



Fan Engagement Case

By: Chia Hsin Hsieh



Agenda



Case Statement & Object



Business Goal



Solving problem

- **Chose Analyze range**
- **Cleaned data and Putted Stopwords**
- **Get the Top Words**



Conclusion



Case Statement & Object

02

<input type="checkbox"/>	 A_Oct2019.csv	82.4 MB	Jan 13, 2021, 1:28 PM
<input type="checkbox"/>	 B_Nov2019.csv	86.1 MB	Jan 13, 2021, 1:28 PM
<input type="checkbox"/>	 C_Dec2019.csv	89.9 MB	Jan 13, 2021, 1:28 PM
<input type="checkbox"/>	 D_Jan2020.csv	99.2 MB	Jan 13, 2021, 1:28 PM
<input type="checkbox"/>	 E_Feb2020.csv	89.2 MB	Jan 13, 2021, 1:28 PM
<input type="checkbox"/>	 F_Mar2020.csv	74.9 MB	Jan 13, 2021, 1:28 PM
<input type="checkbox"/>	 G_Apr2020.csv	49 MB	Jan 13, 2021, 1:28 PM
<input type="checkbox"/>	 H_May2020.csv	50.2 MB	Jan 13, 2021, 1:28 PM
<input type="checkbox"/>	 I_June2020.csv	57.4 MB	Jan 13, 2021, 1:28 PM
<input type="checkbox"/>	 J_July2020.csv	59.8 MB	Jan 13, 2021, 1:28 PM
<input type="checkbox"/>	 K_Aug2020.csv	92.7 MB	Jan 13, 2021, 1:28 PM
<input type="checkbox"/>	 L_Sep2020.csv	65.3 MB	Jan 13, 2021, 1:28 PM
<input type="checkbox"/>	 M_Oct2020.csv	54.4 MB	Jan 13, 2021, 1:28 PM

We were given datasets of NBA fan Tweets engagement from Oct. 2019 to Oct. 2020

Analyze the NBA fan Tweets engagement and helping marketing team understand where to allocate spend.



Business Goal

03



- Maximize the Cost Allocation**
- Expand Market Penetration**
- Increase Revenue**



Solving Problems

04

```
# Source Code  
text      <- read.csv('K_Aug2020.csv',nrows = 10000)  
text      <- read.csv('L_Sep2020.csv',nrows = 10000)  
text      <- read.csv('M_Oct2020.csv',nrows = 10000)
```

The latest data

The trend of attention

The better connect the next month



Solving Problems

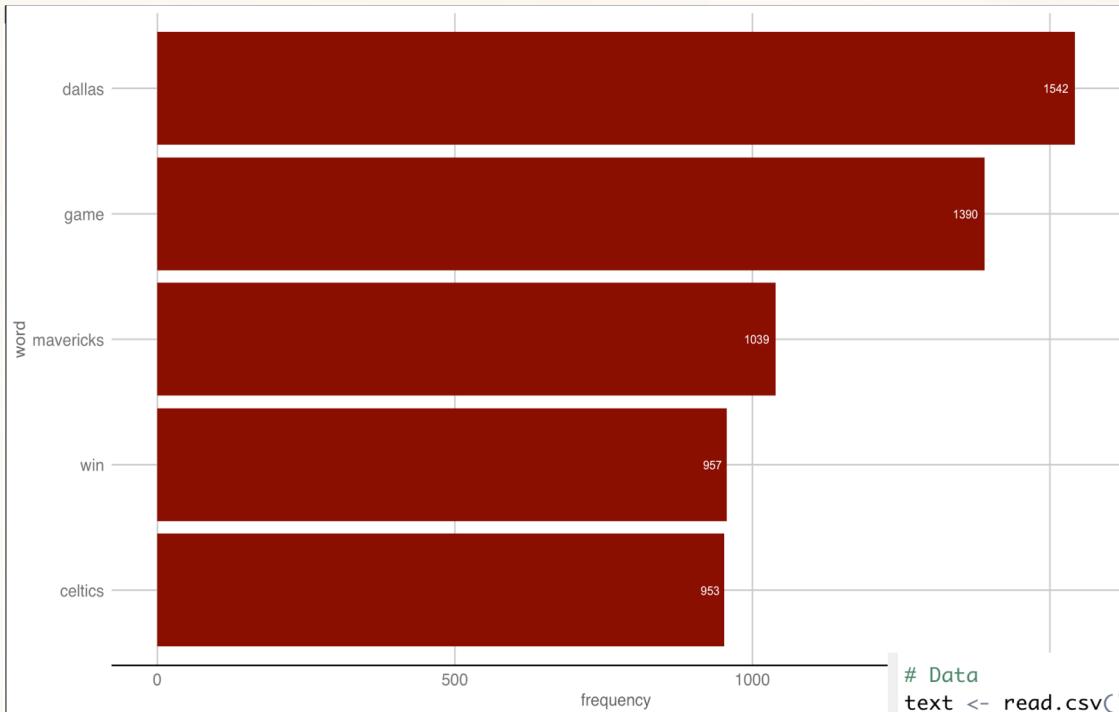
05

```
cleanCorpus<-function(corpus, customStopwords){  
  corpus <- tm_map(corpus, content_transformer(qdapRegex::rm_url))  
  corpus <- tm_map(corpus, removeNumbers)  
  corpus <- tm_map(corpus, removePunctuation)  
  corpus <- tm_map(corpus, stripWhitespace)  
  corpus <- tm_map(corpus, content_transformer(tryToLower))  
  corpus <- tm_map(corpus, removeWords, customStopwords)  
  return(corpus)  
}  
  
#'Stops words  
stops <- c(stopwords('SMART'), 'us', 'won', 'win', 'lose', 'today', 'tonight', 'twitter', 'season'  
      , 'nba', 'score', 'lol', 'amp', 'hawks', 'cavs', 'Atlanta Hawks', 'love')
```

Cleared the data

Putted the Stopwords

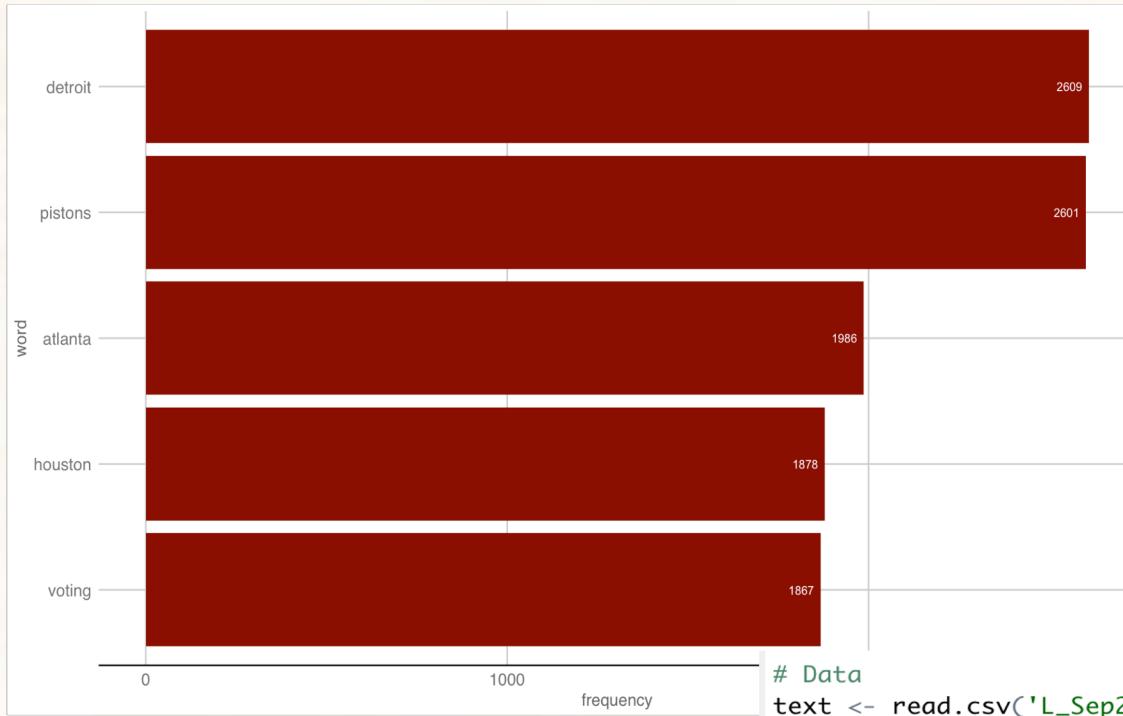
The Top Words of Aug. 2020



```
# Data  
text <- read.csv('K_Aug2020.csv',nrows = 10000)
```

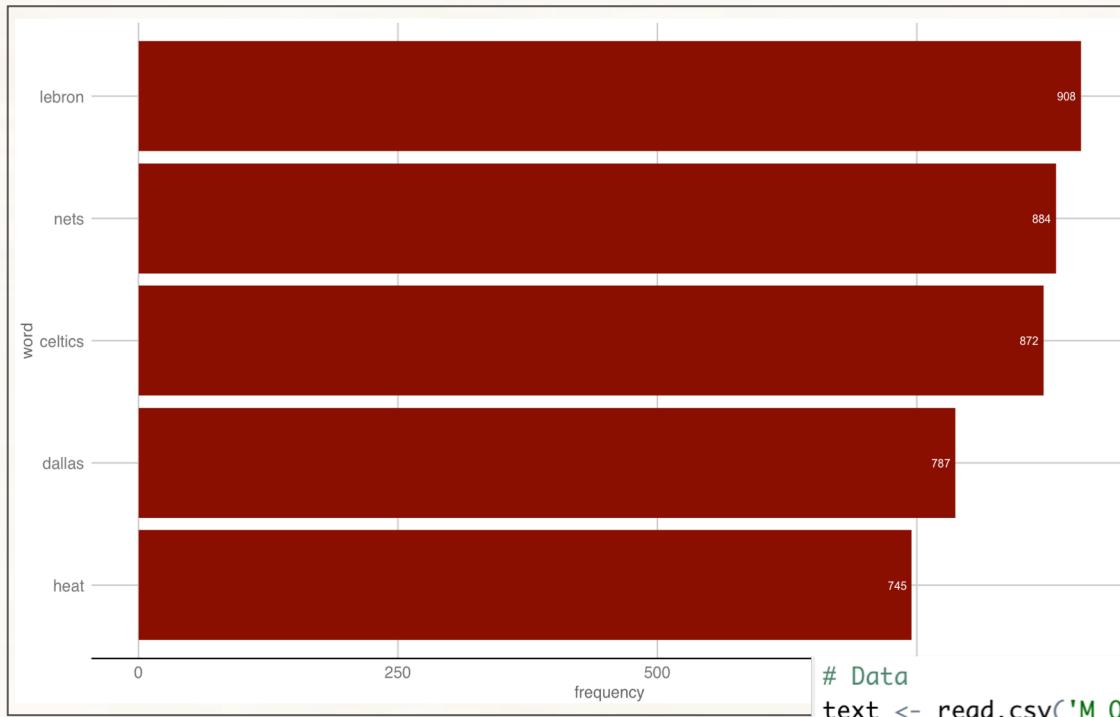
The Top Words of Sep. 2020

07



```
# Data  
text <- read.csv('L_Sep2020.csv',nrows = 10000)
```

The Top Words of Oct 2020



The Top 5 Frequent Result of Each Month

08

Aug. 2020

dallas

game

mavericks

win

celtics

Sep. 2020

destroit

pistons

atlanta

houston

voting

Oct. 2020

lebron

nets

celtics

dallas

heat

The Top 5 Frequent Result of Each Month

08

Aug. 2020

dallas

game

mavericks

win

celtics

Sep. 2020

destroit

pistons

atlanta

houston

voting

Oct. 2020

lebron

nets

celtics

dallas

heat



Conclusion

09

Cooperation with **Celtics** or **Dallas Mavericks**



**THANK YOU FOR
YOUR ATTENTION**

