

T.C.
BİLECİK ŞEYH EDEBALI ÜNİVERSİTESİ
İKTİSADİ VE İDARİ BİLİMLER FAKÜLTESİ



OYUN TANITIMI VIDEO YORUM ANALİZİ

SOSYAL MEDYA VE WEB ANALİZİ FİNAL ÖDEVİ

Alperen ÖZDEMİR

YÖNETİM BİLİŞİM SİSTEMLERİ

Nur Kuban TORUN

Haziran 2024

Özet

Bethesda adındaki bir oyun şirketi 11 Kasım 2011 de çıkardığı “The Elder Scrolls V Skyrim” adlı oyunu ile çok iyi satış rakamları yakaladı ve oyun senelerce konuşuldu. 7 yılın ardından bir devam oyunu olarak 11 Haziran 2018 tarihinde YouTube’da “The Elder Scrolls VI” adlı oyunun 36 saniyelik bir tanıtım videosunu paylaştılar. İlk yıllarda birçok kişi heyecanla oyun hakkında yeni haberler ve gelişmeler bekledi. Yıl 2024 olmasına rağmen hala daha 2018 yılında yayınlanan tanıtım videosundan sonra oyun hakkında başka hiçbir video veya gelişme paylaşmadılar.

Bu projede Youtube üzerindeki "The Elder Scrolls VI – Official Announcement Teaser" videosundaki 72243 yorum çekilmiştir. R programlama dilinde metin madenciliği ve emoji analizi kullanılarak gerçekleştirilen analizlerde en fazla kullanılan kelimeler üzerinden duygu analizi yapılmış ve elde edilen sonuçlara göre insanların, geçen zaman içerisinde oyun hakkındaki pozitif yorumlarının yanı sıra fazlasıyla negatif yorumlarının da eklendiği görülmüştür.

1. Giriş

The Elder Scrolls, Bethesda Game Studios tarafından geliştirilen ve Bethesda Softworks tarafından yayınlanan bir aksiyon-rol yapma video oyunu serisidir. İlk olarak 1994'te piyasaya sürülen bu seri, zengin ve detaylı bir açık dünya deneyimi sunmasıyla tanınır. The Elder Scrolls serisi, derin hikayeler, kapsamlı keşif olanakları ve özgün bir evren ile dikkat çeker.

Başlıca oyunlar şunlardır:

1. **The Elder Scrolls: Arena** (1994): Serinin ilk oyunudur ve geniş bir açık dünya sunar.
2. **The Elder Scrolls II: Daggerfall** (1996): İlk oyunun konseptlerini genişleten ve büyük bir açık dünya sunan devam oyunudur.
3. **The Elder Scrolls III: Morrowind** (2002): Seriyi modernize eden ve eleştirmenler tarafından beğenilen üçüncü oyundur.
4. **The Elder Scrolls IV: Oblivion** (2006): Gelişmiş grafikler ve daha geniş bir oyun dünyası sunar, büyük beğeni toplamıştır.
5. **The Elder Scrolls V: Skyrim** (2011): En popüler ve tanınmış oyunlardan biridir.

Seri, Tamriel adlı hayali bir kıtada geçer ve her oyun bu kıtanın farklı bölgelerini keşfetmenize olanak tanır. Oyunlar, oyunculara geniş bir özgürlük sunarak çeşitli görevleri yerine getirmelerine, farklı karakterler oluşturmalarına ve dünyayı istedikleri gibi keşfetmelerine imkan tanır. The Elder Scrolls serisi, detaylı dünyası, karmaşık hikayeleri ve zengin oynanışı ile RPG türünde önemli bir yere sahiptir.

2. Yöntem

Google Developer hesabından alınan API ile Youtube üzerinden “The Elder Scrolls VI – Official Announcement Teaser” adlı videodan yorumlar çekilmiş, R programlama üzerinde veri temizliği, duygu analizi ve görselleştirme yapılmıştır.

2.1. Metin Madenciliği

Yorum çekme, temizleme ve sıralama işlemleri yapılmıştır.

2.2. Duygu Analizi

Polarite analizi ve duygu analizi yapılmıştır. Grafikler ile “Pozitif” , “Negatif” ve “Nötr” olarak gösterilmiştir.

2.3. Araştırmanın Amacı

Youtube kullanıcılarının videodaki konu hakkında verdikleri tepkileri ölçmek, yaptıkları yorumlarla duygu analizi yapmak amaçlanmaktadır. Ayrıca en çok kullanılan emojileri saptamak amaçlanmaktadır.

3.5. R Programlama Kütüphaneleri

R programlama için kullanılan yüklenen ve çağırılan kütüphaneler Şekil 1’de verilmiştir.

```
1 # Kütüphanelerin yüklenmesi
2 install.packages("tuber")
3 install.packages("httpuv")
4 install.packages("dplyr")
5 install.packages("tidytext")
6 install.packages(c("tm", "SnowballC"))
7 install.packages("stringr")
8 install.packages("ggplot2")
9 install.packages("wordcloud")
10 install.packages("wordcloud2")
11 install.packages("textdata")
12 install.packages("ggthemes")
13 install.packages("sentimentr")
14 install.packages("pander")
15 install.packages("pastecs")
16 install.packages("syuzhet")
17 install.packages("stringi")
18 install.packages("tidyverse")
19 install.packages("janitor")
20 install.packages("emoji")
21 install.packages("readr")
22 install.packages("openxlsx")
23 install.packages("ggtext")
24
25 #Kütüphanelerin çağırılması
26 library(syuzhet) # Metinlerden duygu ve duygu tabanlı olay örgüsü çıkarmak için bir araç sağlar
27 library(tuber) # Youtube API'ye erişim sağlar
28 library(httpuv) # HTTP ve websocket isteklerini doğrudan R içinden işlemek için düşük seviye soket ve protokol desteği sağlar
29 library(dplyr) # Veri manipülasyonu için bir dil sağlar
30 library(tidytext) # Tidy veri ilkelerini kullanarak birçok metin madenciliği görevini daha kolay ve etkili araçlarla tutarlı hale getirir.
31 library(tm) # R içinde metin madenciliği uygulamaları için bir çerçeve sağlar
32 library(SnowballC) # Porter'in kelime kök bulma algoritmasını uygulayan C 'libstemmer' kütüphanesine bir R arayüz sağlar
33 library(stringr) # Karakter dizileri üzerinde çalışmayı kolaylaştıran fonksiyonlar sağlar
34 library(ggplot2) # Grafikleri deklaratif bir şekilde oluşturmak için bir sistem sağlar
35 library(wordcloud) # Gözde kelime bulutları çizmek için bir araç sağlar
36 library(wordcloud2) # Kelime bulutları çizmek için bir araç sağlar
37 library(textdata) # Metinle ilgili veri setlerine kolay erişim sağlar
38 library(ggthemes) # 'ggplot2' için ek temalar, geoms ve ölçekler sağlar
39 library(sentimentr) # İngilizce'deki metin polaritesi duygusunu hızlı bir şekilde hesaplar
40 library(pander) # R nesnelerini Pandoc'un markdown formatına dönüştürmek için minimal ve kolay bir araç sağlar
41 library(pastecs) # Uzun-zaman serilerinin düzenlenmesi, ayrıştırılması ve analizi
42 library(stringi) # Metin verileriyle çalışmayı kolaylaştıran işlevler sunar. Metin manipülasyonu ve analizi için.
43 library(tidyverse) # Veri manipülasyonu, görselleştirme, modelleme ve raporlama gibi işlemler için yardımcı olur.
44 library(janitor) # Veri temizleme ve düzenleme.
45 library(emoji) # Emojileri kullanma ve görüntüleme.
46 library(readr) # R dilinde veri okuma işlemlerini hızlı ve kolay hale getiren bir pakettir.
47 library(openxlsx) # Excel dosyalarıyla çalışmayı sağlayan bir pakettir.
48 library(ggtext) # Veri görselleştirme aracı olan ggplot2 için metin öğelerini zenginleştiren bir pakettir.
```

Şekil 1: Proje için kullanılan R kütüphaneleri

```
36 # Youtube üzerinden yorum çekme ve csv olarak kaydetme
37 app_id <- "[REDACTED].apps.googleusercontent.com"
38 app_secret <- "[REDACTED]"
39 yt_oauth(app_id, app_secret, token = "")
40 get_all_comments(video_id = "gElfIo6uw4g")
41 comments1 <- get_all_comments(video_id = "gElfIo6uw4g")
42 write.csv(comments1, file = "youtubecomments_alperen.csv")
```

Şekil 2: YouTube yorumlarını çekme ve csv olarak kaydetme

```
60 # Yorum verimi okuyorum
61 yorumlar <- read.xlsx("C:\\R_proje\\youtubeyorumlar_alperen.xlsx")
```

Şekil 3: Xlsx formatındaki yorumları okuma

```
47 # Yorum veri çerçevesini seçme
48 yorumlar <- yorumlar %>%
49   dplyr::select(x,textoriginal)
```

Şekil 4: Verisetimi parçalayıp ihtiyacım olan yeri alma

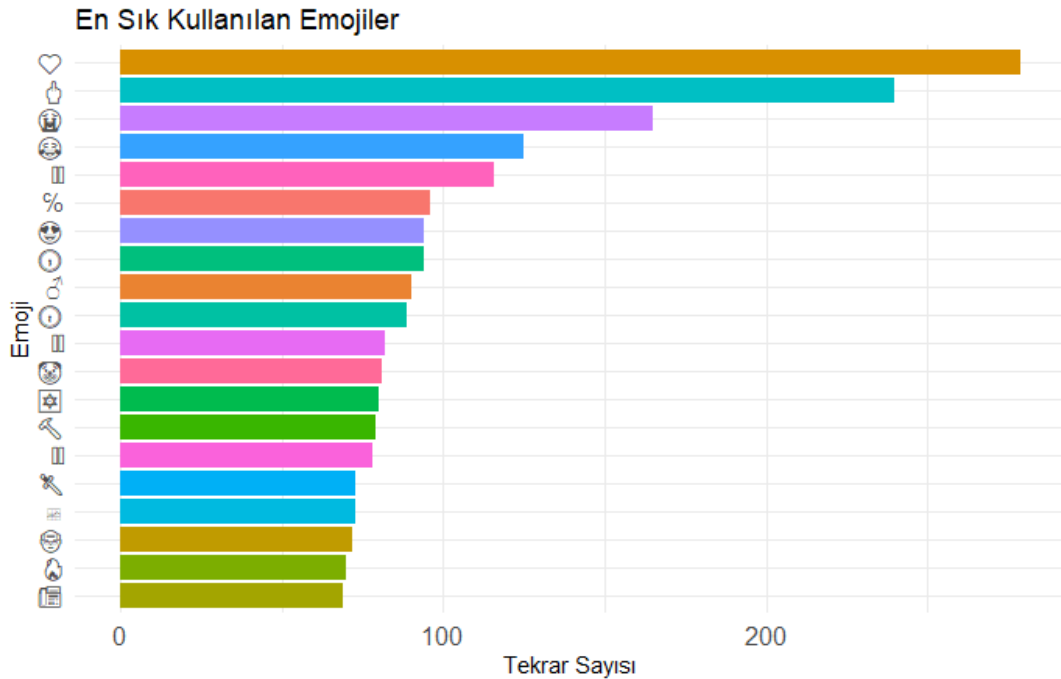

```

126 # video_yorum$textoriginal vektöründeki tüm emoji'leri çıkarır
127 emoji_jler <- stri_extract_all_regex(yorumlar$textoriginal, "\\p{so}")
128
129 # Çıkarılan emoji'lerin her birinin kaç kez tekrar ettiğini hesaplar
130 emoji_tablosu <- table(unlist(emoji_jler))
131
132 # Emoji ve tekrar sayısını içeren bir veri çerçevesi oluşturur
133 emoji_df <- data.frame(emoji = names(emoji_tablosu),
134                       tekrar_sayisi = as.numeric(emoji_tablosu)) %>%
135   # Tekrar sayısına göre azalan sırada düzenler
136   arrange(desc(tekrar_sayisi)) %>%
137   # En sık tekrar eden ilk 22 emoji'yi seçer
138   slice_head(n = 20)
139
140
141 # Değişiklikleri kontrol et
142 print(emoji_df)
143
144 # Barplot oluşturma
145 ggplot(emoji_df, aes(x = reorder(emoji, tekrar_sayisi), y = tekrar_sayisi)) +
146   geom_bar(stat = "identity", aes(fill = emoji), show.legend = FALSE) +
147   coord_flip() +
148   labs(title = "En Sık Kullanılan Emoji'ler",
149        x = "Emoji",
150        y = "Tekrar Sayısı") +
151   theme_minimal() +
152   theme(
153     axis.text.x = element_text(size = 12),
154     axis.text.y = element_markdown(size = 12)
155   )

```

Şekil 7: Bar grafik

Veride kullanılan tüm emoji'leri çıkarıyorum, her birinin kaç kez tekrar ettiğini hesaplayıp tekrar sırasına göre azalan olarak sıralıyorum. En çok kullanılan 20 emoji'yi seçiyorum ve histogram oluşturuyorum.



Şekil 8: Emoji kullanım sıklığının histogram şeklinde gösterilmesi

208	208	DAY #1900 : Dear logbook, another hundred days passed ! I...
209	209	**In Memoriam: Todd Howard Soars into the Aurora Borealis...
210	210	**Developing Story: Todd Howard's Draconic Form Struggle...
211	211	**Breaking News: Todd Howard's Human Form Cracks, Reve...
212	212	I'm pretty sure we will have GT6 before this 😊
213	213	By the time this comes out, I'll have fathered two kids. I was ...
214	214	Oh my... I'm awed...
215	215	I see you comrades, take a seat it's gonna be a while...
216	216	Gamers will always remain oppressed and you cannot tell m...
217	217	#DAY #267. Я из России. Пожал руку преподу сегодня)...
218	218	А вы тоже ждете выхода tes 6 в 2024г?
219	219	After almost 6 years.....Hello? Is there anything else co...
220	220	My tea's gone cold, I'm wondering why I got out of bed at a...

Şekil 9: Veri temizleme işlemi yapılmamış bir örnek

```

77 # Çıkarılacak kelime listem
78 thestoplist <- readLines("C:\\R_proje\\stopwords_alperen.csv")
79 thestoplist <- data.frame(word = thestoplist)
80
81 # Küçük harfe çevirme
82 yorumlar$textoriginal <- tolower(yorumlar$textoriginal)
83
84 # Linkleri ve etiketleri kaldırma
85 yorumlar$textoriginal <- gsub("http[^\ ]*", "", yorumlar$textoriginal)
86 yorumlar$textoriginal <- gsub("@[^\ ]*", "", yorumlar$textoriginal)
87
88 # Emojileri kaldırma
89 yorumlar$textoriginal <- gsub("<..*?>", "", yorumlar$textoriginal)
90
91 # Sayıları kaldırma
92 yorumlar$textoriginal <- removeNumbers(yorumlar$textoriginal)
93
94 # Noktalama işaretlerini kaldırma
95 yorumlar$textoriginal <- str_replace_all(yorumlar$textoriginal, "[[:punct:]]", "")
96
97 # Çıkarılacak kelime listemdeki kelimeleri kaldırma
98 yorumlar <- yorumlar %>% filter(!textoriginal %in% thestoplist)
99
100 # İngilizce alfabeye uygun olmayan kelimeleri kaldırdım
101 for(i in seq_along(yorumlar)) {
102   if(is.character(yorumlar[[i]])) {
103     yorumlar[[i]] <- stringr::str_replace_all(yorumlar[[i]], "[^a-zA-Z ]", "")
104   }
105 }
106
107 # Çıkarılan yerlere NA değeri atama
108 yorumlar$textoriginal[yorumlar$textoriginal == ""] <- NA
109
110 # NA değerlerini kaldırma
111 yorumlar <- na.omit(yorumlar)
112
113
114 # Yorumları kelimelere ayırıp çıkarılacak kelimeleri çıkarma
115 yorumlar <- yorumlar %>%
116   unnest_tokens(word, textoriginal) %>%
117   anti_join(thestoplist)
118
119
120 # Tekrar cümle haline getirme
121 cümle <- yorumlar %>%
122   group_by(x) %>%
123   summarise(word = paste(word, collapse = " "))
124
125 # Kelimelerin tekrar sayısına göre sıralama
126 kelime_tekrar <- yorumlar %>% count(word, sort = TRUE)

```

Şekil 10: Veri temizleme

Verideki linkler, etiketler, emojiler, sayılar, noktalama işaretleri ve İngilizce alfabesine uyumlu olmayan kelimeleri çıkarıp, çıkarılan yerlere NA değeri atıyoruz ve NA değerlerini kaldırıyoruz. 88. Satırda kelimelere ayırıp çıkarılacak kelimeleri çıkarıyoruz. Daha sonra temizlenmiş yorumları kelimelere ayırıp tekrar sıklığına göre sıralıyoruz.

	word	n
1	game	10305
2	like	6147
3	skyrin	5952
4	elder	5135
5	scrolls	5135
6	bethesda	4945
7	dear	4193
8	time	3618
9	logbook	3495

Şekil 11: Temizleniş ve kelimelere ayrılmış veri örneği

```

94 # Tekrar cumle haline getirme
95 cumle <- yorumlar %>%
96   group_by(x) %>%
97   summarise(word = paste(word, collapse = " "))
98
99 # kelimelerin tekrar sayısına göre sıralama
100 kelime_tekrar <- yorumlar %>% count(word, sort = TRUE)

```

Şekil 12: Cümle analizi için kelimelere ayırdığımız veriyi tekrar cümle haline çevirme ve tekrar sayılarına göre sıralama

```

207 # kelime bulutu oluşturma
208 wordcloud2(kelime_tekrar, size = 1, minsize = 0, gridSize = 0,
209   fontfamily = 'Times New Roman', fontweight = 'bold',
210   color = 'random-dark', backgroundcolor = "white", shuffle = TRUE,
211   rotateRatio = 0.4, shape = 'circle', ellipticity = 0.65,
212   widgetsize = NULL, figPath = NULL, hoverFunction = NULL)
213
214
215
216
217 # En çok tekrar eden 30 kelimeyi seçme
218 kelime_tekrar_top30 <- kelime_tekrar %>%
219   top_n(30, n)
220
221 # Histogram oluşturma
222 ggplot(kelime_tekrar_top30, aes(x = reorder(word, n), y = n)) +
223   geom_bar(stat = "identity", fill = "#2ea061") +
224   coord_flip() +
225   labs(title = "En Çok Tekrar Eden 30 Kelime",
226     x = "kelime",
227     y = "Tekrar Sayısı")
228

```

Şekil 13: En çok tekrar eden kelimelerin histogram ve kelime bulutu şeklinde gösterilmesi

```

234 # 'sentiment' fonksiyonu, 'Tekrar_yorum$word' veri çerçevesindeki her kelimenin duygu değerini hesaplar.
235 polarite <- sentiment(cumle$word)
236
237 # 'cbind' fonksiyonu, 'Tekrar_yorum$word' ve 'polarite' veri çerçevelerini birleştirir ve 'tablo' adlı yeni bir veri çerçevesi oluşturur.
238 tablo <- cbind(cumle$word, polarite[,c(3,4)])
239
240 # İstatistikleri hesapla
241 stat.desc(polarite$sentiment, basic=T) %>% pander()
242
243
244
245 # Polarite görseli
246 ggplot(tablo, aes(word_count, sentiment))+
247   geom_point(color="#2ea061")+
248   geom_hline(yintercept = mean(tablo$sentiment), color="red", size=1)+
249   labs(y = "Skor", x = "Kelimelerin Frekansı") +
250   theme_gray()+
251   ylim(-3,3)+
252   labs(caption = "Veri kaynağı: Youtube'da The Elder Scrolls VI - Official Announcement Teaser" adlı videodur.")+
253   theme(plot.caption = element_text(hjust = 0, face = "italic"))

```

Şekil 14: Duygu analizi için kelimelere duygu değeri verme, polarite analizi oluşturma

```

157 # yorumlar içerisindeki veri çerçevesini Bing ile birleştir.
158 Duygu <- yorumlar %>%
159   inner_join(get_sentiments("bing")) %>%
160
161   # Her bir kelimenin ve duyarlılık puanının frekansını hesapla
162   count(word,sentiment)
163
164 # Duygu verisini duyarlılığa göre grupta ve en çok tekrar eden 20 kelimeyi göster
165 Duygu %>%
166   group_by(sentiment) %>%
167   top_n(20) %>%
168   ungroup() %>%
169   mutate(word = reorder(word, n)) %>%
170
171   # Bar grafiği oluştur
172   ggplot(aes(word, n, fill = sentiment)) +
173   geom_col(show.legend = FALSE) + # Barları çiz ve efsaneyi gizle
174   facet_wrap(~sentiment, scales = "free_y") +
175
176   # Eksenleri etiketle
177   labs(y = "",
178        x = NULL) +
179
180   # Grafikte x ve y eksenlerini çevir
181   coord_flip()

```

Şekil 15: Veri çerçevesini bing ile birleştirme, duyarlılık puanı hesaplama ve görselleştirme

```

290 #Yorumlariin yogunluklari
291 bar_grafik <- tablo$sentiment
292 bar_grafik <- as.data.frame(bar_grafik)
293
294 # Notr, negatif ve pozitif bar gorselleri
295 bar_grafik$bar_grafik[bar_grafik$bar_grafik > 0] <- 'Pozitif'
296 bar_grafik$bar_grafik[bar_grafik$bar_grafik == 0] <- 'Notr'
297 bar_grafik$bar_grafik[bar_grafik$bar_grafik < 0] <- 'Negatif'
298
299 # 'bar_grafik' sutununu bir faktor olarak donusturun
300 bar_grafik$bar_grafik <- as.factor(bar_grafik$bar_grafik)
301
302 # Kategorilere gore sayilari hesaplayin
303 a_sayilari <- table(bar_grafik$bar_grafik)
304
305
306 # Renklerin bir vektorunu olusturun
307 renkler <- c("#d33030", "#4b6d65", "#2ea061")
308
309 # Bar grafigini olusturun ve her bir bar icin farkli bir renk belirleyin
310 barplot(a_sayilari, main='Duygu Dagilimi', xlab='kategori', ylab='Sayi', col=renkler)

```

Şekil 16: Yorumların duygu yoğunluklarını hesaplama ve "Pozitif", "Negatif", "Nötr" olarak görselleştirme

```

211 # yorumlar icerisindeki her bir kelimeyi frekansa gore sirala
212 afins <- yorumlar %>%
213   count(word, sort = TRUE) %>%
214
215   # AFINN ile birlestir.
216   inner_join(get_sentiments("afinn"), by = "word") %>%
217   group_by(word) %>%
218
219   # Her bir kelimenin toplam katkisini hesapla.
220   summarize(contribution = sum(n * value)) %>%
221
222   # En yuksek 12 kelimeyi al
223   top_n(10, abs(contribution)) %>%
224
225   # kelimeleri katkiya gore yeniden sirala.
226   mutate(word = reorder(word, contribution))
227
228
229 # 10 farkli renk iceren bir vektor olusturun
230 renklere <- c("brown", "black", "grey", "red", "turquoise", "orange", "yellow", "darkgreen", "green", "#3b5998")
231
232 # Bar grafigini olusturun ve her bir bar icin farkli bir renk belirleyin
233 ggplot(data = afins, aes(word, contribution, fill = word)) +
234   geom_col() + # Barlari ciz
235   scale_fill_manual(values = renklere) + # Belirli renkleri kullan
236   coord_flip() + # Grafikte x ve y eksenlerini cevir
237   labs(y = "kelime sikligi * AFINN puani", x = "kelimeler") # y eksenini etiketle

```

Şekil 17: Yorumları AFINN ile birleştirme ve kendi renklerine göre görselleştirme

Bulgular

Yapılan yorumlar içerisinde en çok tekrar eden kelimeler Şekil 14'te görüldüğü üzere Game(Oyun), Like(Beğenmek), Skyrim(Oyun ismi), Bethesda(Şirketin ismi), Dear(Kibar hitap), Time(Zaman), Logbook(Günce) kelimeleri olmuştur. En çok tekrar eden emojiler ise:



Emojileri olmuştur.

Kelime	Tekrar Sayisi
game	10300
like	6200
skyrim	6000
bethesda	4900
dear	4300
time	3700
logbook	3500
tes	3300
hope	3100
todd	2900
trailer	2600
good	2500
games	2300
release	2100
starfield	2000
high	1800
howard	1500
long	1400
waiting	1400
people	1400
life	1400
gonna	1400
great	1400
released	1400
love	1400
hammerfell	1400
rock	1400
teaser	1400
finally	1400
yes	1400

Şekil 18: En çok tekrar eden 30 kelimenin histogram ile gösterilmesi



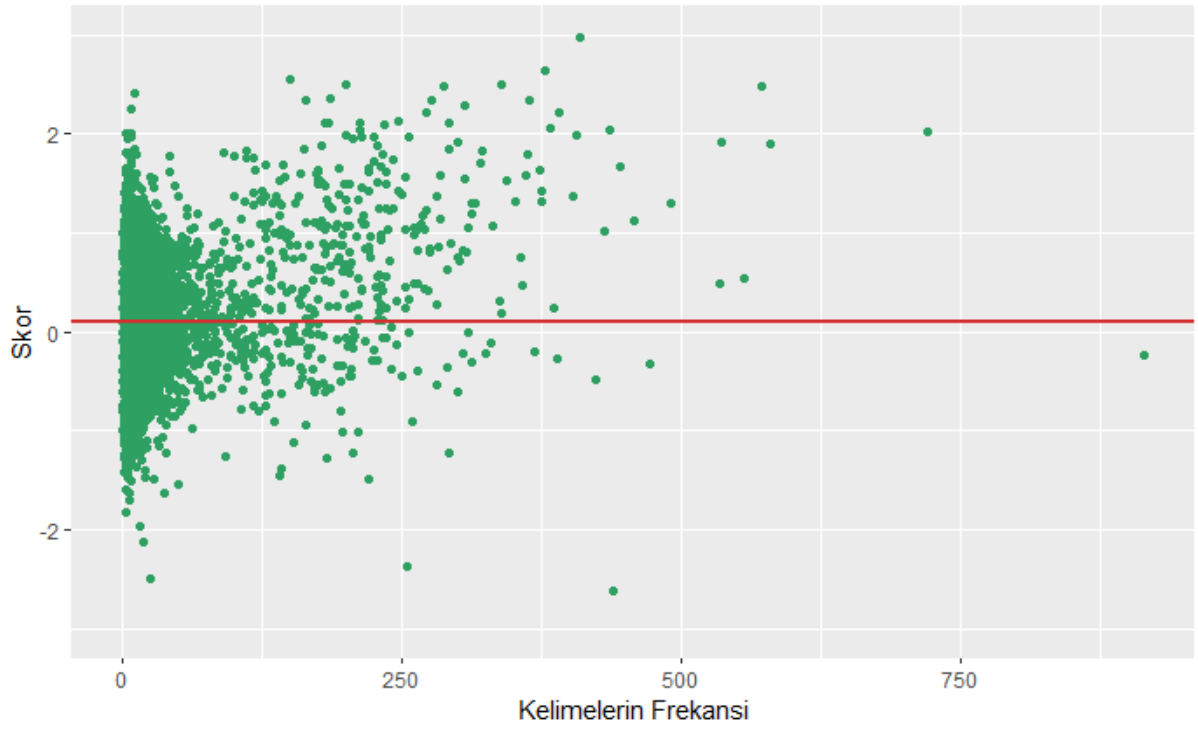
Şekil 19: En çok tekrar eden kelimelerin kelime bulutu ile gösterilmesi

nbr.val	nbr.null	nbr.na	min	max	range	sum	median	mean
64693	24760	0	-2.622	10.37	12.99	6368	0	0.09844

Table: Table continues below

SE.mean	CI.mean.0.95	var	std.dev	coef.var
0.001464	0.00287	0.1387	0.3725	3.784

Şekil 20: İstatistikler

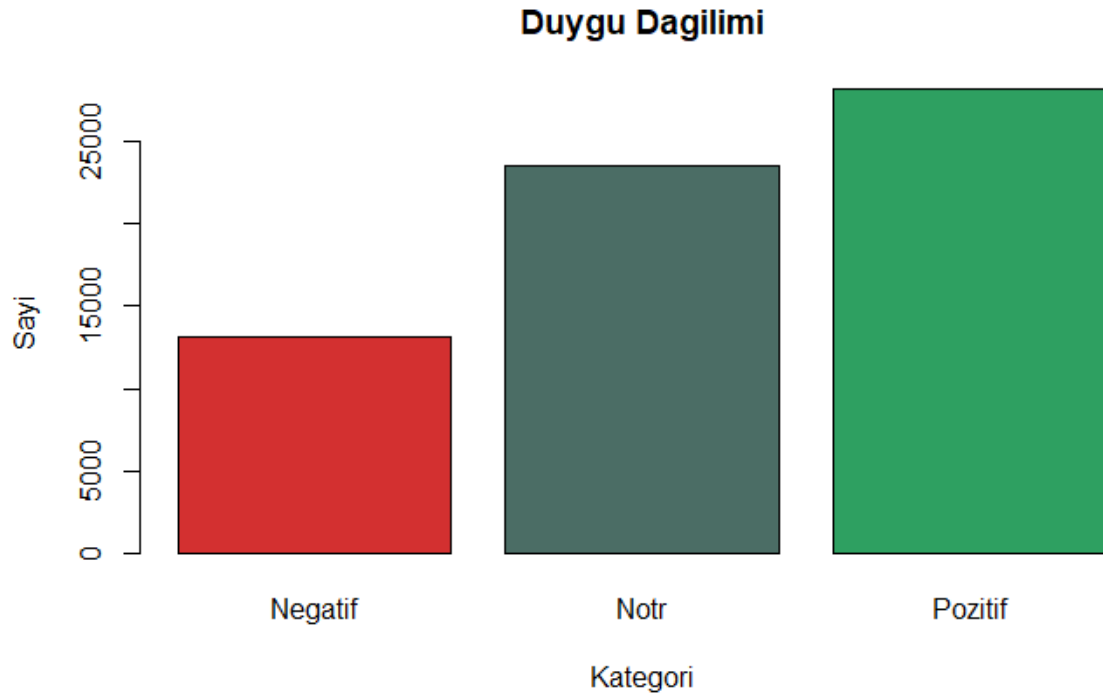


Veri Kaynagi: Youtube'da *The Elder Scrolls VI – Official Announcement Teaser* adlı videodur.

Şekil 21: Polarite görseli



Şekil 22: Negatif ve pozitif kelimelerin görselleştirilmesi



Şekil 23: Duygu dağılımı grafiği



Şekil 24: AFINN ile kelime sıklığını pozitif ve negatif olarak gösterme

Sonuç ve Tartışma

Bethesda Game Studios tarafından geliştirilen ve Bethesda Softworks tarafından 2018 yılında yayınlanan "The Elder Scrolls VI – Official Announcement Teaser" videosu İlk başlarda çok fazla olumlu yorum barındırıyordu stüdyonun oyun ile alakalı güncel bilgiler vermemesi ve ilgisiz davranması sonucunda zaman geçtikçe daha kötü yorumlar yapılmıştır. Emojilerde sıklıkla argo içeren, dalga geçme ile alakalı emojiler atılmıştır. Yıl 2024 olmasına rağmen stüdyo hala daha ilgili oyunla ilgili bir açıklama veya yeni bir video yüklememiş, insanları bilgisiz bırakmaya devam etmiştir. Bunun sonucunda her geçen gün kötü yorum ve emojiler artmaya devam etmektedir.

Kaynaklar

YouTube. (2022). CNET “The Elder Scrolls VI – Official Announcement Teaser”

Adresi:

<https://youtu.be/OkFdqqyI8y4?feature=shared>

Torun N. K. & Şimşek N. (2023) Academicianstudies Sosyal Medya Yasasına Yönelik Kullanıcı Tepkilerinin Çözömlenmesi: YouTube ve Twitter

Adresi: <https://journals.academicianstudies.com/sjissr/article/view/232/154>

Bulut T. (2022). Duygu Analizi Üzerine Bir Vaka Çalışması

Adresi: <https://rpubs.com/tevfik1461/duyguanalizi>