

İstatistik anahtar kelimeli ilanların içeriklerinden öngörü elde etme

Öğrenci: Fatih Ekici

Araştırma Projesi Danışmanı: Doç.Dr Fatih Kızılaslan

Önsöz

Bu bitirme projesinin amacı istatistik bölümü mezunlarına ,öğrencilerine ve istatistik bölümünü tercih etmeyi düşünen adaylar için özel sektörün bekentileri hakkında bilgi verebilmektir. Pek tabi veri analitiği alanında kariyer yapmayı düşünen diğer mühendislik veya fen bölümü öğrencileri içinde faydalı olacağını düşünmekteyim.

Analiz metnini okumadan önce analize ilişkin şahsi eleştirilerimi paylaşmak isterim. Analiz edilecek iş ilanları resmin büyük parçasını kaybetmemek adına pozisyon bazlı olarak değil sadece istatistik keyword'ünün bulunması üzerinden oluşturuldu. Veri üzerinde daha sonrasında istatistik bölümü öğrencilerinin yeteneklerine çok uzak olan veya anlamsız olan iş ilanları elle temizlendi. Bununla beraber veriseti tek bir kaynak(kariyer.net) vasıtasiyla üretilmiştir. Bu durumlar sebebiyle analizde bahsedilen yeteneklere ait frekansların sektörel bazda gerceği tam anlamıyla yansımadığı söylenebilir. Yine de okuyucuya özel sektör ile alakalı yeni bir bakış açısı kazandıracaktır.

Analiz'in teknik tarafında veri kazıma ve model oluşturma aşamaları Python üzerinden yapıldı. Verinin temizlenmesi ve raporun oluşturulması aşamasında ise R dili tercih edildi. El ile yapılması gereken işlemler ve harita grafiği ise Excel üzerinden oluşturulmuştur.

Analizin son aşaması olan kümeleme analizinde çok daha başarılı modellemeler ve yöntemler tercih edilebilir. Ayrıca İş ilanlarında aranan özellikler için NLP yöntemleri örnek olarak Phrase extraction tercih edilerek bilinmeyen çok daha fazla bilgi aydınlatıcı kavuşturulabilir.

Bu projeyi destekleyen ve veri sponsoru olan **Kariyer.net'e**, projeye olan yakışıkları ve geri dönüşleri için Proje danışmanım olan **Doç.Dr. Fatih Kızılaslan'a**, Projenin kümeleme ayağının gerçekleşmesinde inkar edilemez katkısından ötürü bölüm arkadaşım olan **Harun Kürevi**'ye ve teşekkür ederim. Özel olarak **Prof.Dr Esra Akdeniz'e** eğitim hayatında bana olan destekleri ve yön göstericiliği sebebiyle sonsuz teşekkürlerimi belirtmek isterim.

Giriş

Kariyer.net üzerindeki ilanları python üzerinde selenium kütüphanesi ile çekildi. Daha sonrasında excel üzerinden el ile veri analitiği ile alakasız veya istatistik mezunlarının girme şansı olmayan/girmek istemeyeceği ilanları düşürüldü. Toplam veri sayısı 717 den 496 adet ilana düşürülmüştür. Bu işlemin yapılma sebebi +400 adet olan eşsiz pozisyon isimlerini daha kolay bir biçimde üst grplara alabilmek ve daha tutarlı sonuçlar elde etmektir.

Buna ek olarak **pozisyon kısa** adı altında ilanlar belirli başlı niteliklere göre gruplandırılmıştır.

1. veri analisti(iş zekası, veri analitiği, veri bilimci ,dijital reklamcılık sektör veya tittlelarına aitilanlar)
2. finans analisti(finans sektörü hakkında bilgi gerektiren analist pozisyonları)
3. develooper(herhangi bir sektörde yazılım-it-bilgi işlem tarafından pozisyonlar

4. iş analisti(finans analistine benzer şekilde genel anlamda yapılan iş ile alakalı analist pozisyonuna sahip pozisyonlar)
5. insan kaynakları (çoğunlukla ik işleri olmak üzere az sayıda personel ile ilgilenen pozisyonlar)
6. ticaret uzmanı(çoğunlukla evrak işleri, ithalat, ihracat ve satış pozisyonlarında uzmanlık gösteren pozisyonlar)
7. Muhasebe

Bazı ilanlar pek çok pozisyon'a yakın davranışlarından dolayı en yakın seçenek tercih edilmiştir.

Tanımlayıcı bilgiler

Kaç farklı pozisyon vardır ve en çok tekrar edenler hangileridir?

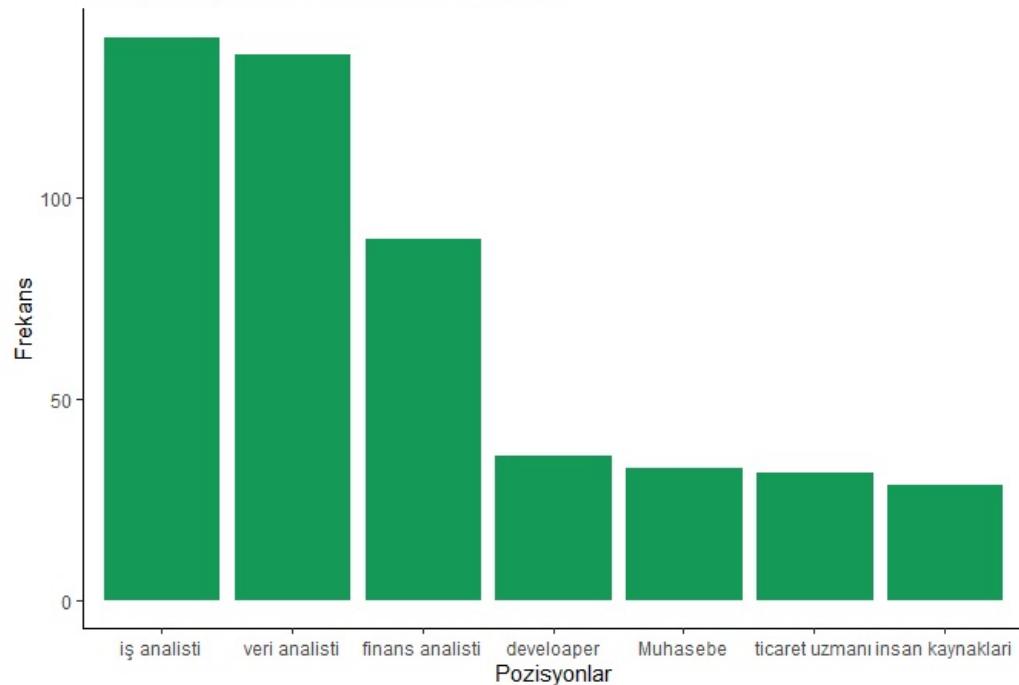
Toplam 394 adet essiz isimli pozisyon bulunmaktadır.

Pozisyon	Frekans
İş Analisti	14
İş Zekası Uzmanı	8
Satin Alma Uzmanı	8
Bütçe ve Raporlama Uzmanı	6
CRM Uzmanı	6
Analiz ve Raporlama Uzmanı	5
İnsan Kaynakları Uzmanı	5
Raporlama Uzmanı	5
Veri Analisti	5
E-Ticaret Uzmanı	4

Gördüğü gibi frekans sayıları benzer işlerin farklı sayılarla belirtilmesi sebebiyle azdır.Kısaltılmış ilanların frekans tablosu aşağıda verilmiştir.

Indirgenmiş pozisyon	Frekans
iş analisti	140
veri analisti	136
finans analisti	90
developoper	36
Muhasebe	33
ticaret uzmani	32
insan kaynakları	29

İndirgenmiş ilanların frekansları tablosu



Beklenildiği üzere istatistik anahtar kelimeli ilanlarda analyst pozisyonlarının ön sıralarda olduğunu, diğer meslek dallarında daha az ön planda olduğunu görmekteyiz. Finans sektörünün tek başına yüksek miktarda analyst barındırdığı görülebilmektedir.

Bu kısım indirmenin bir yan etkisi olarak net yorumlarda bulunulmasını engellemektedir. Bu sebeple sektör ve departmanları inceleyerek daha iyi bir öngörü elde edilmeye çalışılmıştır.

Departman dağılımı ve sektörel dağılım

Kariyer.net ilanlara ait sektör ve departman bilgilerini de sunmakta, bu sayede iş ilanlarının hangi sektör ve departmanda daha çok yoğunlaştığı aşağıda tablo ve grafiklerle gösterilmiştir.

Departman ait frekans tablosu

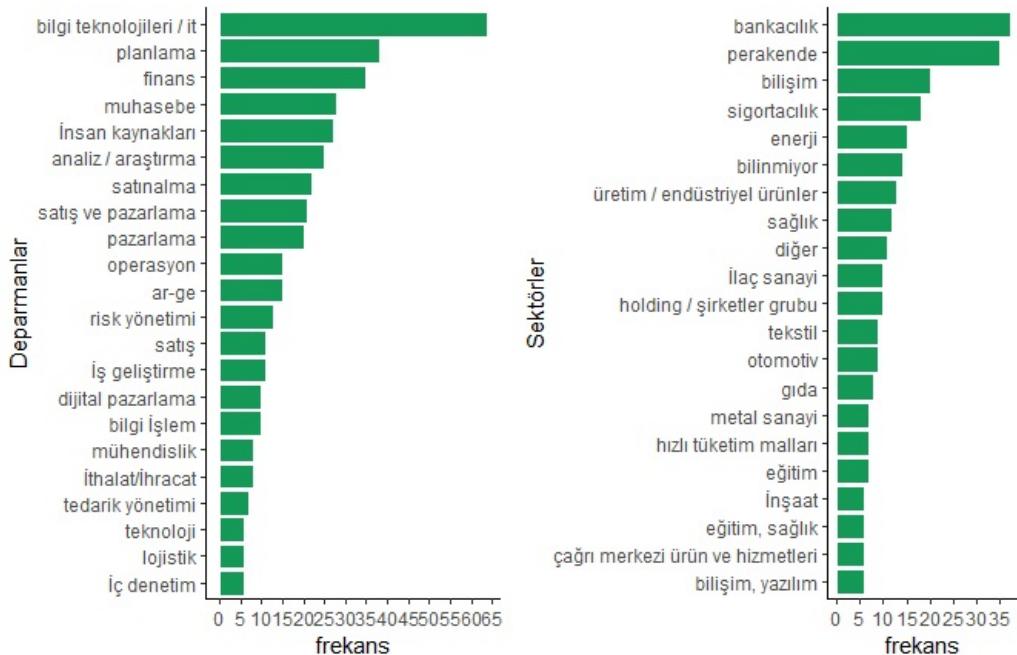
Departman:	frekans
bilgi teknolojileri / it	64
planlama	38
finans	35
muhasebe	28
İnsan kaynakları	27
analiz / araştırma	25
satınalma	22
satış ve pazarlama	21
pazarlama	20
ar-ge	15
operasyon	15
risk yönetimi	13
İş geliştirme	11

satış	11
bilgi İşlem	10
dijital pazarlama	10
İthalat/Ihracat	8
mühendislik	8
tedarik yönetimi	7
İç denetim	6
lojistik	6
teknoloji	6

Sektörlere ait frekans tablosu

Firma Sektörü:	frekans
bankacılık	37
perakende	35
bilişim	20
sigortacılık	18
enerji	15
NA	14
Üretim / endüstriyel ürünler	13
sağlık	12
diğer	11
holding / şirketler grubu	10
İlaç sanayi	10
otomotiv	9
tekstil	9
gıda	8
eğitim	7
hızlı tüketim malları	7
metal sanayi	7
bilişim, yazılım	6
çağrı merkezi ürün ve hizmetleri	6
eğitim, sağlık	6
İnşaat	6

Departman ve sektörlerde göre istihdam istatistikleri



Bilgi teknolojileri/it, planlama ve finans alanı, en çok istihdam sağlayan departmanlardır .Sektörel anlamda ise bankacılık,perakende ve bilişim sektörü istatistik mezunlarına en çok istihdam sağlayan 3 sektördür. Bu sonuç, istatistik mezunlarının gün geçtikçe daha fazla bilgisayar ve yazılım teknolojileri ile iç içe çalıştığını göstermektedir.

Sektörel anlamda en çok istihdamın finans/bankacılık sektörü olduğu gözden kaçmamalıdır.

İş ilanlarının konumları

İş ilanlarının konumlarında İstanbul'un Avrupa ve Anadolu yakaları 2 ayrı yer sayılırken kalan illerde herhangi bir değişiklik yapılmayacaktır.

Ayrıca bazı iş ilanlarında birden fazla konum olması, her bir bölgeyi ekstra sayılabileceği anlamına gelmektedir.Bu sebeple ilan sayısından daha fazla sayıda vilayet bulunmuş olmaktadır.

il	frekans
adana	5
afyon	1
amasya	1
ankara	25
antalya	3
aydın	1
batman	1
bolu	1
bursa	13
çanakkale	1
denizli	2

erzurum

1

eskişehir

5

hatay

1

isparta

1

istanbul

388

izmir

26

kahramanmaraş

1

karabük

1

kayseri

5

kocaeli

21

konya

2

manisa

3

mersin

1

nevşehir

1

sakarya

1

sivas

2

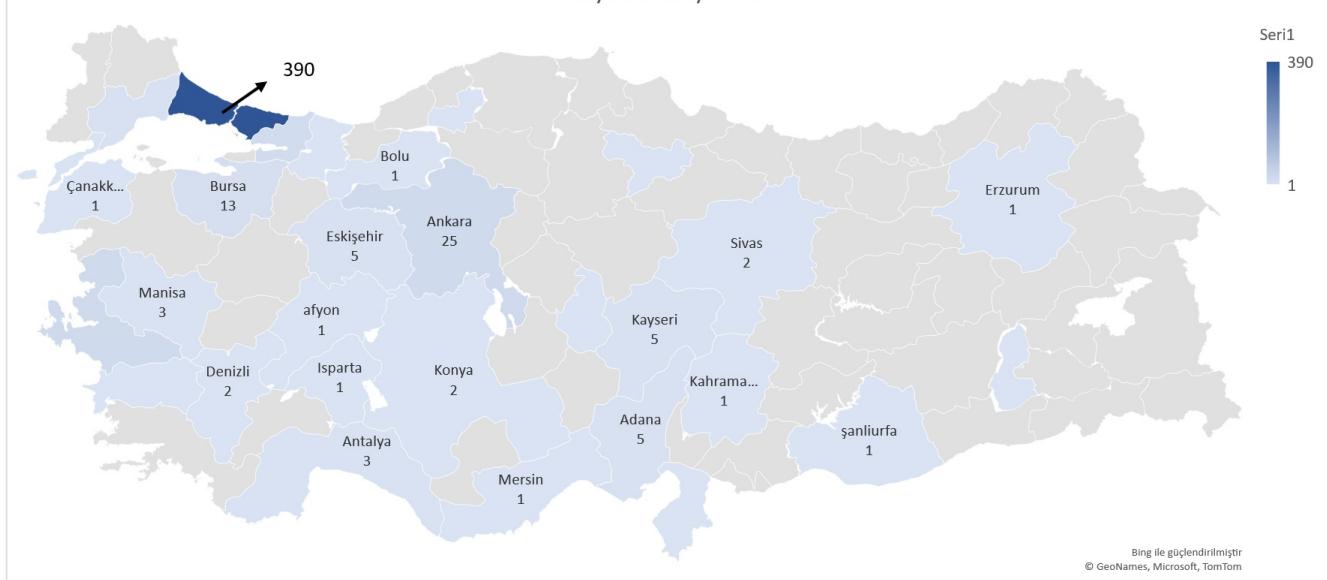
şanlıurfa

1

tekirdağ

5

Toplam çalışılabilir konum sayısı:522
kaynak:Kariyer.net



iş ilanlarının çoğunluğu incelendiğinde ağırlıklı olarak İstanbul'un iş havuzunun çok yüksek bir kısmını barındırdığı görülebilir. Finans,it,arge sektörlerinin çoğunlukla İstanbul'da yoğunlaşması bu sonucun nihai sebebidir.

il

frekans

Bununla beraber Avrupa Yakasının iş imkanı sayısı, asya yakasına göre ~1,6 kat daha fazladır. Avrupa yakasında ikamet etmek iş bulma konusunda kolaylık sağlama ihtimali kuvvetlidir.

Askerlik şartları

Askerlik durum bilgisinde yapıldı, muaf, tecilli, yapılmadı şeklinde 4 farklı kategori bulunmakta. **Yapıldı** ve **muaf** kesinlikle askere gitmeyen bir kişiyi ifade ettiği için bu 2 grup yapıldı olarak kodlanmıştır. **tecilli** ve **yapılmadı** ise beraber değerlendirilmiştir.

```
## [1] "Yapıldı"
## [3] "Yapıldı, Muaf, Tecilli"
## [5] "Muaf"
## [7] "Yapıldı, Tecilli"
## [9] "Yapılmadı"
## [11] "Yapıldı, Muaf, Yapılmadı"
## [13] "Yapıldı, Muaf, Yapılmadı, Tecilli"
```

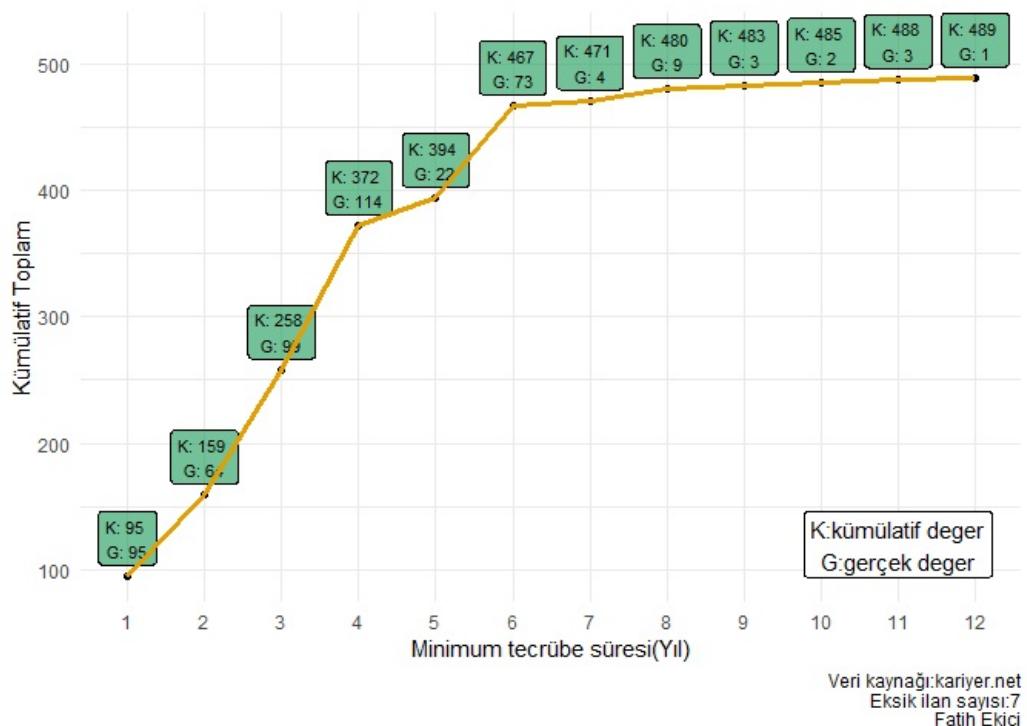
NA
 "Yapıldı, Muaf"
 "Tecilli, Muaf, Yapıldı"
 "Muaf, Yapıldı, Tecilli"
 "Muaf, Yapıldı"
 "Yapıldı, Tecilli, Muaf"

Askerlik Durumu:	freq
tecilli-yapılmadı	242
yapıldı-muaf	35
NA	219

Veride askerlik bilgisine dair büyük bir kısmda herhangi bir bilgi girişi yapılmamıştır. 219 sayıda ilan hakkında bilgi yoktur. Yinede yeterli sayıda ilan askerlikle alakalı bilgi vermektedir. Bilinenin aksine ilanların büyük bir kısmında işverenlerin askerliğin makul bir tecil süresini kabul edilebilir bulduğu görülmektedir.

Tecrübe

Pozisyonlar için gerekli olan tecrübe seviyesi aşağıda belirtilmiştir.

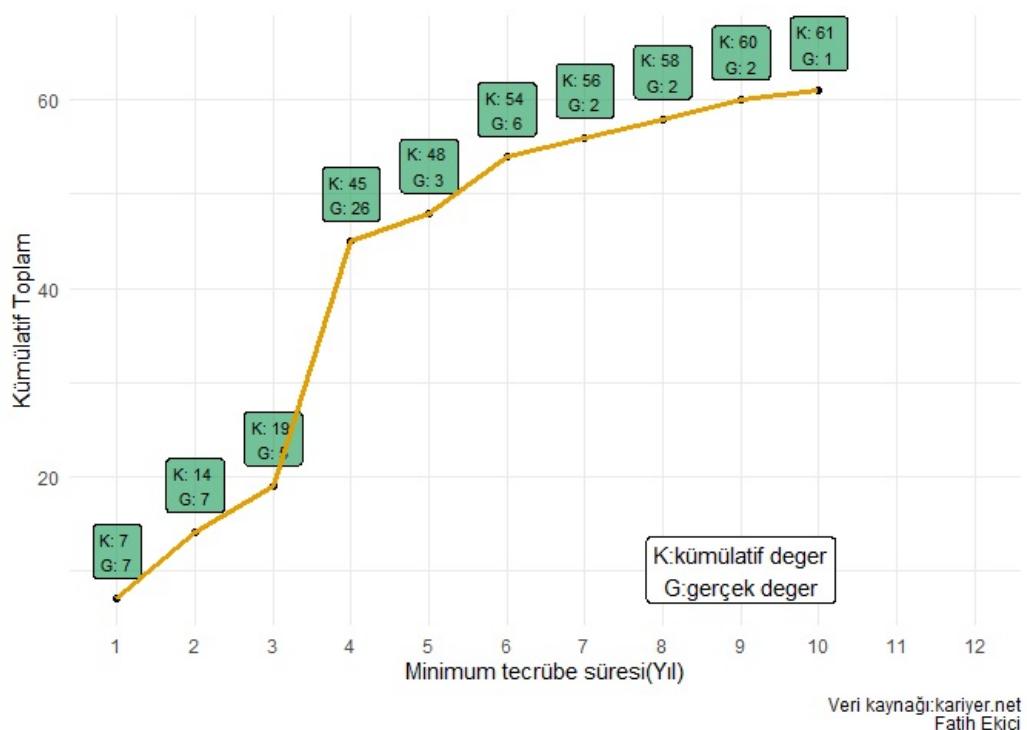


En çok aranan seviyenin mid level pozisyonlar olduğunu görülmüştür. 398 kişi yani 5 yıldan fazla tecrübeye sahip birisi ilanlarının ~%80'nine başvuruda bulunabilmektedir.

Maksimum tecrübe istatistiği yeterli veri olmaması sebebiyle analize dahil edilmemiştir.

Yönetici pozisyonlar

İçerisinde yönetici,müdür,manager kelimelerinin geçtiği ilanlardan yönetici sınıfı için gerekli minimum çalışma süresinin grafiği aşağıda verilmiştir.



İçerisinde **yönetici,müdür,manager** kelimelerinin geçtiği ilanlarının minimum tecrübe süresini incelendiğinde 5.yıldan itibaren ilan sayısında hatırlı sayılar artış görülmüştür. 5 yıllık tecrübe yönetici pozisyonu sayısındaki ivme dikkate alındığında Türkiyedeki yönetici pozisyonları aranan bir özellik olduğu görülmüştür. # İlan içeriklerinin analiz edilmesi

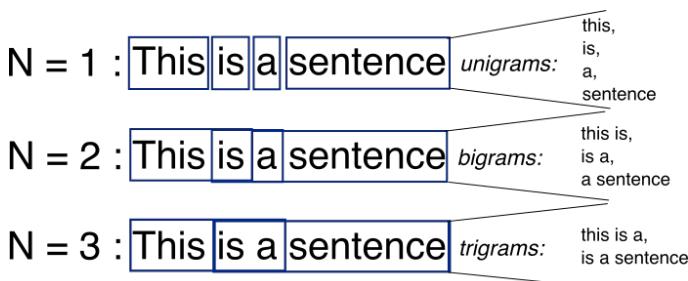
Bu aşamada ilan içeriklerindeki yeteneklerin bulunması ve benzer iş ilanlarının kümelenerek bu ilanlara

ait becerilerde ne tür farklılıklar/benzerlikler olduğu bulunacak ve bulgular analiz edilecektir.

Rapor kapsamında çoğunlukla hard skill olarak geçen teknik yetenekler baz alınmıştır.Hard skills(Teknik yetenekler), Teknik bilgi gerektiren bir iş için adaydan öncelikli olarak beklenen iş bilgisidir. Soft skills(ince yetenekler) ise bir adayın sosyal kabiliyetleri ile alakalı vasıflardır. ikna kabiliyeti, takım çalışmasına uyumluluk, sunum becerisi, değişikliğe adapte olabilme gibi pek çok kabiliyet vardır. İnce yetenekler, adayların iş yerinde yükselebilme şansını yükseltirken aynı zamanda işten alınabilecek verimi artırmasını sağlar bu sayede teknik becerilerinin yanı sıra diğer adaylardan sıyrılabılır. İnce yeteneklerin belirlenmesi bu çalışmanın bir odağı değildir. Çünkü ince yetenekler sektörden bağımsız olarak her zaman geçerlidir. Bununla beraber bir adayın ince yeteneklerini objektif olarak belirlemek oldukça uzun bir süreçtir(yönetici adayları için Assessment centre hizmeti alınması).

Teknik yetenekler

Teknik yetenekler ilanlarda geçmesi en muhtemel beceriler göz önüne alınarak hazırlanmıştır. Burada asıl önemsenen nokta teknik beceriler ve veri analitiğinde kullanılan yazılımlar ve programlar olmuştur. / Aşağıdaki tabloda teknik yetenekler tek tablo halinde veriden regular expression sorguları yardımı ile 1 ve 2'li n-grams yöntemi ile çekılmıştır. N-grams bir metnin istenilen sayıda art arda kelimelerin sıralanması ile oluşur.



Yetkinlikler	Frekans
raporlama	317
ingilizce	304
office	283
excel	181
bütçe&finans	139
sql	116
Sunum Yeteneği	90
bankacılık	85
erp	73
sap	67
insan kaynakları	33
e-ticaret	29
power_bi	26
ver_ambarı	20
r	19

SPSS	19
sas	18
python	17
agile	15
ehliyet	14
jira	14
git	13
etl	12
scrum	11
machine_learning	11
SSI_SSRS	10
ver_gorsellestirme	10
.net	8
kanban	8
c#	7
java	7
cloud servisleri	7
javascript	5
seo	5
iç denetim	5
html_css	4
devops	4
qlikview	3
tableau	3
c++	2
nodejs	2
odi	2
scala	1
docker	1

Tablo'daki niteliklerden bazıları kategorilere göre karşılaştırılmadan incelendiğinde temel becerilerden sonra en çok aranan yeteneğin excel ardından sql bilgisi olduğu fark edilmektedir.

Veri analitiği sektöründe pek çok açık kaynaklı ve yeni ücretli paket yazılımlar çıksa da pek çok firma

farklı gerekçelerle excel'i kullanmaya devam etmekte olduğu görülmüştür. Yaygın kanının aksine excel'e olan hakimiyetin pek çok sektörde aranan özellik olduğu görülmüştür.

Analiz süresince excel programına dair yetkinliğin derecesinin açık bir biçimde belirtilmediği gözden kaçmamalıdır (örnek: "**adaylardan iyi derecede excel bilgisi beklemektedir.**" şeklinde olan ilanlarda, iyi derecede excel bilgisi olan birinden hangi meziyetlere sahip olması beklendiği belli değildir.

SQL bir veritabanından(misalen excel tablosu) kuralara göre sorgular yazarak uygun verileri elde etmemizi sağlayan bir sorgu dilidir. Excel'in ardından veri analitiği veya benzer bir alanda çalışan birisi için en önemli özeliğin SQL dağıtımlarından birisinde iyi derecede bilgi sahibi olmak olduğu görülmüştür.

İş zekası yazılımları ve veri analitiğine özgün programlama dilleri

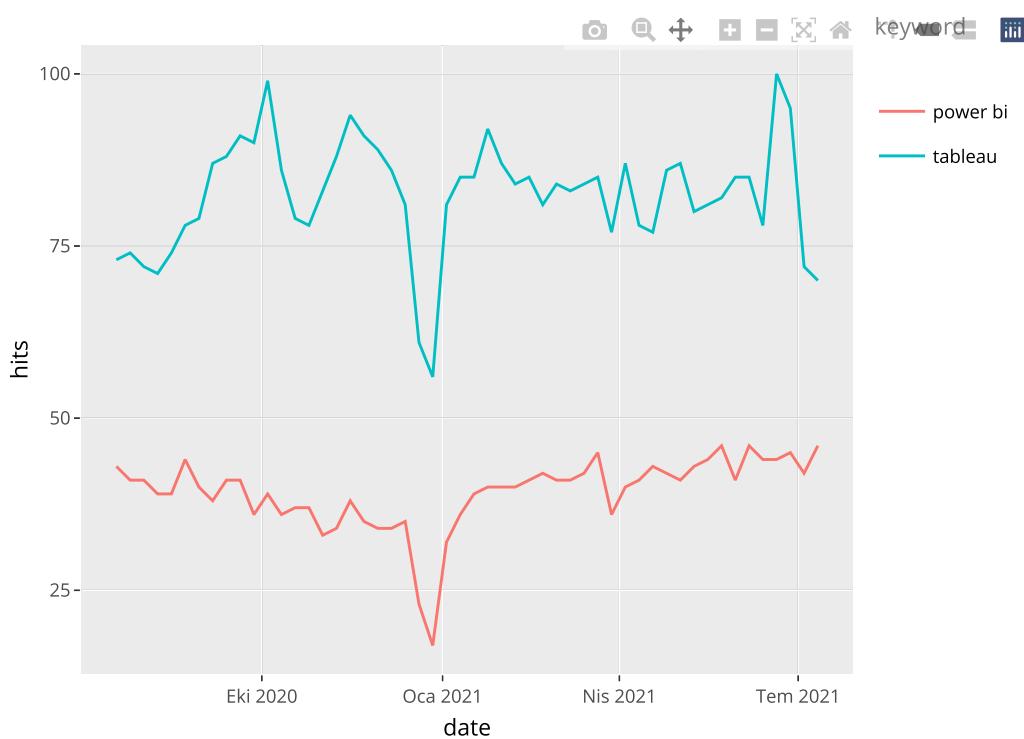
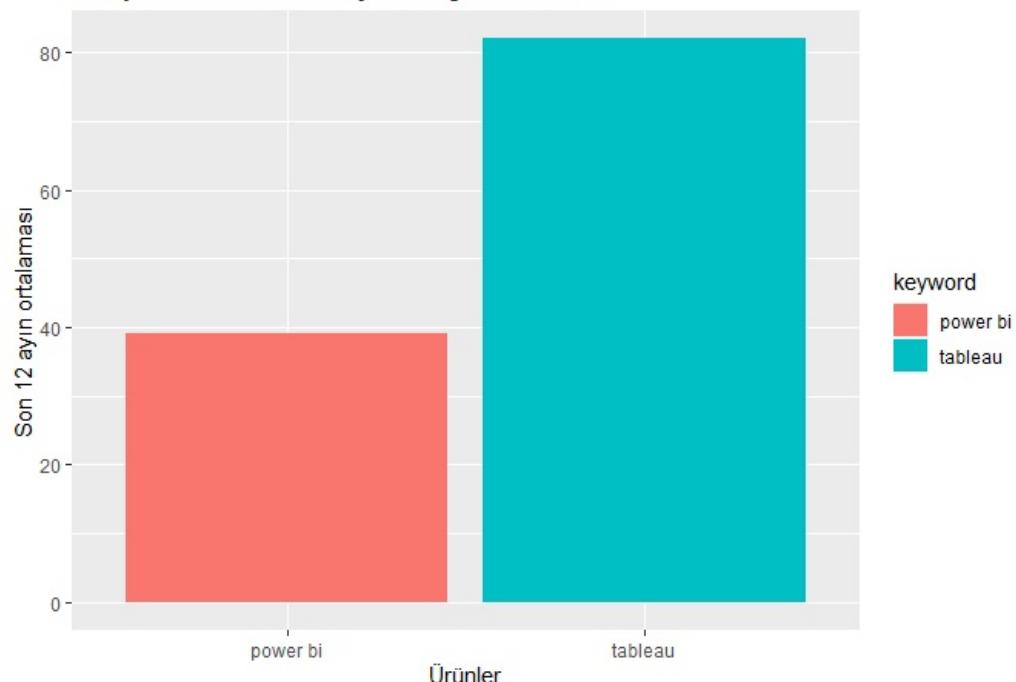
İş zekası, kısaca bir firmanın sahip olduğu veriler ile iş problemlerini çözmeye yönelik süreçlerin tasarlanması ve geliştirilmesi kapsamaktadır. İş zekası tooları bu anlamda birer araçtır, Her firma kendi kriterlerine göre kendisine en uygun aracı seçmektedir. Toplam 28 ilan içerisinde iş zekası, business intelligence kelimeleri geçmektedir.

Belirtmek isterim ki burada programların sayıları tam anlayıla bir pazar payını yansıtmadır. Bu analizde kullanılan veri özellikle iş zekası üstüne tasarlanmadığı için burada verilen ilanlarda geçme sıklığı yüksek güven içermemektedir. Yinede bir öngörü sağlama bağılamında faydalı olduğu düşünülmektedir.

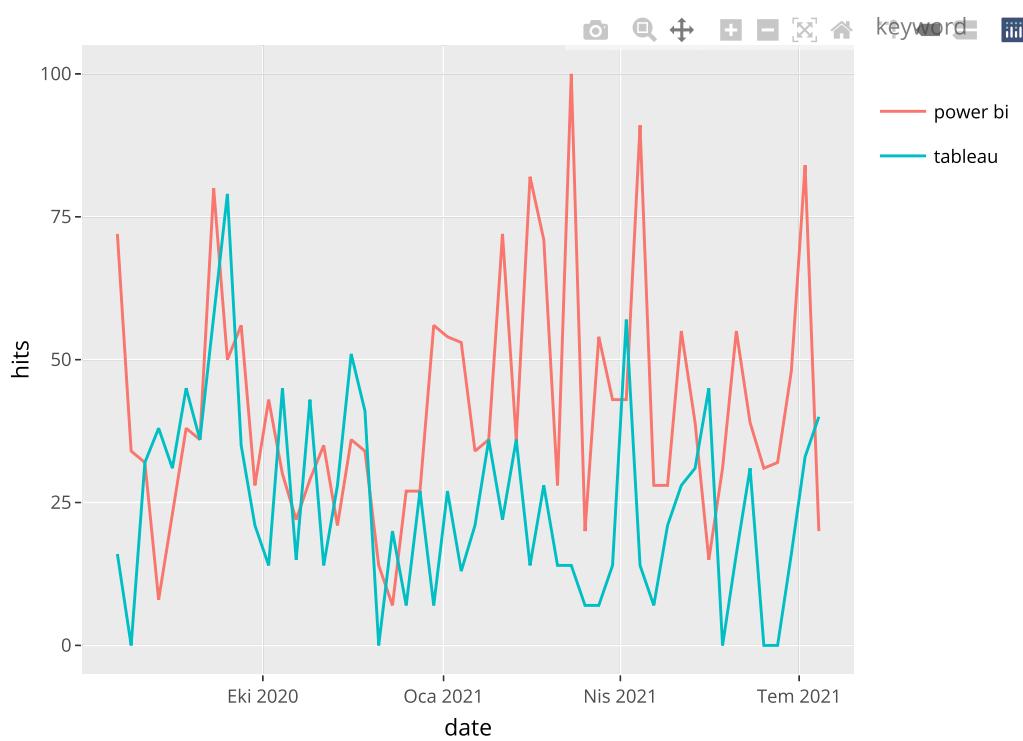
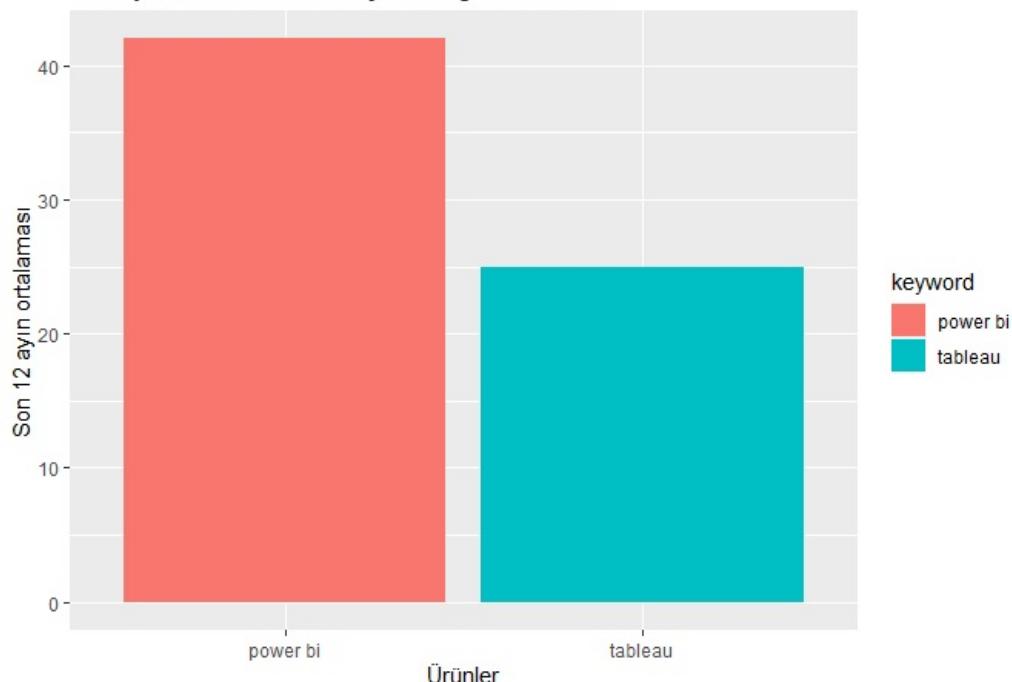
Yetkinlikler	Frekans
power_bi	26
r	19
python	17
SSI_SSRS	10
qlikview	3
tableau	3

Power bi microsoft tarafından sunulan pek çok programlama diline destek sunan çok kullanışlı bir iş zekası yazılımıdır. Microsoft ürünleri ile olan uyumu dolayısı ile Türkiye'de iyi bir pazar payı yakalamıştır. İlanlar üzerinden yapılan karşılaştırmayı daha anlamlı hale getirmek için google trends verileri incelenmiştir.

Dünyadaki arama sonuçlarına göre trendler



Türkiyedeki arama sonuçlarına göre trendler



Google trendleri incelendiğinde, İş zekası pazarı için Türkiye'de tableau ve qlikview'e göre daha fazla popülerdir. Dünyada ise genel arama miktarına göre Tableau Power BI'ye göre daha fazla popülerliğe sahiptir.(qlikview 2 programa'a nazaran aşırı düşük arama sonucu almasından dolayı dahil edilmemiştir.)

Sonuç olarak Power Bi Türkiye pazarında tableau'ya kıyasla daha fazla popüler olduğu, buna nazaran küresel pazarda Tableau'nun daha popüler olduğu sonucuna varılmıştır.

SSIS ve SSRS ise MS-SQL(T-SQL) veri tabanı üzerinde çalışan entegrasyon ve veri raporlama yazılımlarıdır. Microsoft ekosistemi içinde olması sebebiyle diğer microsoft ürünleri ile birlikte uyumlu çalışır ve bu sebeple tercih edilir.

R ve Python

R dili, istatistiksel hesaplamalar yapmak, üst seviye grafikler oluşturmak için yapılmış nesne tabanlı programlama ve fonksiyone programlama dili şematiğine sahip bir dildir. [kaynak](#)

Python ise Google desteği ile Guido van rossum tarafından geliştirilmiş çok amaçlı bir programlama dilidir. R'in esas amacı istatistiksel hesaplama yapmak iken Python çok amaçlı kullanılabilecek kolay öğrenilebilme ve kolay okunabilme amaçlanarak geliştirilmiş bir dildir. Popülerlik açısından karşılaştırıldığında Python uzun bir süredir R diline karşın tercih edilmektedir.[kaynak](#)

Özel sektör tarafından Python, sektör ihtiyaçlarına(esneklik,performans,çeşitlilik) daha iyi cevap verebilirken R, yapılma amacıyla uygun olarak istatistiksel raporlamalar,veri görselleştirme ve özellikle sağlık sektörü veri analitiği tarafında kendisine yer edinmiştir. Subjektif bir tahlilde iki dile de hakimiyet pek çok açıdan avantaj sağlamaktadır. Akademik araştırma yapılması planlanmıyorsa Python daha avantajları olan bir dildir.

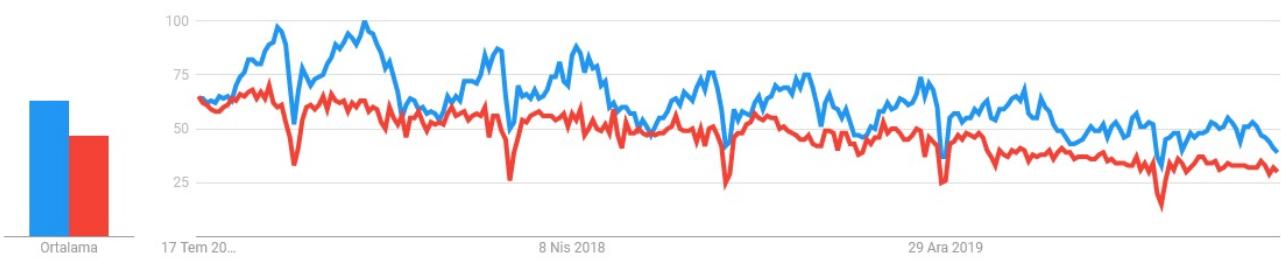
Diğer yazılımlar

Sas

SAS (Statistical Analysis System) veri erişimi, veri yönetimi, veri analizi ve veri sunumu sağlayan istatistiksel analizlerin yapılabildiği bir yazılım paketidir. Veri analizi, veri temizleme veya veri analitiği işlemleri ile ilgili birçok alanda kullanılabilen paket ürün sunuyor.

Spss

SPSS IBM tarafından istatistiksel analizler yapılabilmesi ve istatistiksel raporlarının oluşturulması amacıyla kullanılan akademik sektör ve araştırma & bilgi toplama sektöründe kullanılan popüler bir programdır.



SPSS SAS Institute

kaynak:Google trends

Sas ile karşılaştırmalı Google Trendsleri incelendiğinde 2 programın kullanımının azaldığı görülmektedir.

Java ve Scala

Java, Sun Microsystems mühendislerinden James Gosling tarafından geliştirilmeye başlanmış açık kaynak kodlu, nesneye yönelik, zeminden bağımsız, yüksek verimli, çok işlevli, yüksek seviye, adım adım işletilen bir dildir. Java programlama dili yazılımının pek çok alanında kullanılabilir. Veri analitiği tarafında ise veriyi analiz etmekten çok verinin depolanması ve işlenmesi tarafında kullanılmaktadır.

Scala ise Java gibi bir java virtual machine üzerine çalışıyor olup kullanıcılarının yüksek performanslı veri analiz işlemleri yapmasına fırsat tanıyan programlama dilidir. Bu dil ile geliştirilen en popüler yazılım Apache Spark'tır.

TIOBE index'ine göre Java, C'den sonra en popüler ikinci programlama dilidir. Sonrasında Python programlama dili artan ivme ile gelmektedir.(TIOBE index, programlama dillerinin popülerlik sıralamasını yapan bir kurumdur, liste aylık olarak yenilenmesinden ötürü değişiklikler olabilmektedir)

IT iş akışı ve iş yönetimine ait beceriler

Aşağıdaki yetkinlikler, IT veya yazılım tabanlı ar-ge departmanlarında önemli olan becerilerdir. İlanlar el ile incelendiğinde 36 ilanın yazılım geliştirici pozisyonu olarak kabul edilmiştir. Aşağıda bahsedilen yazılımlar iş için ana beceri olmaktan çok işin yönetimini ve sürecini organize etmesi sebebiyle bahsedilmesi uygun görülmüştür.

Yetkinlikler	Frekans
agile	15
jira	14
git	13
scrum	11
kanban	8

Git

Linux'un yaratıcısı olan Linus Torvalds tarafından geliştirilmiştir. Git bir version kontrol sistemi ve repository yönetme sistemidir. Kullanıcıların metin bazlı işlemlerini log olarak tutar ve aynı anda pek çok geliştiricinin projeye müdahale etmesine izin verir. Bu sayede bir projede olan biteni kontrol etmek kolaylaşmaktadır. Yazılım tarafındaki başlangıç seviyesi ilanlarda talep edilen bir yetkinlik olduğu görülmüştür.

Jira

Jira, Atlassian tarafından geliştirilen, hata izleme ve çevik proje yönetimine izin veren tescilli bir sorun izleme ürünüdür.

Agile/Scrum/Kanban

bir yazılım geliştirme prensipleri ve yönetim zihniyetidir. Scrum ve kanban ise bir proje yönetme biçimidir. Esasen sağladıkları çalışanların en verimli çalışması ve yazılım ürünün en kısa biçimde üretilmesidir. Bir yazılım projesi pek çok farklı departmandan ve çok yüksek sayılı ekiplerle sürdürülüğü için bu projenin yönetimi hayatı rol üstlenmektedir.

Benzer iş ilanlarının kümelenmesi

Analizimizi yaparken benzer iş ilanlarını bulmak için metin üzerinde Sentence transformer ile ilanlar'a ait Metinler vektörel değerler dönüştürülmüş daha sonrasında ise hiyerasyik kümeleme yapılarak 10 ayrı kümeye alınmıştır.

Burada yapılan kümeleme analizi deneysel bir yaklaşımındır. Yapılan kümeleme analizinde daha iyi sonuçlar çıkabilmesi için sentence transformer'in şartlarını daha iyi sağlayacak şekilde modifiye edilmelidir. Yine de deneyimlenmesi amacıyla yapılan kümeleme umut vaat edici sonuçlar ortaya koymaktadır.

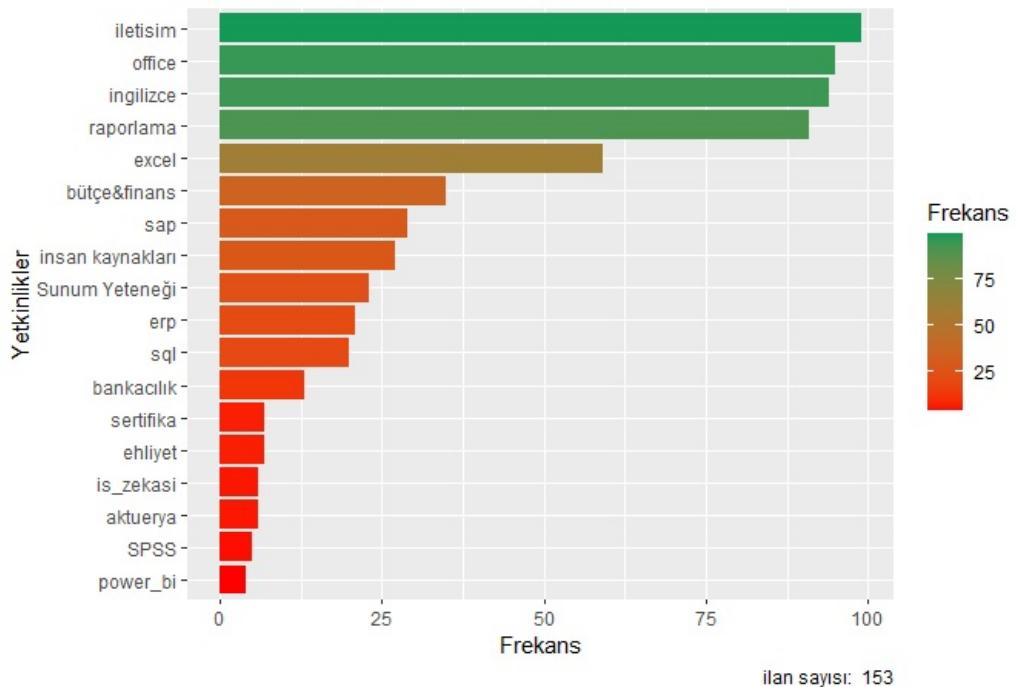
Kümeleme performansı yeterli olmaması sebebiyle kümeler üzerinden yeteneklerin gruplandırılmasına ilişkin bir açıklama yapılmayacaktır. Yetenekler açıklanırken sektör ve departmanlara ilişkin detaylı bilgi paylaşılmıştır. Yine de başarılı bir biçimde kümelenen pozisyonlar ile başarısız kümelenen bazı pozisyon incelenerek kümelenenin performansı analiz edilmiştir.

Dipnot: Kümeleme analizine ilişkin detaylar ve kodlar kaynakça kısmında paylaşılmıştır.

0 numaralı küme



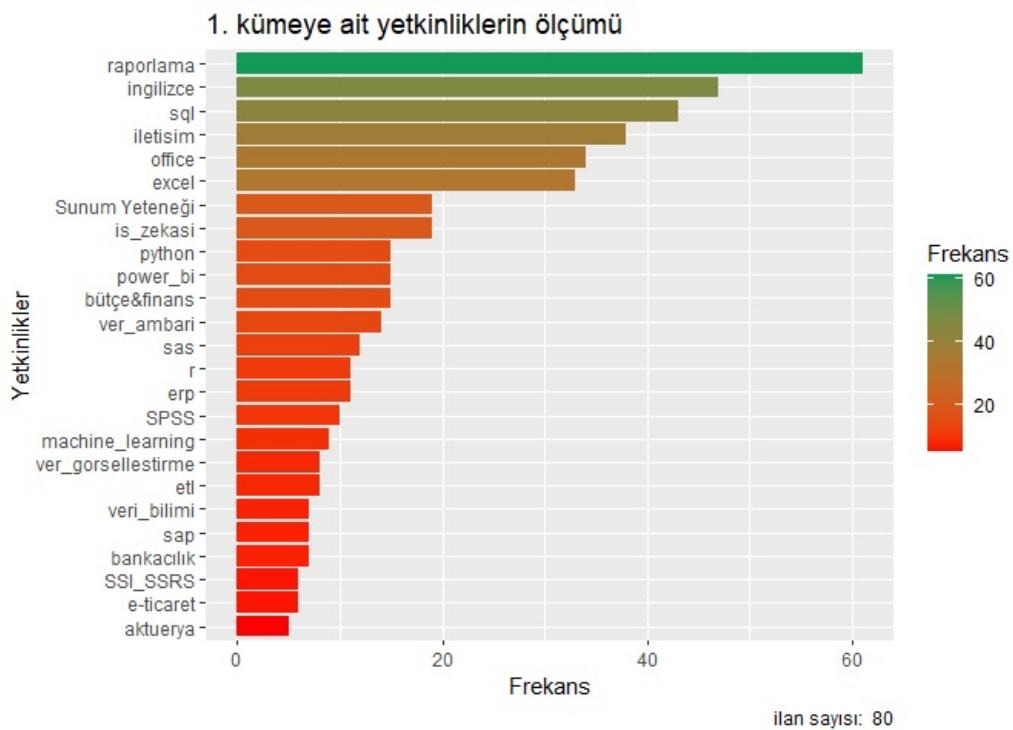
0. Kümeye ait yetkinliklerin ölçümü



Karmaşık metinleri ifade etmenin en kolay yollarından birisi Wordcloud oluşturmaktır. Yukarıda da görüldüğü gibi insan kaynakları-planlama-raporlama-satış-iş geliştirme ve bir çok pozisyonu bir arada görüyoruz. Model ne yazıkki insan kaynakları ilanlarını aynı bir pozisyon olarak yorumlayamadı ve 152 ilanı tek bir kategoride toparladı.

0 nolu küme de iş zekası tooları ve programlama dillerine dair yetenekleri çok alt sıralarda görmekteyiz. 0 nolu küme çoğunlukla finans insan kaynakları vb sektörlerde iş geliştirmeye dair departmanları kapsamaktadır. Sap programlarının yoğun bulunması buna örnektir.

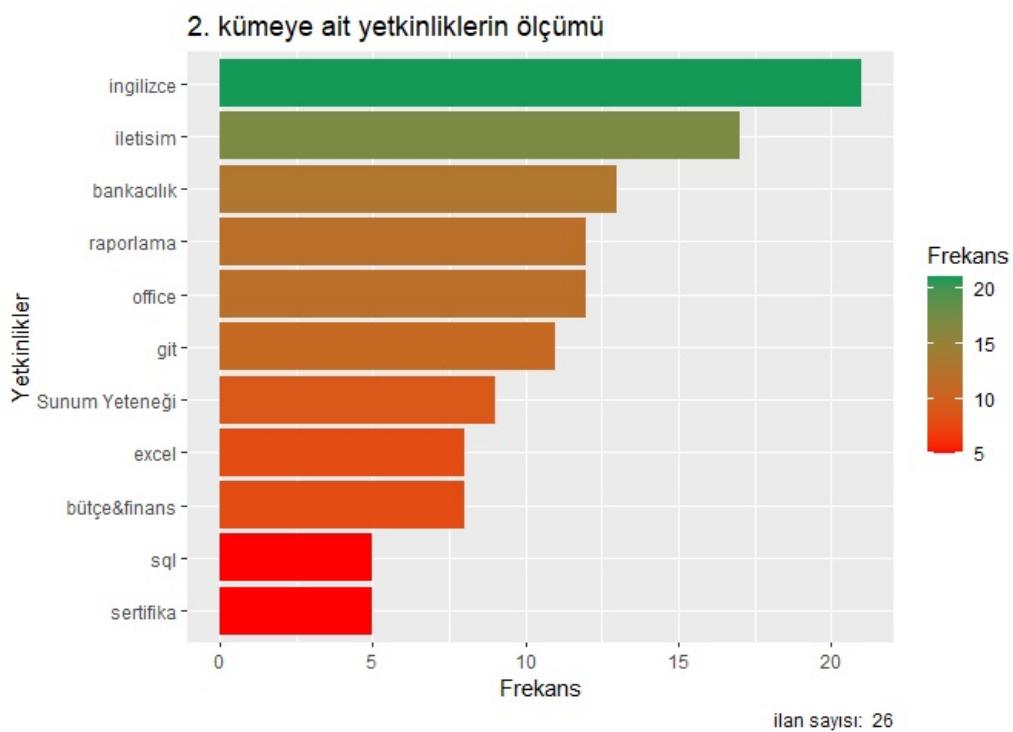
1 numaralı küme



1 numaralı kümeye çok açık bir biçimde veri analitiği ilanlarını kapsamaktadır. Sqş dilinin excel'den önde olması ve diğer programlama dilleri ve veri analitiğine ilişkin keywordlerin bulunması bunu açıklayabilmektedir.

2 numaralı kümeye

yoneticisi
uzmani
yardimcisi

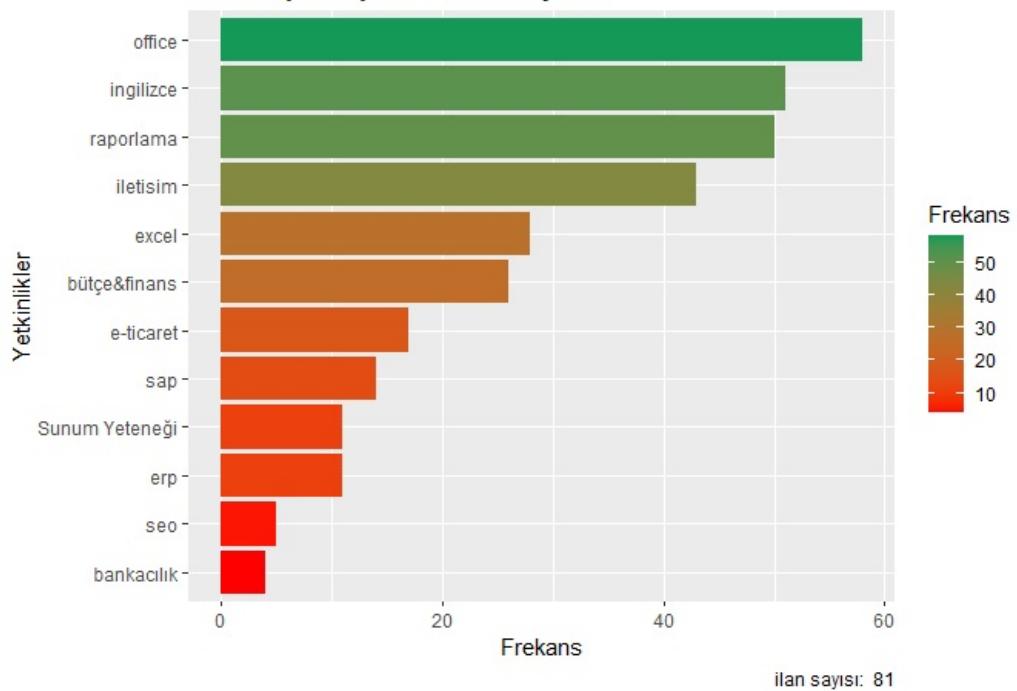


2 numaralı kümenin pek anlamlı kümelendiğini düşünmüyorum. Git bir yazılım geliştirme ürünü iken herhangi bir başka yazılım geliştirme ürünü görünmemektedir. Bu kümeye ilişkin net bir yorum yapılamaz.

3 numaralı kümeye ait yetkinliklerin ölçümü



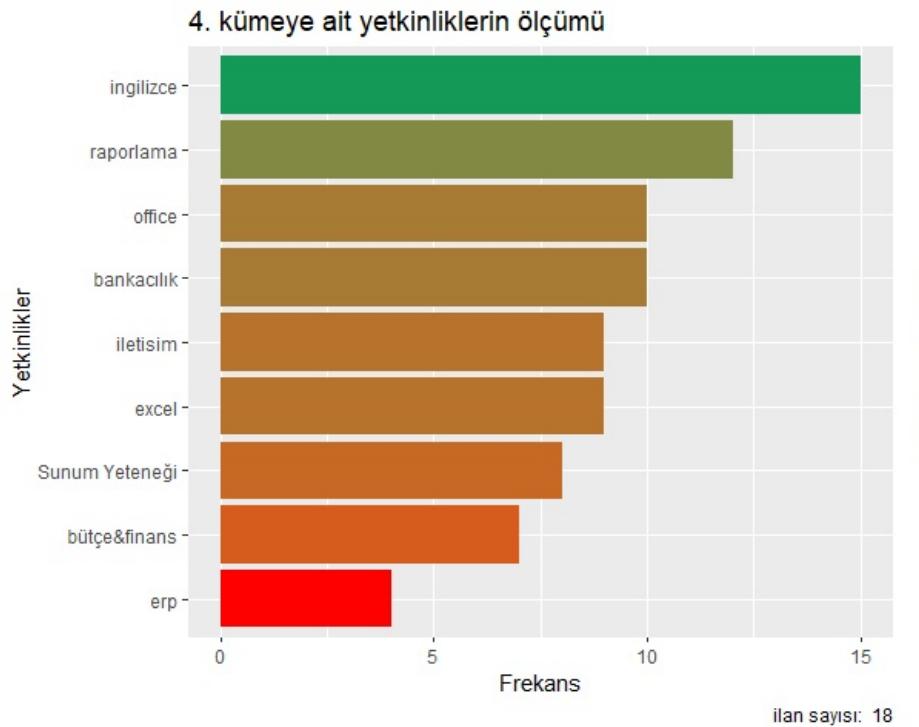
3. kümeye ait yetkinliklerin ölçümü



3. küme e-ticaret pazarlama satış alanlarındaki ilanları içermektedir.

4 numaralı küme

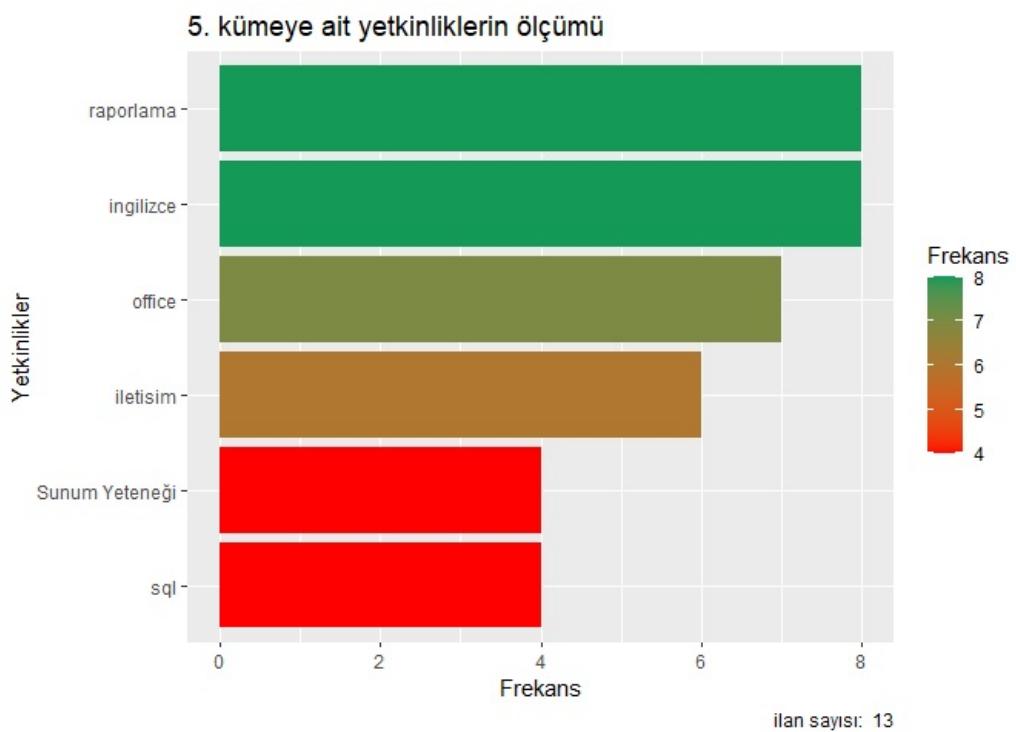
uzmani risk kredi yoneticisi



4.küme hakkında da pek bir yorum yapılabileceği söylenemez. Az sayıda bankacılık üzerine olduğu görünmektedir fakat becerilerine dair bilgiye saptanmamıştır.

5 numaralı küme

uzmani



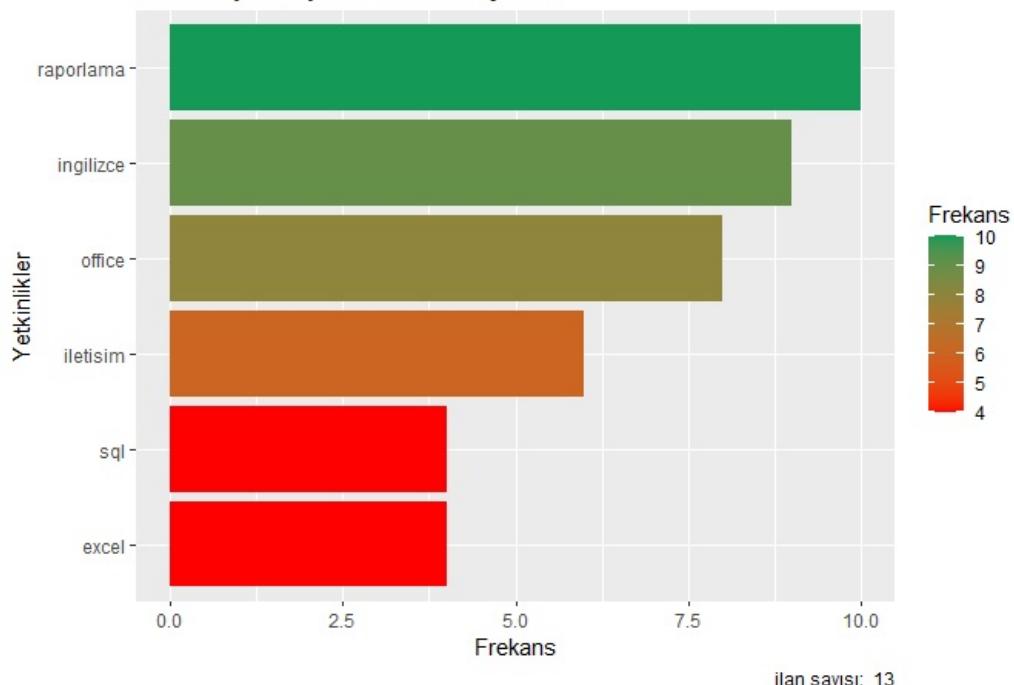
Herhangi bir yorum yapılmaya uygun bulunmamıştır.

6 numaralı küme

uzmani

raporlama sorumlusu

6. kümeye ait yetkinliklerin ölçümü

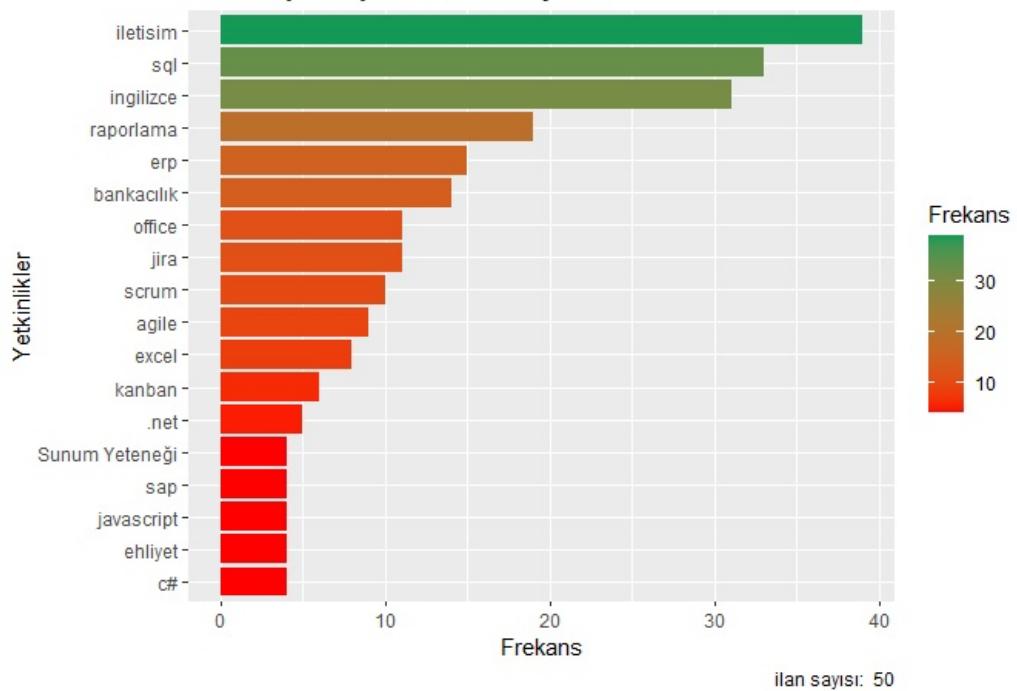


Herhangi bir yorum yapılmaya uygun bulunmamıştır.

7 numaralı küme



7. kümeye ait yetkinliklerin ölçümü

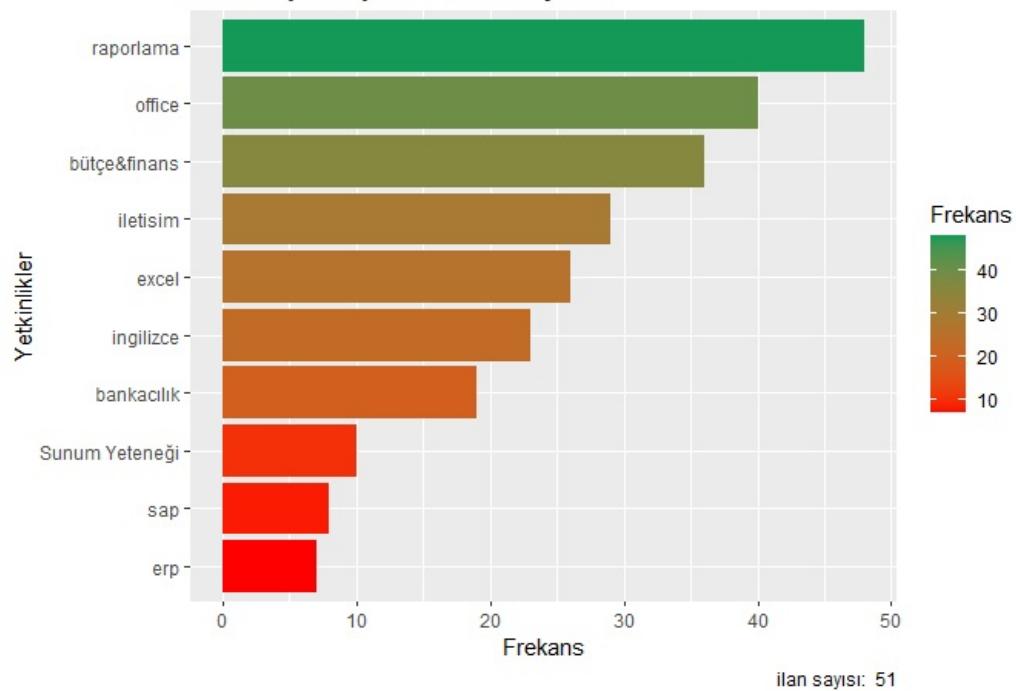


7 numaralı küme çok açık bir biçimde yazılım geliştirme ilanlarını içermekte veya yazılım-arge iş ilanlarını içermektedir.

8 numaralı küme

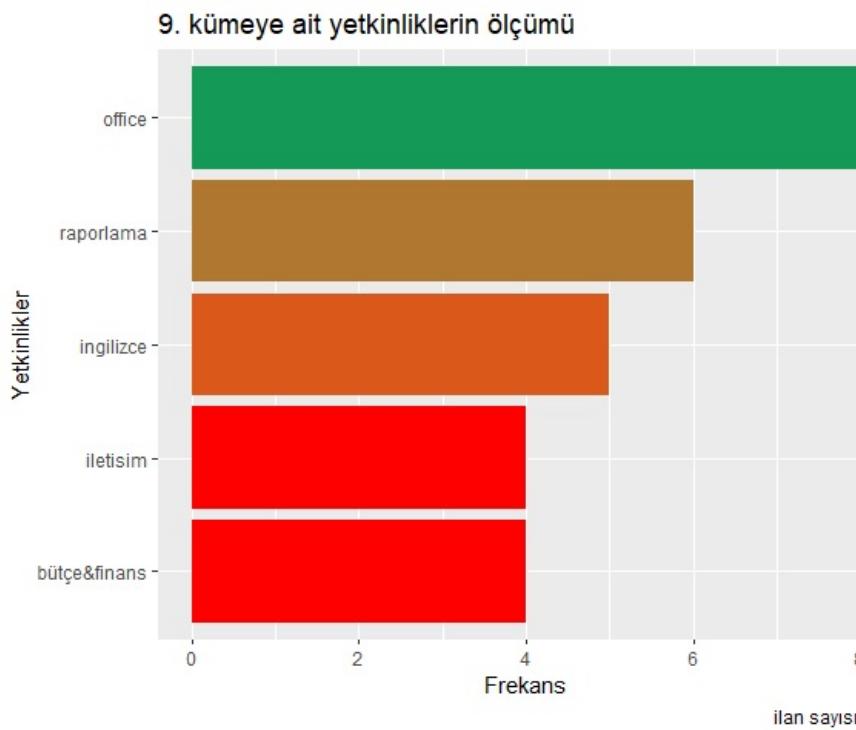


8. kümeye ait yetkinlıkların ölçümü



8 Numaralı küme de bir bütçe finans kısaca ekonomi ağırlıklı ilanları içermektedir.

9 numaralı küme



9 numaralı küme analiz edilmeye uygun bulunmamıştır.

Kümeleme analizinin değerlendirilmesi

Kümeleme analizi metinlerin elle taranması ile oluşan yetkinlik ve anahtar kelimeler ile bazı noktalarda başarılı sonuçlar vermişken ilanların hepsini anlamlı bir biçimde yorumlayamamıştır. Analiz daha kompleks yöntemlerle güçlendirildiği takdirde benzer iş ilanlarına ulaşmak daha mümkün görünmektedir.

Sonuç

- Yapılan analizler sonucunda istatistiğin ve istatistikçilerin pek çok farklı departmanda ve sektörde kabul edildiği ortaya konulmuştur. Bu sonuç istastistik bölümünün mezunlarına çok farklı alanlarda çalışabilme kabiliyeti sunduğunu göstermektedir.
- Yeni teknolojilere olan hakimiyetin iş ilanlarında göze çarptığı gözlemlenmiştir.
- İş ilanlarının büyük bir çoğunluğunun İstanbulda olması, adayların İstanbulda ikâmet etmelerinde önem arz edebilir. Pandemi dolayısıyla uzaktan çalışma olanaklarındaki artış bu zaruriyetiümüzdeki yıllarda azaltabilir.
- Adaylara en çok istishdam sağlayan sektörlerin bankacılık, finans, e-ticaret ve yazılım sektörü olduğu gözlemlenmiştir.
- Askerliğin tecilli olmasını pek çok firma tarafından ilk kriter olarak kabul edilmediği gözlenmiştir. Bu durum adayın tecil süresi ve başvurduğu işin kıdem seviyesi ile ilişkili olduğu gözden kaçırılmamalıdır.
- Yönetici pozisyonlarında 4. ve 5. yıllarında artış gözlemlenmiştir.
- İşverenler tarafından en çok aranan kabiliyetlerin İngilizce becerisi, Excel, SQL ve sektör tecrübesi olduğu gözlemlenmiştir.
- İncelenen iş ilanlarında Bilgisayar Mühendisliği, Endüstri Mühendisliği, Ekonometri mezunlarının aynı şekilde kabul edildiği gözlemlenmiştir.