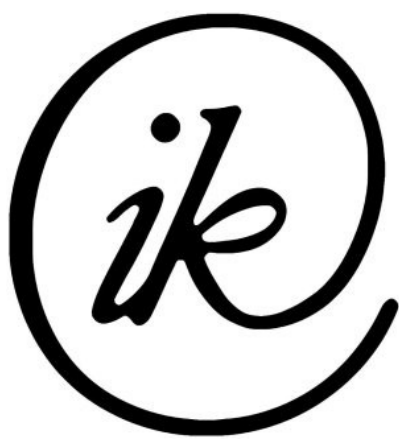




Tudáskezelő rendszerek labor

## **Twitter Super Bowl elemzés**



Készítette : Huzinecz Erik

Neptun kód : SI2T2E

Félév : 2016/2017/1

## 1. Bevezető

A feladat során megvizsgáljuk azt, hogy egy nagy twitter adathalmazból milyen értékes adatokat lehet kinyerni egy esemény vagy termékre vonatkozóan. Az alapvető feladat az, hogy a meglévő adatokat feldolgozzuk, valamint valamilyen algoritmus alapján rendszerezük az adatokat.

## 2. Eszközök és adathalmaz

### 2.1. Eszközök és technológiák

Kutatás és fejlesztés során felhasznált eszközök:

- [Netbeans IDE](#)
- Google Chrome
- NodeJs
- Javacript
- Twitter nagy adathalmaz CSV formátumban

### 2.2. Adathalmaz

A feladat során az ELTE Információs Rendszerek Tanszéke által begyűjtött twitter adatokkal dolgoztam, ebből is azokkal a tweetekkel, amelyek 2013.január és 2016.április közöttiek, valamint szerepel a tweet szövegében az #SB hashtag.

## 3. Történeti háttér

A Super Bowl a fő amerikaifutball-bajnokság, az NFL döntője. Az NFL két konferenciája, az AFC és az NFC győztese játssza egymással. Az USA-ban kisebb fajta nemzeti ünnepnek számít az esemény, milliók nézik a televízióban. Az amerikai futball egyre nagyobb nemzetközi népszerűsége okán a világ többi táján is egyre több nézője van, bár Európában az

időeltolódás miatt hajnalra csúszik a mérkőzés közvetítése, mivel az Egyesült Államokban a reklámok szempontjából oly fontos vasárnap esti főműsoridőben kezdik a mérkőzést. Az első Super Bowlt 1967-ben rendezték a két rivális amerikai futball liga (AFL, NFL) győzteseinek között, ekkor még AFL-NFL World Championship Game néven. Az AFL, és az NFL 1970-es egyesülése óta nevezik a mérkőzést hivatalosan is Super Bowlnak. [1]

Mivel nagyon sokan kedvelik a sportágot, így számomra is megfelelő feladatnak bizonyult az, hogy kutatást végezzek a twitter adatok segítségével. Az USA-ban a twitter felhasználók száma nagyon magas, ami még egy fontos szempont. Úgy vélem, hogy a magas twitter felhasználás, valamint a sportághoz való ragaszkodás alapján értékes adatok nyerhetők ki a twitter adathalmazból.

Az egyes csapatokra való hivatkozást tweetek alapján fogom elemezni.

#### 4. Vizsgálati szempontok

Arra a következtetésre jutottam, hogy a tweeteket dátum, valamint tweet szöveg alapján elemzem. Mivel twitter bejegyzésekről van szó, így feltételezem, hogy a győztes csapatnak szurkoló felhasználók tweetjeiben szerepel valamely a csapatra utaló hashtag.

## 5. Kutatás és hashtagek

Az előző fejezetben arról beszéltem, hogy valamilyen szabályszerűség alapján szűrjük ki a tweeteket. Kutatást végeztem a 2014, 2015, 2016-os super bowl kapcsán, a résztvevő csapatok kapcsán. A kutatás arra irányult, hogy megtudjam a csapatok beceneveit, rövidített neveit, amiből egy hashtag halmazt tudok felállítani az egyes csapatokra vonatkozóan.

### 5.1 Kutatás eredménye

Érintett csapatok és a hozzá tartozó hashtagek, évekre bontva:

2014

- [Seattle Seahawks](#) (győztes)
  - #SEAHAWKSWin
  - #seahawks
  - #GoSeahawks
  - #Goseahawks
  - #goseahawks
  - #Seahawks
  - #SEAHAWKS
  - #Seattle
- Denver Broncos
  - #BroncosWin
  - #GoBroncos
  - #Broncos
  - #Denver
  - #DENVER

2015

- [New England Patriots](#) (győztes)
  - #Pats
  - #PATS
  - #PatsWin

- #PatriotsWin
- #GoPatriots
- #Patriots
- #PATRIOTS
- #NewEngland
- #NEP
- Seattle Seahawks
  - #SEAHAWKSWin
  - #seahawks
  - #GoSeahawks
  - #Goseahawks
  - #goseahawks
  - #Seahawks
  - #SEAHAWKS
  - #Seattle

2016

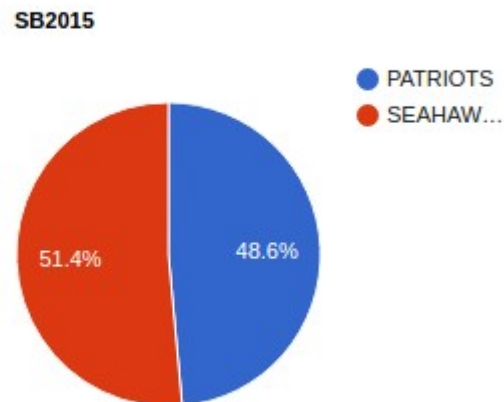
- Carolina Panthers
  - #PanthersWin
  - #pantherswin
  - #panthers
  - #Panthers
  - #Carolina
- [Denver Broncos](#) (győztes)
  - #BroncosWin
  - #GoBroncos
  - #Broncos
  - #Denver
  - #DENVER

## 6. Grafikonok

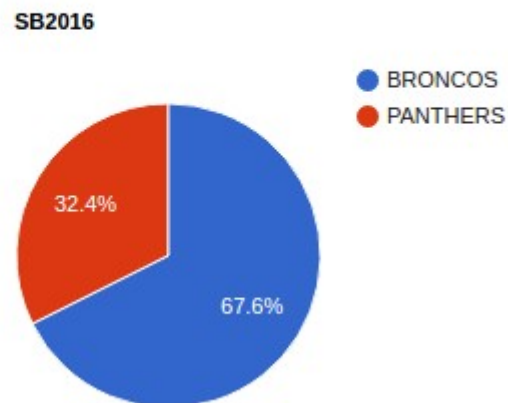
### 6.1. Grafikonok super bowlok szerint

A következő grafikonokon látható az, hogy az egyes csapatok hashtagjeit hányan használták az érintett **super bowl**t megelőző **3 hónapban**

2015, győztes : PATRIOTS



2016, győztes : BRONCOS



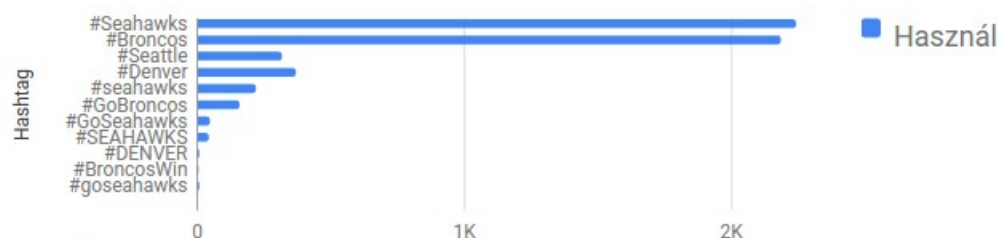
## 6.2. Grafikon hashtag szerint

A következő grafikonokon látható az, hogy az hányan használták az egyes hasztegeket supwer bowlokra bontva az érintett **super bowl**t megelőző **3 hónapban**

SB2015



SB2014



SB2016



## 7. Összefoglaló

A kutatás során bepillantást nyerhettem a twitter, valamint konkrétan a super bowl esemény rejtelseibe. Az eredménytermék, amit készítettem könnyen átalakítható más, egyéb témájú twitter adathalmazt vizsgáló programmá. Az eredmény, amit a program készít egy json fájl, amit bármely programnyelven könnyen fel lehet dolgozni.