sermaye analitiğinin hedefi, devletin işgücü kaynağını verimliliğini ve performansını arttırmak olduğu anlaşılmıştır. Dolayısıyla işgücünün performansı mikro ölçekte doğrudan İK alanında bir konu olduğu için, *devlet analitiğinde ulusal beşeri sermaye analitiği kapsamı ve amaçları oluşturulurken İK analitiği literatüründen faydalanılması gerekir*. Bu bağlamda, bu analitiğinin daha iyi anlaşılması için alt bölümlerde İK analitiği ve büyük veri konularında bilgiler paylaşılmıştır ve İK analitiğinden makro düzeyde devlet analitiğinin bir unsuru olarak tanımlanan ulusal beşeri sermaye analitiğinin kapsam ve unsurları açıklanmıştır.

1.3.1.1. İK Analitiklerinin Tarihçesi

İnsan kaynakları, büyük verinin kullanılabilceği hızla büyüyen ve yeni fırsatlar yaratan bir alan haline gelmiştir. Ancak İK alanında tahminlemeler, İK süreçlerinin depolanması ve beşeri sermayenin analizi 2010'ların sonunda başladı. Son yıllarda yapılan Deloitte araştırmasına raporunda, firmaların sadece %8 ila %10'u İK'da tahminleme analizlerinden faydalanmaktadır. İK bölümünde yapılandırılmış veriler olduğu gibi, yapılandırılmamış ve hiç kullanılmamış yüksek miktarda veri olduğu düşünülmektedir (King Jr, 2018).

İnsan kaynakları analitiğinde, Isson ve Harriott (2016) ile King Jr'ın (2018) derledikleri çalışmalar öne çıkmaktadır ve önemli temel görüşler barındırmaktadır. King Jr'a (2018) göre veri analitiği sayesinde, İK verilerinden değerli bilgiler çıkarılabilir ve beşeri sermayenin yapılandırılmamış verileri ile, iş performansını arttırmak, geliştirmek ve

yetenek yaşam döngüsünün her aşamasını optimize etmek için analizler yapılabilir. Dolayısıyla, “İK analitiği, bir işletmenin beşeri sermaye varlığını geleceğe dönük tahminler üretmek amacıyla farklı veri kaynaklarıyla entegre ederek geliştirdiği analizler bütünüdür.” şeklinde tanımlanabilir (Isson ve Harriott, 2016).

Tablo 1: Literatürdeki İnsan Kaynakları Analitiği İsimlendirmeleri

İngilizce Adı	Türkçesi
People Analytics	İnsan Analitiği
Human Resource Analytics	İnsan Kaynakları Analitiği
Human Capital Analytics	Beşeri Sermaye Analitiği
Labor Analytics	İşgücü Analitiği
Workforce Analytics	İşgücü Analitiği
Talent Analytics	Yetenek Analitiği
Talent Management Analytics	Yetenek Yönetimi Analitiği
Employee Analytics	İşçi Analitiği

İnsan Kaynakları Analitiği literatüründe, Tablo 1’de özetlendiği gibi bir kavram karışıklığı vardır. Huselid (2018) bu konuya değinmiştir. Akademik çalışmalarda ve sektörde bu isimlendirmede *insan*, *insan kaynakları*, *işgücü*, *beşeri sermaye*, *işçi* ve *yetenek* gibi bir çok başlık atılmıştır. Uluslararası literatürde eş anlamlı kelimeler, yakın anlamlı ifadeler ve birbirini kapsayan kavramlar ile tek bir kavram üzerine inşa edilmemiştir. Ancak Türkçe literatürde en sık ifade İK Analitiği olduğu gözlemlenmiştir. Huselid’e (2018) göre, bu kavramlardaki varyasyonun nedeni sözcüklerin anlamıyla ve kültürel etkileriyle ilgilidir.

Sonuç olarak, mikro ölçekteki firmalar özelinde yapılan analitik çalışmaları için Türkçe literatürün aşına olduğu “*İnsan Kaynakları Analitiği (İK analitiği)*” ifadesi daha uygun bulunmuştur. İnsan kaynakları analitiği; işletmelerin beşeri sermayenin etkinliğini arttırmak için çalışan yeteneklerini kazanmak ve iş gücü performansını yükseltmek üzere İK departmanı odaklı analitik faaliyetler bütünüdür şeklinde tanımlanabilir.

Isson ve Harriott tarafından, insan kaynakları analitiğinde beşeri sermayenin karşılaştığı sorunlar 7 bölümde incelenmiştir. Ve bu bölümler, bu analitikle ilgili çalışma çerçevesini oluşturmuştur (Isson ve Harriott, 2016). Literatürde ve özel sektörde bu çerçeve kabul görmüş bir modeldir. Bu çerçevenin bölümleri, İK faaliyetlerinin genel bir yaşam döngüsü olarak ifade edilebilir. Bu bölümler aşağıdaki gibidir (Isson ve Harriott, 2016):

- İşgücü planlaması (*Şirketin şu anda kaç yeteneğe ihtiyacı vardır? Rekabet etmek ve kazanmak için şirket bu yeteneğe/yeteneklere ihtiyaç duyar*)
- Yetenek kaynağı (*Adaylar için nereye bakılacak?*)
- Yetenek kazanımı (*Kim işe alınacak?*)
- Yetenek bağlılığı ve Yetenek işe alım ve Yetenek alıştırma (*Bağlılık programı ve yeni başlayan programının çalışan performansı ve sadakati üzerindeki etkisi nedir?*)
- Yetenek performansı yönetimi ve Yeteneğin yaşam boyu değer (*Yüksek veya düşük performansı sağlayan nedir?*)
- Yeteneklerin korunması (*Ayrılma riski olan parlak yetenekler kimlerdir ve neden?*)
- Yeteneğin zindeliği, Yeteneğin beden sağlığı ve Yeteneğin güvenliği (*Sağlık programının üretkenlik ve sadakat üzerindeki etkisi nedir?*)

Bu çerçevede ve daha önceki tanımlamalardaki yoğun yetenek kavramı vurgusu önemlidir. Bu durumun, Alayoğlu'nun (2010) da İK alanında yeni dönem diye nitelendirdiği, *Yetenek Yönetimi Yaklaşımı* etkisidir. Bu bağlamda, İK analitiği veya beşeri sermayenin analizi, yetenek yönetimi (YY) ile ilgilidir. Alayoğlu'na (2010) göre YY'nin işletmelerin rekabet stratejilerinden, işgücü performanslarına, İK'dan tüm organizasyonun bütününe ilgilendirecek kadar geniş bir alanı kapsayan bir sürü tanımı yapılmıştır.

Alayoğlu'nun görüşlerine göre ve bu yaklaşımdan hareketle, yetenek yönetim sistemleri (TMS) önemli bir konudur. Bu bağlamda, İK analitiği literatüründe yetenek yönetim sistemleriyle ilgili bilgiler bulunmaktadır. Isson ve Harriott (2016) kitabında, 2005 yılında ilk TMS'nin kurulduğunu ifade etmiştir. Bu sistemlerin, işe alım, performans yönetimi, öğrenme ve geliştirme ve ücret yönetimi gibi YY'deki temel süreçlerinin otomasyonu ve gelişimini sağlayan entegre sistemler olduğunu söylemektedir. İK bölümünde depolanan ve kullanılan çalışanlara dair birçok iş, performans ve yetenek verilerini depolayan ve günümüzde, bu sistemlerin sosyal medya verileriyle birlikte diğer dijital ayak izi ve yetenek davranışı verilerini de işleyebilecek sistemler haline geldiğini aktarmıştır.

Sonuç olarak İK alanı öncülüğünde büyük verinin de süreçlere dahil edildiği veri analitiği uygulamalarına İK analitiği denildiği anlaşılmaktadır. Bu analitiğin amaçları 3 konuya ayrılabilir. Bu açıdan, amaçlardan birincisi *şirket stratejilerine katkıda bulunmak ve yöneticilerin karar almalarını kolaylaştırıcı destekler vermek*, ikincisi *çalışan performansını yükseltmek ve bu performansa katkı verecek yetenekleri keşfetmek* ve son olarak üçüncüsü ise, *beşeri sermayenin şirket politikalarında etkinliğini arttırmak ve işgücü verimliliğini stratejik olarak şirket ölçeğinde başarmak* şeklinde özetlenebilir

1.3.1.2. İK Analitiklerinde Büyük Verinin Kullanımı

Veri analitiği, verimliliği ve etkinliği arttırmak, karar vermeye yardımcı olmak için bilgiler elde etmek üzere veri edinmek ve çözümlenmek için tasarlanmış süreçler ve faaliyetler bütünü olarak tanımlanır (Le Roux, 2012). Bu bağlamda büyük veri analitiği, iş analitiği, İK analitiği, devlet analitiği gibi analitik sözcüğünün, tasarlanan süreç, iş veya faaliyet alanına ya da amacına göre şekillenerek kavramların çoğaldığı görülmüştür.

Günümüz bilgi çağında, üretilen veri miktarı ve türleri katlanarak çoğalmakta ve bu nedenle yeni analiz tekniklerine ihtiyaç artmaktadır. Büyük veri yaklaşımlarında, büyük verinin boyutlarını tanımlarken bu boyutlara yeni eklemeler yapılırsa dahi, Demirel ve diğerlerine (2019) göre büyük verinin hacmi, çeşitliliği ve hızı büyük veri kavramının özünü ifade etmektedir. Nihayetinde büyük veri, kendisinin ölçümünü zorlaştıran özelliklere sahip bir veri bilimi kavramıdır.

Sonuçta, bu analitik yeteneğin kazanımı için, büyük verinin işletmelerin ve devletlerin her iş koluna entegre ettikleri, keşfi tamamlanmış bir teknoloji olarak henüz edilemiyor. Dolayısıyla İK analitiği ve büyük veri araştırmaları, bilimsel araştırmalar kapsamında bugünün şartlarında kritik boşluklara sahiptir. Bu konuda, kısıtlı insan kaynakları analitiği ve devlet yönetimindeki büyük veri örnek olabilir (Isson ve Harriott 2016, Richards, 2017).

1.3.1.3. Ulusal Beşeri Sermaye Analitiğinin Kapsamı

Ulusal beşeri sermaye analitiğinin kapsamı belirlenirken mikro ölçekteki İK analitiği unsurları ve *devlet analitiği* amaçları dikkate alınmıştır. Bu nedenle, ulusal beşeri sermaye analitiği çerçevesi oluşturulurken *İK analitiği çerçevesi* referans alınmıştır. Bu bağlamda ulusal beşeri sermaye analitiği de İK analitiğiyle benzer süreçlerden oluşmaktadır. Isson ve Harriott (2016) ile King Jr (2018), İK analitiğiyle ilgili uygulamalar hakkında kapsamlı bilgiler vermişlerdir.

İK analitiği, büyük veri ile muazzam bir değişikliğe uğrayacaktır. 90'larda doğan pazarlama devrimine benzer, beşeri sermaye ve YY için bu devrimin geldiğini düşünüyor. YY'nin optimizasyonunda büyük verinin itici bir güç olacağı tahmin ediliyor. Isson ve Harriott (2016: 67), 7 bölümden oluşan yetenek yönetimi döngüsü referans alarak İK analitiğini açıklamışlardır. Bu unsurları şu şekilde ifade edilmiştir:

1. İşgücü planlaması
2. Kaynak bulma
3. Satın alma / işe alım
4. İşe alıştırma, kültüre uyum ve görevlendirme
5. Çalışan karmaşası ve elde tutma
6. Performans değerlendirme ve geliştirme ve çalışanın yaşam boyu değeri
7. Çalışan zindeliği, sağlığı ve güvenliği

Bu YY döngüsünün her aşamasında analitik yapılabilir. İşbu YY döngüsünde, analitikleri çalıştırırken IMPACT'ten yararlanılmaktadır. IMPACT modeli, veri analizi yapılması için oluşturulmuş temel bir iş analizi çerçevesidir (Isson ve Harriott, 2013). Buradan hareketle oluşturulan beşeri sermaye analitiği (İK analitiği) bölümleri Şekil 1'de gösterilmiştir.



Şekil 1: Beşeri Sermaye Analitiği Çerçevesi

YY döngüsündeki *kaynak bulma* adımının analitik süreçteki karşılığı yeteneklerin belirlenmesi ve keşfedilmesi olarak kabul edilebilir. Çünkü işgücüyle ilgili planlama aşaması, yöneticilerinin stratejik hedefleriyle ilgili üst seviye bilgi içermektedir. Şekil 1'deki yetenek kazandırma bölümü ise, YY'de belirlenen yetenekler organizasyonun iş ilanına gelen başvurular değerlendirildikten sonraki mülakat, iş sözleşmesi ve yeteneğin organizasyona katılmasını ifade eder. Bu bağlamda, *yetenek keşfetme* olarak adlandırılan bölüm organizasyonun dış dünya ile iletişim kurulan bir faaliyeti olarak görülmektedir.

Bilindiği üzere, büyük veri kavramında dış kaynakların rolü önemlidir. Dolayısıyla, yetenek keşfetme ve bulma aşamalarında büyük veriden yararlanılabilir. Bu konuda, iş ilanlarının büyük veri olarak analiz edilmesiyle ilgili örnekler literatürde mevcuttur (Isson ve Harriott, 2016, King Jr, 2018). Bu bağlamda, ulusal düzeyde beşeri sermaye analitiği gerçekleştirilebilir için iş ilanlarının ulusal düzeyde analitiği *devlet analitiği* kapsamında değerlendirilebilir. Bu analitik ile, literatürdeki örneklerden yola çıkarak şu yorum yapılabilir: İş ilanlarının ulusal düzeyde analitiği ile; makro ölçekte ise işsizlik, ekonomik kalkıma, göç planlaması, şehirleşme, eğitim gibi sorunlu alanlarda karar almayı kolaylaştırıcı ve olağan performansı değerlendirebilecek bir uygulama geliştirilebilir.

Daha önceki bölümlerde, devlet analitiğinin BAE uygulamalarının beşeri sermaye yönetiminin makro politika olarak değerlendirildiğini ve özellikle işgücü piyasalarını ve göçmen politikalarında kullanıldığı aktarılmıştı. Nihayetinde iş ilanları üzerine makro ölçekte analizler gerçekleştirmek, iş ilanlarıyla devlet analitiği çalışmak, BAE örneğindeki makro düzeydeki beşeri sermaye araştırmaları gibi geleceğe yön verecek stratejik uygulamalardan biri olacaktır.

Nihayetinde, iş ilanlarının analiz edilmesi ulusal beşeri sermaye analitiğinin bir konusudur. Ve iş ilanları mikro ölçekte yetenek keşfetme ve bulma analitiğinde değerlendirilen bir konu olması sebebiyle, makro ölçekteki devlet analitiğinde de iş ilanları analiz edilirken benzer bir yol haritası izlenmesi gerekmektedir. Dolayısıyla iş ilanlarının analizi, yetenek keşfetme ve bulma analitiğinin bir konusu olarak devlet analitiğinde ulusal beşeri sermaye analitiği kapsamındadır.

Yetenek Keşfetme Analitiği: Yetenek kaynağı oluşturmak üzerine analitik çalışmalar için literatürdeki araştırmalar çok kısıtlıdır. Bu konuyla ilgili literatür yalnızca, Isson ve Harriott'ın (2016) genel olarak İK analitiği üzerine çıkardıkları kitabın bir bölümünde ve 2018 yılında referans olarak bu araştırmayı alan King Jr'ın yapısal olmayan verilerin analizleri üzerine derlediği kitapta geçmektedir. Dolayısıyla, birbirine referans olan iki kaynak dışında yetenek keşfedilirken iş ilanlarının analiz edilmesine dair henüz başka bir bilimsel bir örnek uygulama bulunamamıştır.

Ulusal beşeri sermaye analitiği kapsamında ifade edilen yetenek keşfetme analizlerinin konumlandırılması için araştırmacıların (Isson, Harriott ve King Jr) görüşleri referans alınmıştır. Buna ek olarak, bu araştırmada tanımlanan devlet analitiği çerçevesindeki makro bakış açısıyla birleştirilmiştir. Mikro ölçekte, yetenek kaynağı edinme olarak geçen kavram, şirket beşeri sermayesine katılımı ile devam etmiştir. Ancak makro bir ölçekte kavramın daha kapsayıcı olması ve bu yeni makro yaklaşımda yapılan faaliyetlerin öne çıkması için ismi doğrudan *yetenek keşfetme analitiği* şeklinde ifade edilmiştir. Yetenek keşfetme analitiği kavramında dikkat edilmesi gereken bir diğer nokta ise, yetenek olarak ifade edilen terimin bir uzmanlık, bir kariyer yeteneği, bir beceri olduğudur.

Bu bağlamda, Isson ve Harriott (2016) ile King Jr (2018) çalışmalarında yetenek kavramı için, bazı süreçlerde bir uzmanlık veya beceriyi ifade ederken bazı süreçlerde ise yeteneğin sahibi olan bireyi izah ettiği anlaşılmıştır. Bu nedenle, ulusal beşeri sermaye için yetenek keşfetme analitiği, uzmanın veya bireyin değil yalnızca uzmanda veya bireyde olması gereken ya da sahip olunan uzmanlıkları, yetenek ve becerileri, kariyer tecrübelerini keşfetmeyi ve bunlardan bilgi elde etmeyi amaçlar ve yürütür.

İK bölümü YY'yi mikro ölçekte, işgücü planlamasına göre ilana nasıl ve ne zaman çıkacağını, çıktığı ilanın performansını öğrenmek için kullanır. Bu yüzden YY ilan odaklı bir süreçtir denilebilir. Dolayısıyla makro bir araştırmada devlet kurumları, iş ilanları üzerinden işgücü piyasalarının planlanması için değerlendirmeleri YY teknikleriyle yapılabilir. Ancak bu iş, İK'nın iş ilanlarıyla ve temel bilinen metrikleriyle ilgili kısmıdır. Bu alan, büyük veri çalışması olması nedeniyle, bu ilanlardan daha çok bilgi edinilebilir.

Isson ve Harriott'ın (2016), iş ilanlarının büyük veri olarak kullanılmasını mikro ölçekte İK'nın iş başvurularındaki ilanın performansını ölçme odaklı incelemiştir. Bu nedenle ilan içeriğinin doğruluğu, doğru yeteneği ifade eden kavramlar, ilanın doğru konumlanması gibi metrikler ile İK performansını artırma ve şirket stratejilerine beşeri sermayede doğru bir yatırımla birlikte işletme performansını yükseltmeyi kurgulamışlardır. Ancak makro anlamda değerlendirildiğinde, mikro yaklaşımın aksine ulusal anlamda iş ilanlarının analizi bu kadar kısıtlı bilgiler içermediği varsayılmıştır. İş ilanları, büyük veri olarak değerlendirilmeli ve yalnızca insan kaynakları metrikleriyle sınırlandırılmamalıdır. İşte bu bakış açısıyla, iş ilanlarının kapsamı ulusal anlamda genişletilebilir. Bu açıdan, beşeri sermayenin makro ölçekte hangi metrikleri değiştirdiğine bakılmalıdır.

Isson ve Harriott (2016) iş ilanlarının analiziyle ilgili metin analizleri, makine öğrenimi ve yapay zekadan faydalanılarak uygun/başarılı özgeçmişi öngörmenin (tahminleme) yapmanın mümkün olduğunu ifade etmişlerdir. Böylece firmanın hedeflerine göre işgücü stratejileri bağlamında, arz-talep denklemini tahmin etme seviyesine ulaşılacağını söylemişlerdir. Bu araştırma yapılırken, Isson ve Harriott'un *arz-talep tahminleme* fikri kritik derecede önemli bulunmuştur ve makro araştırmalar için gelecek vizyonunu oluşturacak, kilidi açacak anahtar sözcükler bunların olabileceği düşünülmüştür.

Bu bağlamda, yetenek keşfetme teknikleriyle iş ilanları üzerine yapılacak devlet analitikleri, *makro politikalarda tahminleme* yapılırken değerli bilgiler verebileceği fikrini doğurmaktadır. Dolayısıyla, ilanlar üzerinden özel sektörün arz-talep denklemini makro bir bakış açısıyla devlet kurumları inceleyebilir. Bu analizlerdeki bulgular, işgücü piyasasına yönelik stratejik planlarda kritik bilgiler verebilir. Böylece işsizlik, ekonomik kalkınma, eğitim, göç planlaması gibi önemli kamusal sorunlara yardımcı metrikler elde edilecektir. Bunlarla birlikte, kamu kurum ve kuruluşlarına mevcut durumu ölçme ve gelecek politikalarla ilgili tahminleme yapma fırsatı, öngörü yeteneği kazandıracaktır.

Nihayetinde makro yaklaşım olarak, devlet analitiği için ulusal beşeri sermaye analitiği ve yetenek keşfetme analizleriyle süreci ifade eden bir model ortaya çıkarılmıştır. Bu model Şekil 2’teki gibidir:



Şekil 2: Yetenek Keşfetme Analizlerinin Devlet Analitiği Hiyerarşisi

BÖLÜM 2: ARAŞTIRMANIN YÖNTEMİ

Bu bölümde, araştırma yönteminde faydalanılan metin madenciliği ve coğrafi bilgi sistemleri yaklaşımları hakkında bilgiler paylaşılmıştır. İlk kısımda, devlet analitiği araştırması için, metin madenciliği ve CBS, bu tezin yöntemindeki temel teknik yaklaşımlar olarak ele alınmıştır. İkinci bölümde ise, araştırmayla ilgili uygulama modeli, araştırma konusu için verilerin toplanması ve analitik yazılıma verilerin hazırlanması ile ilgili süreçler anlatılmıştır.

2.1. Araştırma Yöntemindeki Yaklaşımlar

Araştırma yöntemi belirlenirken, incelenecek veriler üzerinde teknikler uygulanır. Bu araştırmada, büyük veri analitiği yaklaşımıyla, devlet analitiği çerçevesinde ifade edilen ulusal beşeri sermaye analitiği kapsamında yetenek keşfetme analizleri için iş ilanlarının metin bazlı analizleri ve ilanlardaki konum bilgilerinden bölgesel analizler yapılacaktır. Dolayısıyla, metin madenciliği ve CBS, bu araştırmanın temel yöntemlerini oluşturmuştur.

2.1.1. Metin Madenciliği

Metin madenciliği, daha önceden bilinmeyen bilgilerin keşfedilmesi için farklı metin kaynaklarından otomatik olarak bilgi çıkarma işi olarak tanımlanır. Dolayısıyla metin madenciliğinin amacı, henüz kimsenin yazamadığı ve kimsenin bilmediği bir bilgi, dolayısıyla hiç daha önceden bilinmeyen bir bilgiyi çıkarmak için yapılan madencilik yöntemidir (Hearst, 2003).

2.1.1.1. Metin Madenciliğinin Tarihçesi

Grimes (2007), metin analizlerinin internetin kullanımının yayılmasına, bilgi sistemlerinin ve analizlerin ilerlemesine paralel uzun bir tarihi olduğunu söylemektedir. Bu bağlamda, metin madenciliği kavramının ilk olarak 1990'ların sonlarında ortaya çıktığı bilinmektedir.

Metinlerle yapılan veri madenciliğinin ilk yıllarında, metin kaynağı “sözcük torbası” olarak görülmüştür ve bu yaklaşıma göre, metinlerde yer alan bir sonraki sözcüğü bulmak için “n-gram” adı verilen, bir olasılığa dayalı dil modeli olarak metin madenciliğinde

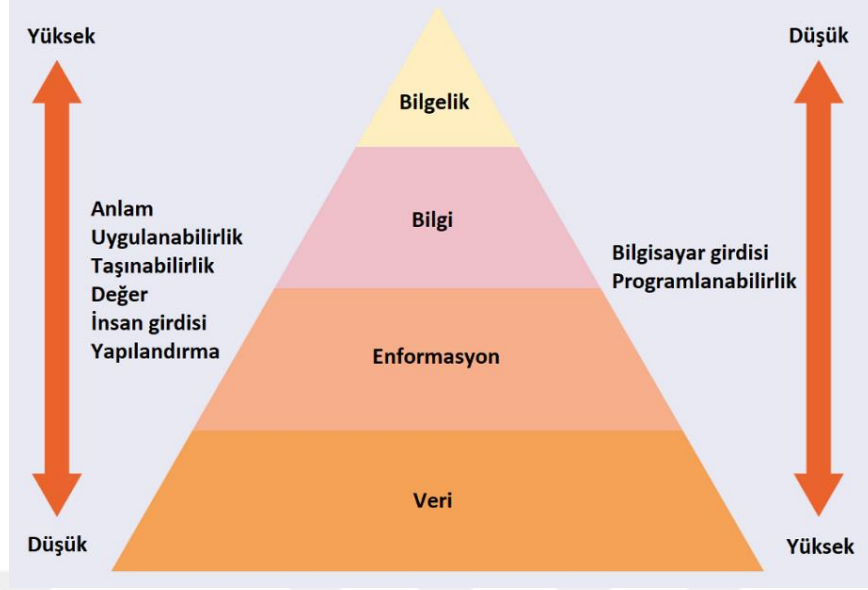
sıkça kullanılmaktadır. Ancak metin madenciliği kavramının ortaya çıkışından daha eski bir geçmişe dayanmaktadır. Aslında metinlerin analiz edilmesi iş zekası kavramıyla paralel ilerlemiştir (Grimes, 2007).

Grimes'in (2007) yorumuna göre, 1960'larda başlayan yönetim bilişim sistemleri ve bu alandaki iş zekası uygulamalarında belgelerin soyutlanması ve kodlanması amacıyla sözcük desenlerinden yararlanılmıştır. Bu nedenle, iş zekası ve dolayısıyla iş analitiğinin sayısal verilerin ayıklanması yerine metinlerin ayıklanması, kümelmesi ve sınıflandırılmasına odaklandığı görülmüştür.

Yönetim bilişim sistemlerinin 90'lara kadarki süreçte bir yazılım alanı olarak görülmesinin nedeni, ilişkisel veri tabanı sistemlerindeki ilerlemenin sayısal veriler üzerinde olmasındandır. Çünkü, metinler "*yapılandırılmamış*" verilerdir ve işlenmesi zorluklar içerir. Ancak sayısal odaklı yaklaşım, sonraki yıllarda analitik araçlar ve tekniklerin (raporlama, OLAP, veri madenciliği, ETL ve veri ambarları) benimsenmesiyle metin analizini kolaylaştırmıştır (Grimes, 2007). Nihayetinde, metin madenciliğinin ve metin analizlerinin yapılandırılmamış ve/veya yarı yapılandırılmış veriler içeren büyük veri kavramına kadar geliştirilen uzun bir literatürü ve köklü bir geçmişi vardır.

2.1.1.2. Metin Madenciliği Kavramları

Veri: Bocij ve diğerlerine (2008) göre, veri, işlenip bilgiye dönüştürülene kadar değeri çok az olan ya da hiç değeri olmayan ham gerçekler veya gözlemlerdir. İlişkisiz verilerin bir anlam ifade etmediği ve genellikle "*gürültü*" oluşturduğu tanımlanır. Bu nedenle, veriler bir bağlam içine yerleştirildiğinde anlam kazanırlar. Verinin ham halinden, bağlam içine yerleştirilerek anlam kazandırıldığı hiyerarşiyi açıklamak için DIKW (Data, Information, Knowledge, Wisdom) piramidi oluşturulmuştur. Bu hiyerarşi Şekil 3'te gösterilmiştir:



Şekil 3: DIKW Piramidi

Verinin bu hiyerarşisine göre, bir bilgisayar girdisi olduğu ve bilgisayardaki bir takım veri işleme faaliyetleriyle enformasyon aşamasına geçildiği anlaşılmaktadır. Bu enformasyon daha sonraki adımda, insan faktörüyle kullanılmak amaçlı bilgiye dönüştürülmektedir. Nihayetinde, verinin bu bilgisayarlı sistemler tarafından geliştirilen yazılım teknikleriyle yapılandırıldığı ve ilişkilendirildiği yöntemlerden biri de metin madenciliği metotlarıdır.

Veri Bilimi: Bilgisayar bilimleri, analizler ve dilbilim, ekonometri, sosyoloji gibi diğer disiplinlerden gelen teknikler ve araçlarla giderek bütünleştiği, birlikte yorumlandığı bir dünyaya doğru gidilmektedir. Yapılandırılmış, yarı yapılandırılmış ve yapılandırılmamış verilerden bilgi ve öngörü elde etmek için bilimsel yöntemleri inceleyen disiplinler arası bir bilim dalıdır. Veri madenciliği ve büyük veri kavramları bu alanda değerlendirilir (Dhar, 2013).

Veri Madenciliği: Veri madenciliğiyle ilgili literatürde birçok tanım yapılmıştır. Bu tanımlardan birkaçı aşağıda verilmiştir (Friedman, 1998):

- Veri madenciliği, potansiyel olarak yararlı ve nihayetinde verilerdeki sabit kalıplar altında geçerli romanın tanımlanması için anlaşılması zor bir işlemdir (Fayyad).

- Veri madenciliği, daha önce bilinmeyen anlaşılabilir ve eyleme geçirilebilir bilgilerin büyük veri tabanlarından çıkarılması ve önemli iş kararları almak için kullanılması işlemidir (Zekulin).
- Veri Madenciliği, bilgi keşif sürecinde veri içindeki önceden bilinmeyen ilişkileri ve örüntüleri ayırt etmek için kullanılan bir dizi yöntemdir (Ferruza).
- Veri madenciliği, verilerdeki avantajlı kalıpları keşfetme sürecidir (John).
- Veri madenciliği, bilinmeyen ve beklenmedik bilgi kalıpları için geniş veri tabanlarına baktığımız bir karar destek sürecidir (Parsaye).

Veri madenciliği, bu tanımlarda ifade edildiği üzere ve bu faaliyetlerin “madencilik” olarak adlandırılması nedeniyle veriler bütünüünün bir bilgi dağı olarak görüldüğü ve bu eylemin bu dağıın altındaki cevherleri ortaya çıkarmak olduğı anlaşılmaktadır. Diğer bir deyişle, buzdağının görülmeyen tarafını keşfetme işi denilebilir. Dolayısıyla metinsel içeriklerin depolandığı bir veri kaynağında bu metinlere yönelik bir bakışta, yüzeysel olarak elde edilemeyen bilgileri ayıklama işi, metin madenciliğidir.

Büyük Veri: Günümüzde en popüler kavramlardan biri olan Büyük Veri’nin kesin bir tanımı yoktur. Lohr’a (2013) göre, bu kavramın mucidi ABD’li bilgisayar bilimcisi John R. Mashey’dir. İlk kez Mashley (1997)’de ifade edilen Büyük Veri teriminin, buradan doğduğı düşünölmektedir. Büyük veri, özellikle insan davranışı ve etkileşimleri ile ilgili örüntüleri, eğilimleri ve birliktelikleri ortaya çıkarmak için hesaplanarak analiz edilmiş son derece büyük veri setleridir (Oxford Online Dictionary, 2020). Büyük verinin literatürde özelliklerine yer verilmiştir. Bunlar 3V olarak bilinen hacim, hız ve çeşitlilik boyutlarıyla birlikte kapsamlılık, kararlılık, dizinsellik, ilişkisellik, uzantısallık ve ölçeklenebilirlik gibi yeni kavramlardır (Kitchin ve McArdle, 2016). Bu bağlamda, büyük verinin hacim, hız ve çeşitlilik boyutlarıyla kısaca bir tanımı; yapılandırılmış ve yapılandırılmamış türde çok çeşitliliğe sahip, hacim olarak büyük alanları kaplayan ve bir akış hızına sahip olan veya zamanla büyüyen karmaşıklıkta kolaylıkla analiz edilemeyen veri setleridir.

İlişkisel Veri Tabanı: Günümüzde oldukça popüler olan ilişkisel veri tabanı sistemlerinin çıkış noktası Codd’un 1970’te önerdiği ilişkisel veri modeline dayanmaktadır (Codd, 2002). Bu modelin temel mantığı, verilerin normalizasyon kuralları çerçevesinde yapılandırılarak ilişkilendirilmesidir. Normalizasyon, verilerin

korunmasını ve verilerden oluşabilecek artıkları ve tutarsızlıkları kaldırmak amacıyla daha esnek bir veri tabanı inşa etmek için oluşturulmuş kurallardır (Microsoft, 2020). Dolayısıyla, bu ilişkisel veri tabanları, metin madenciliğinin daha esnek raporlama yeteneklerine nazaran hız ve performans odaklıdır. Sonuç olarak, metin madenciliği için diğer veri depolarından da yararlanılabilir. Bu yapılar; ilişkisel olmayan veri tabanları, veri ambarları ve farklı formatlardaki dokümanlar, videolar ve ses kayıtları biçiminde çok çeşitli dijital kayıtlardır.

Veri Ambarı: Normalizasyon kuralları ile oluşturulan sade, esnek ve hızlı bir veri tabanı tablolarından oluşturulan raporlar ve analizler eski yöntemlerle kısıtlıdır. Ancak veri ambarları daha fazla raporlar ve analizler almak için hazırlanmış veri tabanı sistemleridir. Bocij ve diğerlerine (2008) göre, veri ambarları karar desteği sunmak için bilgi üretmek amaçlı güncel ve geçmiş verileri bir arada depolayan veri tabanı organizasyonlarıdır. Bu nedenle, yazılımsal olarak ilişkisel veri tabanı tablolarından farklı değıllerdir fakat amaçları nedeniyle normalizasyon kuralları dışında kalan verileri depolayabilirler. Dolayısıyla, veri ambarları klasik ilişkisel veri tabanı tablolarına göre daha geniş bir raporlama ve analiz olanağı sunar.

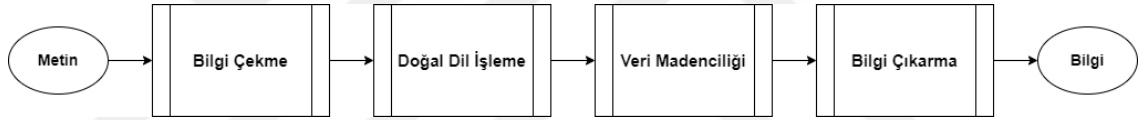
Veri ambarı sistemleri oluşturulurken iki ana yaklaşım vardır. Bu yaklaşımlar ETL (Ayıkla, Dönüştür, Yükle) ve ELT (Ayıkla, Yükle, Dönüştür) olarak özetlenebilir. Bu yaklaşımlar, DIKW piramidindeki verinin bilgiye dönüşme süreciyle doğrudan ilgilidir. Bunların yanı sıra, normalizasyon teknikleriyle birlikte Çevrimiçi Analitik İşleme (OLAP), Çevrimiçi İşlem İşleme (OLTP) gibi sistemler veri ambarı teknolojisinin ilerlemesine katkı vermektedir (Bocij ve diğerleri, 2008).

Metin Madenciliği Algoritmaları: Algoritmaların hedeflediği faaliyetler çerçevesinde birtakım görevler bulunmaktadır. Ghosh ve diğerlerine (2012) göre, metin madenciliği algoritmalarının görevleri; kategorizasyon, kümeleme, kavram madenciliği, bilgi alma ve bilgi çıkarma olarak özetlenebilir. Metin madenciliğinde kullanılan algoritmalarından bazıları Tablo 2’de verilmiştir.

Tablo 2: Popüler Metin Madenciliği Algoritmaları (Intellspot, 2020)

K-Means Kümeleme	Naive Bayes Sınıflandırıcı	K-En Yakın Komşu (KNN)	Destek Vektör Makineleri (SVM)	Karar Ağaçları
Genelleştirilmiş Doğrusal Modelleme	Yapay Sinir Ağları	Birliktelik Kuralları	Genetik Algoritmalar	Lantent Dirichlet Allocation (LDA)
Apriori Algoritması	PageRank Algoritması	Negatif Olmayan Matris Ayrıştırma	Minimum Açıklama Uzunluğu	Boosting Algoritmaları
Sıralama Algoritmaları	Bölme Algoritmaları	Bulanık Kümeleme	Hiyerarşik Kümeleme	Aglomerasyon Algoritmaları

Metin Madenciliğinin Aşamaları: Metin madenciliği yöntemi sürecinin aşamaları aşağıdaki Şekil 4’te gösterilmiştir (Ghosh ve diğerleri, 2012):

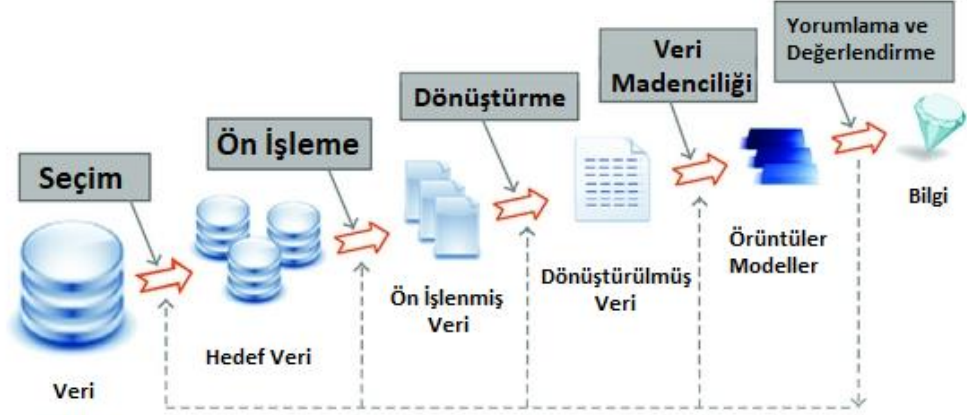


Şekil 4: Metin Madenciliği Sürecinin Aşamaları

Yukarıdaki şekilde görüldüğü gibi metin madenciliği, metin verisinin ham bir biçimde bilişim sistemine yüklendiği (bilgi çekme), ardından ön işleme süreci olarak doğal dil işleme (NLP) yöntemleriyle işlendiği, sonra veri madenciliği teknikleriyle metinden enformasyon elde edilerek en son aşamada bu enformasyonun yorumlanarak bilgi çıktısının kazanıldığı bir süreçtir. Burada NLP metotlarının veri madenciliği öncesinde uygulanması metin madenciliğini diğer madencilik uygulamalarından ayıran kritik bir noktadır. Bu bağlamda, NLP bilgisayarlı sistemlerin insan dilini anlayabilmesi için gerçekleştirilen analizlerin olduğu süreçtir ve yapay zeka alanındaki en eski ve en zor problemleri barındırır (Ghosh ve diğerleri, 2012).

Metin madenciliğinin NLP adımından sonra gelen veri madenciliği süreci, metinlerin analizleri ve yorumlanacak çıktıları ilgilendiren bir diğer önemli kısımdır. Metin madenciliği algoritmaları, NLP ve veri madenciliği süreçlerinde kullanılabilir. Bu

bağlamda, veri madenciliği süreçleri aşağıda Şekil 5’te gösterilmiştir (Fayyad ve diğerleri, 1996).



Şekil 5: Veri Madenciliği Sürecinin Aşamaları

2.1.2. Coğrafi Bilgi Sistemleri

Coğrafi Bilgi Sistemleri (CBS), karar desteği için arazi, altyapı, üretim kaynakları, çevre yönetimi gibi konularda mekansal analiz ile kentsel ve bölgesel kalkınma planlamasında giderek kabul gören uygulamalardan biri olmuştur. CBS, coğrafi verilerin hazırlanarak analiz edildiği, görüntülediği ve yönetildiği karar destek sistemleridir. (Fedra, K. ve Reitsma, 1990). Bu nedenle, özellikle kamu yöneticilerinin karar desteği olarak genellikle kullandıkları araçlardan biri haline gelmiştir.

CBS ile ilgili literatürde çeşitli tanımlar mevcuttur. “CBS, konumla ilişkilendirilmiş verileri toplayan, depolayan, işleyen ve sunan bir yönetim bilişim sistemidir.” denilebilir veya “CBS, haritacılığın veri tabanı ve istatistiksel analiz geliştirildiği bir bilgisayarlı sistemdir.” şeklinde tanımlanabilir (Dawsen, 2011).

Diğer bir KDS ve CBS uygulamalarına bakış açısını ise Zhu (2013), CBS'nin haritalama ve görselleştirme yetenekleriyle karar desteği yeteneklerini olarak ikiye ayırarak yapmıştır. Zhu'ya göre CBS, verilerin haritalanmasını ve kendine özgü temel analizleri sunan bir sistemdir. Ancak KDS olarak beklenen analitik yetenekler daha farklıdır. Bu nedenle, KDS için geliştirilen analitikler coğrafi veya mekansal problemlerin çözümlerine yönelik modeller içermelidir. Dolayısıyla bu iki ayrı bölümde değerlendirilen CBS ve KDS bir üst kavramda birleşirler ve bu kavram Mekansal Karar Destek Sistemi'dir.

Bu bağlamda, CBS yaklaşımlarının ve literatürdeki gelişmelerin CBS için KDS çerçevesindeki konumlandırılmasının farklı bakış açılarına sahip olduğu bilinmelidir. Bir başka görüş, Eldrandaly (2010), KDS'leri ve bu çerçevede CBS'yi açıklamıştır. Eldrandaly'a göre KDS için 3 kategori belirlemiştir. Bunlar, *Çok Kriterli Karar Verme Teknikleri, Uzman Sistemler ve Coğrafi Bilgi Sistemleridir*. Zhu'nun yaklaşımına benzer olarak, Eldrandaly CBS'nin analitik performansının artırılması için uzman sistemlerle ve çok kriterli karar verme teknikleri ile entegrasyonunun yapılabileceğini aktarmıştır.

Buradaki yorumlardan bir çıkarım yapılacak olursa ve bu bağlamda, Türkiye'deki CBS uygulamaları incelendiğinde CBS'nin analitiğinin geliştirilmesi yöntem araştırmalarında kritik bir konu olarak görülmüştür. Bununla birlikte, CBS'nin ve bu karar desteği alanının yalnızca verilerin haritalanmasını ve görselleştirme süreçlerini kapsamadığı, burada da analizlerin çok önemli rol aldığı anlaşılmaktadır.

Nihayetinde, CBS karar desteği niteliği olan, kamu yöneticilerin aşına olduğu bir sistemdir. CBS, verilerdeki mekansal bilgileri kullanarak, haritalama ve konum bazlı analizler gerçekleştirme gibi kendisine özgü yeteneklere sahiptir. Bu nedenle, CBS özellikle kamu yönetimine stratejik planlamada karar alıcılara destek vermek için olumlu katkılar üretir. Sonuç olarak, devlet analitiği uygulamalarında CBS'nin rol alması elzem bir konudur.

2.2. Araştırma Modelinin Ve Veri Setinin Hazırlanması

Önceki bölümlerde araştırmanın yöntemi verilmiştir. Bu bölümde, araştırmayla ilgili uygulama modeli, araştırma konusu için verilerin toplanması ve analitik yazılıma verilerin hazırlanması ile ilgili süreçler anlatılacaktır.

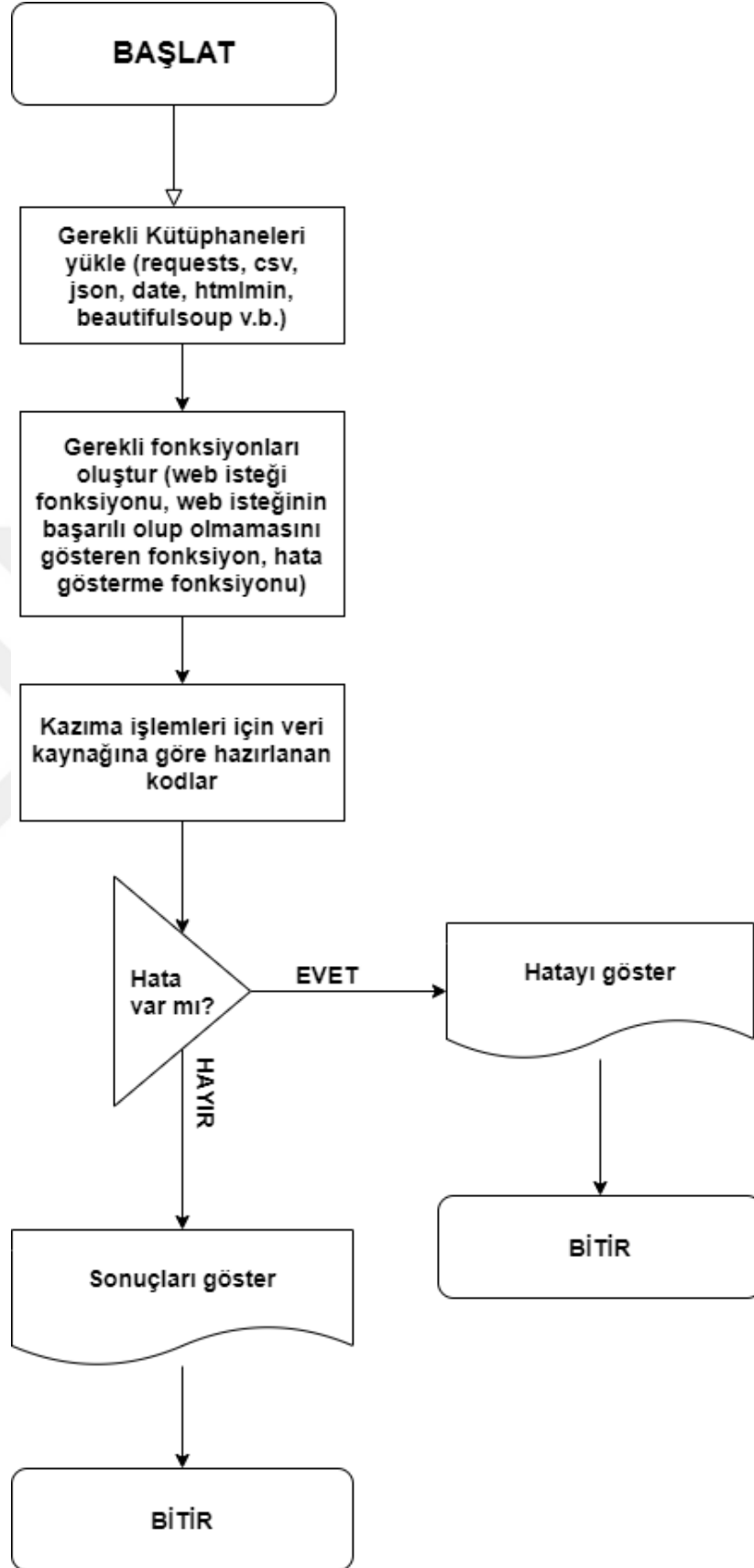
2.2.1. İş İlanlarının İnternet Üzerinden Toplanması

Bilimsel araştırmanın yöntemine göre veri toplama işi değişiklik gösterebilir. Büyük veri analitiği ve makro düzey bir kapsamı olan bu araştırma binlerce iş ilanına ihtiyaç vardır. Bu nedenle, klasik veri toplama yöntemlerinin yerine modern teknikler araştırılmıştır. Bu ölçekte, Mitchell (2018)'in ifade ettiği internetten veri kazıma (web kazıma) yöntemini, büyük miktarda verileri hızlı ve işlemek için çok güçlü bir teknik olarak bilir.

Veri kazıma için programlama bilgisi gerekir. Python programlama dili bu teknik için kolaylıklar barındırmaktadır. Tez araştırmasında, verilerin kaydedilmesi ve metin madenciliği uygulaması için öncesi ön temizleme (verileri paket yazılıma uygun formata getirmek) için Python dilinden yararlanılmıştır. Bu ön temizleme işlemi, *veri madenciliği sürecinde ifade edilen NLP adımı değildir*. Araştırmadaki Doğal dil işleme süreci Wordstat yazılım paketinde yapılmaktadır. İş ilanları belirlenen araştırma planına göre periyodik olarak, Kariyer.Net internet sitesinden hazırlanan Python betikleri ile indirilmiştir.

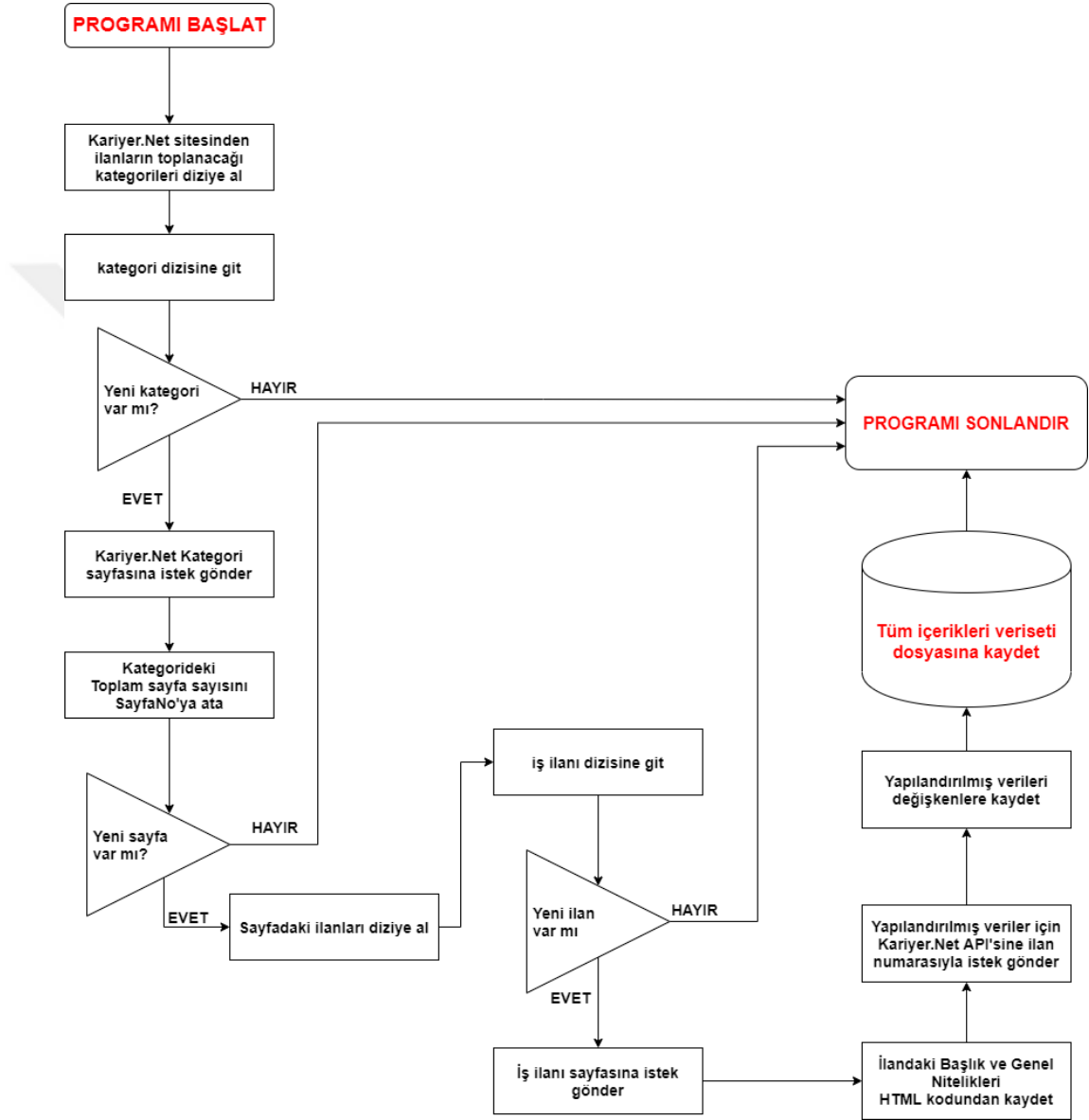
Python dilinin, kendine özgü kısıtları ve problemleri vardır. Bu nedenle, hazırlanan betiğin işlem sırasında kontrol edilip tekrar doğru çalışır hale getirilmesi için kodların güncellenmesi gerekebilir. Mitchell (2018), kitabında anlattığı üzere API'ler (Uygulama Programlama Arayüzü) veri toplayıcı botlar hazırlamak için uygun formatlar sunar. API'ler güncel hayatta, yazılımlarda kullanılan web hizmetleri olarak tanınırlar.

Kariyer.Net'teki internet sitesi ve sitedeki iş ilanlarının yapısı incelendiğinde yapılandırılmış ve yapılandırılmamış olarak farklı türde veriler yayınlandığı tespit edilmiştir. Bu bağlamda, yapılandırılmış verileri Kariyer.Net API'lerinden, yapılandırılmamış veriler için ise sayfa kaynağından HTML kodları parçalanarak kazıma işlemleri gerçekleştirilmiştir. Veri kazıma yönteminin temel algoritmasını ifade eden akış şeması Şekil 6'daki gibidir:



Şekil 6: Veri Kazıma Algoritmasının Akış Şeması

Bu algoritma Python dili ile internetteki verilerin toplanması için genel bir çerçeve oluşturmaktadır ve aynı zamanda diğer programlama dillerine de uyarlanabilir. Şekildeki algoritmada veri kazımalarının yapılacağı siteyle ilgili kod bölümü bir başka işlem olarak ifade edilmiştir. Bu bağlamda Kariyer.Net için kurulan kodlamanın algoritması Şekil 7'deki gibi gerçekleştirilmiştir:

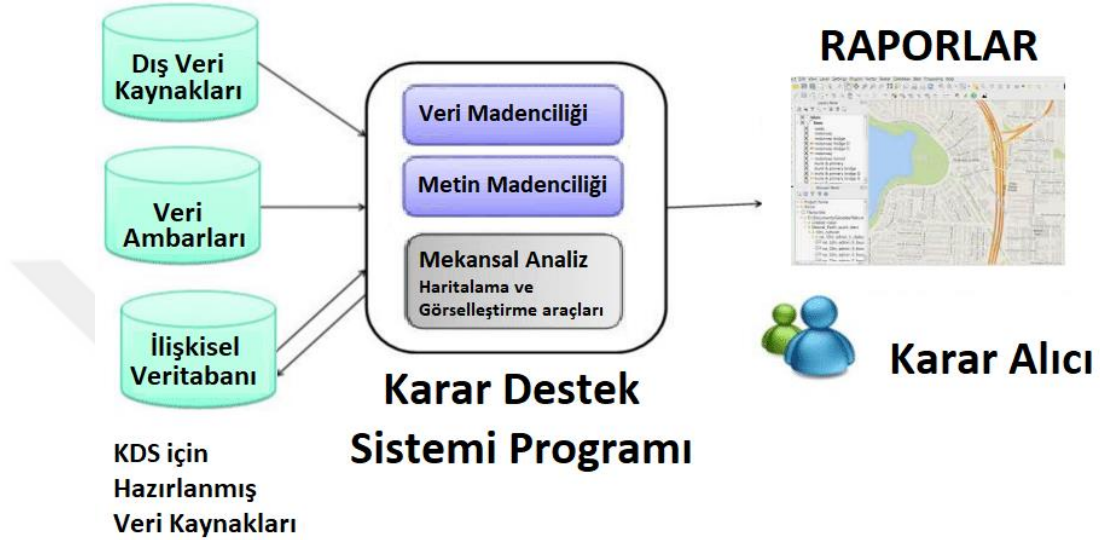


Şekil 7: Kariyer.Net Veri Kazıma İşlemleri Akış Şeması

Şekil 6 ve Şekil 7'deki akışa göre, veri seti hazırlanırken indirme işlemleri, Kariyer.Net üzerinden 2020 yılı Ocak ve Şubat aylarında iki defa veri indirme işlemi gerçekleştirilmiştir.

2.2.2. Araştırmanın Modeli

Araştırma için veri kazıma ile veriler elde edilmiş ve veri seti hazırlanarak Wordstat yazılımıyla ampirik bir çalışma yapılmıştır. Dolayısıyla veri seti, modelde büyük veriyi temsil etmektedir. Wordstat ise önerilen modelde, KDS uygulamasının metin madenciliği ve CBS desteğini sunan yazılımdır. KDS, Şekil 8’deki gibi bir model kullanmaktadır.



Şekil 8: Araştırmanın KDS Modeli

Sonuç olarak, iş ilanları veri kazıma teknikleriyle indirilmiştir ve ön temizleme aşamaları Python betikleri ve Excel ile gerçekleştirilmiştir. Sonuç olarak, araştırmanın KDS modeli Şekil 8’deki gibidir ve KDS yazılımı Wordstat olmuştur. Provalis Research (2020), Wordstat paket yazılımıyla metin analizleri ve buna entegre olarak mekansal analizler yapabilmek için GIS Wiever adında bir başka paketi bir arada verebilmektedir. Dolayısıyla, Wordstat araştırmacıların kolaylıkla mekansal analizler için CBS ve doküman analizleri için metin madenciliği yöntemlerini işleyebilecekleri ve bir arada kullanabilecekleri paket bir programdır (Akbıyık, 2019).

2.2.3. Veri Setinin Analizlere Hazırlanması

Veriler CSV formatında Excel dokümanı olarak saklanmıştır. Ancak veriler CSV formatında kaydedildiğinde, ilk ham veri setinde Wordstat programına uyumsuz kirli ve gürültülü veriler bulunmaktadır. Bu nedenle, veri setini analiz programına almadan veri setinin temizlenmesi gerekmiştir.

Bu temizleme süreci, uygulama modelinde (Şekil 8) KDS öncesi bir işlem olarak görülmelidir ve daha önce de belirtildiği üzere, KDS çerçevesindeki veri madenciliğinin NLP süreciyle karıştırılmamalıdır. Nihayetinde, analizlerin hesaplanacağı paket yazılıma veri seti alınmadan önce, verilerdeki gürültünün giderilmesi için CSV formatındaki dosyada Python ile hazırlanmış betikler ve Excel yazılımı üzerinde düzeltmeler yapılmıştır.

İlk olarak Ocak ve Şubat aylarında toplanan verilerdeki karakter hataları ve veri setindeki sütunlar arası kaymalar ve karakter hataları Python betikleriyle ve Excel’de bul ve değiştir yöntemiyle düzeltilmiştir. Ardından kirli verileri bu şekilde temizlenmiş iki veri dosyasındaki veriler, alt alta bir Excel dosyasında toplanmıştır. Veri setinde ilan numarası anahtar alan olarak belirlenmiştir. Konum bilgileri hariç yapılandırılmış verileri kayıt altına almak üzere, öznitelik alanları oluşturulmuş, ardından ilan numarasına göre tekrar eden kayıtlar temizlenmiştir.

Kariyer.Net verilerinde, ilanların ülke, şehir, ilçe ve mahalle bilgileri tek bir alanda birleştirilmiş halde tutulmaktadır. İlanın konumunu belirleyecek olan bu bilgi için, öncelikle Türkiye ölçeğinde il bazlı çalışmak uygun bulunmuştur. Dolayısıyla verilerdeki yabancı ülke ve ilçe ile daha detaylı mahalle bilgileri şehir bilgisi sütunundan kaldırılmıştır. Fakat bu temizlik işlemi, ilanın metin analizlerinin yapılacağı içeriğinde yer alan konum bilgileriyle ilgili değildir.

Bu bağlamda, yapılandırılmamış konum bilgilerindeki gürültünün giderilmesinin ardından ilanlardaki birleşik halde yer alan şehir bilgileri ayrıştırılarak her şehir yeni satırlar oluşturacak şekilde dağıtılmıştır. Böylece verilerdeki kirlilik ve gürültü gibi seti bozan yapılar giderildikten sonra şehir bilgilerine göre genişletilmiştir.

Haritalama için gereken enlem ve boylam bilgileri, Wordstat tarafından alınabilmektedir (Akbıyık, 2019). Ancak veri setinde enlem ve boylam bilgisinin hazır halde bulunmasının pratik uygulamada avantaj sağlayabileceği gerekçesiyle araştırmacı tarafından hazırlanmıştır. Bu işlem için kullanılan Python betiği ile veri kazıma örneği olarak Şekil 9’da paylaşılmıştır:

```

BASE_URL = "http://py4e-data.dr-chuck.net/json?key=42&address="

country_names = ["Adana","Adiyaman","Afyonkarahisar","Ađrı","Aksaray","Amasya","Ankara","Antalya","Ardahan","Artvin",
,"Aydın","Balıkesir","Bartın","Batman","Bayburt","Bilecik","Bingöl","Bitlis","Bolu","Burdur","Bursa",
,"Çanakkale","Çankırı","Çorum","Denizli","Diyarbakır","Düzce","Edirne","Elazığ","Erzincan","Erzurum",
,"Eskişehir","Gaziantep","Giresun","Gümüşhane","Hakkâri","Hatay","Iğdır","Isparta","İstanbul","İzmir",
,"Kahramanmaraş","Karabük","Karaman","Kars","Kastamonu","Kayseri","Kırıkkale","Kırklareli","Kırşehir",
,"Kilis","Kocaeli","Konya","Kütahya","Malatya","Manisa","Mardin","Mersin","Muğla","Muş","Nevşehir",
,"Niğde","Ordu","Osmaniye","Rize","Sakarya","Samsun","Siirt","Sinop","Sivas","Şanlıurfa","Şirnak",
,"Tekirdađ","Tokat","Trabzon","Tunceli","Uşak","Van","Yalova","Yozgat","Zonguldak"]

scraped_data = []

with open("koordinatlar.csv", "w", newline="") as csvfile:
    writer = csv.writer(csvfile, delimiter=',', quotechar='"')

    headers = ["Konum", "Enlem", "Boylam"]
    writer.writerow(headers)
    csvfile.close()

x=0

for brand in brand_names:
    BRAND_URL = BASE_URL + brand
    x=x+1
    r = requests.get(BRAND_URL)
    data = r.json()

    data = data['results'][0]['geometry']['location']

    lat = "*" + str(data['lat']) + "*"
    lng = "*" + str(data['lng']) + "*"

    print(str(x) + ". şehir: " + brand + "\n")

    sehir = "*" + brand + "*"

    row = [sehir, lat, lng]

    scraped_data.append(row)

with open("koordinatlar.csv", "a", newline="", encoding="utf-8") as csvfile:
    writer = csv.writer(csvfile, delimiter=',', quotechar='"')

    for row in scraped_data:
        writer.writerow(row)
        scraped_data.clear()
    csvfile.close()

```

Şekil 9: Enlem Ve Boylam Bilgileri İndirme Betiđi

Şehirlerin enlem ve boylam bilgilerini Dr. Severance'in açık kaynak olarak sunduđu internet sitesindeki CBS servisleri ile json formatında <http://py4e-data.dr-chuck.net/json> adresinden edinilebilmektedir. Şekil 9'daki betik ile adres bilgileri bir CSV formatında verisetinde kullanılmak için kaydedilmiştir. Aynı şekilde iş ilanları Şekil 9'daki betikten farklı olarak Kariyer.Net internet sitesi üzerinden araştırma sürecinde belirlenmiş tarihlerde HTML ve JSON formatındaki kayıtlardan süzülerek indirilmiştir. JSON ve HTML verilerini parçalama ve CSV formatında saklamak için hazırlanan betikte JSON ve Richardson (2020) tarafından paylaşılan BeautifulSoup4 kütüphaneleri kullanılmıştır.

Bu konuda karşılaşılan zorluklardan biri, bazı iş ilanlarında firma bilgilerinin bulunmamasından dolayı ilan içeriklerinin boş indirilmesi olmuştur. Makro anlamda bir araştırma sebebiyle bütüncül bir yaklaşım olarak, analiz edilecek tüm iş ilanlarına erişilmek istenildiđi için eksik indirilen ilanlar tespit edilip tekrardan yeni betikler

hazırlanmıştır. Şekil 10'da firma bilgisi olmayan ilanların nasıl toplandığına dair hazırlanmış örnek Python kodu verilmiştir. Gelecekteki çalışmalar için kütüphanelerin ve sayfa yapılarının güncel sürümlerine uygun bir şekilde betikleri oluşturmaları tavsiye edilir.

İndirilen veriler tek bir CSV dosyasında birleştirilmiştir. Bu veri setinde şehir ile enlem ve boylam eşleşmesi gerçekleştirilerek veri kaynağı dosyasında konum bilgileri Wordstat ortamına hazır hale getirilmiştir. Bu yaklaşımın sebebi, GISViewer yazılımına analizler aktarılırken Şehir bilgisinden konum bilgilerinin hızlı bir şekilde projede işlemektir. Araştırmacılar Wordstat'ın etkinliğini ve performansını ölçmek için kendi tercihlerine göre verileri hazırlarken Wordstat üzerinden de enlem boylam bilgileri oluşturabilirler.

```
for page in non-firm:

    say=say+1

    ilan_url = BASE_URL + page
    print(" "+str(say)+" . ilan no: " + page)

    raw_html = simple_get(ilan_url)
    ilan = BeautifulSoup(html_minify(raw_html), "html.parser")

    genel_nitelikler = ilan.find("table")
    if genel_nitelikler != None:
        genel_nitelikler = str(ilan.find("table"))
        genel_nitelikler = genel_nitelikler.replace("<br>"," ")
        genel_nitelikler = genel_nitelikler.replace(","," ")
        genel_nitelikler = re.sub('<[^<]+?>', '', genel_nitelikler)

    genel_nitelikler = str(genel_nitelikler)
    deneme = genel_nitelikler.find("itelikler:")
    if deneme != None:
        nitelik = genel_nitelikler.split("itelikler:")
        if len(nitelik)>1:
            row = [page,nitelik[1]]
        else:
            row = [page,nitelik[0]]

    scraped_data.append(row)

    with open("genelnitelik-yok.csv", "a", newline="") as csvfile:
        writer = csv.writer(csvfile, delimiter=';', quotechar='')

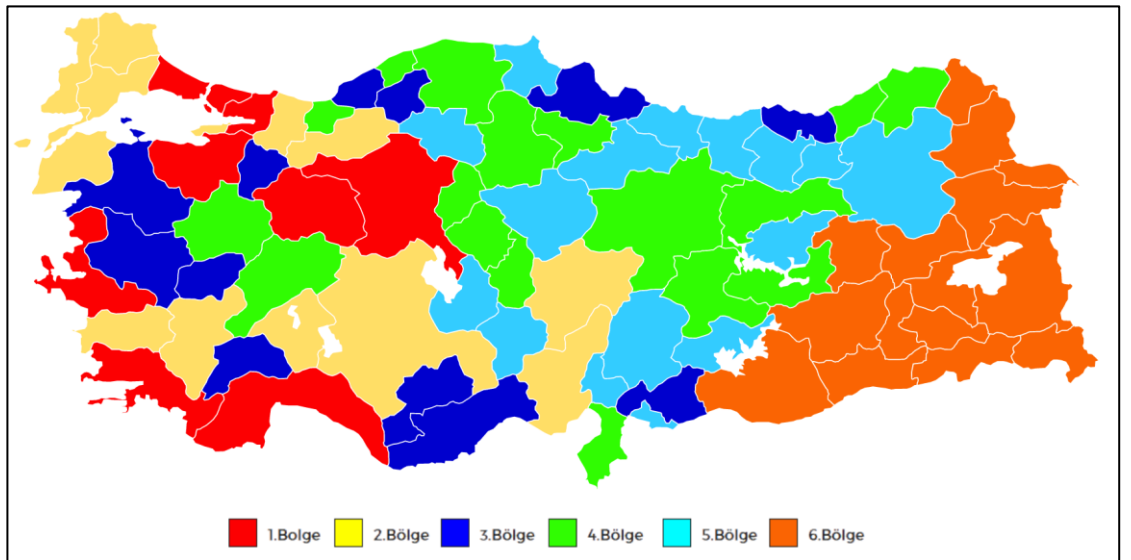
        for row in scraped_data:
            writer.writerow(row)
        scraped_data.clear()
    csvfile.close()
```

Şekil 10: Firma Bilgisi Olmayan İlan İçeriklerini İndirme Betiği

Veri setindeki bütünlük sağlandıktan sonra, karakter hataları hazırlanan Python betikleri ile giderilmiş ve son bir eşleşme olarak yatırım teşvik bölgelerindeki dağılımlar için, bölge numaraları atamaları yapılmıştır. Sonuç olarak, Wordstat yazılımına aktarılmak üzere Excel dosya formatında (xlsx) iş kayıtlar tek bir dosyada kayıtlar tutulmuştur.

Birleştirme işleminin ardından, veri setini homojen hale getirmek ve normalize etmek için Türkçe dili dışında yayınlanmış İngilizce, Almanca ve Fransızca ilanlar kaldırılmıştır. Bir sonraki temizlik aşamasında iş ilanlarının yayınlanma tarihleri 2019 yılı öncesi olan iş ilanları kaldırılmıştır. Nihayetinde 1 Ocak 2019 - 13 Şubat 2020 tarihleri arasında yayınlanmış, Türkçe içeriklere sahip 81 il bazında 48426 adet iş ilanı metin madenciliği analizleri ve coğrafi analitikler gerçekleştirmek için Wordstat yazılıma aktarılmıştır.

Analizler için metin madenciliği çerçevesinde iş ilanlarının başlık, içerik, konum ve tarih bilgilerinin yanı sıra tezin temel hedeflerine uygun olarak, son olarak coğrafi analizleri zenginleştirmek üzere veri setine son aşamada eklenen yatırım teşvik bölgeleri boyutu kullanılmıştır. Bu konuda, yatırım teşvik bölgelerini belirlemek üzere, T.C. Resmi Gazete’de 7 Ağustos 2019 tarihinde yayınlanmış Cumhurbaşkanlığı Kararnamesi (1402 Karar Sayılı Yatırımlarda Devlet Yardımları Hakkında Kararda Değişiklik Yapılmasına Dair Karar, 2019) baz alınarak güncellenmiş olan 2020 yılı Türkiye Bölgesel Yatırım Teşvik Haritası kullanılmıştır (Denge Müşavirlik, 2020). Bölgesel yatırım teşviklerinin haritadaki dağılımı aşağıdaki şekilde gibidir:

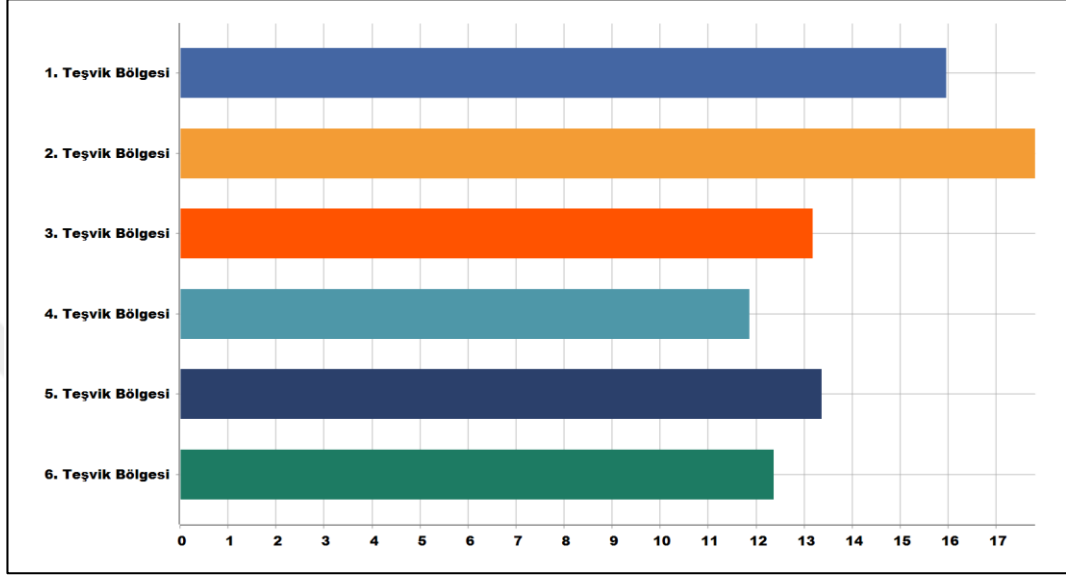


Şekil 11: İllere Göre Yatırım Teşvik Bölgeleri Haritası

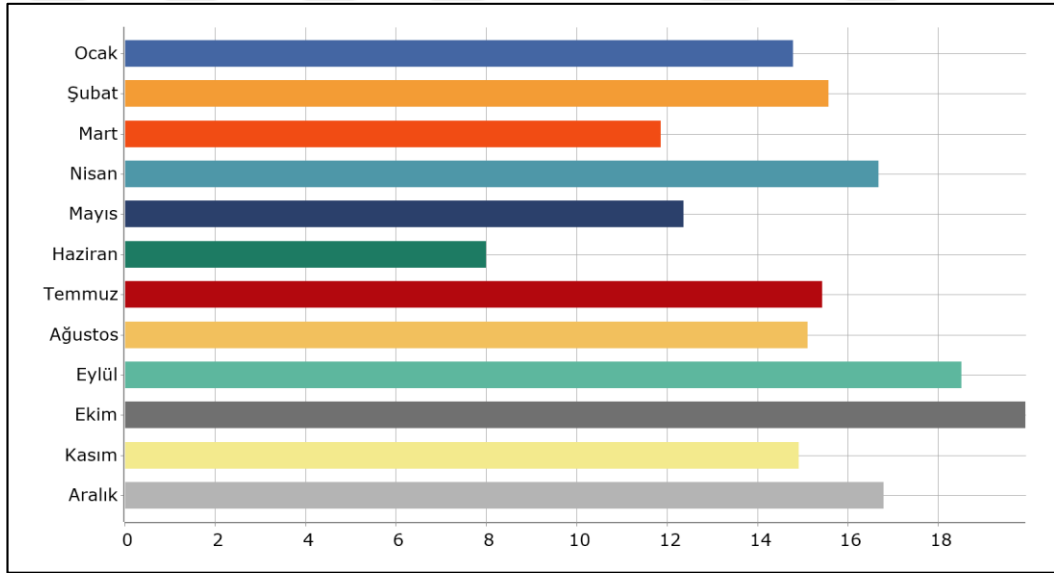
Sonuç olarak bu tez çalışmasında analiz edilmek üzere, yatırım teşvik bölgeleri eklenmiş veri setinde *Başlık, Genel Nitelikler, Şehir, Enlem, Boylam, Yatırım Teşvik Bölgesi, Çalışma Şekli, Pozisyon Seviyesi, İlan Yayınlanma Tarihi* şeklinde 9 nitelik alanı bulunmaktadır.



Bu amaçla frekansı 1000'den fazla olan ve yatırım teşvik bölgeleri düzeyindeki ve/veya 12 aylık dönemdeki dağılımları anomali gösterdiği tespit edilen; *Analitik, Bilgisayar, Elektronik, Gayrimenkul, Gıda, Kalite, Lojistik, Mağazacılık, Muhasebe, Mühendislik, Otomotiv, Perakende, Sağlık, Üretim, Yazılım ve Yenilikçi* kelimeleri seçilmiştir.



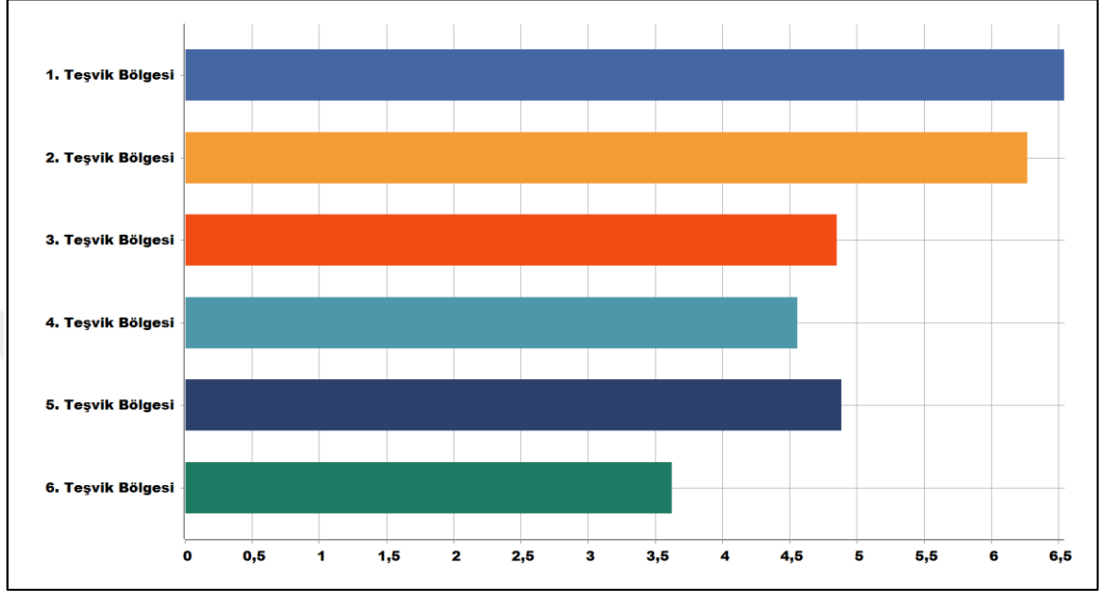
Şekil 13: Analitik – Yatırım Teşvik Bölgesi Dağılımı



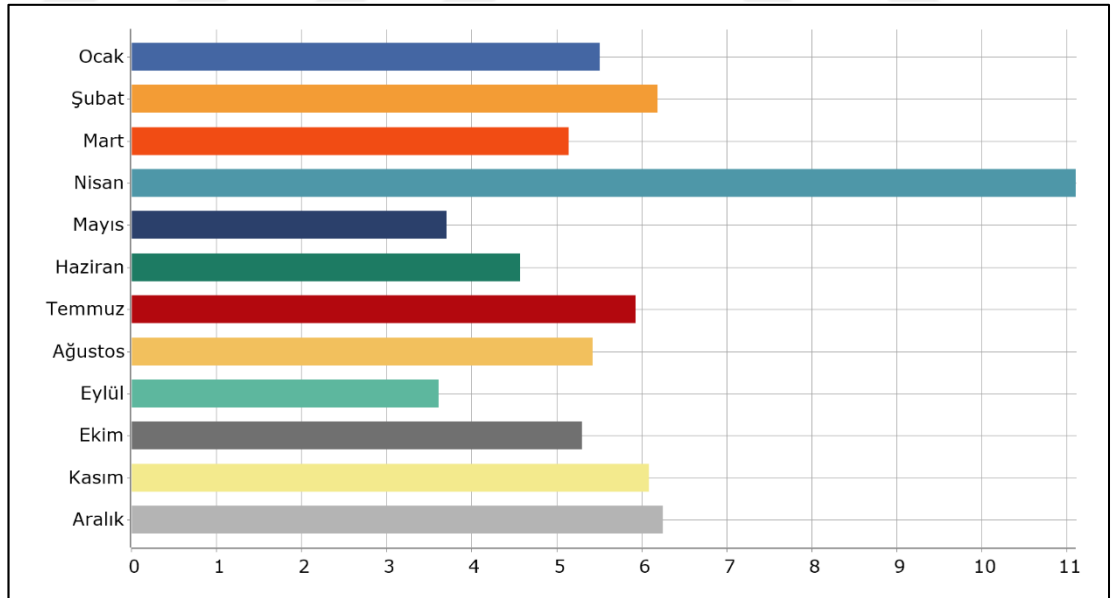
Şekil 14: Analitik – Aylık Dağılımı

Şekil 13'e göre; Analitik teriminin 2. Bölge'deki şehirlerde trendinin yüksek olduğu bulunmuştur. Buna göre; bu niteliğin 1. Bölge'den diğer bölgelere doğru geçişkenlik göstermeye başladığı anlaşılabilir. Şekil 14'teki bulgulara göre; yaz başına doğru bir

azalış olmasına rağmen tüm aylarda dağılım göstermesi, dönemsel olarak farklılık göstermeyen popüler bir terim olduğunu ifade ediyor. 4. Bölge, 5. Bölge ve 6. Bölge frekans düzeylerinin diğer bölgelere yakınlığı ve aylık dağılım, Analitik teriminin ülke genelindeki yakın dönemdeki trend bir nitelik terimi olduğunu söylemektedir.



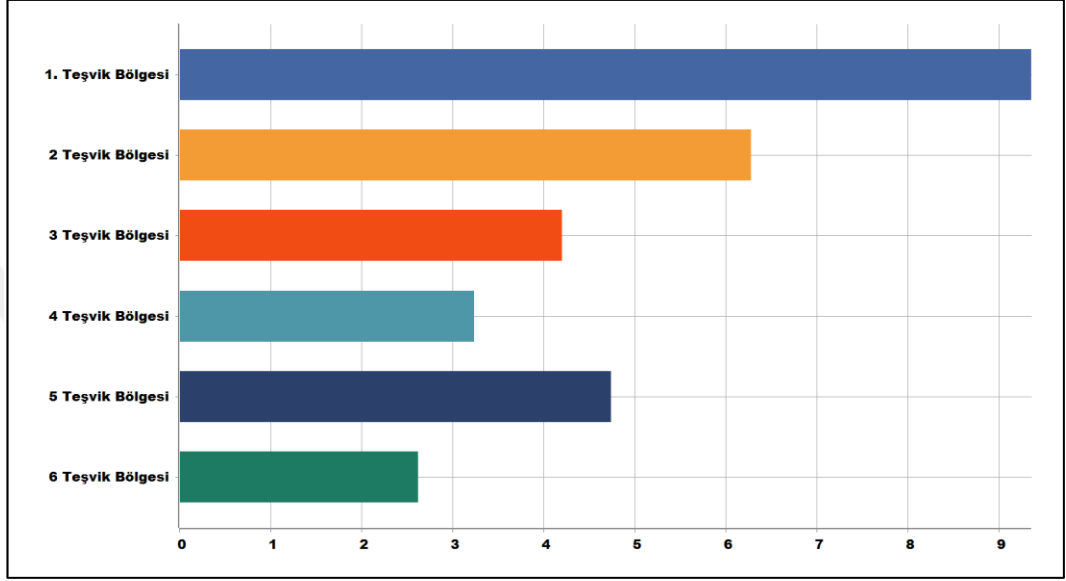
Şekil 15: Bilgisayar – Yatırım Teşvik Bölgesi Dağılımı



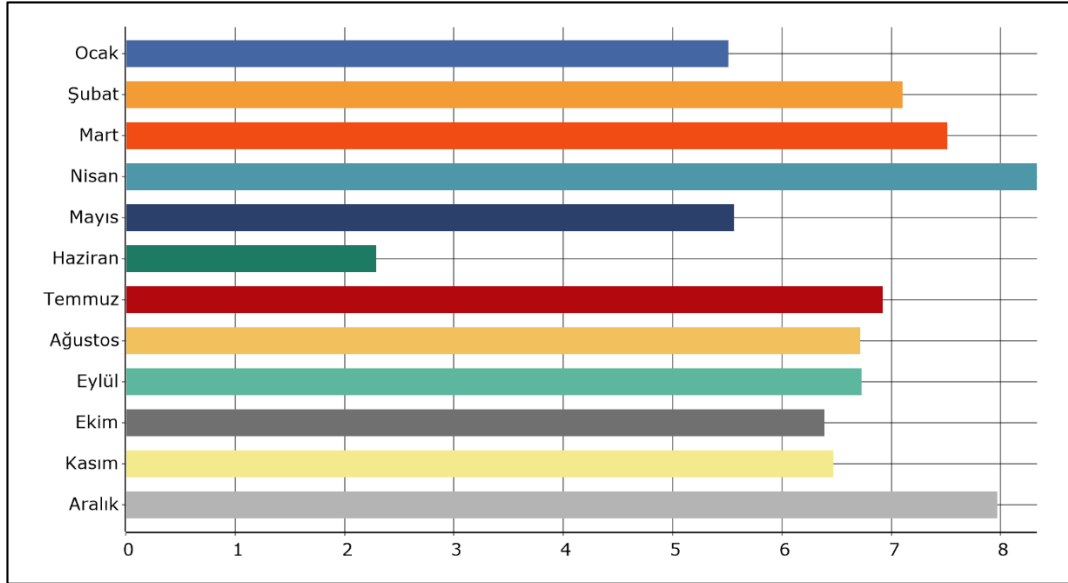
Şekil 16: Bilgisayar – Aylık Dağılımı

Şekil 15'e göre; Bilgisayar teriminin yatırım teşvik bölgelerinin düzeyine göre paralel bir azalış sergilediği görülmektedir. Nüfusun ve istihdamın yüksek olduğu şehirlerde bu niteliğe talebin daha çok olduğu anlaşılmaktadır. Şekil 16'ya göre, Bilgisayar niteliğinin

Nisan ayında yüksek bir frekansı olup, yaz aylarında diğer dönemlere göre talebinin düştüğü anlaşılmaktadır. Yaz döneminde ise Temmuz ayında tekrar bir yükseliş gösterdiği bulunmuştur. Dolayısıyla, bu nitelik için, işgücü planlamasının bahar ortalarında yapıldığı ve firmaların güz dönemine bu yeteneği kazandırmak istedikleri tahmin edilmektedir.



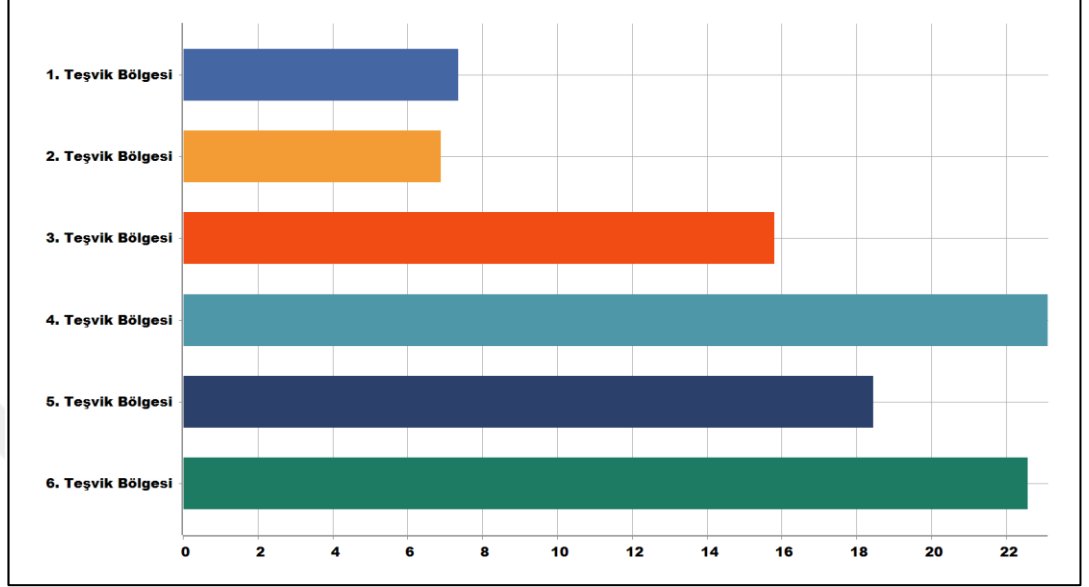
Şekil 17: Elektronik – Yatırım Teşvik Bölgesi Dağılımı



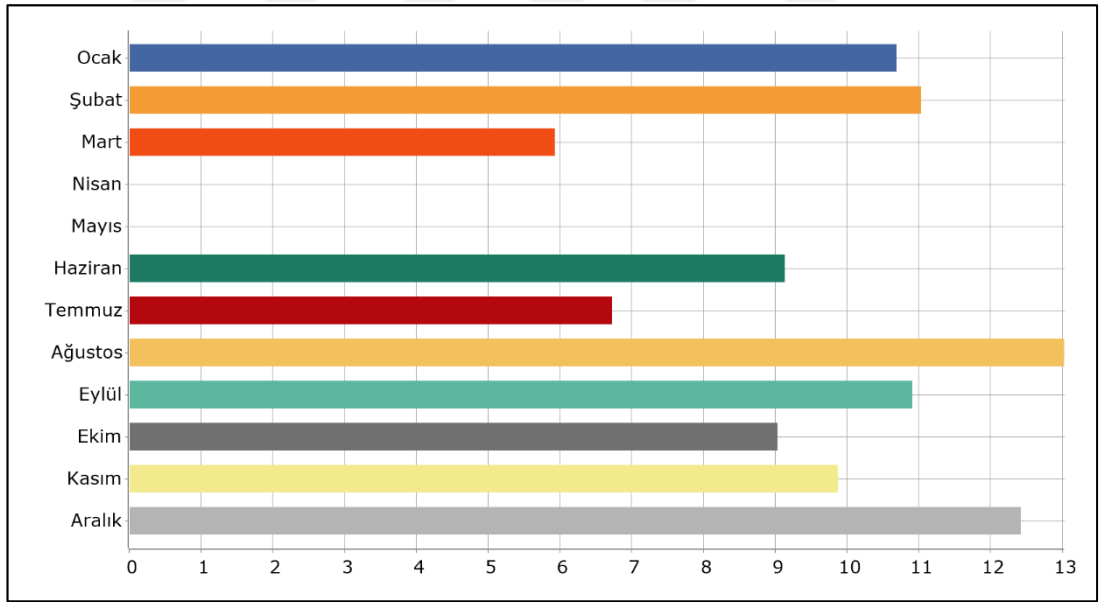
Şekil 18: Elektronik – Aylık Dağılımı

Şekil 17 ve Şekil 18’deki Elektronik niteliği dağılımları, yatırım teşvik bölgesi düzeyinde iş arayanlar için özellikle 1. Bölgenin cazibe yarattığı, işverenler açısından ise Haziran

ayı hariç bu niteliğin sıkça arandığı görülmektedir. Diğer bir teknolojik ifade olan, Bilgisayar terimi ile kıyaslandığında Elektronik niteliğini talebinin az olduğu dikkat çekmiştir.



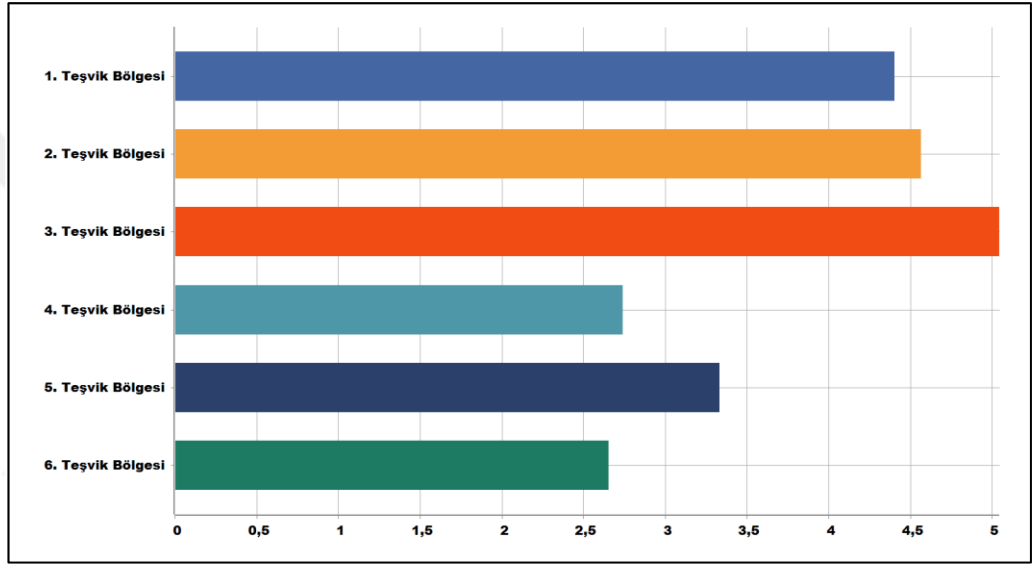
Şekil 19: Gayrimenkul – Yatırım Teşvik Bölgesi Dağılımı



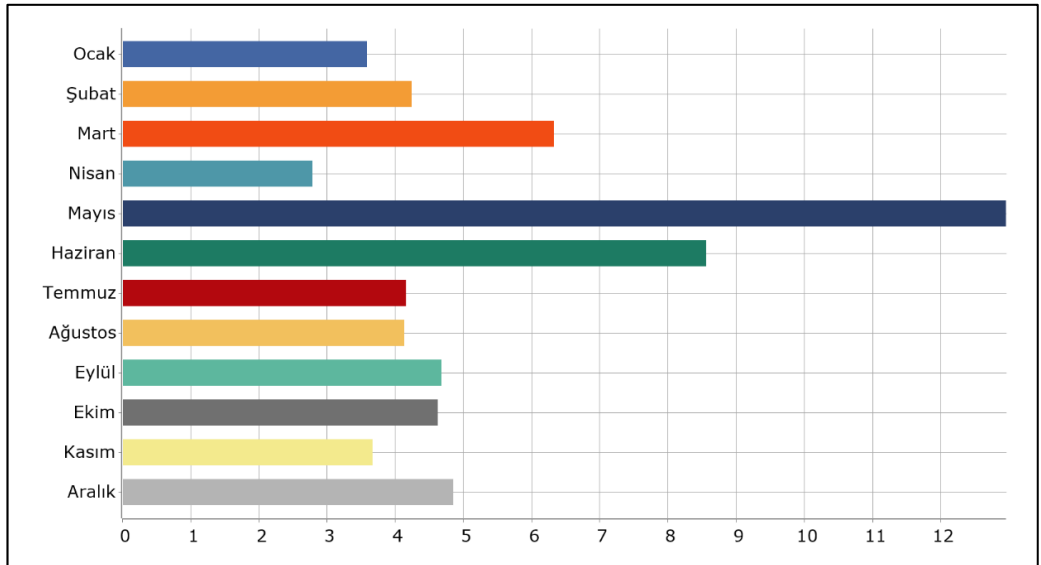
Şekil 20: Gayrimenkul – Aylık Dağılımı

Gayrimenkul terimi, Şekil 19'daki bölgesel dağılıma göre, düşük seviye yatırım teşvik bölgelerinde bir çekim alanı oluşturmuştur. Bu ölçekte, nüfusun ve istihdamın düşük olduğu az gelişmiş kırsal bölgelerde trend gösteren bir nitelik olduğu anlaşılmaktadır.

Şekil 20'deki Gayrimenkul dağılımı, Nisan ve Mayıs aylarında bu niteliğe hiçbir talep olmadığını göstermektedir. Veri setinin en son 2020 yılı Şubat ayındaki kayıtları bulundurduğu ve ilanların ortalama 60 ila 90 gün yayında kaldığı göz önüne alındığında, Gayrimenkul iş ilanlarının devir döngüsünün 1 yıllık zaman diliminde çok hareketli olduğu anlaşılmaktadır. Bu sonuç, Gayrimenkul niteliğinin popülaritesini gösterdiği gibi yatırım teşvik bölgeleriyle birlikte değerlendirildiğinde ülke genelinde sıkça talep edilmesinin yanında büyükşehirlerin aksine, kırsalda daha ciddi bir talep olduğunu ifade etmektedir.



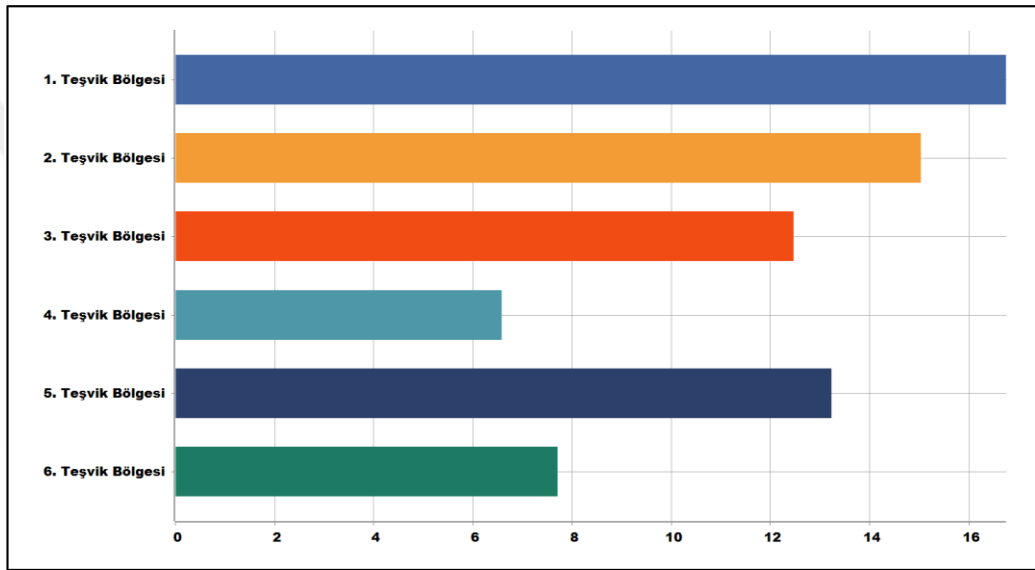
Şekil 21: Gıda – Yatırım Teşvik Bölgesi Dağılımı



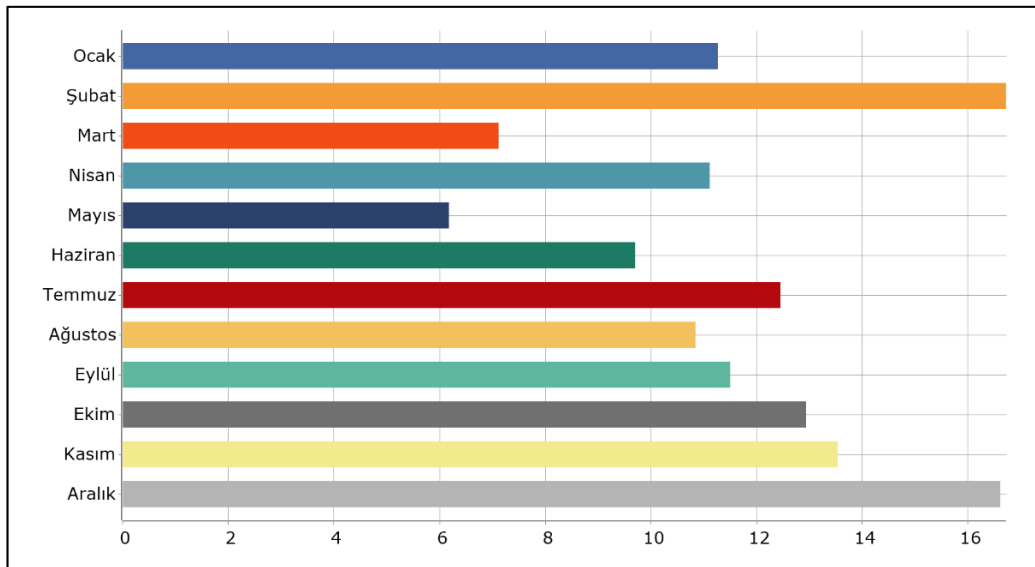
Şekil 22: Gıda – Aylık Dağılımı

Gıda terimi frekansı incelendiğinde Şekil 21'deki yatırım teşvik bölgelerindeki dağılımının 3. Bölge ve hemen ardından 2. Bölge şehirlerin olduğu görülmektedir. Buna göre, Gıda niteliğiyle ilgili istihdamın işgücü istihdamının ve sanayinin bulunduğu ancak gelişmişlik seviyesi olarak ilk sıralarda yer almayan iller olduğu anlaşılmaktadır. Bu anlamda, Gıda niteliği gelişmekte olan şehirlerde talep görmektedir denilebilir.

Şekil 22'de dikkat çekici Mayıs ayındaki trend ise, bu talebin bahar sonunda gerçekleştirildiğini göstermektedir. Bu işgücü planlaması, Bilgisayar niteliğinin Nisan ayındaki trendine benzemektedir.



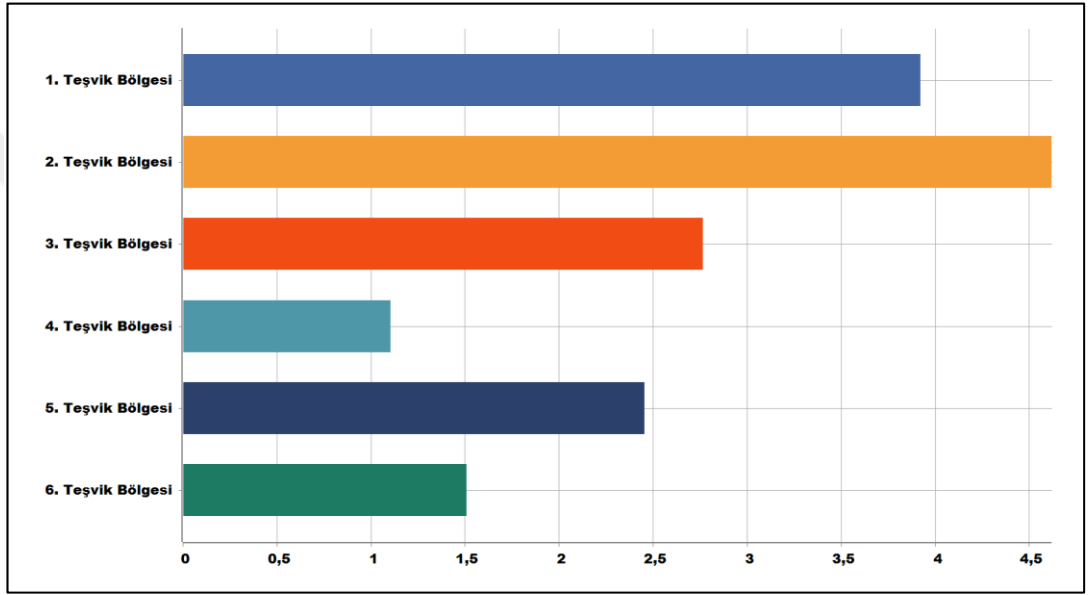
Şekil 23: Kalite – Yatırım Teşvik Bölgesi Dağılımı



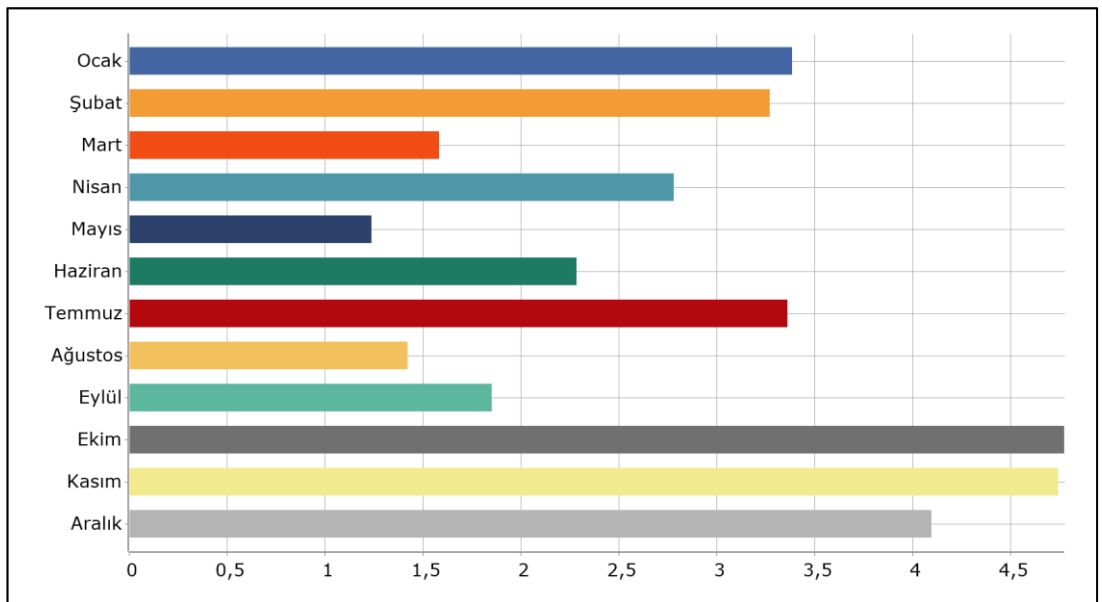
Şekil 24: Kalite – Aylık Dağılımı

Kalite teriminin Şekil 23'teki yatırım teşvik bölgeleri dağılımında 5. Bölge'deki trendi dikkat çekicidir. 4. Bölge düzeyindeki talep azlığı, 4. Bölgedeki kentlerin komşu illerinde 3. ve 5. Bölge kentlerinin daha büyük olması ile açıklanabilir. Dolayısıyla, Kalite niteliğinin nüfusla ilişkisi olduğu anlaşılmaktadır.

Öte yandan Şekil 24'teki Aralık ve Şubat başta olmak üzere kış aylarındaki talep yoğunluğu, Bilgisayar ve Gıda gibi niteliklerin işgücü planlama dönemlerinin aksine yıl başına göre ayarlandığını göstermektedir.



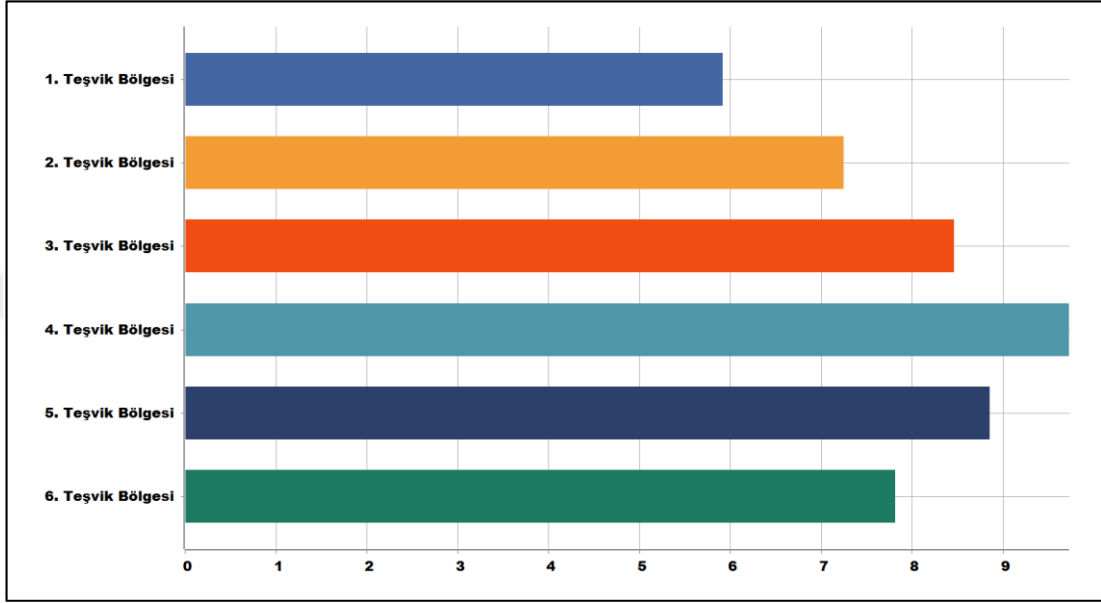
Şekil 25: Lojistik – Yatırım Teşvik Bölgesi Dağılımı



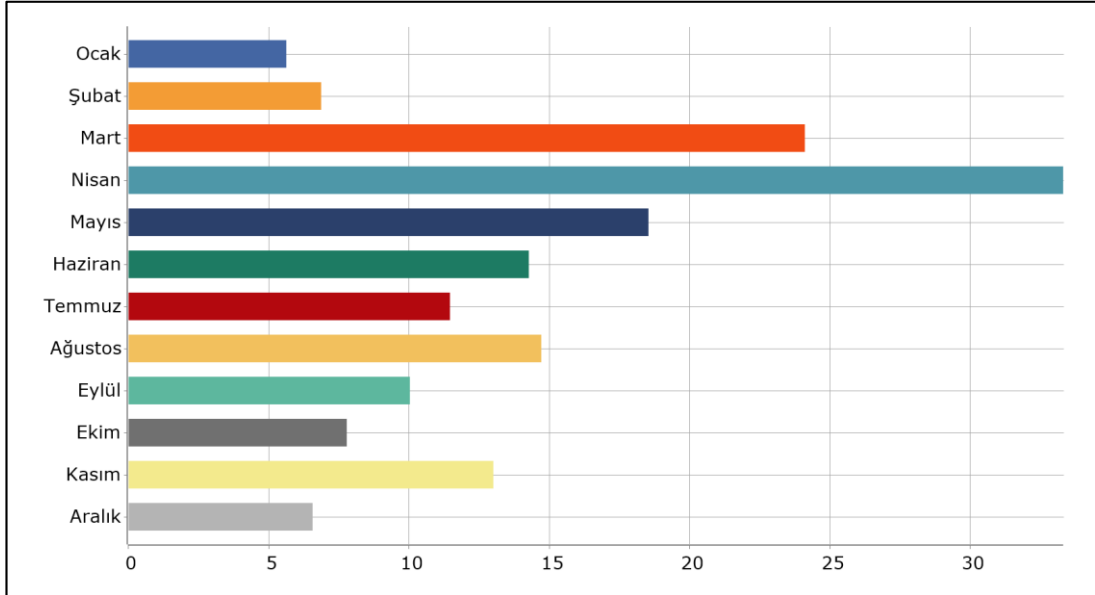
Şekil 26: Lojistik – Aylık Dağılımı

Lojistik teriminin, Şekil 25’te başta 2. Bölgede trendinin yüksek olduğu görülmüştür. Buna göre, metropollerin bağlayan büyükşehirlerde, endüstri ve tarım ekonomisi büyük, gelişen illerde Lojistik niteliği arzı görülmüştür.

Şekil 26’daki bulgular, Lojistik niteliği üzerine işgücü planlamalarının güz dönemi sonunda daha çok yapıldığını göstermektedir.



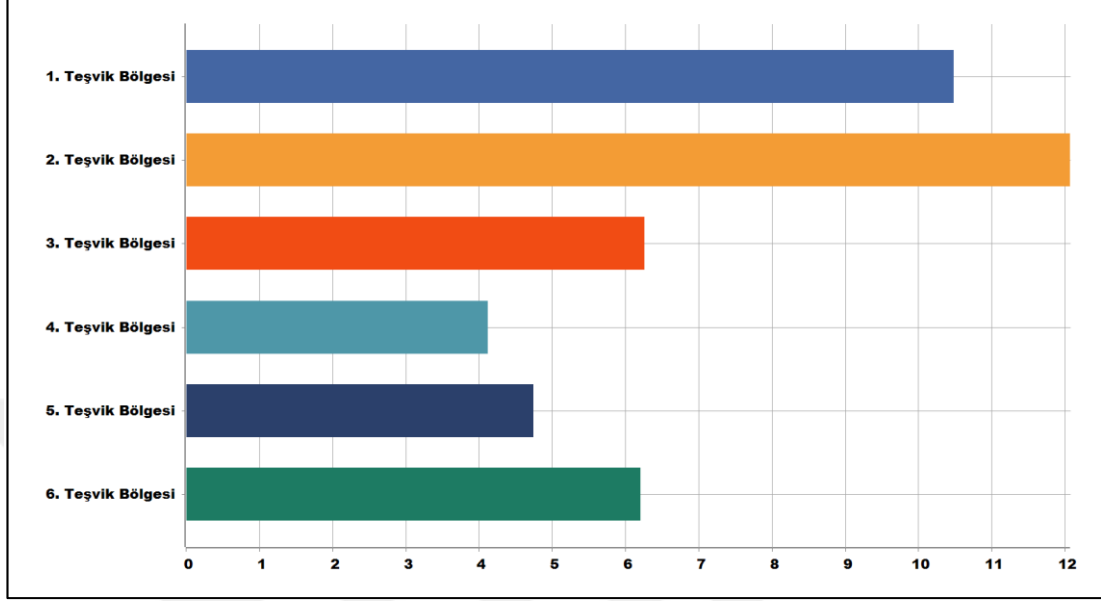
Şekil 27: Mağazacılık – Yatırım Teşvik Bölgesi Dağılımı



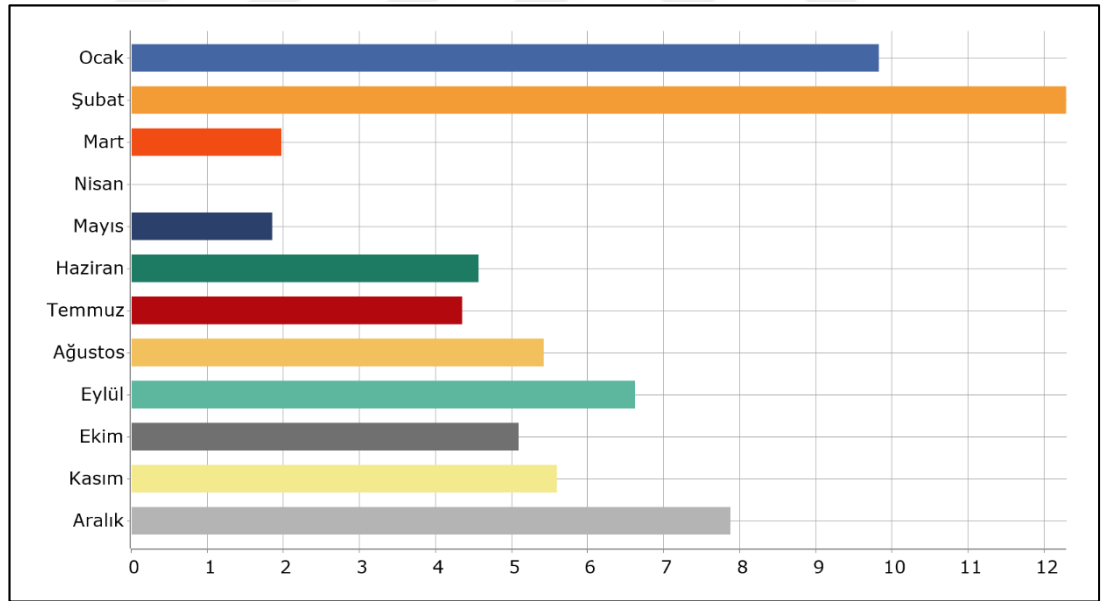
Şekil 28: Mağazacılık – Aylık Dağılımı

Şekil 27’ye göre, Mağazacılık terimi 4. Yatırım teşvik bölgesinde daha çok aranmaktadır.

Şekil 28'deki aylık dağılımda ise bahar başı ve ortalarında işgücü planlamaları yapıldığı görülmektedir. Dolayısıyla Mağazacılık niteliği, büyükşehirlere yayılmış, sanayinin ve nüfusun oluşturabildiği bir istihdam niteliği olarak bulunmamıştır.



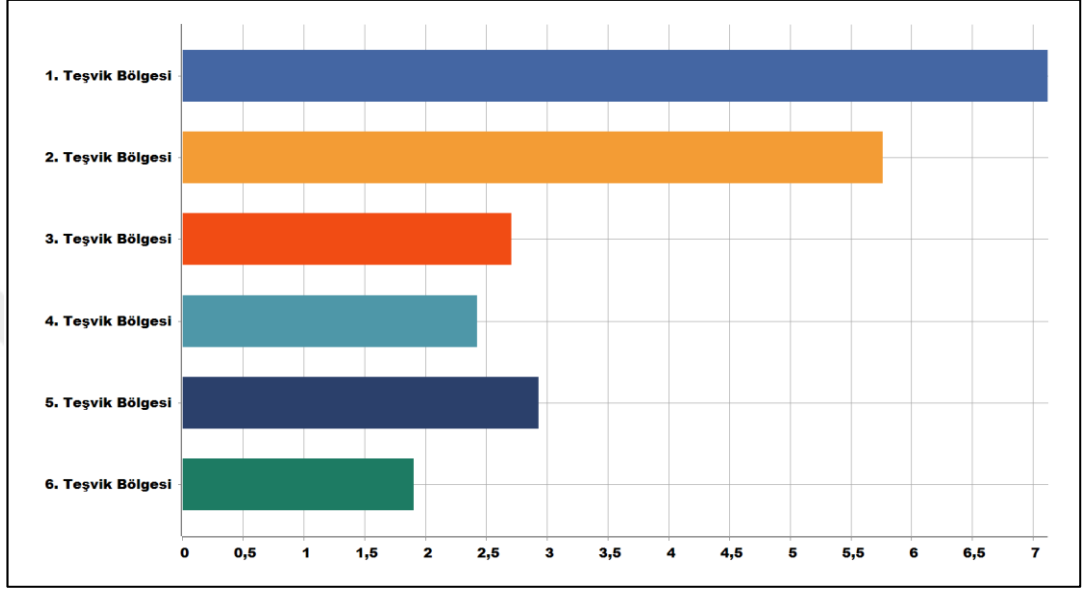
Şekil 29: Muhasebe – Yatırım Teşvik Bölgesi Dağılımı



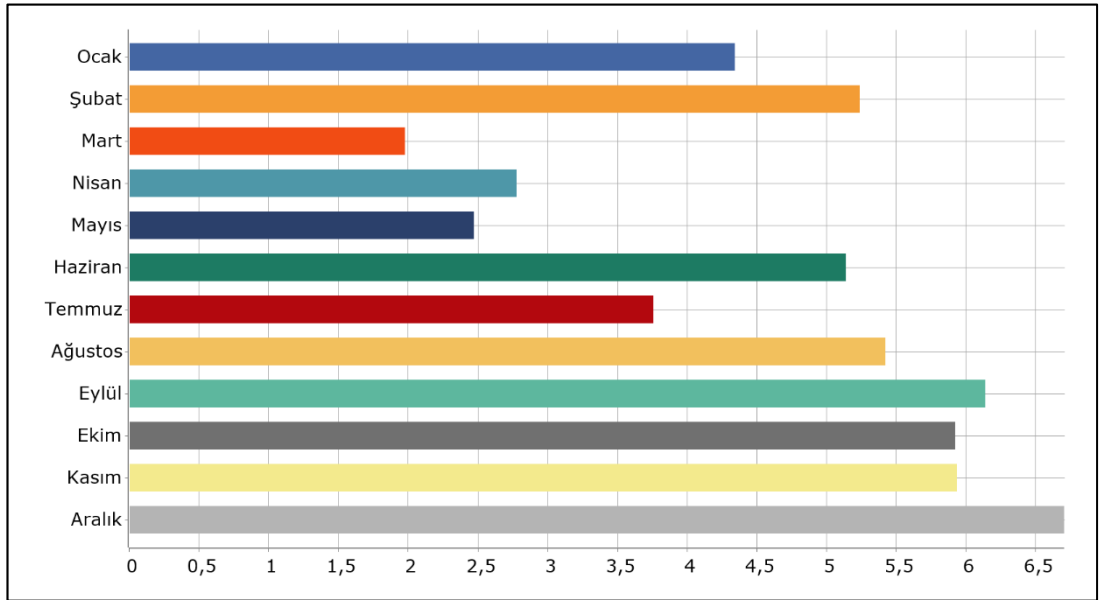
Şekil 30: Muhasebe – Aylık Dağılımı

Şekil 29'daki Muhasebe teriminin yatırım teşvik bölgelerindeki dağılımı, Lojistik ile benzerlik göstermektedir. Fakat Lojistik bulgularına nazaran, nüfus yoğunluğunun çekim merkezi oluşturmada daha kritik bir rol oynadığı anlaşılmaktadır.

Şekil 30'daki Muhasebe niteliğinin aylık dağılımı, Gayrimenkul niteliğindeki benzer başka bir ilan devir hızına işaret etmektedir. Öyle ki, Ocak ve Şubat aylarındaki ayrışma ve ortalama olarak kabul edilmiş 60 ila 90 günden daha az görülen Nisan ayı iş ilanı yokluğu, bu nitelikteki ilanların yakın zamanda yayınlanıp, sıkça işveren tarafından takip edildiğini göstermektedir.



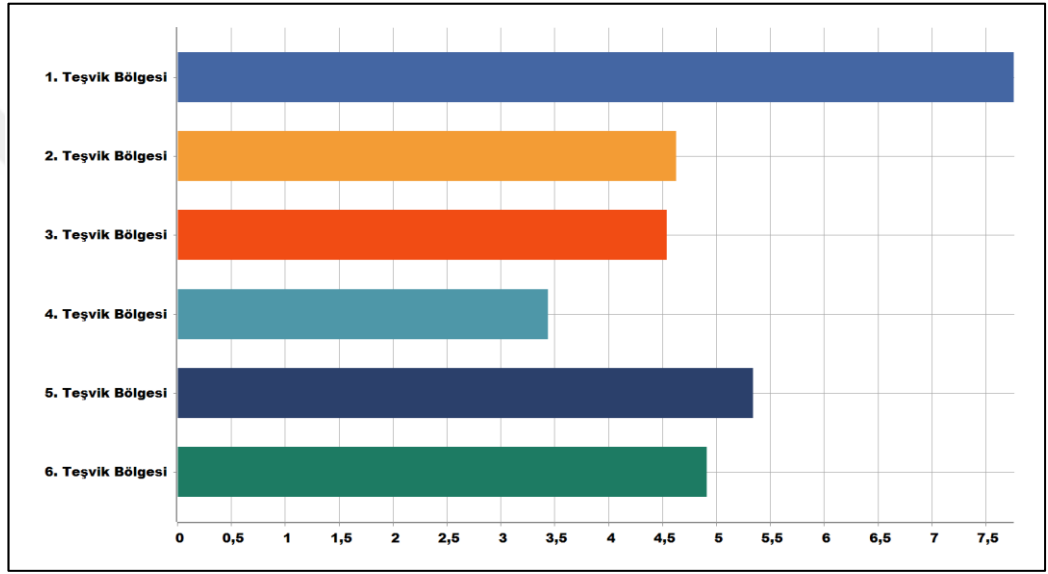
Şekil 31: Mühendislik – Yatırım Teşvik Bölgesi Dağılımı



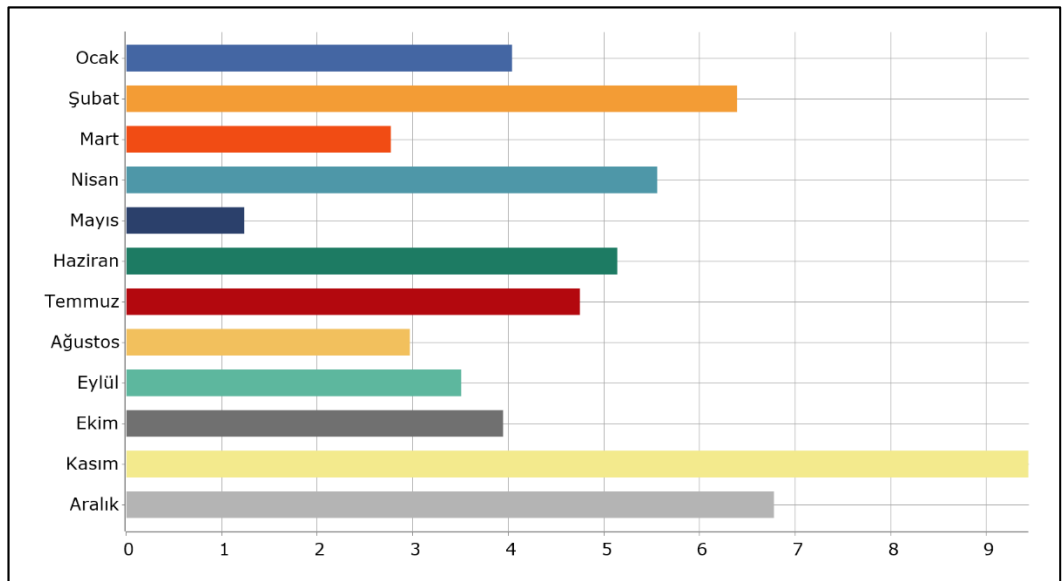
Şekil 32: Mühendislik – Aylık Dağılımı

Şekil 31'deki Mühendislik teriminin yatırım teşvik bölgeleri dağılımı incelendiğinde, sanayinin ve ekonominin güçlü olduğu ve istihdam olan şehirlerden az gelişmiş şehirlere göre talebin düştüğü görülmektedir.

Şekil 32'deki aylık dağılımlar incelendiğinde, her mühendisliğin farklı dönemlerde işgücü planlamasına dahil olacağı tahmin edildiği için tüm yıla yayılmış olması beklenen bir bulgudur. Fakat bahar aylarındaki talebin düşüklüğü mühendislik niteliğine yönelik işgücü planlamalarının işverenler tarafından bu dönemlerde daha az yapıldığı ve genel anlamda ilan devir hızının yüksek olduğu tahmin edilmiştir.



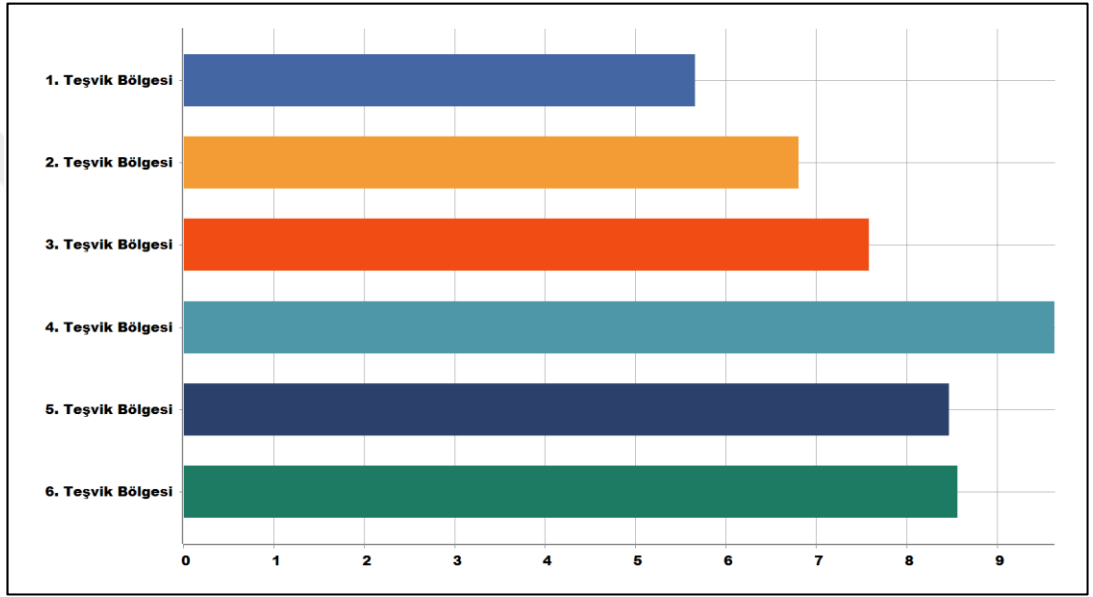
Şekil 33: Otomotiv – Yatırım Teşvik Bölgesi Dağılımı



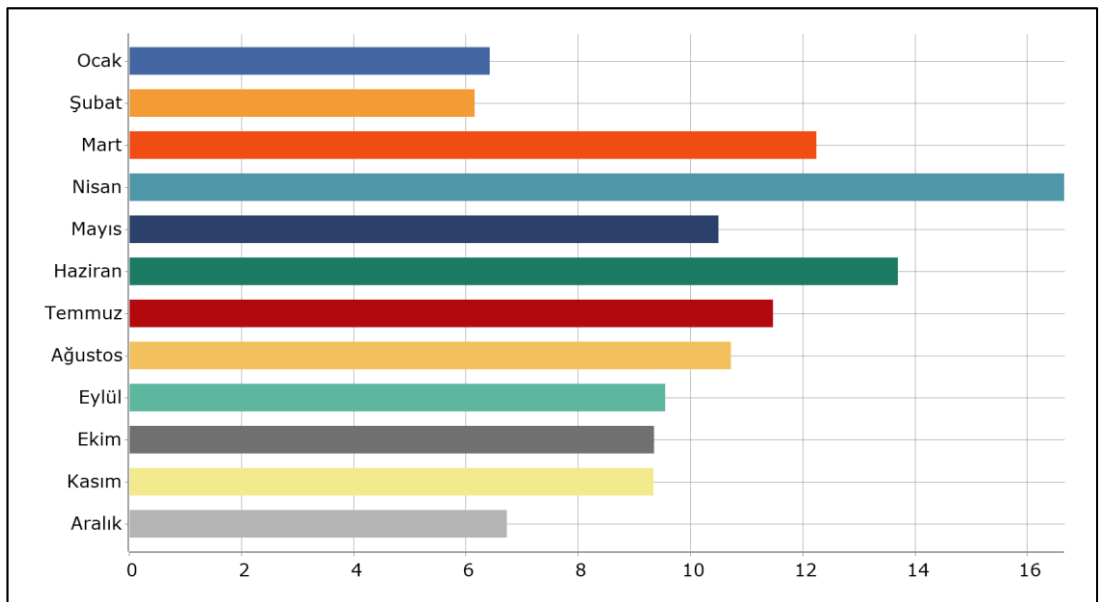
Şekil 34: Otomotiv – Aylık Dağılımı

Otomotiv terimi Şekil 33'te yatırım teşvik bölgelerine göre dağıtıldığında, 1. Bölgenin arzının yüksekliği göze çarpmaktadır. Ancak diğer bölgelerde neredeyse eşit dağılım göstermiştir. Bu bağlamda, 1. Bölgenin otomotiv endüstrisi açısından bu niteliği talep ettiği fakat diğer bölgelerde ise yan sanayi veya satış-pazarlama yönüyle tüketicilere yönelik bayii sisteminin bu arzı oluşturduğu anlaşılmıştır.

Şekil 34'teki aylık dağılımda ise firmaların Otomotiv niteliğini yılbaşından önce işgücüne katılımını sağlamak istedikleri tahmin edilmektedir.

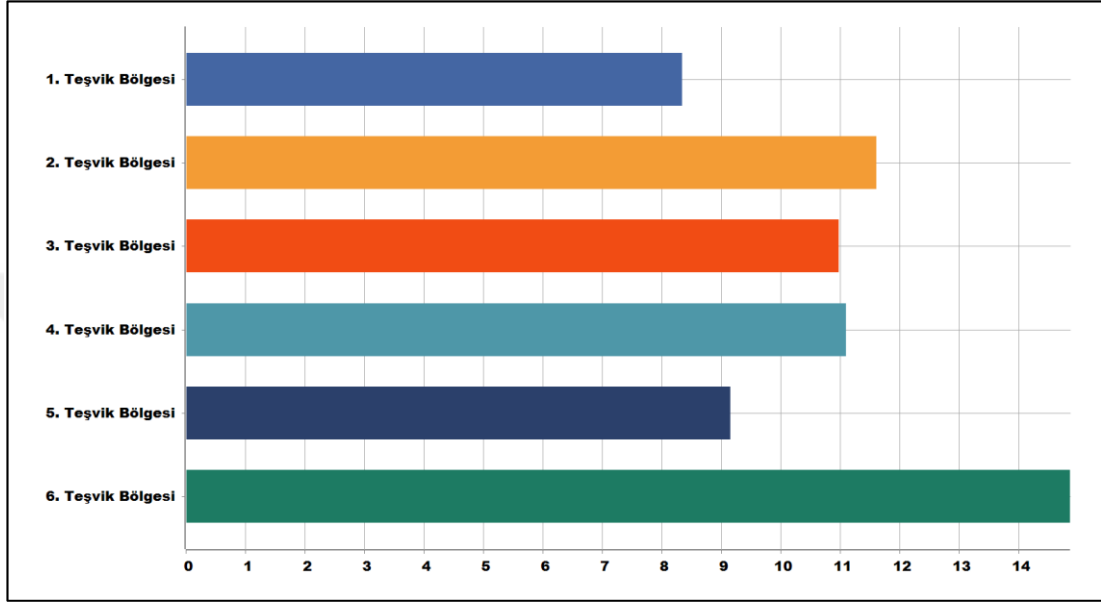


Şekil 35: Perakende – Yatırım Teşvik Bölgesi Dağılımı

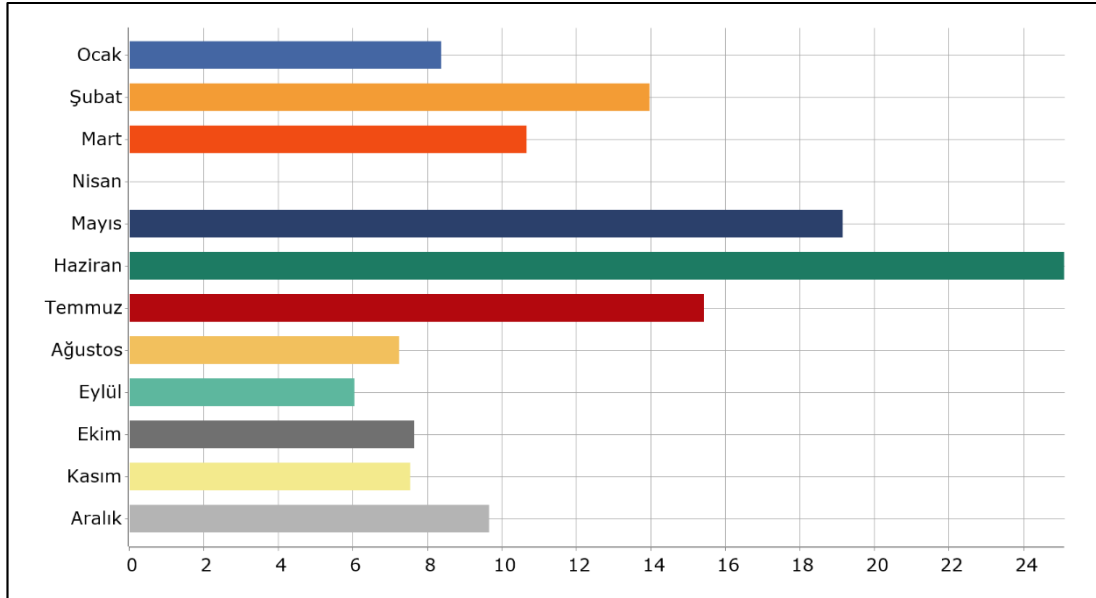


Şekil 36: Perakende – Aylık Dağılımı

Perakende teriminin Şekil 35 ve Şekil 36'daki yatırım teşvik bölgeleri ile aylık dağılımları incelendiğinde Mağazacılık bulgularıyla paralellik gösterdiği görülmektedir. Neredeyse aynı dağılımları gösteren bu iki terim karşılaştırıldığında fark ortaya koymak için detaylı incelenmesi gerekmektedir. Aynı sektörün terimleri olduğu bilinen bu iki terimin benzer sonuçlar ortaya çıkarması şaşırtıcı olmayan bir durumdur.



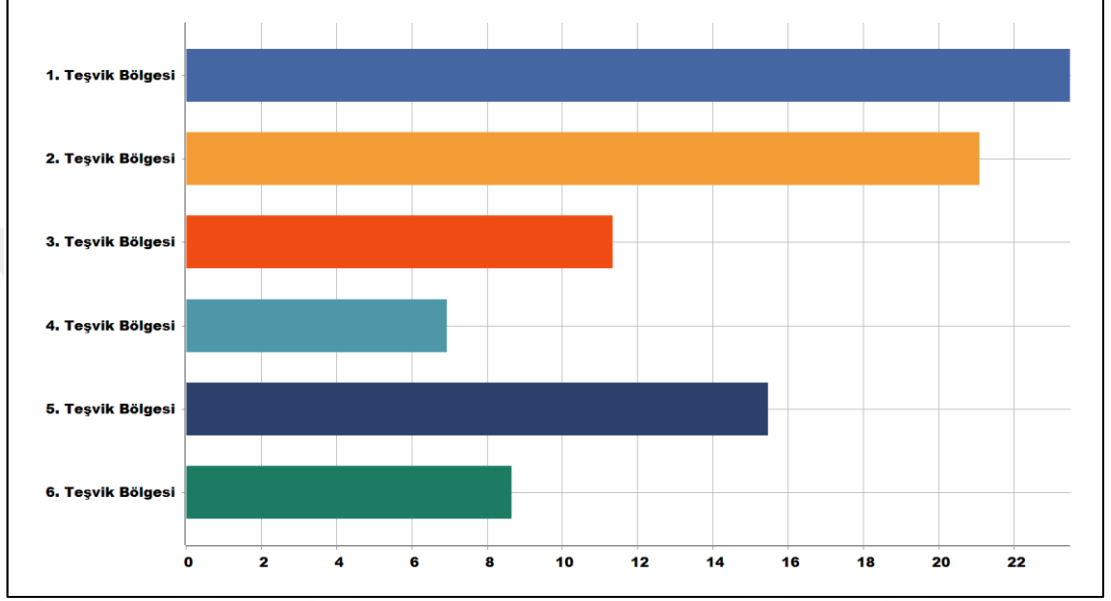
Şekil 37: Sağlık – Yatırım Teşvik Bölgesi Dağılımı



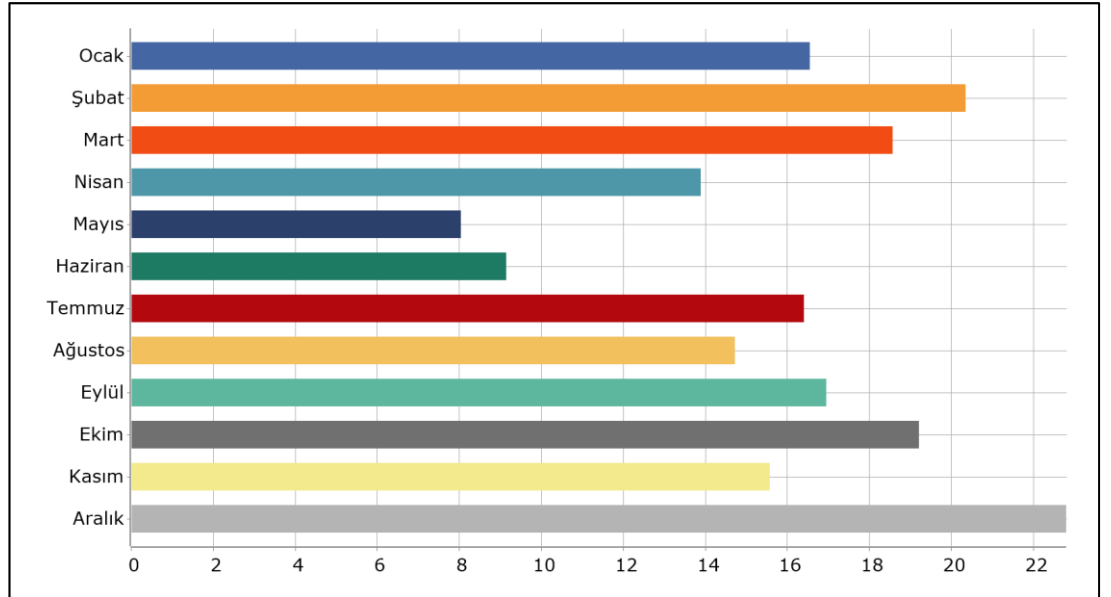
Şekil 38: Sağlık – Aylık Dağılımı

Sağlık terimi Şekil 37'deki yatırım teşvik bölgelerindeki dağılımında 6. Bölgede aranan nitelik olarak keşfedilmiştir. Güneydoğu Anadolu bölgesi illerini ve Doğu Anadolu

bölgesinin kırsal kesimini kapsayan bir bölge olarak, bu 6. bölge Türkiye Cumhuriyeti devletinin ekonomik kalkınma alanında sorunlu bir coğrafi bölgesi olarak tanımlanabilir. Ayrıca Şekil 38'e göre, ortalama 60 ila 90 günden daha az olduğu tahmin edilen devir hızı ve çoğu bulguların aksine yaz başındaki işgücü planlaması takvimi, kamu politikalarında detaylıca ele alınması gereken bir başlık olarak Sağlık niteliğini ortaya çıkarmaktadır.

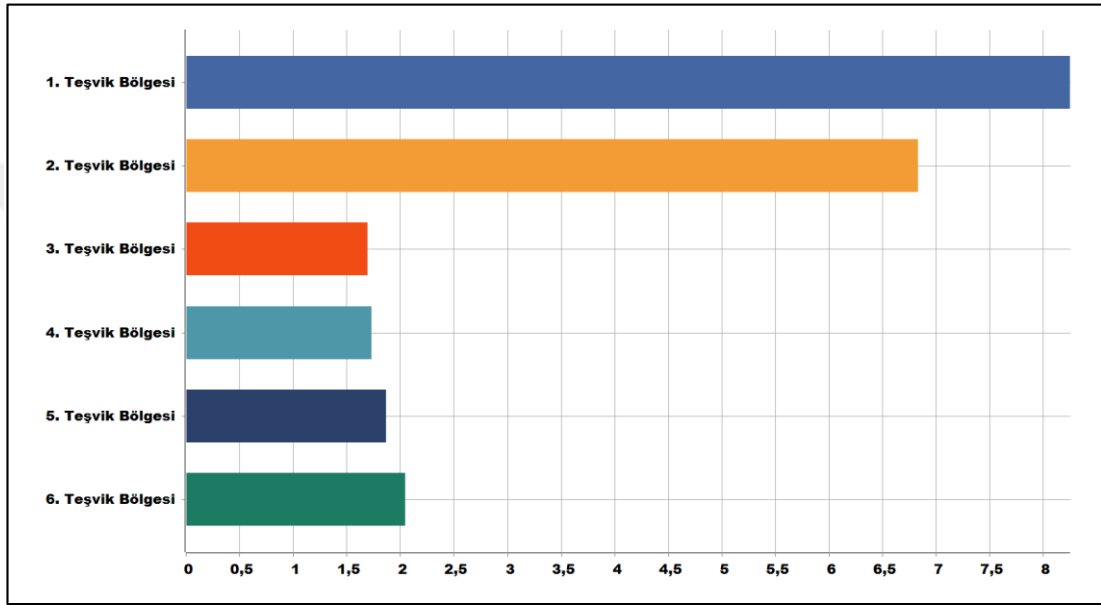


Şekil 39: Üretim – Yatırım Teşvik Bölgesi Dağılımı

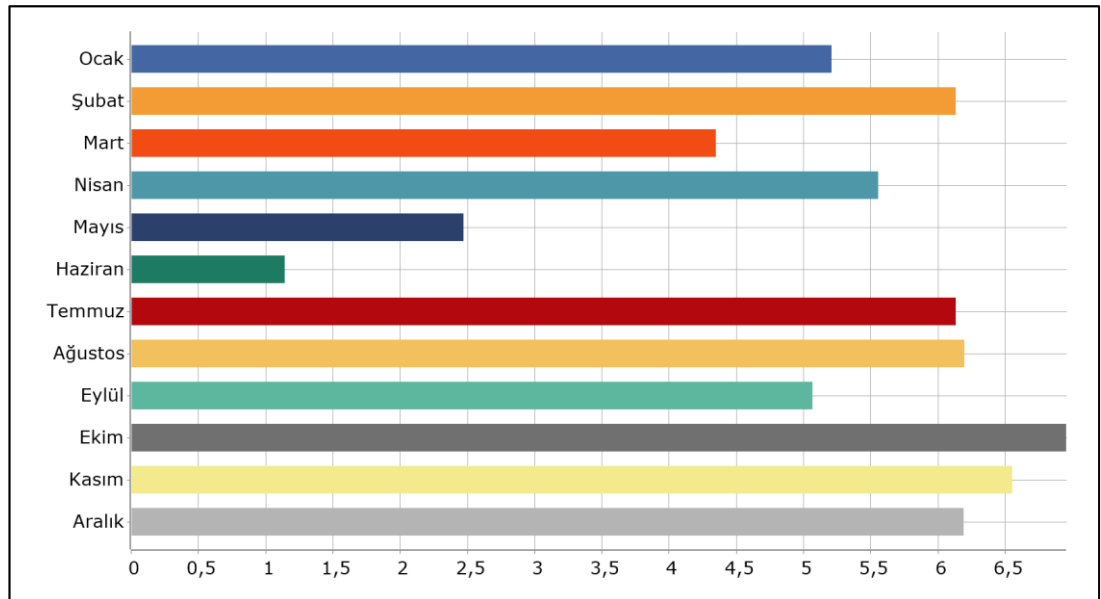


Şekil 40: Üretim – Aylık Dağılımı

Şekil 39'daki yatırım teşvik bölgeleri dağılımında, Üretim terimi beklendiği gibi başta 1. Bölge olmak üzere talep edilen bir nitelik olarak kırsala gidildikçe azalmaktadır. Üretim teriminin, Şekil 40'taki aylık dağılımı incelendiğinde yaz başına doğru talebinin düşmesi ve yılbaşı öncesi talep artışı genel insan kaynakları planlamalarında beklenen bir durumdur. Dolayısıyla Üretim niteliği, üretimin yani sanayinin ve tarımın olduğu kentlerle doğrudan ilişkili bulunmuştur. Bu nedenle keşfedilen bu Üretim teriminin, üretim sektörünü ifade eden birebir ilişkili bir nitelik olduğu anlamı da çıkarılabilir.

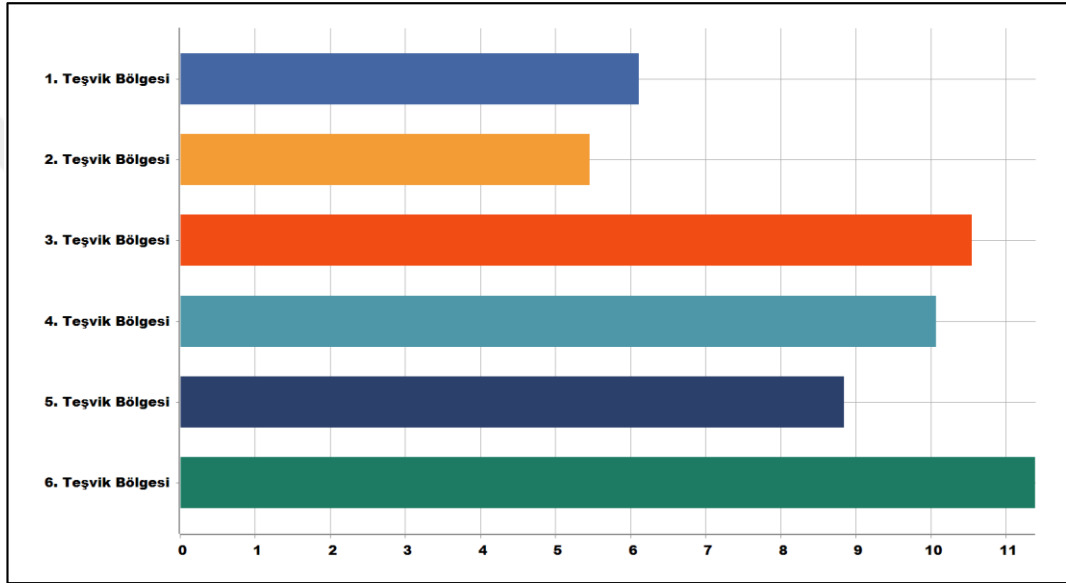


Şekil 41: Yazılım – Yatırım Teşvik Bölgesi Dağılımı

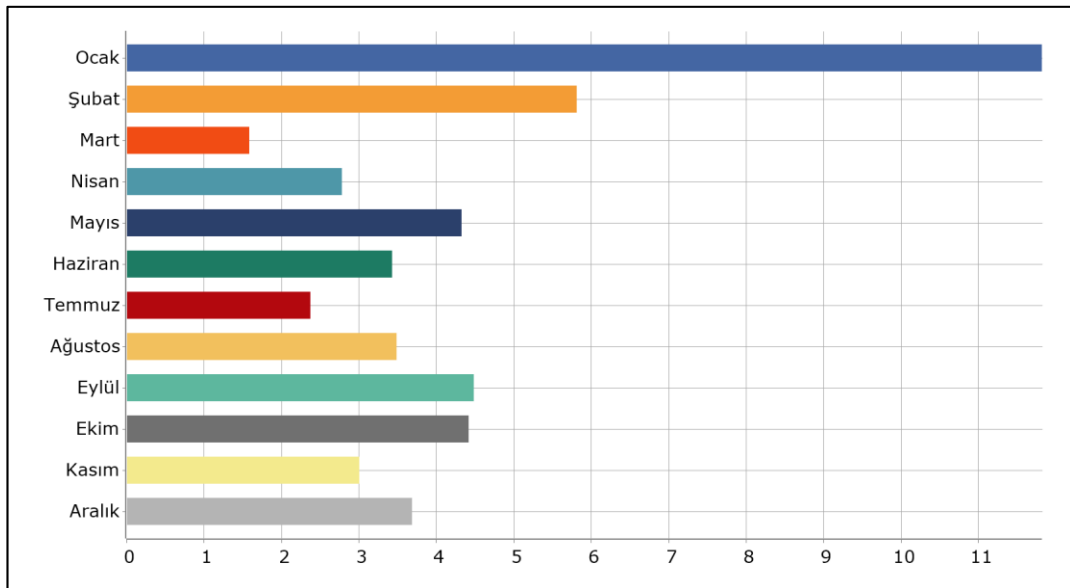


Şekil 42: Yazılım – Aylık Dağılımı

Şekil 41'deki Yazılım terimi dağılımına bakıldığında, 1. ve 2. bölgelerde direkt diğer yatırım teşvik bölgelerine göre ayrılmıştır. Buna göre, yazılım niteliğinin istihdam gücü yüksek ve sanayisi gelişmiş kentlerde büyük bir işverenler tarafından arz ve talep edildiğini göstermektedir. 1. yatırım teşvik bölgesinin liderlik etmesi, henüz ilgili kentlerde bu talebin doygunluğa erişmediği anlamına gelebilir. Şekil 42'de Yazılım niteliği talebinin yüksekliği her aya yakın dağılmış frekanslardan anlaşılabilir. Ancak Nisan sonu düşüşe geçen trendin, firmaların yazılım işgücü planlamalarına yaz aylarında başladıklarını göstermektedir.



Şekil 43: Yenilikçi – Yatırım Teşvik Bölgesi Dağılımı



Şekil 44: Yenilikçi – Aylık Dağılımı

Şekil 43'te Yenilikçi teriminin yatırım teşvik bölgelerindeki dağılımı gösterilmiştir. Buna göre Yenilikçi niteliği kırsaldan büyükşehirlere doğru bir trend göstermiştir. Şekil 43'teki bu dağılım, metropoller ve istihdamın yüksek olduğu kentlerden az gelişmiş şehirlere doğru ters bir ilişki ortaya koymuştur. Buna ek olarak, Şekil 44'teki aylık dağılımda ise tüm yıla yayılmış bu nitelik yılbaşından itibaren yükseliş gösterse dahi, veri setinin hazırlandığı tarihlerde trendi oldukça yüksek artış göstermiştir.

Şekil 12'deki kelime bulutunda Bahçeşehir sık bulunan terimlerden biri olduğu görülmüştür. Bahçeşehir Okulları'nın iş ilanları öğretmenlik iş ilanlarını veri setinin hazırlandığı dönemlerde sıkça yayınlamıştır ve Yenilikçi terimi özellikle bu iş ilanlarının içerisinde sıkça geçmektedir. Dolayısıyla Yenilikçi terimindeki Şekil 43'teki ve Şekil 44'teki anomali olarak tespit edilmiş dağılımlar böyle açıklanabilir. Bu bağlamda, öğretmenlik alanında Yenilikçi terimi aranan bir niteliktir. Nihayetinde burada tespit edilen ayırışan bulgu, araştırmacı tarafından beklenen bir durumun eseridir.

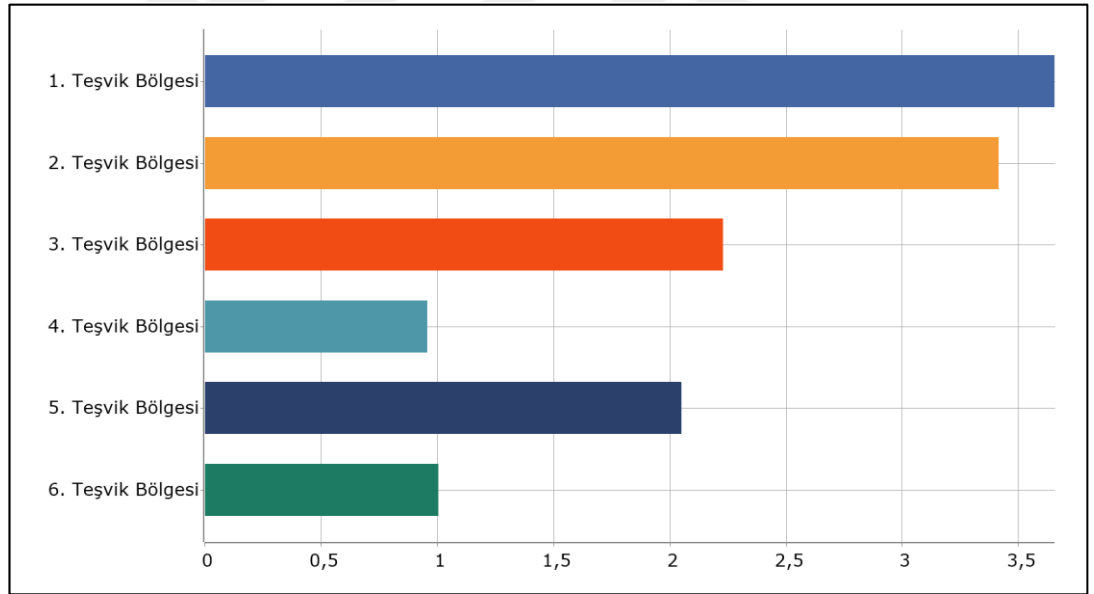
3.2. Sık Kullanılan Kelime Grupları

Oluşturulmuş metin analizi modeline göre, sık geçen kelime gruplarının oluşturduğu frekans tablosundan elde edilen kelime bulutu olarak Şekil 45'de gösterilmiştir.

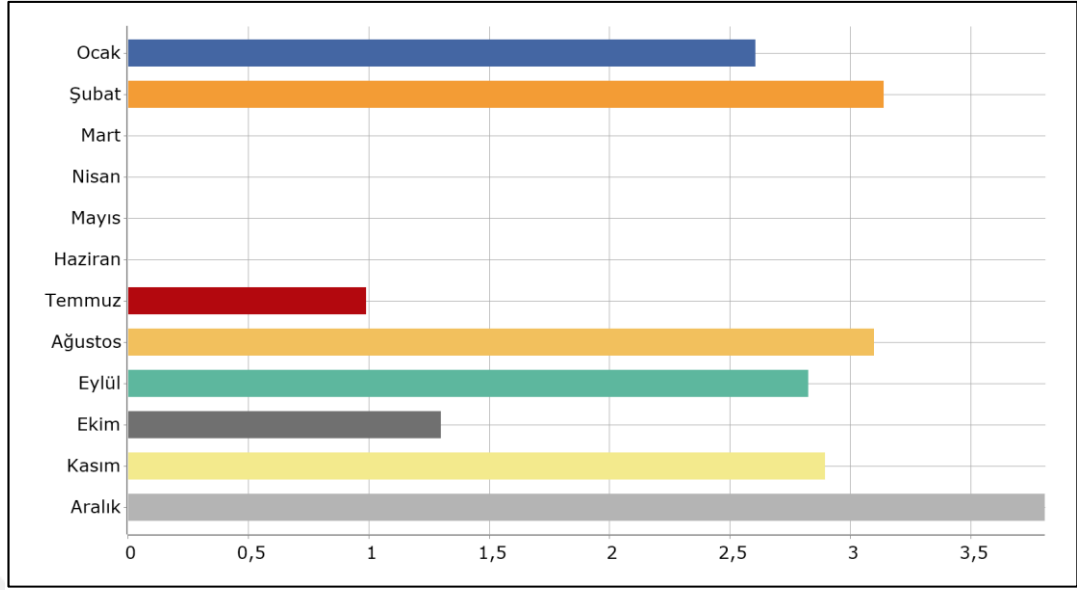


Şekil 45'deki bulgulara göre, iş ilanlarında en sık geçen kelime grupları Şekil 12'deki kelimelerin oluşturduğu frekans tablosuna göre hazırlanmış kelime bulutundaki bulgulara benzer, başvuruda bulunacak başta takım çalışmalarına uyumu ve erkek adayların askerlik durumu, okunan veya mezun olunan bölümler, seyahat engeli veya sürücü belgesi sahip olma durumları ve bunların ardından mesleki bilgi, yetenek veya tecrübelerini genel anlamda ifade eden sözcüklerden oluştuğu görülmüştür.

Keşfedilen kelime gruplarının 10000 sözcükte bulunma sıklıkları yatırım teşvik bölgeleri düzeyindeki ve 12 aylık dönemdeki dağılımları incelenmek istenmiştir. Bu amaçla frekansı 500'den daha yüksek olan ve yatırım teşvik bölgeleri düzeyindeki ve/veya 12 aylık dönemdeki dağılımları anomali gösterdiği tespit edilen kelime grupları seçilmiştir. Belirlenen bu kelime grupları şu başlıklar ile ifade edilebilir: *AR-GE, Dış Ticaret, Elektrik-Elektronik, E-Ticaret, Gayrimenkul Değerleme, İngilizce Bilen, İnsan Kaynakları, İş Sağlığı ve Güvenliği, Part-Time, Satın Alma, Sosyal Medya.*

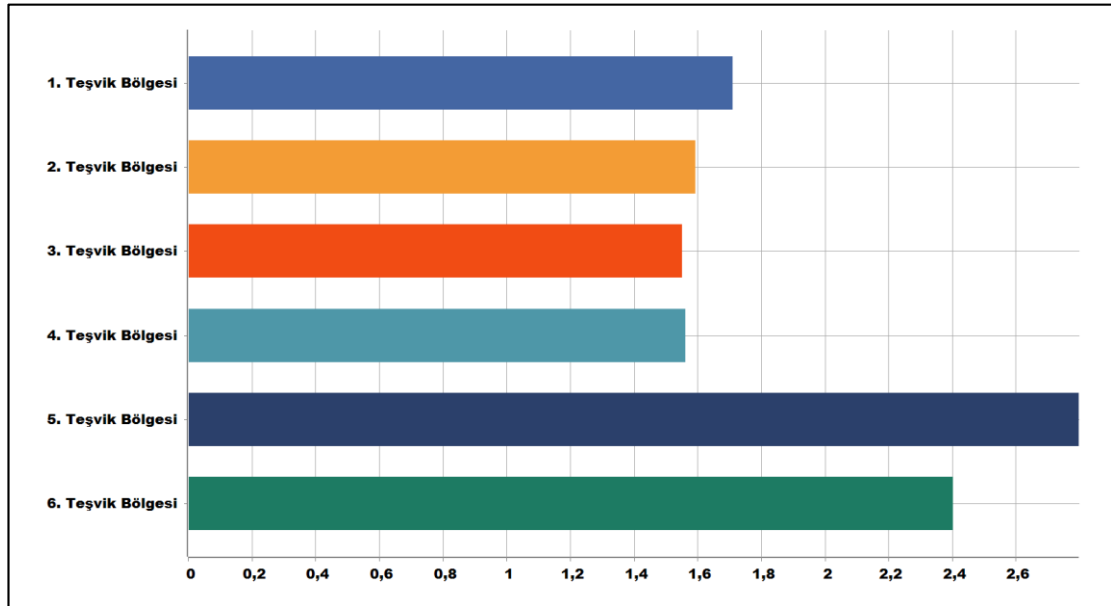


Şekil 46: AR-GE – Yatırım Teşvik Bölgesi Dağılımı

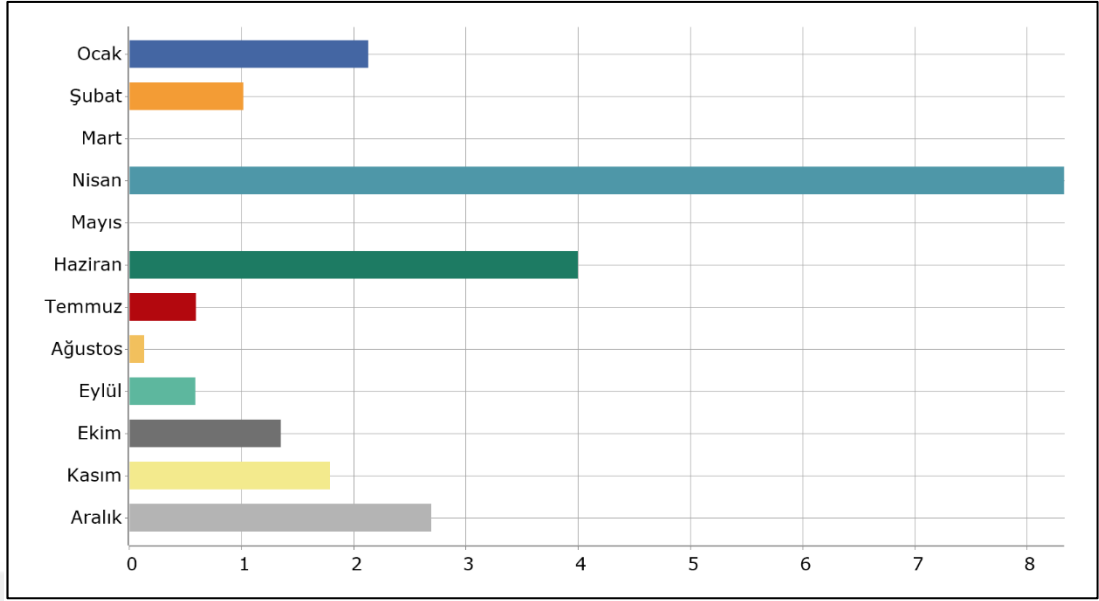


Şekil 47: AR-GE – Aylık Dağılımı

Şekil 46'ya göre AR-GE niteliği metropollere, sanayinin ve istihdamın güçlü olduğu şehirlerde arandığı anlaşılmaktadır. Bununla birlikte Şekil 47'deki aylık dağılım incelendiğinde, AR-GE niteliği Mart-Haziran arası ilanlarda rastlanmamaktadır. 60 ila 90 günlük bir ortalamanın üzerinde olan bu bulgu, diğer ayların dağılımına ve sıklık değerinin düşük olması nedeniyle, istihdamın az olduğu ve bu nitelik için olası bir ihtiyaçta kısa süreli ilana çıktıkları anlaşılmaktadır. Böylece talebin de çabuk karşılandığı yorumu yapılabilir.

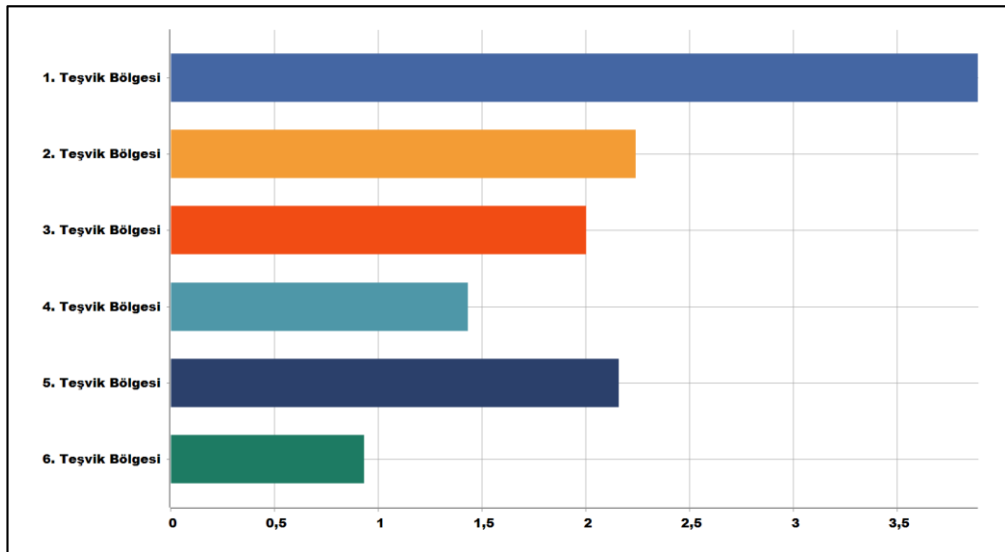


Şekil 48: Dış Ticaret – Yatırım Teşvik Bölgesi Dağılımı

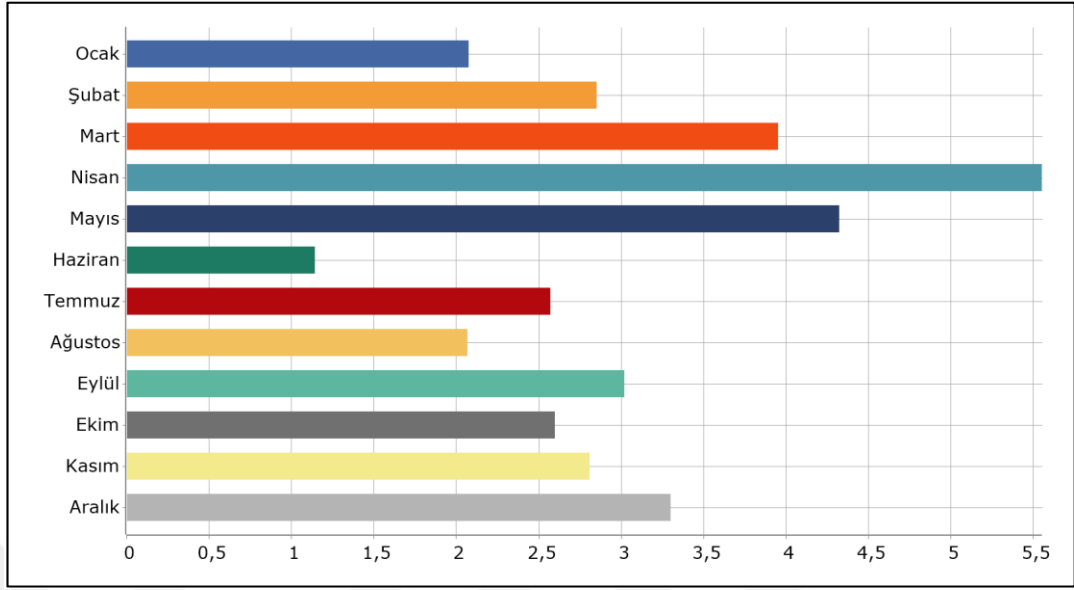


Şekil 49: Dış Ticaret – Aylık Dağılımı

Dış Ticaret kelime grubu niteliği için yatırım teşvik bölgeleri dağılımı Şekil 48’de incelendiğinde, 5. ve 6. bölgelerin iş ilanlarında diğer bölgelere göre daha sık aradıkları bir özellik olduğu görülmüştür. Şekil 49’daki aylık bazda bu nitelik incelendiğinde dağınık bir grafik oluşturmuştur ve bu nedenle mevsimsel veya dönemsel etkilerden söz edilememiştir. Kırsal ve kalkınmaya açık bölgelerdeki sıklığının yüksekliği ve aylık olarak düzensiz bir yaklaşım göstermesi, Dış Ticaret niteliğini kamu politikalarında özellikle ihracat ölçeğindeki kritik rolü nedeniyle önemi bir bulgu olarak değerlendirilmesine yol açmıştır.

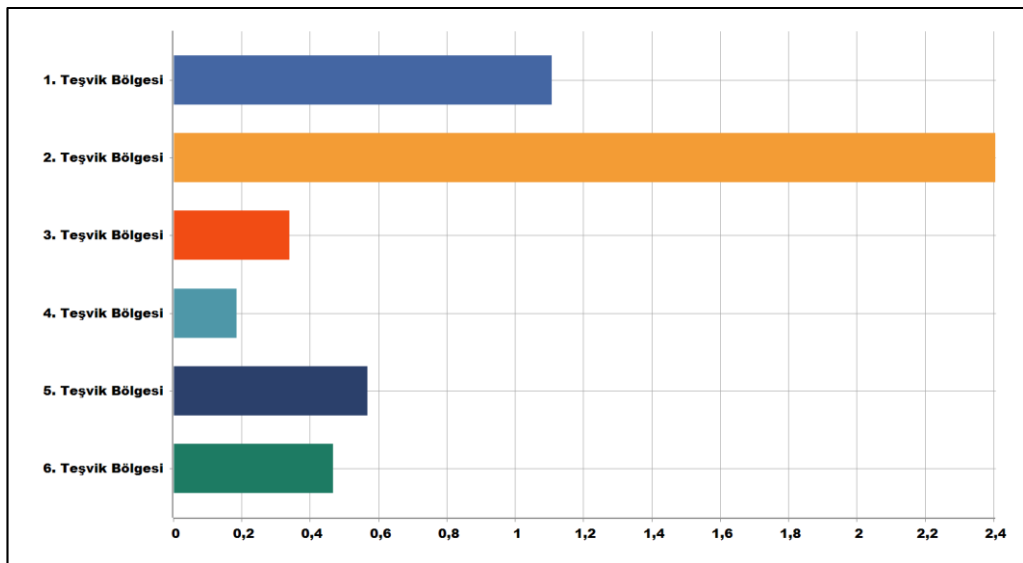


Şekil 50: Elektrik-Elektronik – Yatırım Teşvik Bölgesi Dağılımı

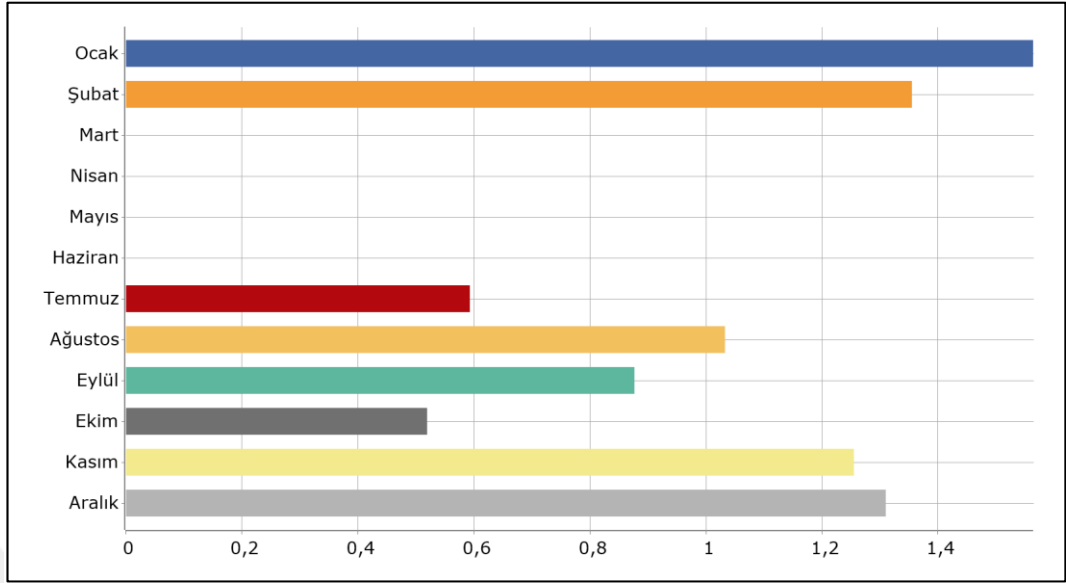


Şekil 51: Elektrik-Elektronik – Aylık Dağılımı

Elektrik-Elektronik kelime grubunun Şekil 50'deki yatırım teşvik bölgelerine göre dağılımı incelendiğinde, frekans analizlerindeki Şekil 17'deki Elektronik terimi dağılımıyla benzerlik göstermektedir. Dolayısıyla bölgesel dağılım ile ilgili benzeri coğrafi yorumlar yapılabilmektedir. Ancak Şekil 51'deki aylık dağılım ile Elektrik teriminin Şekil 18'deki aylık dağılımı kış aylarında ayrışmaktadır. Kış döneminde Elektrik-Elektronik öbeğinin Elektronik dağılımına göre daha düşük bir seyir izlemesi, Elektronik kavramı özelinde bu trendin aşağı yönlü bir eğilim göstermeye başladığı izlenimi uyandırmıştır.

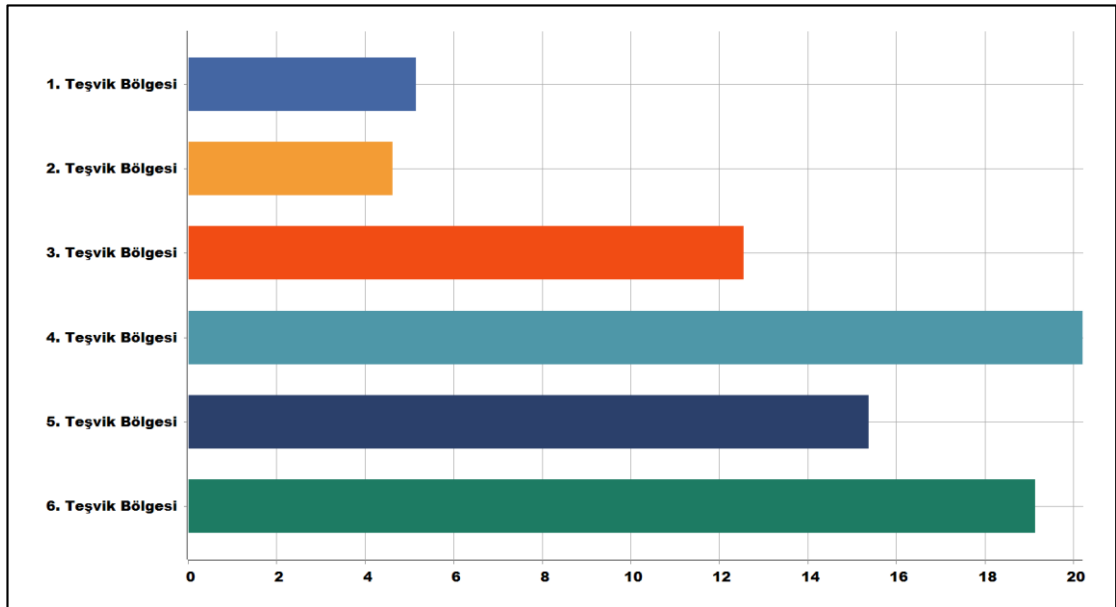


Şekil 52: E-Ticaret – Yatırım Teşvik Bölgesi Dağılımı

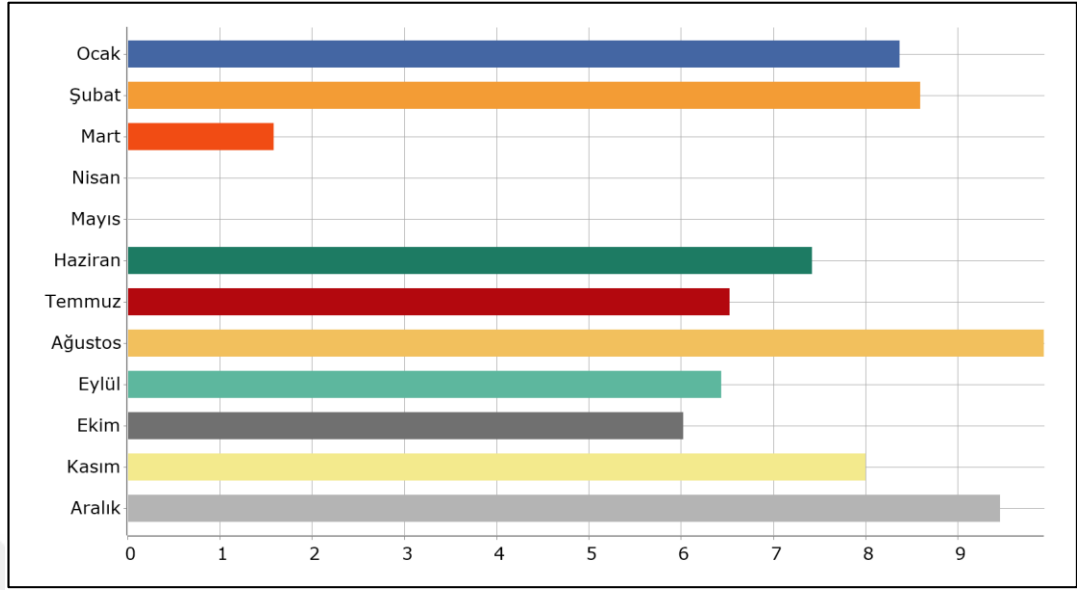


Şekil 53: E-Ticaret – Aylık Dağılımı

E-Ticaret kelime grubunun yatırım teşvik bölgeleri düzeyindeki Şekil 52’de gösterilen dağılımında, 2. bölgedeki trendi dikkat çekici olmuştur. Bu trendi 1. bölgesinde şehirler izlemiştir. Dolayısıyla E-Ticaret odaklı istihdamın özellikle bölgelerin özellikleri ile değerlendirildiğinde, yüksek teknoloji içermeyen ve/veya yan sanayi merkezli olduğu anlaşılmaktadır. Aylık dağılım ölçeğindeki Şekil 53’te, 90 günü aşan Mart-Haziran arası oluşturduğu boşluk ve yaz dönemi düşük trend eğilimi E-Ticaret niteliğini, istihdam açısından ilan dönüşümünün hızlı ve seyrek çevrildiği bir alan olarak tahmin edilmiştir.

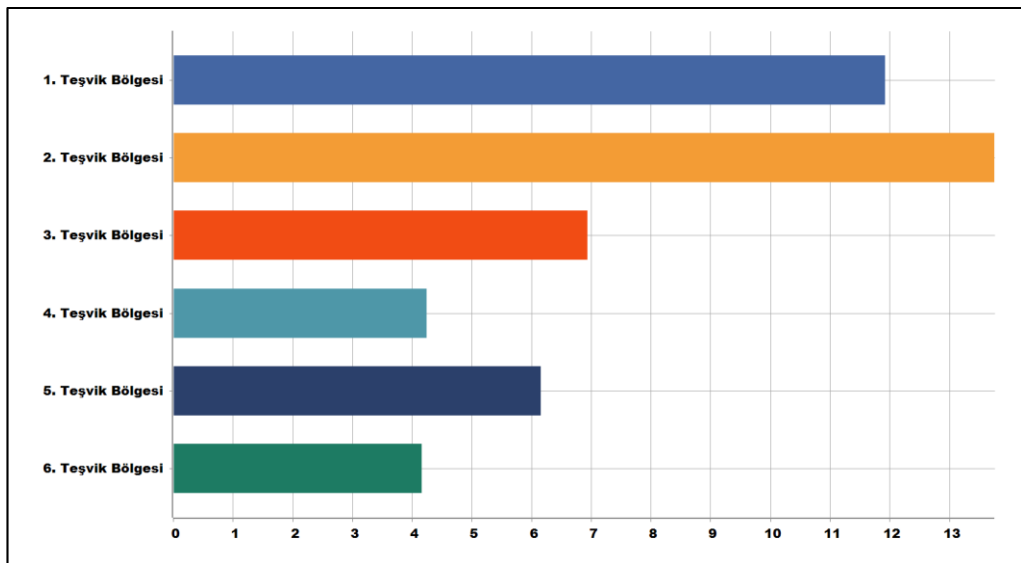


Şekil 54: Gayrimenkul Değerleme – Yatırım Teşvik Bölgesi Dağılımı

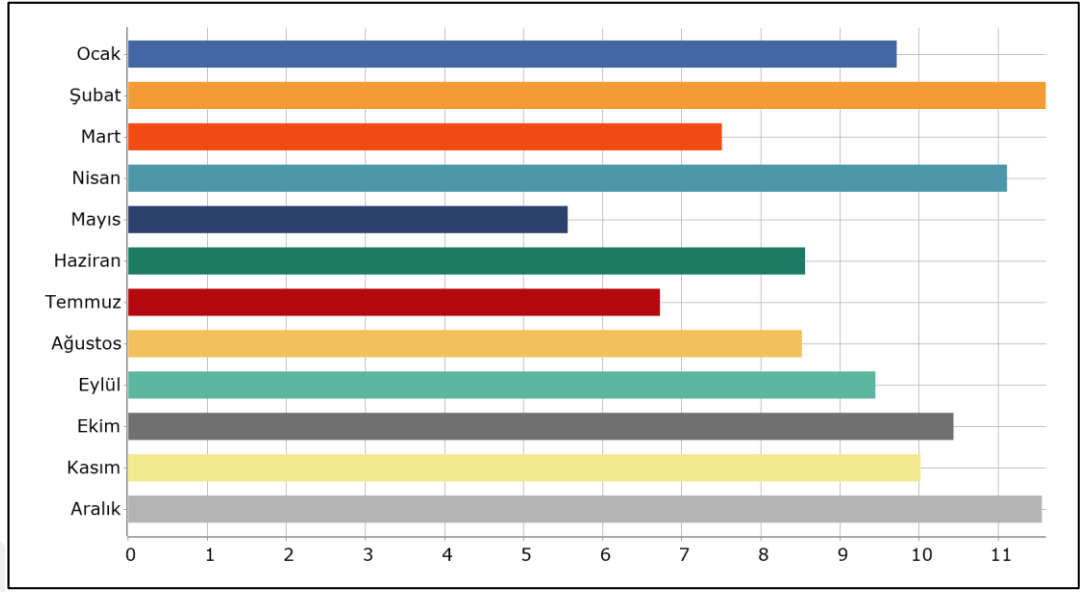


Şekil 55: Gayrimenkul Değerleme – Aylık Dağılımı

Gayrimenkul Değerleme kelime grubu için yatırım bölge teşvikleri dağılımı Şekil 54’te, daha önceki Şekil 19’da gösterilen frekans analizindeki Gayrimenkul dağılımına paralel ve yine Şekil 55’teki aylık dağılımlar incelendiğinde ise Şekil 20’de gösterilen dağılım ile benzerlik gösterdiği görülmüştür. Mart ve Temmuz ayları özelinde farklılık gösteren dağılımların Gayrimenkul teriminin doğrudan inşaat sektörünü hedefleyen ilanlardan etkilendiği, aynı şekilde Gayrimenkul Değerleme öbeğinin ise üniversitelerin genel tüm lisans programlarını kapsayan uzmanlığı ifade etmesi nedeniyle olduğu kanısı yaratmıştır.

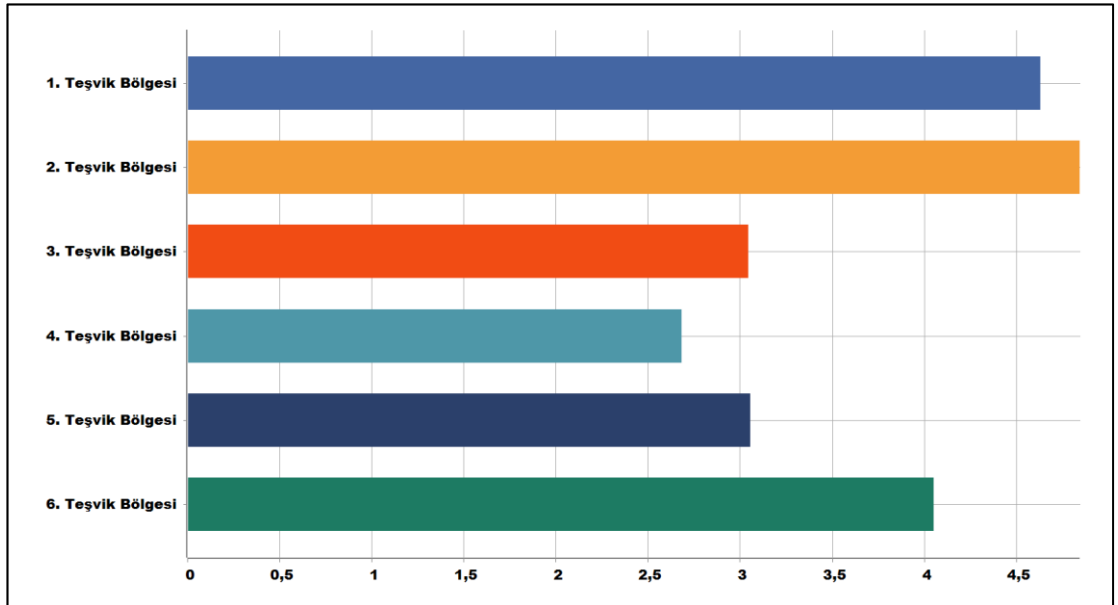


Şekil 56: İngilizce Bilen – Yatırım Teşvik Bölgesi Dağılımı

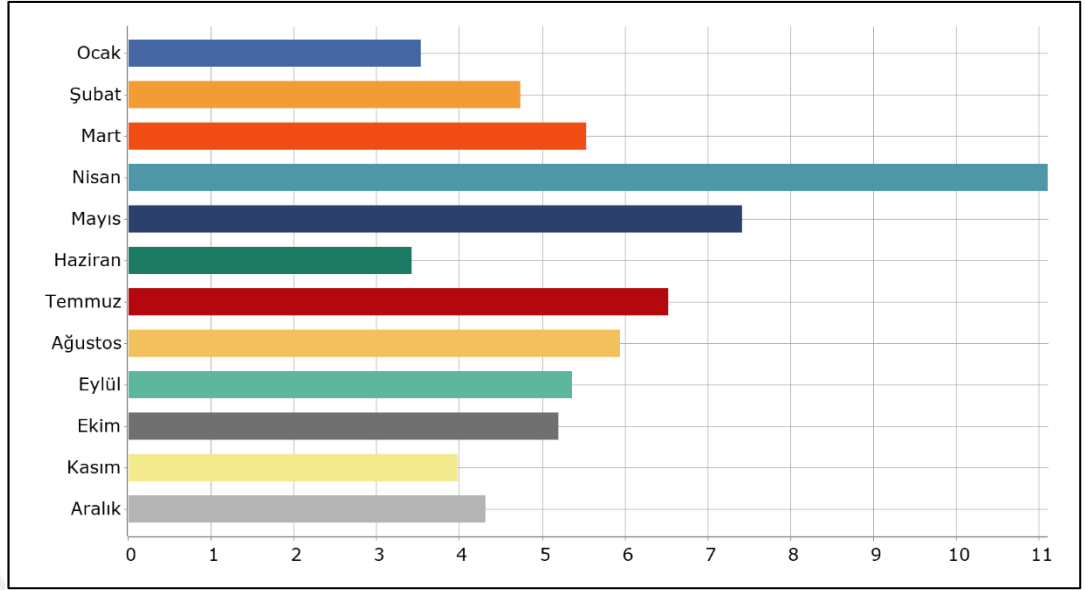


Şekil 57: İngilizce Bilen – Aylık Dağılımı

İngilizce Bilen kelime grubunun, yatırım teşvik bölgeleri üzerinde Şekil 56'daki gibi dağıldığı görülmüştür. Bölgelerdeki dağılım, frekans değerleriyle birlikte incelendiğinde başta büyükşehirler olmak üzere ülke genelinde iş ilanlarında sıkça aranan bir nitelik olduğunu göstermektedir. Bununla birlikte Şekil 57'deki aylık dağılımlara bakıldığında, tüm aylarda trendi yüksek bir kelime öbeği olarak karşılaşılmıştır. Yaz başındaki düşük trendin insan kaynakları uzmanlarının işgücü planlamaları takvimi nedeniyle olduğu düşünülebilir.

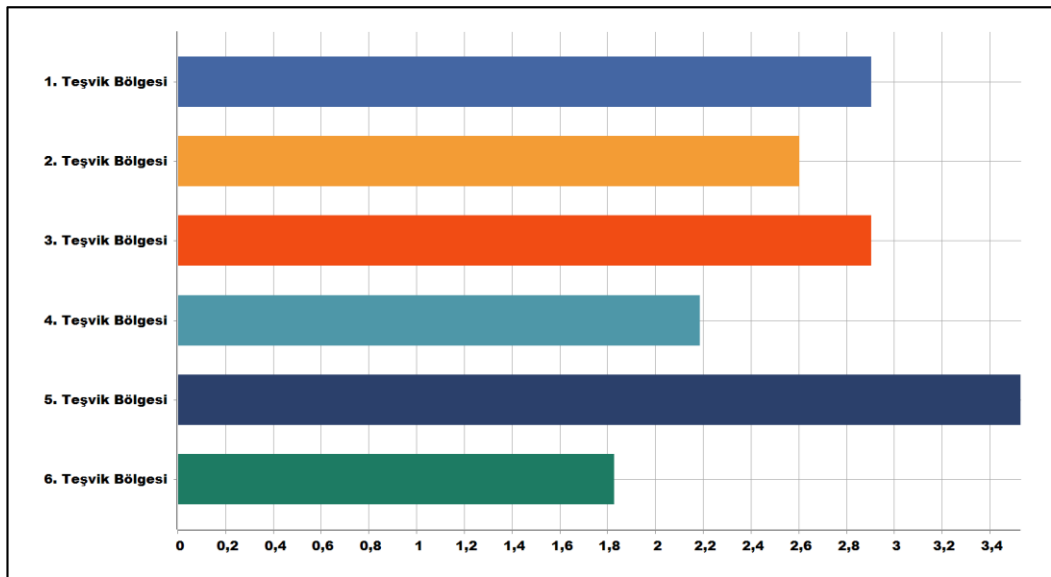


Şekil 58: İnsan Kaynakları – Yatırım Teşvik Bölgesi Dağılımı

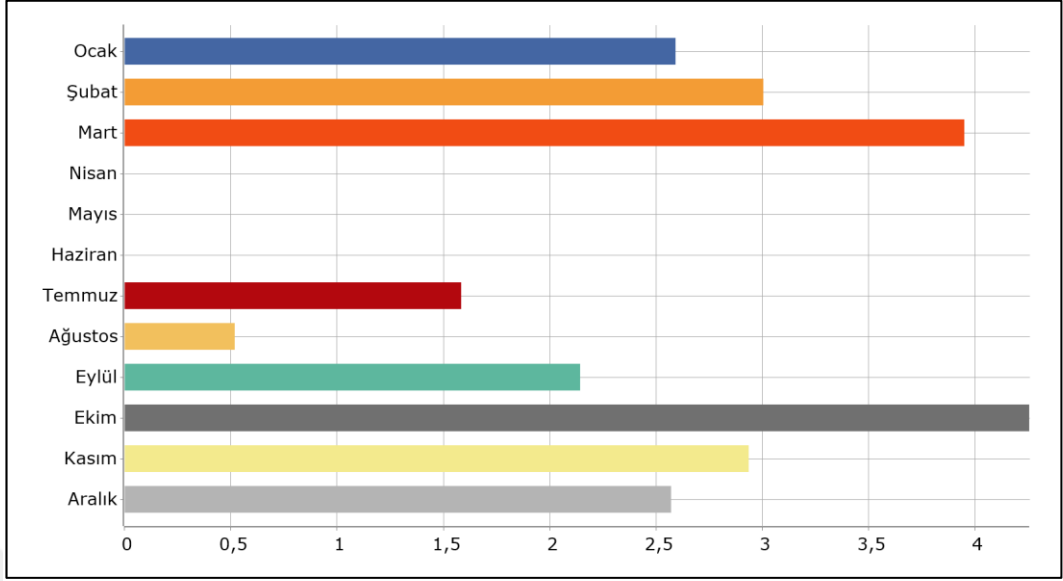


Şekil 59: İnsan Kaynakları – Aylık Dağılımı

İnsan Kaynakları kelime grubu, Şekil 58’deki bölgesel dağılımda frekans değeri yüksek olan ve ülke genelinde 1. ve 2. bölgelerde aranan trend bir alan olmakla birlikte, diğer bölgelerde de talebinin çok olduğu bir kavram olarak görülebilir. İnsan Kaynakları niteliği için oluşan talebin aylık dağılımda Nisan ayında artışı Şekil 59’da görülmüştür. Önceki analizlerde genel işgücü planlama takviminin başlangıcı olarak kabul edilebilecek Haziran ayı ve çevresinin hemen öncesinde yer alan bu trend, İK alanına yönelik istihdam dönemini açıklayıcı olmuştur ve alanın iş takvimine yönelik sebep sonuç ilişkisini ortaya koymuştur.

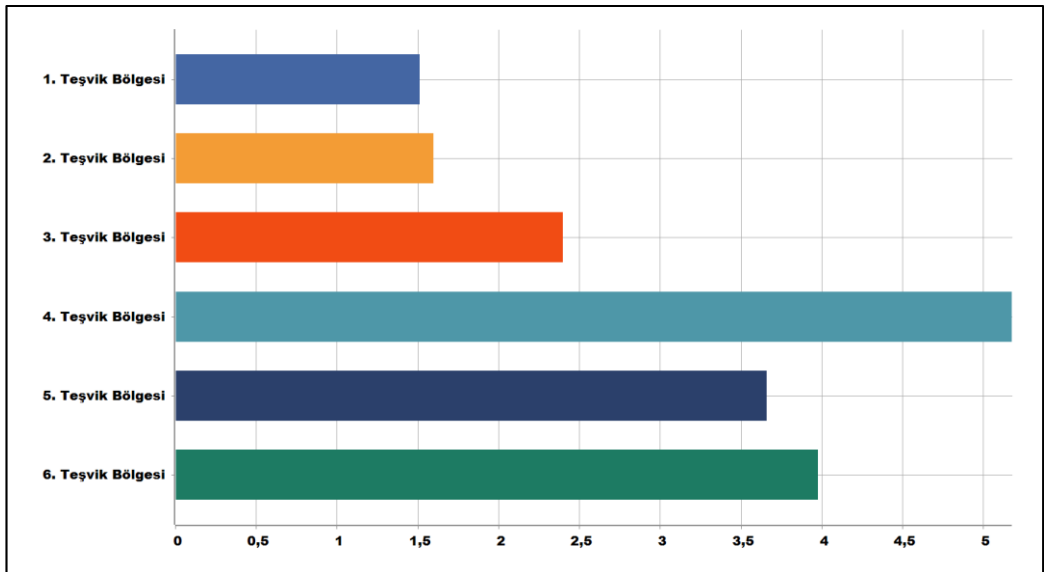


Şekil 60: İş Sağlığı ve Güvenliği – Yatırım Teşvik Bölgesi Dağılımı

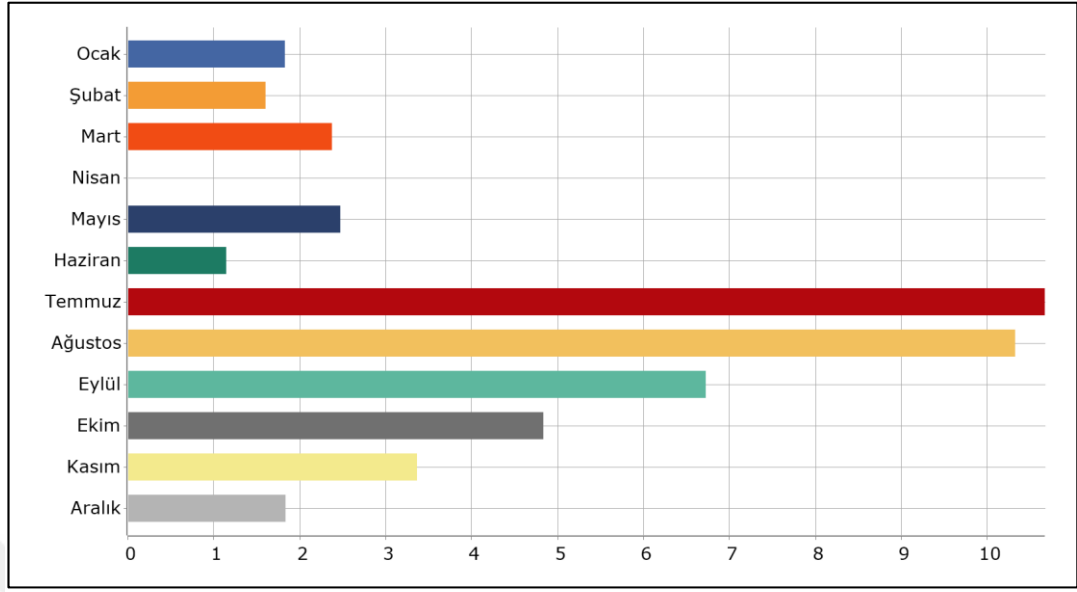


Şekil 61: İş Sağlığı ve Güvenliği – Aylık Dağılımı

İş Sağlığı ve Güvenliği dördü bir kelime grubu olarak yüksek trendiyle dikkat çekicidir. Bu bağlamda, yatırım teşvik bölgelerindeki dağılımı Şekil 60’ta incelendiğinde ülke geneline yaygın bir nitelik olarak görülmektedir. Aylık dağılım ölçeğinde ise bu niteliğin Şekil 61’deki bulguları, bahar aylarında bulunmamakla birlikte yaz döneminde de düşük seyretmiştir. İlan devir döngüsü hızlı olduğu öngörülen kavramın bahar ve yaz mevsimleri öncesinde istihdam sağlandığı ifade edilebilir. Dolayısıyla popülerliğini koruyan bir alan olarak, dönemsel olarak arz-talep edilmeyen bir uzmanlık şeklinde tahmin yürütülebilir.

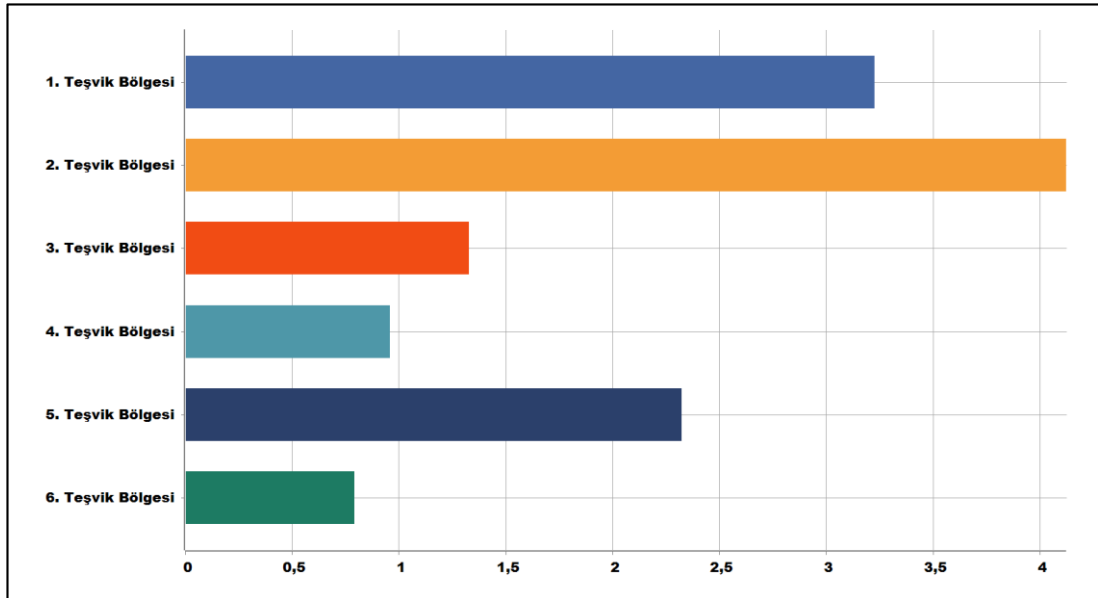


Şekil 62: Part-Time – Yatırım Teşvik Bölgesi Dağılımı

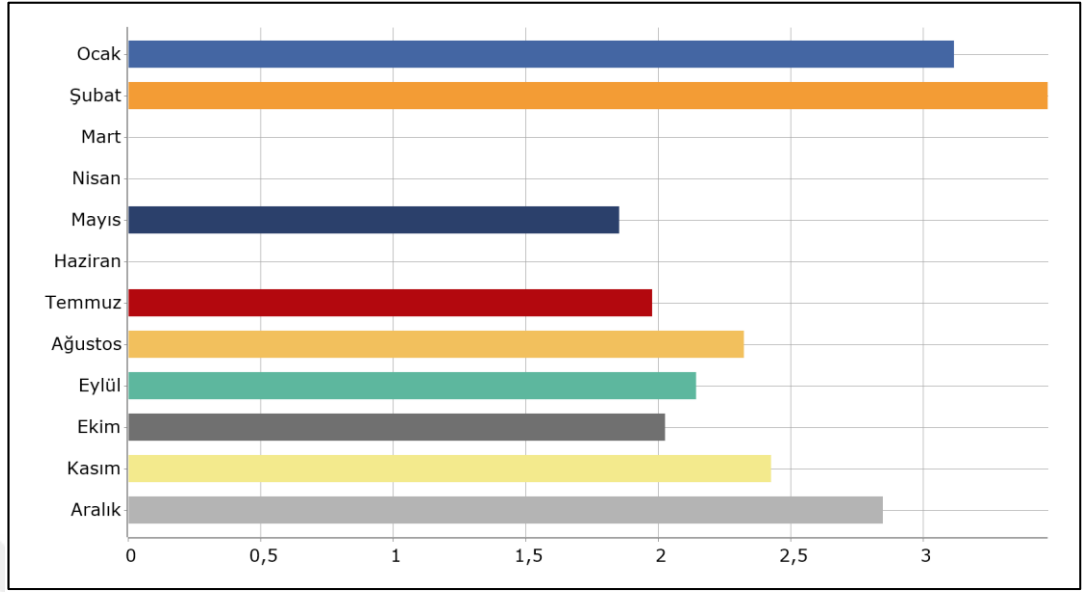


Şekil 63: Part-Time – Aylık Dağılımı

Part-Time kelime grubu, yatırım teşvik bölgelerinde Şekil 62’deki gibi dağılmıştır. Bu kavram, daha küçük kentlerde yoğunlaşmıştır. Bununla birlikte Şekil 63’te görüldüğü üzere, aylık dağılımında Temmuz ve Ağustos aylarında trendin yükseldiği bulunmuştur. Part-Time sürekli bir çalışma şeklini ifade etmemesi nedeniyle yaz dönemindeki istihdamın azalması ve tatil dönemi etkisiyle bu trendin arttığı anlaşılmaktadır. Ayrıca Nisan ayında bulunmaması veri setinin hazırlandığı dönemle ilişkilendirilebilir, dolayısıyla kavramın süreklilik arz etmeyen ilanlarda geçtiği anlaşılmaktadır.

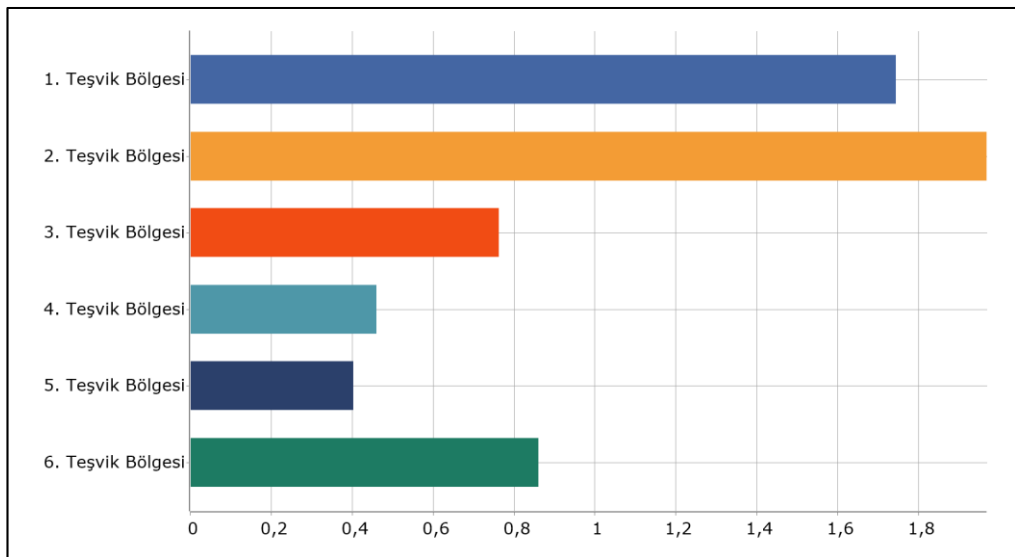


Şekil 64: Satın Alma – Yatırım Teşvik Bölgesi Dağılımı

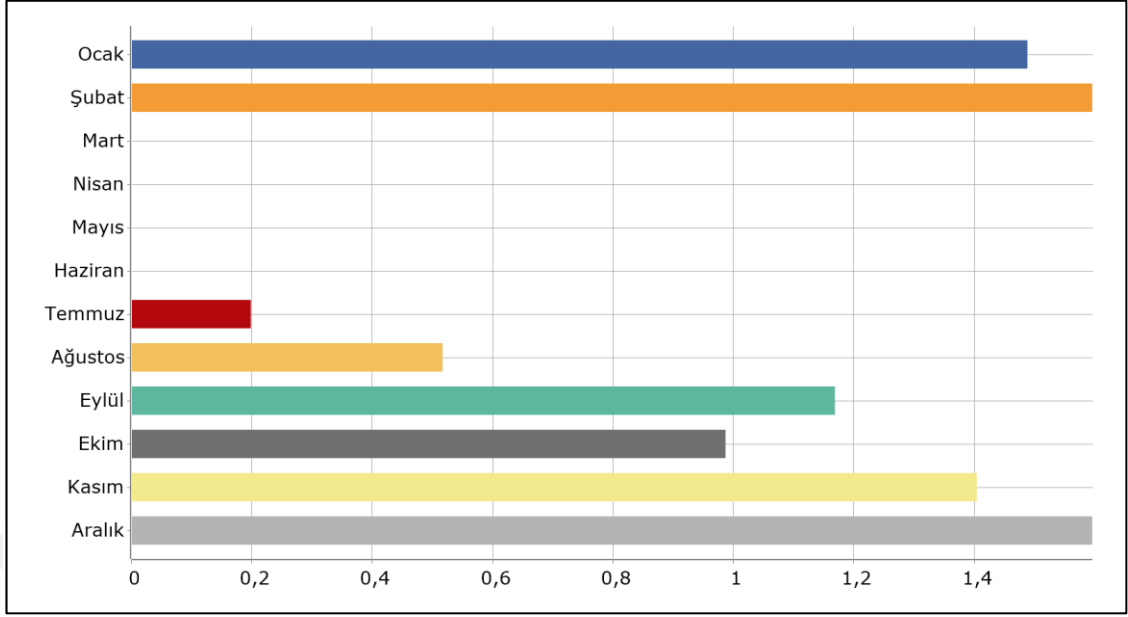


Şekil 65: Satın Alma – Aylık Dağılımı

Şekil 64’te Satın Alma kelime grubunun yatırım teşvik bölgelerindeki dağılım gösterilmiştir. Bölgesel dağılımları trendi yüksek Şekil 25’te gösterilen Lojistik kavramına paralellik göstermiştir. Fakat Şekil 65’teki aylık dağılımları Lojistik özelindeki Şekil 26’daki dağılıma benzememektedir. Dolayısıyla bölgesel trendin Lojistik teriminde olduğu gibi büyükşehirlerin çekim merkezi olduğu bir nitelik olduğu Satın Alma içinde söylenebilir. Ancak Mart, Nisan ve Haziran ayındaki sıklık gözlemlenmemiş olması, trend olan bu niteliği içeren ilanların devir hızının yüksek olduğunu ifade etmektedir.



Şekil 66: Sosyal Medya – Yatırım Teşvik Bölgesi Dağılımı



Şekil 67: Sosyal Medya – Aylık Dağılımı

Şekil 66’da Sosyal Medya kelime grubunun yatırım teşvik bölgelerindeki dağılım gösterilmiştir. Bu kavram için, önceki analizlerdekine benzer bir şekilde büyük kentlerin daha çok ilgilendiği bir nitelik olduğunu göstermiştir. Sosyal Medya kavramının Şekil 67’teki aylık dağılımı incelendiğinde, sonbaharla birlikte yükselen trendin baharda sona erdiği görülmektedir. Frekansın seyrek olması sebebiyle henüz diğer incelenen nitelikler kadar trendi yüksek olmayan bu alana yönelik istihdamın kış aylarında olduğu görülmektedir.

3.3. Çıkarılmış Varlık İsimleri

Türkçe iş ilanları üzerine gerçekleştirilen metin madenciliği analizlerinden varlık isimleri elde edilmiştir. Toplamda 21679 varlık ismi çıkarılmış olup, bunlardan toplam sayısına göre sıralandığında ilk 100 tanesi Tablo 3’te gösterilmiştir. Bu varlıkların, İK terminolojisi için metin madenciliğinde analizlerinin özellikle NLP aşamalarında kullanılabilecek terimler olduğu anlaşılmaktadır. Bunların bir kısmı işin tanımı, bir kısmı işlerin niteliği ve bir diğer kısmı ise firmanın misyon ve vizyonunu ifade ettiği görülmüştür. Bu nedenle İK alanında iş ilanları konusunda bir standartizasyona ve ortak bir iş ilanı terminolojisi gereksinimine, Türkçe metin madenciliği düzeyinde sözcüklerden daha başarılı anlamsal çıkarımlar yapabilen NLP tekniklerine ihtiyaç olduğu görülmüştür.

Tablo 3: Çıkarılmış Varlık İsimleri

Varlık İsmi	Toplam	Benzersiz
TANIMI	30672	2478
Tercihen	24432	10904
Is	23449	4664
Erkek	19119	18686
Türkiye	17331	13939
Office	17012	834
Satis	16348	7031
MS Office	14013	13328
İngilizce	12989	10518
Müşteri	12686	9819
Üniversitelerin	11760	4020
İletişim	10985	10405
Mühendisliği	9617	54
Takim	9579	8675
Ekip	9368	8870
Nitelikler	9282	1264
Eğitim	9009	1767
İstanbul	8752	3627
İnsan	7425	5178
Genel	7366	771
İyi	7295	6824
Seyahat	7205	6781
B.	6863	5674
Analitik	6012	5863
Planlama	5860	4236
Magaza	5116	2798
Üretim	4596	2684
Esnek	4595	4532
Uzmanı	4592	157
GAYRİMENKUL	4586	1022
Meslek	4399	217
Değerleme	4373	248
ANKARA	4011	2936
Askerlik	4001	3836
Aktif	3883	3404
Kalite	3880	1688
Öğretim	3861	4
Yogun	3836	3611
Anadolu	3799	477
Bölge	3772	506
Teknik	3741	1606
Dis	3668	2642
Yeni	3656	2996
Yüksek	3606	706
Elektrik	3601	1280

Tablo 3: Çıkarılmış Varlık İsimleri (Devamı)

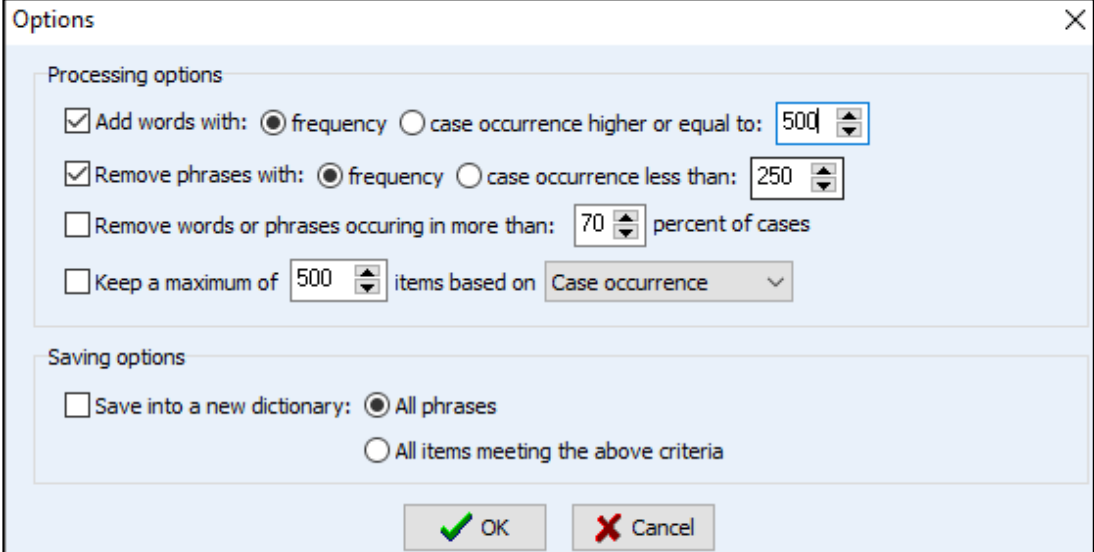
Varlık İsmi	Toplam	Benzersiz
Ürün	3518	2493
Koleji	3446	18
Avrupa	3405	1652
Makine	3396	1219
Sağlık	3362	818
Alanında	3357	3210
Sorumlu	3325	2298
Muhasebe	3316	1476
Bahçesehir	3196	99
Pazarlama	3170	1804
Lisans	3169	1501
Izmir	3090	2438
Tüm	3080	1813
Saha	3048	1252
Sorumluluk	2957	2566
Üniversite	2900	1921
İşletme	2899	1873
Dinamik	2862	2803
Endüstri	2624	445
Bilgisayar	2601	1278
Müdürü	2577	80
Yönetim	2514	434
Elektronik	2489	938
Diksiyonu	2475	2412
Kurumsal	2474	976
Kimya	2423	759
Magazacılık	2414	1752
Temsilcisi	2391	8
Perakende	2353	1708
Otomotiv	2352	1292
Kişisel	2348	1683
Sürücü	2341	1957
Antalya	2339	1806
Ofis	2339	862
BURSA	2321	1900
Proje	2291	1480
Mevcut	2270	1832
Lise	2249	1449
Group	2249	678
A.S.	2241	2241
Grubu	2232	48

Tablo 3: Çıkarılmış Varlık İsimleri (Devamı)

Varlık İsmi	Toplam	Benzersiz
Yurt	2220	1693
Öğretmenliği	2218	69
Lisesi	2210	126
Çalışma	2187	1323
Sorumlusu	2139	151
Özel	2104	246
İlaç	2102	1135
Pedagojik	2055	1907
İnsaat	2038	416
İletisimi	2029	1978
Günlük	2011	1593
Yenilikçi	2005	1952
İkna	1997	1782
Yönetimi	1966	88

3.4. Çapraz Tablo Analizleri

Bu bölümde, sık kullanılan kelimeler ve kelime gruplarında elde edilmiş çapraz tablo analizlerine yer verilmiştir. Daha önceden oluşturulan modelin üzerine çapraz tablo analizleri için Şekil 68'deki gibi model tekrardan geliştirilmiştir. Frekans ve kelime grubu analizleri için gerçekleştirilen modelin üzerinde, yeni çapraz tablo analizi modelinin sıklık değeri 500 üzeri sözcükler ve sıklık değeri en az 250 olan kelime grupları olacak şekilde yeniden kısıtları verilmiştir.



Şekil 68: Çapraz Tablo Analizi Kısıtları

Oluşturulan modele bu yeni kısıtlar işlendikten sonra 6. yatırım teşvik bölgesine göre çapraz tablo analizi yapılmış ve bu analizden 2 boyutlu benzeşme alanı grafiği oluşturulmuştur. Daha sonra bu analizlerden en çok ve en az geçen kelimeleri gösteren bir tablo ayrıca alınmıştır.

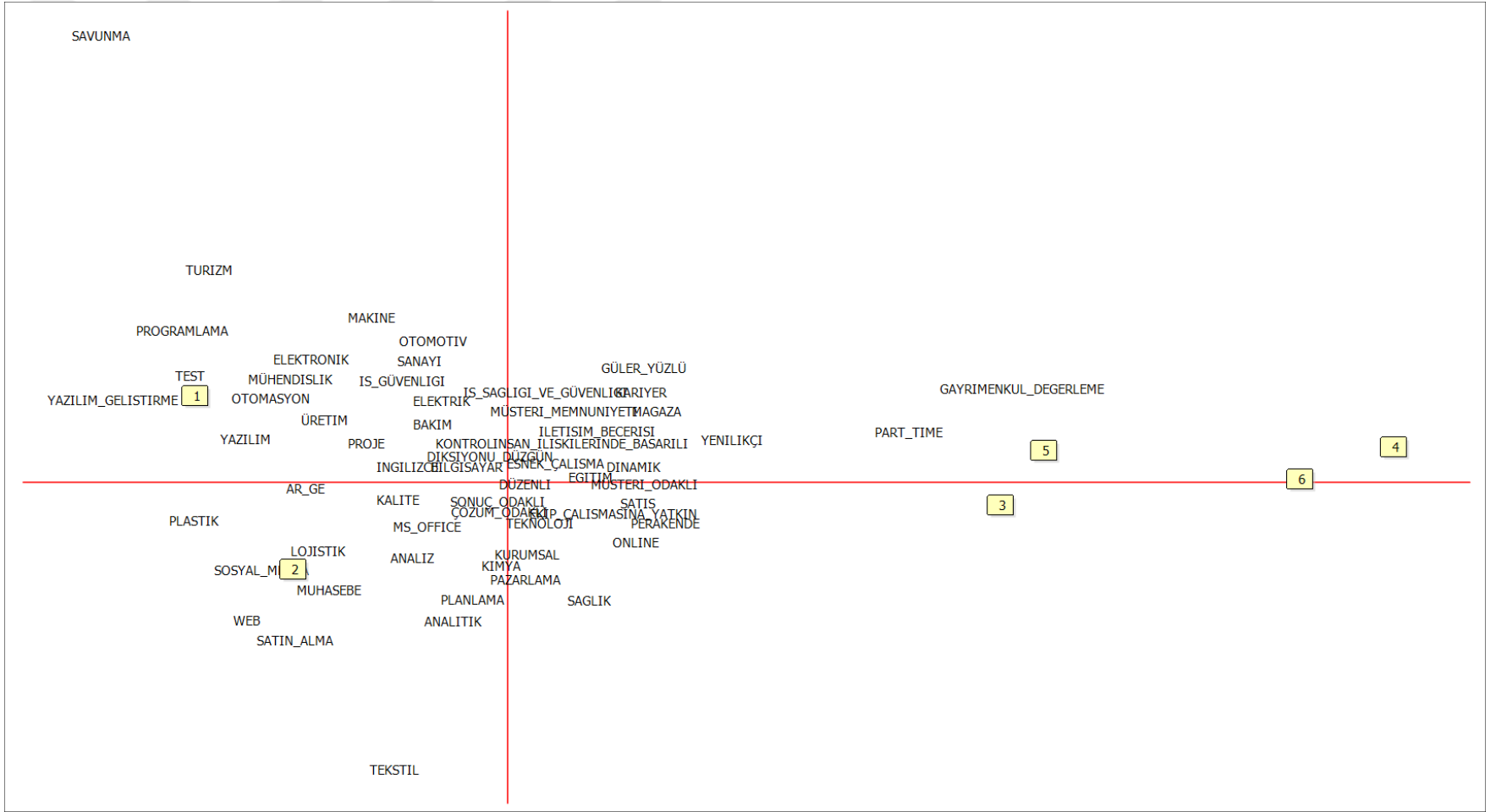
Çapraz tablo analizde keşfedilmiş kelime ve kelime gruplarına yönelik detaylı analizlerde coğrafi analizler gerçekleştirilmek istenmiştir. Bu bağlamda, CBS olarak Wordstat'a entegre olarak kullanılan GISViewer üzerinde keşfedilmiş kelime ve kelime gruplarının ısı haritaları oluşturularak ile coğrafi analizler elde edilmiştir ve bu analizler yorumlanmıştır.

Tablo 4: Çapraz Tablo Analizi

	1. Bölge	2. Bölge	3. Bölge	4. Bölge	5. Bölge	6. Bölge
ANALITIK	840	1086	147	172	181	106
ANALIZ	1590	1803	199	256	343	144
AR_GE	747	695	79	52	112	28
BAKIM	2123	1971	251	362	459	179
BILGISAYAR	1062	1021	158	224	240	88
DIKSIYONU_DÜZGÜN	1303	1234	223	281	298	136
DINAMIK	2276	2234	466	836	784	415
DÜZENLİ	1723	1685	310	420	448	228
EGITIM	3275	3003	810	999	831	510
EKIP_ÇALIŞMASINA_YA TKIN	2221	2180	454	746	684	415
ELEKTRİK	1559	1206	228	250	285	133
ELEKTRONİK	951	681	69	85	120	39
ESNEK_ÇALIŞMA	2073	1978	331	599	601	274
GAYRIMENKUL_DEGER LEME	1049	937	445	1101	840	534
GÜLER_YÜZLÜ	1225	915	226	439	385	189
İLETİSİM_BECERİSİ	1011	888	183	326	293	153
İNGİLİZCE	1847	1686	212	279	334	157
İNSAN_ILİSKİLERİNDE_ BASARILI	1291	1152	279	378	385	185
İS_GÜVENLİĞİ	420	326	43	51	82	34
İS_SAGLIĞI_VE_GÜVEN LİĞİ	460	432	82	108	169	46

Tablo 4: Çapraz Tablo Analizi (Devamı)

	1. Bölge	2. Bölge	3. Bölge	4. Bölge	5. Bölge	6. Bölge
KALITE	2518	2327	331	308	533	174
KARIYER	1188	1139	296	426	409	211
KIMYA	627	725	125	144	150	91
KONTROL	1756	1536	245	328	383	172
KURUMSAL	1392	1639	242	413	348	202
LOJİSTİK	801	941	98	60	134	42
MAGAZA	1926	1989	473	816	711	318
MAKİNE	1675	1083	144	177	294	94
MS_OFFICE	830	875	112	144	167	78
MUHASEBE	1636	1896	159	183	189	137
PLASTİK	427	420	12	19	32	7
PROGRAMLAMA	364	242	12	10	22	8
PROJE	2051	1767	185	295	281	129
SAGLIK	1231	1669	259	428	373	305
SANAYİ	827	566	72	135	133	67
SATIN_ALMA	659	843	47	52	127	22
SATIS	7630	7958	1780	2850	2768	1420
SAVUNMA	385	136	2	3	12	4
SONUÇ_ODAKLI	1772	1752	306	383	419	221
SOSYAL_MEDYA	356	401	27	25	22	24
TEKNOLOJİ	696	776	110	199	223	90
TEKSTİL	567	1014	125	127	91	40
TEST	1004	742	35	30	73	20
TURİZM	1087	620	74	52	38	24
WEB	614	735	54	35	56	21
YAZILIM	1277	1087	51	91	94	55
YAZILIM_GELİSTİRME	409	303	9	3	8	2
YENİLİKÇİ	1085	1003	287	481	454	293
ÇÖZÜM_ODAKLI	1120	1103	184	255	251	143
ÜRETİM	4612	4089	392	360	794	234



Şekil 69: Benzeşme Alanı Grafiği

Tablo 5: En Çok Ve En Az Geçen Nitelikler

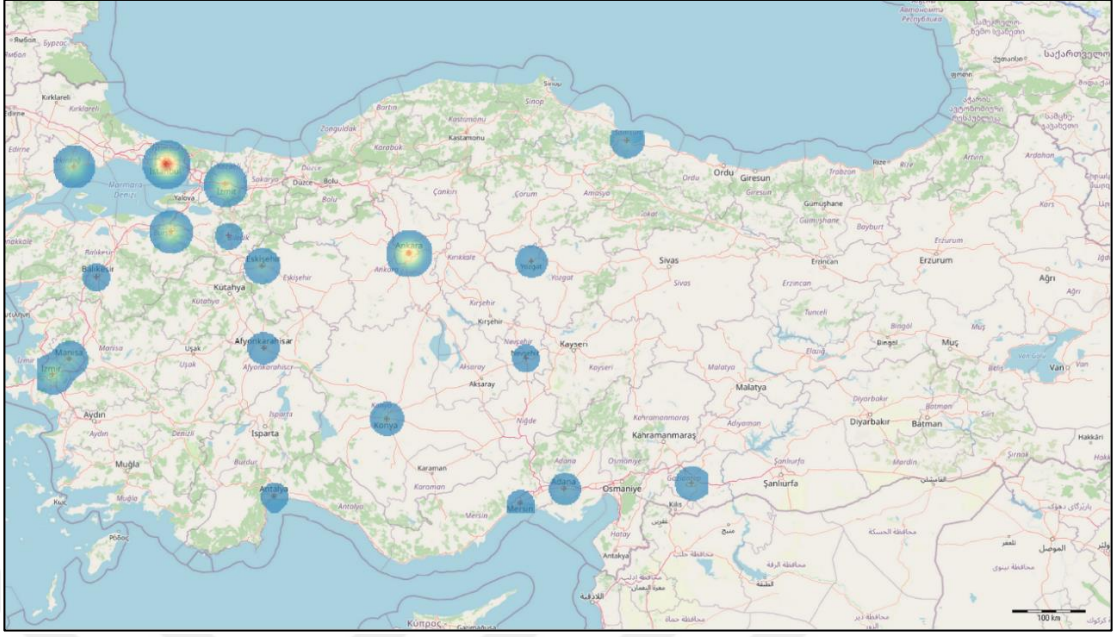
1. Bölge	1. Bölge	2. Bölge	2. Bölge	3. Bölge	3. Bölge	4. Bölge	4. Bölge	5. Bölge	5. Bölge	6. Bölge	6. Bölge
Çok	Az	Çok	Az	Çok	Az	Çok	Az	Çok	Az	Çok	Az
TURİZM	GAYRIMENKUL_DEGERLEME	TEKSTİL	GAYRIMENKUL_DEGERLEME	EGİTİM	ÜRETİM	GAYRIMENKUL_DEGERLEME	ÜRETİM	GAYRIMENKUL_DEGERLEME	TURİZM	GAYRIMENKUL_DEGERLEME	ÜRETİM
ÜRETİM	SATIS	MUHASEBE	GÜLER YÜZLÜ	GAYRIMENKUL_DEGERLEME	YAZILIM	PART_TIME	KALİTE	SATIS	MUHASEBE	YENİLİKÇİ	TEST
SAVUNMA	PART_TIME	SATIN_ALMA	YENİLİKÇİ	SATIS	TEST	SATIS	TEST	PART_TIME	YAZILIM	SATIS	KALİTE
MAKİNE	SAGLIK	WEB	SATIS	KARİYER	MUHASEBE	MAGAZA	MUHASEBE	YENİLİKÇİ	TEST	SAGLIK	TURİZM
TEST	MAGAZA	LOJİSTİK	PART_TIME	MAGAZA	PROJE	YENİLİKÇİ	YAZILIM	MAGAZA	PROJE	PART_TIME	SATIN_ALMA
YAZILIM	PERAKENDE	ANALİZ	EGİTİM	YENİLİKÇİ	PLASTİK	DİNAMİK	LOJİSTİK	DİNAMİK	YAZILIM_GELİSTİRME	EKİP ÇALIŞMASINA_YATKIN	PROJE
ELEKTRONİK	YENİLİKÇİ	ÜRETİM	DİNAMİK	İNSAN_ILISKILERINDE_BASARILI	SATIN_ALMA	GÜLER YÜZLÜ	TURİZM	İS_SAGLIGI_VE_GÜVENLİĞİ	WEB	DİNAMİK	YAZILIM
MÜHENDİSLİK	PAZARLAMA	ANALİTİK	MAGAZA	EKİP ÇALIŞMASINA_YATKIN	OTOMASYON	MÜŞTERİ OD AKLI	WEB	MÜŞTERİ OD AKLI	TEKSTİL	ONLINE	WEB
PROJE	TEKSTİL	PAZARLAMA	KARİYER		SAVUNMA	PERAKENDE	SATIN_ALMA	GÜLER YÜZLÜ	SOSYAL_MEDYA	PERAKENDE	AR_GE
YAZILIM_GELİSTİRME	DİNAMİK	SOSYAL_MEDYA	SAVUNMA		YAZILIM_GELİSTİRME	EKİP ÇALIŞMASINA_YATKIN	AR_GE	KARİYER	ÜRETİM	EGİTİM	LOJİSTİK
PROGRAMLAMA	MÜŞTERİ OD AKLI	PLASTİK	MAKİNE		MÜHENDİSLİK	KARİYER	MAKİNE	PERAKENDE	PLASTİK	KARİYER	ELEKTRONİK
OTOMASYON	ONLINE	YAZILIM	İNSAN_ILISKILERINDE_BASARILI		MAKİNE	EGİTİM	YAZILIM_GELİSTİRME	İNSAN_ILISKILERINDE_BASARILI	MÜHENDİSLİK	MÜŞTERİ OD AKLI	MAKİNE
OTOMOTİV	KARİYER	PLANLAMA	OTOMOTİV		ELEKTRONİK	İLETİŞİM_BECERİSİ	ANALİZ	EKİP ÇALIŞMASINA_YATKIN	SAVUNMA	GÜLER YÜZLÜ	MÜHENDİSLİK
SANAYİ	EKİP ÇALIŞMASINA_YATKIN	SAGLIK	İLETİŞİM_BECERİSİ		PROGRAMLAMA	ESNEK ÇALIŞMA	PLASTİK	ESNEK ÇALIŞMA	PROGRAMLAMA		TEKSTİL
ELEKTRİK	KURUMSAL	KURUMSAL	EKİP ÇALIŞMASINA_YATKIN		BAKIM		ELEKTRONİK		ELEKTRONİK		YAZILIM_GELİSTİRME
KALİTE	ANALİTİK	AR_GE	MÜŞTERİ OD AKLI		ANALİZ		İNGİLİZCE		İNGİLİZCE		PLASTİK

Tablo 4'teki çapraz tablo analizi bulguları 6 yatırım teşvik bölgesine iki boyutlu konumları benzeşme grafiği ile Şekil 69'da gösterilmiştir. Savunma, Tekstil, Gayrimenkul Değerleme ve Part-Time gibi nitelikler eşsiz varlıklar olarak grafiğin çevresine yayılmıştır. Turizm, Programlama, Makine, Otomotiv, Satın Alma, Web, Analitik, Sağlık ve Plastik gibi sektörlerle ilgili veya uzmanlıkları özelleştiren terimler merkezin etrafında konumlandırılmıştır. Fakat en merkezi bulunan noktada, tüm ilanlarda ortak tespit edilen varlıklar Düzenli, Müşteri Odaklı, Sonuç Odaklı, Çözüm Odaklı, İngilizce Bilen gibi sıralanmıştır. Şekil 69'daki benzeşme grafiği, İnsan Kaynakları alanında iş ilanlarında yapısal bir standartta oluşturulması ve değerlendirilebilmesi için uzmanlara ve araştırmacılara bakış açısı ve model geliştirmeye kolaylaştırıcı bir görsel anlatım sunduğu görülmektedir. Tablo 5'te ise gösterilmiş çok ve az düzeyde bulunan kavramlar verilmiştir. Bu kavramlar karşılaştırıldığında, ülke genelindeki 1. ve 2. bölgenin ekonomik, endüstriyel ve teknolojik kalkınmanın lokomotifi olduğu, taşranın üretim ve ilgili yan uzmanlık alanlarından uzak nitelikler aradığı bu karşılaştırılmalı analizlerde daha net görülmektedir.

3.5. Coğrafi Analizler

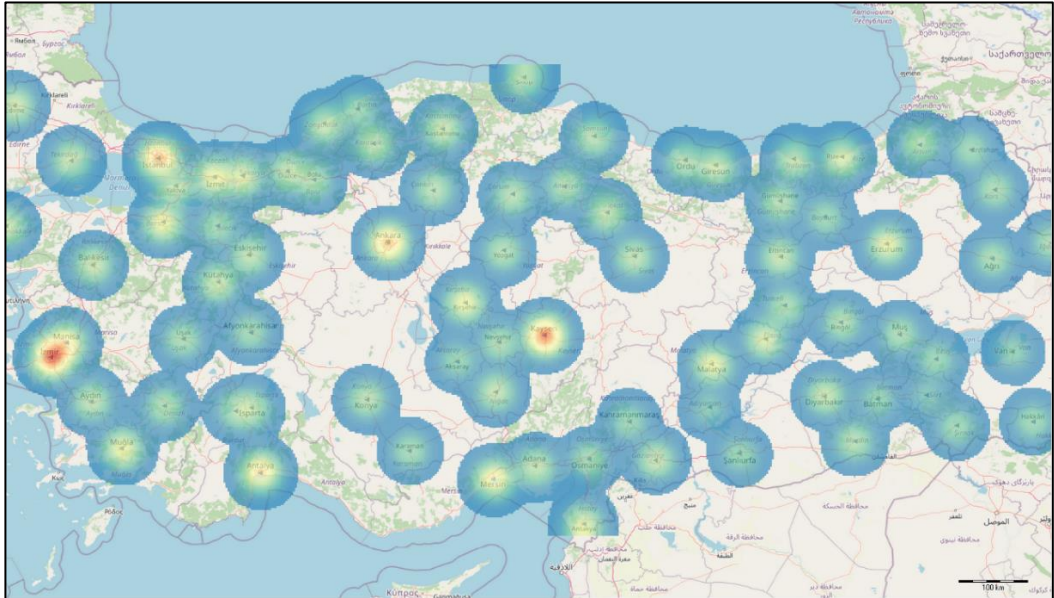
GISViewer ile coğrafi analizleri için, çapraz tablo analizinde tespit edilen *AR-GE*, *Gayrimenkul Değerleme*, *İş Sağlığı ve Güvenliği*, *Lojistik*, *Mağaza*, *Otomasyon*, *Otomotiv*, *Part-Time*, *Satın Alma*, *Savunma*, *Sosyal Medya*, *Tekstil* ve *Turizm* ifadelerinin ve *Yazılım* ile birleştirilerek *Yazılım Geliştirme* ifadesinin sıcaklık haritaları oluşturulmuştur.

Coğrafi analizlerde, çapraz tablo analizinden seçilmiş ifadelerin veri setinde ilanların şehir isimleriyle eşleştirilmiş koordinat bilgileri olan enlem ve boylam değerleriyle işaretlemeleri yapılarak sıcaklık haritaları oluşturulmuştur. Fakat buradaki ısı değeri, frekans analizlerinde olduğu gibi 10 bin vakada geçen kelime sıklığı değildir. Isıyı oluşturan faktör bu haritalarda iş ilanlarının kendileridir. Dolayısıyla çapraz tablo analizinde ilgili kelime veya kelime öbeğinde aynı koordinatlara sahip ne kadar fazla iş ilanı var ise renk kırmızıya doğru dönecektir. Bu nedenle, kırmızıya yakın olan işaretli bölgelerde işverenlerin ilgili nitelikle ilgili yoğun bir şekilde iş ilanı paylaştığı anlaşılabilir.



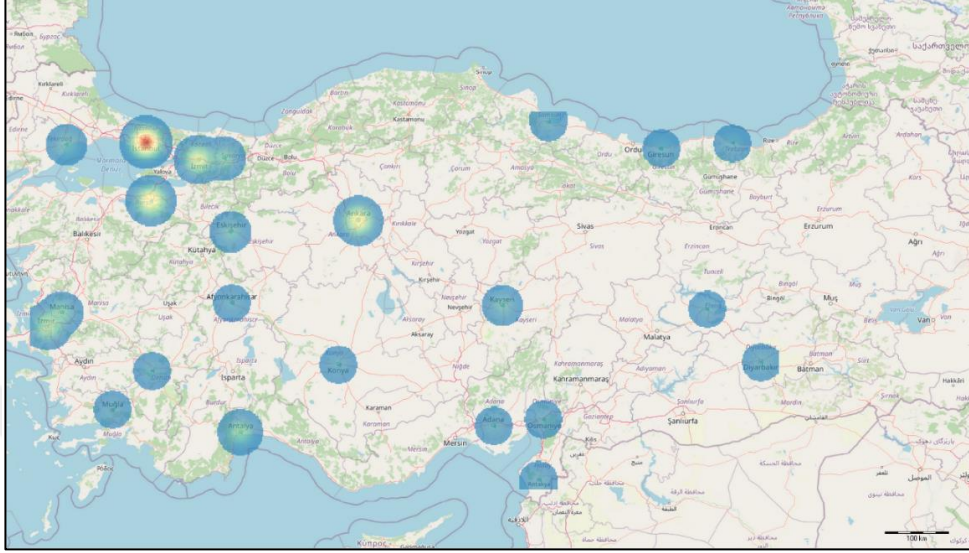
Şekil 70: AR-GE Sıcaklık Haritası

Şekil 70'e göre AR-GE niteliğinde oluşan sıcaklık haritasında, Marmara bölgesi ve Ankara başta olmak üzere sanayinin geliştiği kentlerde yoğunluk olduğu anlaşılmaktadır. Özellikle Otomotiv ve Savunma sanayi faaliyetlerinin etkisi burada direkt göze çarpmaktadır. Bununla birlikte, büyük organize sanayi bölgeleri olan kentler hariç, liman kentleri olan Mersin ve Samsun gibi şehirlerin de AR-GE ile ilgili ilişkisi olduğu görülmektedir.



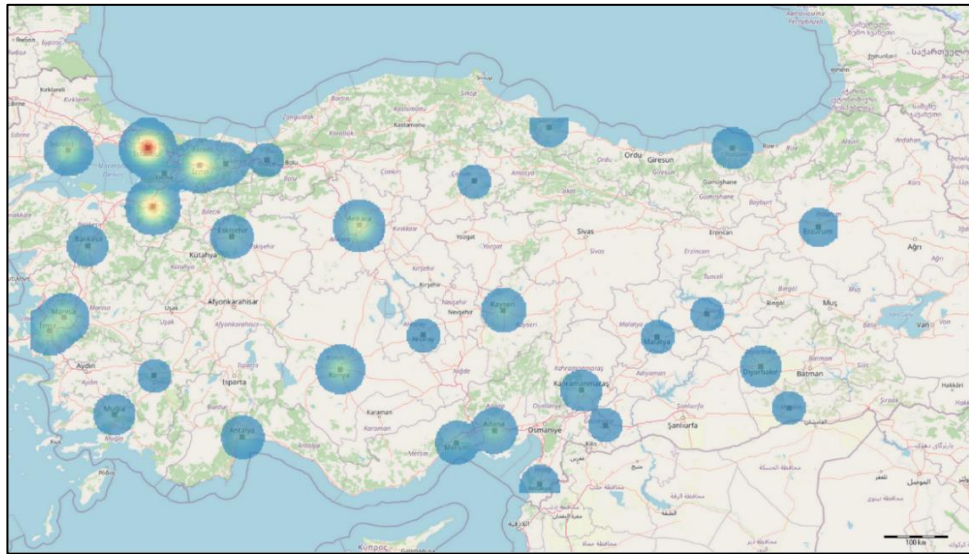
Şekil 71: Gayrimenkul Değerleme Sıcaklık Haritası

Şekil 74'teki sıcaklık haritası, Mağazacılık niteliğiyle ilgili sıcaklık haritasını göstermektedir. Burada mağazacılık sektörüne yönelik neredeyse 81 ilin tamamına yakınında iş ilanı olduğu görülmektedir. Mağazacılık niteliğinin birçok alanda olduğu gibi İstanbul'da daha fazla istihdam edilmek istendiğini anlaşılmaktadır.



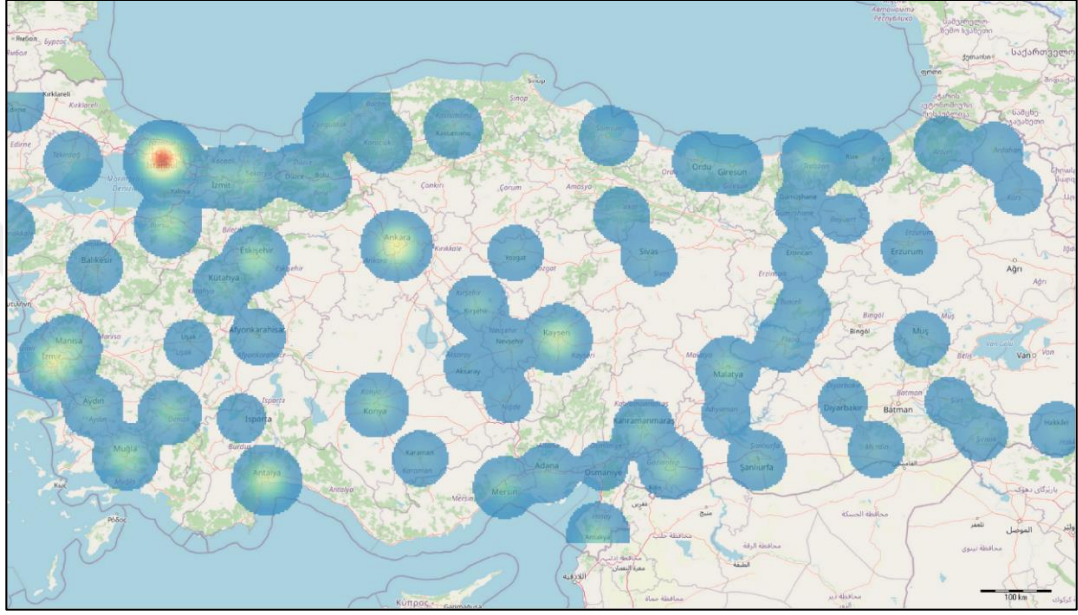
Şekil 75: Otomasyon Sıcaklık Haritası

Otomasyon niteliği için Şekil 75'te oluşan sıcaklık haritası görülmektedir. Otomasyon niteliğinin oluşan haritadan, özellikle İstanbul, Doğu Marmara Bölgesi ve Ankara'da daha çok arz-talep edilmek istendiği anlaşılmaktadır. Bu grafikte oluşmuş sıcaklık değerleri hem endüstriyel faaliyet yerleri, hem nüfus yoğunluğuyla ilişkili gibi gözükmektedir.



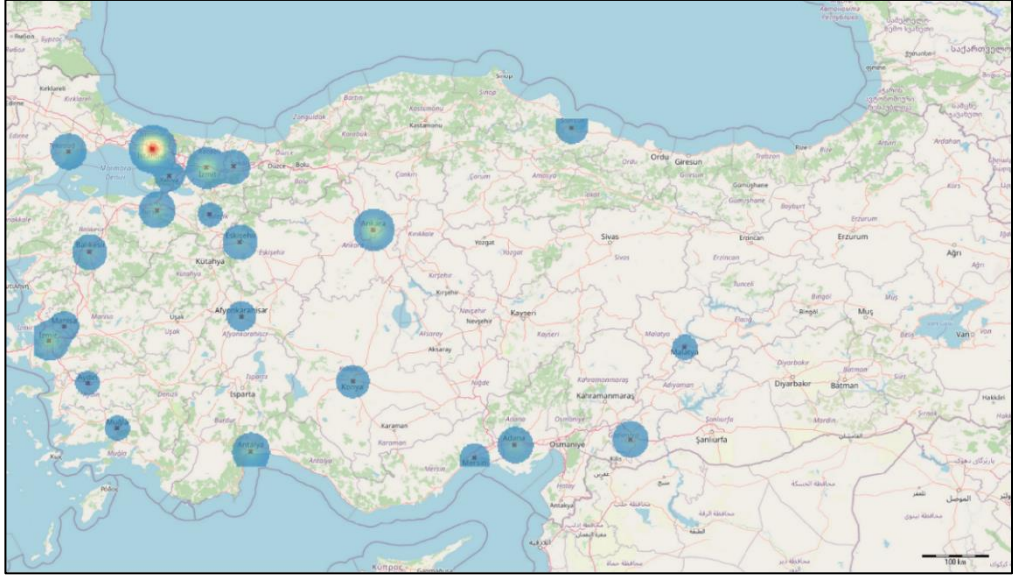
Şekil 76: Otomotiv Sıcaklık Haritası

Şekil 76’da Otomotiv niteliğiyle ilgili sıcaklık haritası gösterilmiştir. Otomotiv niteliğinin kırmızı renkli yoğun olan bölgelerde doğrudan Otomotiv üretimi yapan fabrikaların olduğu bölgeyle ilişkisi haritada ortaya çıkarılmıştır. Ancak daha az ısı göstermesine rağmen Türkiye geneline yayılmış bu ilan ısı haritası, otomotiv satış bayii ile ilişkisi olabileceği tahmini yaptırmaktadır. Bu konuda, frekans analizinde Otomotiv terimi için Şekil 33 ve Şekil 34 ile ilgili yapılan yorumu bu haritanın da desteklediği görülmektedir.



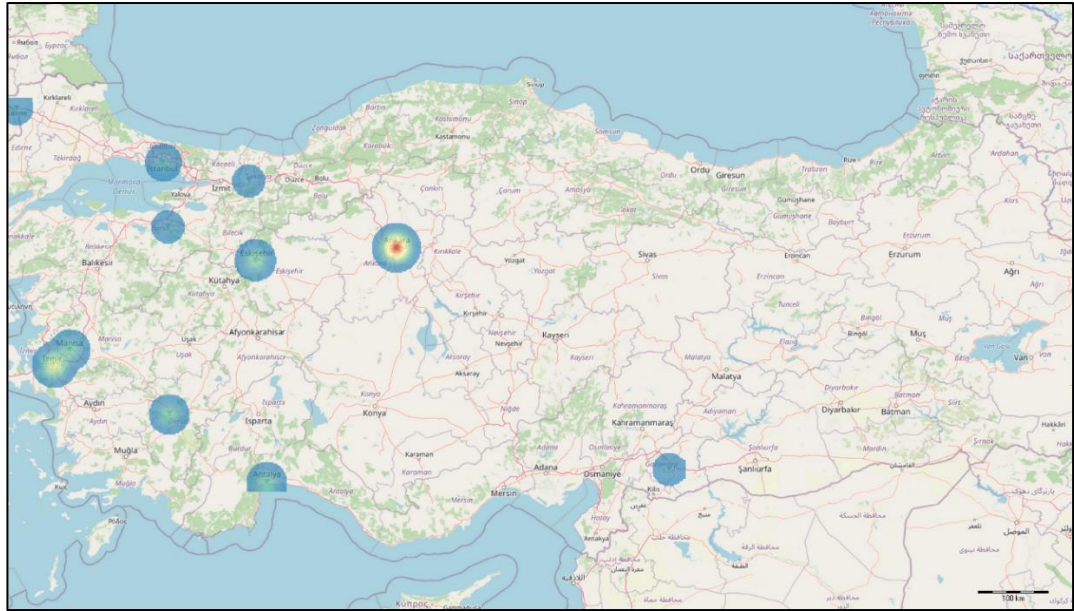
Şekil 77: Part-Time Sıcaklık Haritası

Şekil 77’deki sıcaklık haritasında Part-Time öbeğinin Türkiye haritasındaki yayılımı gösterilmiştir. Part-Time çalışma şeklinin, tüm ülkeye ciddi bir şekilde yayıldığı ve çalışanların böyle önemli bir oranda istihdam edilmek istendiği anlaşılmaktadır. Başta İstanbul ve ardından başkent Ankara’da oluşan yoğunluk, öncelikle turizm bölgelerine doğru azalmış ve daha sonra diğer illere doğru genişlemiştir. Bu niteliğin ısı oluşturduğu şehirler, Şekil 63’te yaz dönemi yükseldiği bulunan istihdamı destekleyici görülmektedir.



Şekil 78: Satın Alma Sıcaklık Haritası

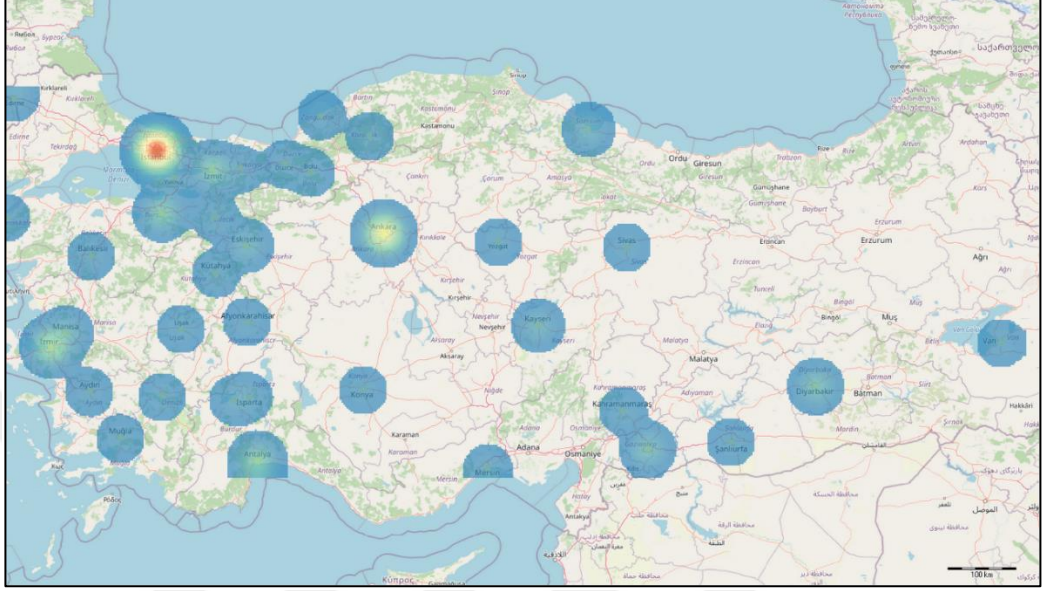
Şekil 78’de Satın Alma niteliğinin sıcaklık haritası verilmiştir. Bu haritada, hemen hemen üretim yapılan şehirlerde Satın Alma niteliği arandığı görülmektedir. Ancak Otomasyon Sıcaklık Haritası ile AR-GE sıcaklık haritası arası bir yerde kaldığı görülmektedir. Bu durum, ülke için henüz gelişmekte olan bir uzmanlık alanı olduğu kanısı oluşturmaktadır.



Şekil 79: Savunma Sıcaklık Haritası

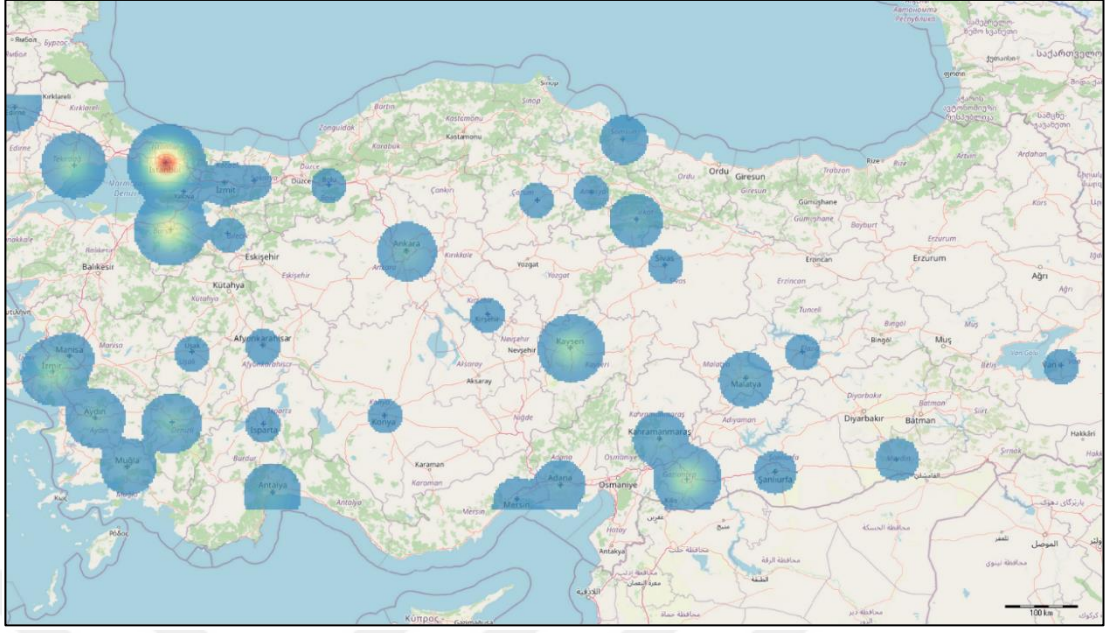
Şekil 79’daki Savunma niteliği sıcaklık haritası, savunma sanayi şirketleri ve girişimlerinin bulunduğu şehirlerde iş ilanlarının çıktığını ve yeni bir alan olduğunu göstermektedir.

Şekil 80’de Sosyal Medya öbeğinin sıcaklık haritası, büyükşehirlerde istihdam edilmek istenen bir nitelik olduğunu göstermektedir. Ancak İstanbul’daki yoğunluk, İstanbul firmalarının bu alanda çok ciddi işgücü ihtiyaçları olduğunu anlatmaktadır.

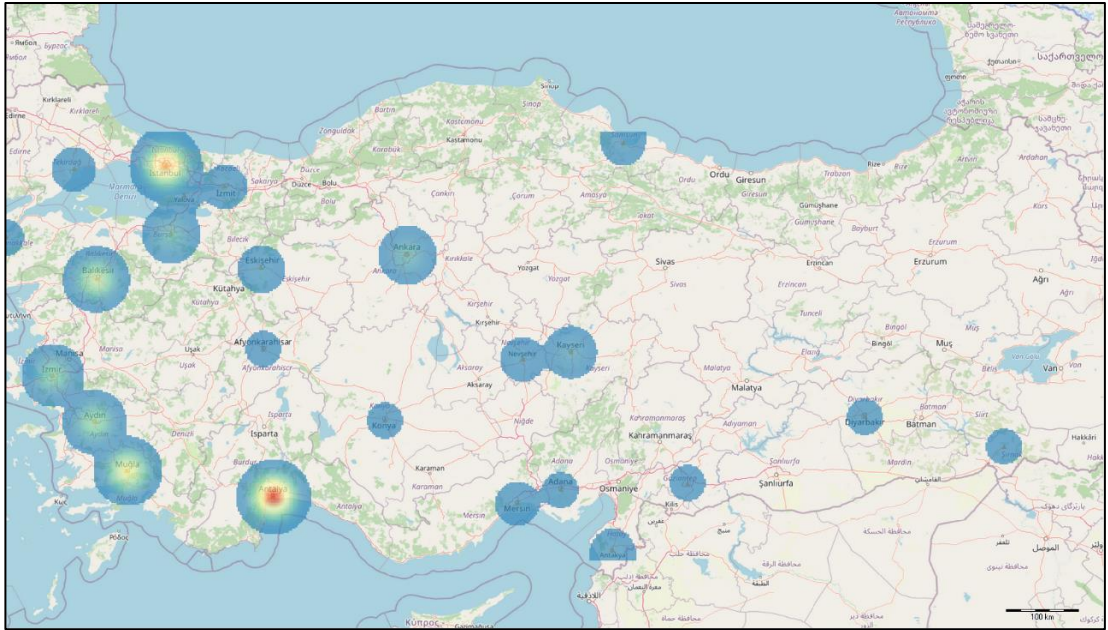


Şekil 80: Sosyal Medya Sıcaklık Haritası

Şekil 81’de Tekstil niteliğinin sıcaklık haritası verilmiştir. Tekstil yönünden lider şehirlerden biri olarak görülen Denizli ve çevresi sıcaklık oluşturmasına rağmen Gaziantep ve Kayseri gibi bu sektörün yeni çekim merkezleri olduğu görülmüştür. Ayrıca İstanbul ve Bursa tekstil sektörü konusunda işgücü talep-arz yoğunluğunun en çok yaşandığı şehirler olarak bulunmaktadır.

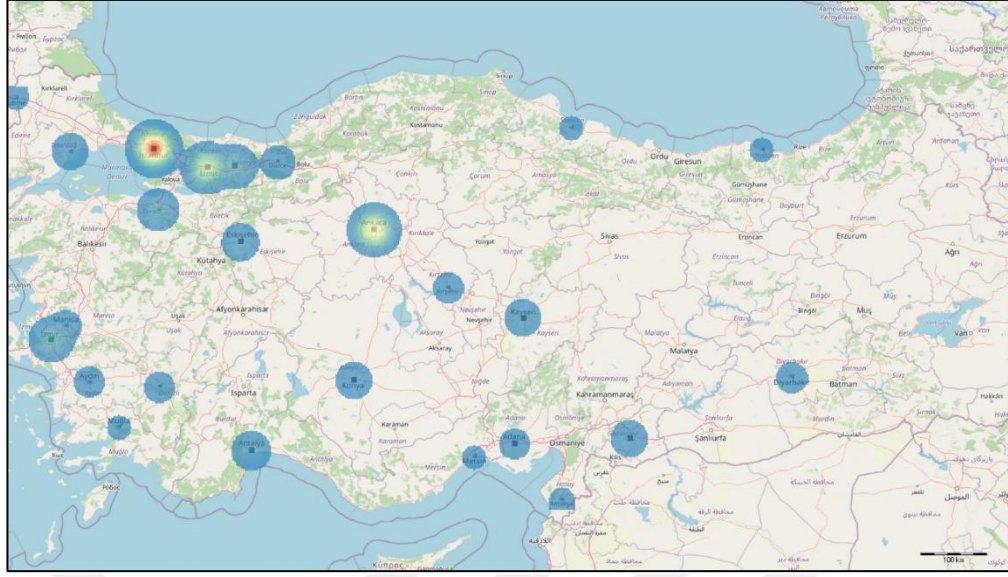


Şekil 81: Tekstil Sıcaklık Haritası



Şekil 82: Turizm Sıcaklık Haritası

Şekil 82'deki sıcaklık haritasında Turizm teriminin yayılımı verilmiştir. Antalya ilinin bu konuda en çok iş ilanı veren şehir olması beklenen bir durumdur. Büyük turistik merkezlerde turizm niteliği arayışı olması normaldir. Fakat Doğu Karadeniz Bölgesi turistik faaliyetlerin yoğun olduğu bir bölge olarak ve Erzurum gibi önemli bir kış turizm merkezinin hiç ısı oluşturmaması kritik bir bulgu olarak görülebilir.



Şekil 83: Yazılım ve Yazılım Geliştirme Sıcaklık Haritası

Şekil 83'te Yazılım ve Yazılım Geliştirme terimleri birleştirilerek oluşturulmuş sıcaklık haritası gösterilmektedir. Teknolojinin ilerlediği günümüzde yazılım üzerine olan bu iki niteliğin oluşturduğu yoğunluk İstanbul merkezli olduğu anlaşılmaktadır. Bu yoğunluğu Ankara takip ederken, diğer şehirlerde iş ilanlarının azaldığı görülmektedir.

Şekil 83'teki bulguların, Şekil 75'teki Otomasyon Sıcaklık Haritası ile bulunduğu şehirler itibariyle kısmen benzerlik gösterdiği görülmektedir. Bununla birlikte oluşan ısı miktarları ise Satın Alma gibi çok fazla değildir. Bu nedenle, yazılım ile ilgili niteliklerinin sanayileşme, nüfus gibi etkilerden olumlu etkilendiği ve yeni bir alan olduğu tahmini yapılabilir.

Sonuç olarak, iş ilanları ulusal beşeri sermayeyi temsil eden bir veri seti olarak devlet analitiği yaklaşımıyla değerlendirilmiştir. Bu bağlamda, Şekil 2'de gösterilen yetenek keşfetme analizleri ile iş ilanlarındaki nitelikler çıkarılmak istenmiştir.

Böylece, iş ilanlarına metin madenciliği uygulanarak iş ilanlarında öne çıkan nitelikler tespit edilmiştir. Niteliklerin, 12 aylık dönemde ve yatırım teşvik bölgelerine göre dağılımları incelenmiş ve CBS aracılığıyla haritaları çıkarılmıştır. Analizlerden elde edilen bulgular, kamu politikalarında karar desteği veren yorumlar getirmiştir. Bu konuda, örneğin işgücü piyasası, işsizlik, göç ve eğitim politikaları gibi kamu yönetimi konularında uygulanabilir bir KDS ortaya koyulmuş, istihdam üretme ve istihdamı yönetme gibi başlıklar için faydalı bilgiler çıkarılmıştır.

SONUÇ VE ÖNERİLER

Bu çalışmada, devlet analitiği çerçevesinde beşeri sermayenin analizi yapılmak istenmiştir. Sermaye kuramlarında ifade edilen işgücü kavramıyla da aynı anlama gelen beşeri sermaye analitiği, makro bir bakış açısıyla incelenmiştir. Bu doğrultuda, firmaların insan kaynakları analizinde kullandığı yetenek keşfetme yaklaşımı ile iş ilanları incelenmiştir. İnsan kaynakları analizlerinde büyük veri olarak değerlendirilebilen iş ilanlarına metin madenciliği ve coğrafi bilgi sistemi yöntemleri uygulanmıştır.

Araştırma için iş ilanları veri seti için, Kariyer.Net sitesinde yer alan 1 Ocak 2019 ve 13 Şubat 2020 tarihleri arasında yayınlanmış Türkçe iş ilanlarından internet kazıma yöntemiyle bir kayıt dosyası oluşturulmuştur. Bu veriler *yatırım teşvik bölgeleri* ve *koordinat bilgileriyle* zenginleştirilmiştir. Veri setinin hacminin ve çeşitliliğinin bu şekilde arttırılmasıyla, toplamda 48 bin 426 satırdan oluşan Türkçe iş ilanı verisi metin madenciliği ve coğrafi bilgi sistemlerinde analiz edilmek üzere hazırlanmıştır.

Çalışkan ve Akbıyık'ın iş ilanlarına yönelik araştırması ve Akbıyık'ın metin analizlerinin sosyal bilimlerdeki uygulanmasıyla ilgili kitabı, araştırmanın deneysel çalışmasının temel dayanağı olmuştur (Çalışkan ve Akbıyık 2018; Akbıyık, 2019). Akbıyık'ın kitabında açıkladığı *Wordstat* ve *GIS Viewer* yazılımlarıyla metin madenciliği ve coğrafi analizleri yapılmış ve böylece araştırmanın bulguları ortaya çıkarılmıştır.

Araştırmanın bulgularının verildiği frekans analizleri ve çapraz tablo analizinde, *iş niteliği, iş tecrübesi, yetenek, meslek veya sektör adları* gibi tür olarak farklı anlam içeren terimlerin iç içe geçtiği görülmüştür. Karmaşık iş ilanı yapısı, firma yöneticilerinin ve insan kaynakları uzmanlarının iş ilanlarında yapısal bir standart kullanmadıklarını göstermiştir. Bu durum, uzmanlar ve araştırmacılar için *insan kaynakları alanında bir veri sözlüğü* oluşturmanın önemli bir ihtiyaç olduğunu göstermiştir.

Araştırma bulgularında, homojen olarak dağılan genel niteliklerin yanı sıra, anomali gösteren nitelikler de tespit edilmiştir. En sık geçen nitelikler, kelime bulutları ile gösterilmiştir. *12 aylık dönemde ve yatırım teşvik bölgelerinde anomali gösteren niteliklerin 10 bin vakada geçme sıklıkları gösterilmiştir.* Sonra bu analiz modeli çapraz tablo analizi için yeniden geliştirilmiştir. Çapraz tablo analizinden çıkarılan niteliklerden coğrafi analizler için *sıcaklık haritaları* oluşturulmuştur.

Niteliklerin dönemsel ve yatırım teşvik bölgelerindeki dağılımları, harita üzerindeki ısı yayılımları incelenebilmiş ve diğer niteliklerle ilişkileri yorumlanmıştır. Böylece niteliklerin ne kadar arz-talep edildiği ve bu niteliğin yatırım teşvik bölgesine göre istihdam edilmesinin olumlu veya olumsuz etkilenip etkilenmediğine bakılmıştır. Haritadaki oluşan ısı değerleri ile niteliklerin yatırım teşvik bölgesi ölçeğinin yanında nüfus, sanayileşme veya coğrafi konum gibi başlıklarla ilgili bir ilişkisi olup olmayacağı konusunda fikirler vermiştir.

Analizin yorumları ve araştırmada edinilen fikirler, devlet analitiği çerçevesinde bazı sonuçlar ortaya çıkarmıştır. İş ilanlarının işgücü ve istihdam sağlamak için kullanılan bir araç olması nedeniyle, makro ölçekte **Türkiye'nin istihdam politikasının bir fotoğrafı** gösterilmiştir. Bu analizlerden, devlet analitiğinin *makro politik karar verme ve bu politikaların değerlendirilmesi hususunda stratejik bilgiler* ortaya çıkardığı anlaşılmıştır.

Böylece, **devlet analitiği çerçevesinde makro politikalarda kullanılmak üzere metin madenciliği ve coğrafi bilgi sistemleri gibi çeşitli teknolojilerden nasıl yararlanılabileceği gösterilmiştir.** Araştırmada, kamu politikalarında istihdam sağlama ve işsizlik gibi sorunların çözümü için faydalı bilgiler üreten bir yöntem sunulmuştur. Araştırma bulgularından, *işgücü niteliği ve piyasası, eğitim, göç, kalkınma politikaları, sanayileşme, şehirleşme, yatırım teşvikleri gibi stratejilerle ilişkili çıktılar* sunulmuştur. Dolayısıyla büyük verinin ve büyük veri olarak iş ilanlarının devlet analitiğinde nasıl kullanılabileceğiyle ilgili önemli bir yol haritası oluşturmuştur.

Bu araştırmanın sonucunda, devlet analitiği yaklaşımında beşeri bilimlere ilgilendiren farklı disiplinlerden ve tekniklerden yararlanılabileceği öğrenilmiştir. *Bu yönüyle, Türkçe literatürün yanı sıra uluslararası literatürün genellikle çalışma alanlarından da ayrılan ampirik bir araştırma haline gelmiştir.* Dolayısıyla, bu çalışma başta *beşeri sermaye* ölçeğinde devlet analitiği çalışmalarının gelecekte geliştirilmesine kılavuz niteliğindedir.

Her araştırmanın kendine özgün teorisinde ve yönteminde kısıtları mevcuttur. Teknolojinin kullanıldığı her araştırma yönteminde olduğu gibi bu çalışmada da çeşitli yazılım ve donanım imkanlarının getirdiği kısıtlar olmuştur. Araştırmacı tarafından tez çalışmasının teorik ve metodolojik kısıtları belirlenmiştir. Bu kısıtların yarattığı zorluklara yönelik, gelecek çalışmalar için bazı öneriler paylaşılmıştır.

Araştırmada analizler için veri seti hazırlanırken ve bu veriler analiz edilirken çeşitli zorlukların farkına varılmıştır. Bu zorluklardan biri, sosyal bilimlerdeki araştırmalarda Türkçe metin madenciliği teknikleri üzerine yapılan bilimsel çalışmaların ve metin madenciliği tekniklerini kullanmak üzere geliştirilmiş yazılımlarının donanımsal ve özellikle yazılımsal yeterliliğiyle ilgilidir.

Metin madenciliğinde NLP yöntemlerinden biri olan Lemmatization, Akbıyık (2019)'a göre *anlamsal kök bulma*, kelimelerin türleri belirlenmesinde kullanılan yazılım algoritmalarını içerir. Niteliklerin belirlenmesini ve analizleri kolaylaştırıcı bir teknik olarak, Türkçede bu metodun dilbilimsel zorlukları vardır. Bu açıdan, Türkçe metin madenciliğinin, yabancı dillerdeki kadar başarılı bir çözümü henüz bulunmamaktadır. Bu konuda, bilgisayar bilimleri alanındaki gelişmeler takip edilmelidir.

Analizler gerçekleştirilirken tespit edilen bir başka kısıt ise, büyük verinin boyutlarıyla ilgili olmuştur. Bu araştırmada, ulusal beşeri sermaye analitiğine yönelik yatırım teşvik bölgeleri gibi veri setinde zenginleştirmeler yapılmış ancak buna benzer büyük verinin hacmini ve çeşitliliğini arttırıcı uygulamalar analiz çeşitliliğini arttıracak ve analizleri geliştirecektir. Büyük veri boyutunun bir zorluk doğurduğu anlaşılan bu çalışmada, yatırım teşvik bölgeleri gibi analitiğin kendi kapsamına yönelik zenginleştirmelerin yorumlamada analiz çeşitliliğinde olumlu katkıları görülmüştür.

Bu tezde verilen, devlet analitiği çerçevesi ve henüz araştırılmamış alt modellerinin uygulanması, geliştirilmesi ve test edilmesi gelecek çalışma konuları arasındadır. Tezdeki devlet analitiği alt modelleri, farklı nitelikteki veriler ile geliştirilmeye ve test edilmeye açıktır. Devlet analitiği kavramında ifade edilen bu modeller, yönetim bilişim sistemlerinde yeni çalışılacak alanlar çıkarmıştır. Böylece gelecek çalışmalarda hem beşeri bilimlere hem bilişim teknolojilerine yönelik katkılar verilir.

Bu tezde, devlet analitiği çerçevesi ile bu çerçevenin alt çerçeveleri ve modelleri oluşturulmuş ancak yalnızca beşeri sermayeye yönelik yetenek çıkarımları gerçekleştirilmiştir. Dolayısıyla gelecek çalışmalarda beşeri sermayeye yönelik birçok gelecek çalışma konuları belirlenebilir. Örneğin, *yetenek kazandırma* olarak adlandırılmış analitik çerçeve modellenerek beşeri sermaye analitiğinin çerçevede gösterilmiş bir sonraki adımı test edilebilir. Daha somut bir ifadeyle, iş ilanlarından ve özgeçmişlerden

keşfedilen yetenekler ile yapay zeka (bilgisayar öğrenmesi) teknikleri kullanılarak eşleşmeler yapılabilir ve böylece ulusal bazda niteliklerin verimli bir şekilde kullanılabilmesi için yeni KDS önerileri oluşturulabilir. Bu düşünceye paralel olarak akademik ölçekte üniversiteler de öğrencilerinin niteliklerini bu konularda ampirik çalışmalarda kullanabilir.

İnsan kaynaklarına yönelik analitik çalışmalar mikro ölçekte firmaların araştırma ve geliştirme yapabilecekleri çok geniş ve yeni bir alandır. Bu çalışmada da görüldüğü üzere, makro ölçekte beşeri sermaye boyutuyla yapılacak çalışmalar kamu yönetiminde kullanılabilir bir çok keşfedilecek yeni analitikler ve araştırma konuları ortaya çıkarmıştır. Dolayısıyla, geçmiş çalışmalar özellikle ABD ölçeğinde siyaset bilimi ve sosyal medya analitikleri üzerinde yoğunlaşsa da bunun devlet analitiği kavramında bu tezde sermaye kuramıyla geliştirilen çerçevede ulusal sosyal sermaye analitiğine karşılık geldiği anlaşılmaktadır. Buradan hareketle, Türkiye özelinde araştırmacılar sosyal medya ve siyasi analitik çerçevelerini oluşturmak veya geliştirmek üzere bu tez çalışmasını rehber alabilirler.

Bu araştırma konusunda, beşeri sermaye analitiğinin kamu yönetiminde karar desteği oluşturması için iş ilanları verileri yatırım teşvik bölgeleri ile zenginleştirilmiş ve böylece devletin yatırım politikalarıyla ilişkileri keşfedilmiştir. Gelecek çalışmalarda, büyük veri (iş ilanları) yatırım teşvik bölgeleri gibi örneğin eğitim politikalarına yönelik eğitim teşvikleriyle ilişkilendirilebilir ve bu sayede kamu yönetiminde eğitim politikalarında karar desteği üretebilen modeller geliştirilebilir. Bunun gibi birçok örnek gelecek araştırma konusu, büyük verinin hacminin ve çeşitliliğinin (yatırım teşvik bölgelerinin eklenmesi gibi) arttırılması ile sağlanabilir.

Bu tez, devlet analitiğinin anlaşılması ve çerçevesinin ortaya koyulması ile birlikte gelecek araştırma konularına akademik, özel sektör ve kamu yönetimi ölçeklerinde çeşitli yenilikçi fikirler vermektedir. Araştırmanın uygulamasında beşeri sermayeye yönelik çerçevenin ulusal bazda ele alınması ve ortaya karar destek sistemi modeli önerisi mikro ölçekteki beşeri sermayenin nasıl kullanılabildiğiyle doğrudan ilişkili olmuştur. Bu yüksek lisans tezi KDS açısından akademide, özel sektörde ve kamu yönetiminde beşeri sermayenin alt modelleri ve diğer sermaye türlerine yönelik analitik modeller olmak üzere büyük verinin gelecekte araştırılabileceği büyük resmi göstermiştir.

KAYNAKÇA

- 1402 Karar Sayılı Yatırımlarda Devlet Yardımları Hakkında Kararda Değişiklik Yapılmasına Dair Karar (2019). Resmi Gazete (Sayı: 30855). Erişim adresi: <https://www.resmigazete.gov.tr/eskiler/2019/08/20190807-2.pdf> (Erişim Tarihi: 1/06/2020)
- Advanced Technology Investment Company. (2020). *About Us* Erişim adresi: <https://www.linkedin.com/company/advanced-technology-investment-company> (Erişim Tarihi: 1/06/2020)
- Akbıyık, A. (2019). *Sosyal Bilimlerde Metin Madenciliği Uygulamaları Wordstat Uygulamaları*. Sakarya: Sakarya Yayıncılık.
- Alayoğlu, N. (2010). İnsan kaynakları yönetiminde yeni dönem: yetenek yönetimi. *Gazi Üniversitesi Ticaret ve Turizm Eğitim Fakültesi Dergisi*, (1), 68-97.
- Albayrak, N. ve Abdioğlu, Z. (2020). Beşeri Sermaye Tahmini: Gizli Değişken Yaklaşımı. *Yönetim Bilimleri Dergisi*, 18(36), 225-254.
- Australian Capital Territory Government. (2020). *ACT Government Analytical Laboratory* Erişim adresi: <https://advanced.jhu.edu/academics/centers/center-for-advanced-governmental-studies/> (Erişim Tarihi: 1/06/2020)
- Bocij, P., Greasley, A. ve Hickie, S. (2008). *Business information systems: Technology, development and management*. Pearson education.
- Bourdieu, P. (1986). Forms of capital. In J. Richardson. *Handbook of Theory and Research for the Sociology of Education* (New York, Greenwood), 241-258.
- Codd, E. F. (2002). A relational model of data for large shared data banks. In *Software pioneers* (ss. 263-294). Springer, Berlin, Heidelberg.
- Çalışkan, K. ve Akbıyık A. (2018). Türkiye İşgücü Piyasalarında İş Analisti Ve Sistem Analist Kavramlarının Algılanması: İş İlanlarına Yönelik Bir İçerik Analizi. *Sosyoekonomik Boyutlarıyla İnovasyon*, 485.
- Davenport, T. H. ve Jarvenpaa, S. L. (2008). *Strategic use of analytics in government*. Washington, DC: IBM Center for the Business of Government.
- Dawsen, C. J. (2011). *Geographic Information Systems*. Nova Science Publishers.
- Demirel, D., Das, R. ve Hanbay, D. (2019). Büyük veri üzerine perspektif bir bakış. In *2019 International Artificial Intelligence and Data Processing Symposium (IDAP)* (ss. 1-9). IEEE.
- Denge Müşavirlik. (2020). *İllere Göre Türkiye Yatırım Teşvik Haritası* Erişim adresi: <https://dengemusavirlik.com/illere-gore-turkiye-yatirim-tesvik-haritasi> (Erişim Tarihi: 1/06/2020)

- Dhar, V. (2013). Data science and prediction. *Communications of the ACM*, 56(12), 64-73.
- Eldrandaly, K. A. E. M. S. (2010). Spatial decision making: an intelligent GIS-based decision analysis approach. *Germany*.
- Emiratisation. (2020). Erişim adresi: <https://en.wikipedia.org/wiki/Emiratisation> (Erişim Tarihi: 1/06/2020)
- Fayyad, U., Piatetsky-Shapiro, G. ve Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI magazine*, 17(3), 37-37.
- Fedra, K. ve Reitsma, R. F. (1990). Decision support and geographical information systems. In *Geographical information systems for urban and regional planning* (ss. 177-188). Springer, Dordrecht.
- Friedman, J. H. (1998). Data Mining and Statistics: What's the connection?. *Computing Science and Statistics*, 29(1), 3-9.
- Ghosh, S., Roy, S. ve Bandyopadhyay, S. K. (2012). A tutorial review on Text Mining Algorithms. *International Journal of Advanced Research in Computer and Communication Engineering*, 1(4), 7.
- Gonzalez, G., Karoly, L. A., Constant, L., Goldman, C. A. ve Salem, H. (2008). *Facing human capital challenges of the 21st century: Education and labor market initiatives in Lebanon, Oman, Qatar, and the United Arab Emirates* (Vol. 786). Rand Corporation.
- Grimes, S. (2007). A brief history of text analytics. *BeyeNetwork*. Erişim Adresi: <http://www.b-eye-network.com/view/6311> (Erişim Tarihi: 20/10/2015)
- Healy, T. ve Côté, S. (2001). *The Well-Being of Nations: The Role of Human and Social Capital. Education and Skills*. Organisation for Economic Cooperation and Development, 2 rue Andre Pascal, F-75775 Paris Cedex 16, France.
- Hearst, M. (2003, 17 Ekim), What is text mining, *SIMS, UC Berkeley*. Erişim Adresi: <https://people.ischool.berkeley.edu/~hearst/text-mining.html> (Erişim Tarihi: 1/06/2020)
- Huselid, M. A. (2018). The science and practice of workforce analytics: Introduction to the HRM special issue. *Human Resource Management*, 57(3), 679-684.
- Intellspot. (2020). *Text Mining Algorithms List: Text Classification Categorization Clustering* Erişim adresi: <http://intellspot.com/text-mining-algorithms/> (Erişim Tarihi: 1/06/2020)
- Isson, J. P. ve Harriott, J. S. (2013). Who Cares about Data?. *Win with Advanced Business Analytics*, 77.

- Isson, J. P. ve Harriott, J. S. (2016). *People Analytics in the Era of Big Data: Changing the Way You Attract, Acquire, Develop, and Retain Talent*. John Wiley & Sons.
- Johns Hopkins University. (2020). *Center for Advanced Governmental Studies / Advanced Academic Programs | Johns Hopkins University* Erişim adresi: <https://advanced.jhu.edu/academics/centers/center-for-advanced-governmental-studies/> (Erişim Tarihi: 1/06/2020)
- Kelly, P. (2016). Predicting the future: climate change and public health. *ACT Population Health Bulletin, Volume 6, Issue 2*. 28-30.
- Khattab, K. ve Tyagi, R. K. (2017). Preparing for Analytics: The Dubai Government Excellence Program. In *Big Data and Analytics Applications in Government* (pp. 57-71). Auerbach Publications.
- King Jr, M. L. (2018). The Power of UDA in Human Capital Management. *Unstructured Data Analytics*, 191.
- Kitchin, R., ve McArdle, G. (2016). What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets. *Big Data & Society*, 3(1), 2053951716631130.
- Le Roux, Y. (2012). Privacy and Data Analytics. *CA Technology Exchange: Insights from CA Technologies*, 3(3), 17-23.
- Lohr, S. (2013). The origins of 'Big Data': An etymological detective story. *The New York Times*. Erişim Adresi: <https://bits.blogs.nytimes.com/2013/02/01/the-origins-of-big-data-an-etymological-detective-story/> (Erişim Tarihi: 1/06/2020)
- Mashey, J. R. (1997, Ekim). Big data... And the next wave of infrastress. In *Computer Science Division Seminar, University of California, Berkeley*.
- Microsoft. (2020). *Description of the database normalization basics* Erişim adresi: <https://docs.microsoft.com/en-us/office/troubleshoot/access/database-normalization-description> (Erişim Tarihi: 1/06/2020)
- Mitchell, R. (2018). *Web scraping with Python: Collecting more data from the modern web*. " O'Reilly Media, Inc."
- Mubadala Investment Company. (2020). Erişim adresi: https://en.wikipedia.org/wiki/Mubadala_Investment_Company (Erişim Tarihi: 1/06/2020)
- Mubadala. (2020). *Advanced Technology Investment Company (ATIC) to become Mubadala Technology* Erişim adresi: <https://www.mubadala.com/en/news/advanced-technology-investment-company-atic-become-mubadala-technology> (Erişim Tarihi: 1/06/2020)
- Navitism. (2020). Erişim adresi: [https://en.wikipedia.org/wiki/Nativism_\(politics\)](https://en.wikipedia.org/wiki/Nativism_(politics)) (Erişim Tarihi: 1/06/2020)

- Oji, K. (2018). *Business models and talent management practices of Japanese multinational companies: 1970-2018* (Doctoral dissertation, Massachusetts Institute of Technology).
- Oxford Online Dictionary. (2020). *Meaning of Big Data by Lexico*. Retrieved from https://www.lexico.com/definition/big_data (Eriřim Tarihi: 1/06/2020)
- Özcan, B (2011). *Sosyal Sermaye ve Ekonomik Kalkınma*, (Doktora Tezi). İstanbul Üniversitesi/Sosyal Bilimler Enstitüsü, İstanbul.
- Öztürk, N. (2006). Ekonomide devletin deęiřen rolü. *Amme İdaresi Dergisi*, 39(1), 17-38.
- Provalis Research. (2020). *Geospatial Intelligence meets Text Analytics - QDA Miner & WordStat* Eriřim adresi: <https://provalisresearch.com/solutions-2/applications/gis-text-analytics/> (Eriřim Tarihi: 1/06/2020)
- Richards, G. (Ed.). (2017). *Big Data and Analytics Applications in Government: Current Practices and Future Opportunities*. CRC Press.
- Richardson, L. (2014). Beautiful Soup <https://www.crummy.com/software>. (Eriřim Tarihi: 1/06/2020)
- Seyidođlu, H. (2002). Ekonomik Terimler Ansiklopedik Sözlük, 3. Baskı, İstanbul.
- The National Human Resources and Employment Authority. (2020). *About Us* Eriřim adresi: <https://www.linkedin.com/company/-tanmia-the-national-human-resources-and-employment-authority> (Eriřim Tarihi: 1/06/2020)
- Tran, B., Straka, P., Falster, M. O., Douglas, K. A., Britz, T., & Jorm, L. R. (2018). Overcoming the data drought: exploring general practice in Australia by network analysis of big data. *Medical Journal of Australia*, 209(2), 68-73.
- Zhu, X. (2013). GIS and spatial decision support. In *Geographic Information Systems* (ss. 1-33). Nova Science Publishers.

ÖZGEÇMİŞ

Burak Buldu, 1989 yılında Antalya’da doğmuştur. İlköğretimini Antalya’nın Kumluca ilçesindeki Mimar Sinan İlköğretim Okulu’nda bitirmesinden sonra, yine aynı ilçede lise öğretimini Kumluca Anadolu Lisesi’nde tamamlamıştır. 2017 yılında Sakarya Üniversitesi İşletme Fakültesi Yönetim Bilişim Sistemleri Bölümü’nde lisans eğitimini tamamlamıştır. Yüksek lisans eğitimini, 2018 yılında başladığı Sakarya Üniversitesi İşletme Enstitüsü Yönetim Bilişim Sistemleri programında yapmaktadır.

