Analysis

#Read Data  
suppressMessages(library(data.table))  
#Read File  
news = fread("OnlineNewsPopularity.csv", stringsAsFactors=F)  
dim(news)

## [1] 39644 61

#Extract publication date from URL  
news[,pub\_date:= as.Date(gsub(".\*mashable.com/(.\*/.\*/).\*","\\1",url))]

## url  
## 1: http://mashable.com/2013/01/07/amazon-instant-video-browser/  
## 2: http://mashable.com/2013/01/07/ap-samsung-sponsored-tweets/  
## 3: http://mashable.com/2013/01/07/apple-40-billion-app-downloads/  
## 4: http://mashable.com/2013/01/07/astronaut-notre-dame-bcs/  
## 5: http://mashable.com/2013/01/07/att-u-verse-apps/  
## ---   
## 39640: http://mashable.com/2014/12/27/samsung-app-autism/  
## 39641: http://mashable.com/2014/12/27/seth-rogen-james-franco-will-live-tweet-the-interview/  
## 39642: http://mashable.com/2014/12/27/son-pays-off-mortgage/  
## 39643: http://mashable.com/2014/12/27/ukraine-blasts/  
## 39644: http://mashable.com/2014/12/27/youtube-channels-2015/  
## timedelta n\_tokens\_title n\_tokens\_content n\_unique\_tokens  
## 1: 731 12 219 0.6635945  
## 2: 731 9 255 0.6047431  
## 3: 731 9 211 0.5751295  
## 4: 731 9 531 0.5037879  
## 5: 731 13 1072 0.4156456  
## ---   
## 39640: 8 11 346 0.5290520  
## 39641: 8 12 328 0.6962963  
## 39642: 8 10 442 0.5163551  
## 39643: 8 6 682 0.5394933  
## 39644: 8 10 157 0.7019868  
## n\_non\_stop\_words n\_non\_stop\_unique\_tokens num\_hrefs num\_self\_hrefs  
## 1: 1 0.8153846 4 2  
## 2: 1 0.7919463 3 1  
## 3: 1 0.6638655 3 1  
## 4: 1 0.6656347 9 0  
## 5: 1 0.5408895 19 19  
## ---   
## 39640: 1 0.6847826 9 7  
## 39641: 1 0.8850575 9 7  
## 39642: 1 0.6441281 24 1  
## 39643: 1 0.6926605 10 1  
## 39644: 1 0.8461538 1 1  
## num\_imgs num\_videos average\_token\_length num\_keywords  
## 1: 1 0 4.680365 5  
## 2: 1 0 4.913725 4  
## 3: 1 0 4.393365 6  
## 4: 1 0 4.404896 7  
## 5: 20 0 4.682836 7  
## ---   
## 39640: 1 1 4.523121 8  
## 39641: 3 48 4.405488 7  
## 39642: 12 1 5.076923 8  
## 39643: 1 0 4.975073 5  
## 39644: 0 2 4.471338 4  
## data\_channel\_is\_lifestyle data\_channel\_is\_entertainment  
## 1: 0 1  
## 2: 0 0  
## 3: 0 0  
## 4: 0 1  
## 5: 0 0  
## ---   
## 39640: 0 0  
## 39641: 0 0  
## 39642: 0 0  
## 39643: 0 0  
## 39644: 0 1  
## data\_channel\_is\_bus data\_channel\_is\_socmed data\_channel\_is\_tech  
## 1: 0 0 0  
## 2: 1 0 0  
## 3: 1 0 0  
## 4: 0 0 0  
## 5: 0 0 1  
## ---   
## 39640: 0 0 1  
## 39641: 0 1 0  
## 39642: 0 0 0  
## 39643: 0 0 0  
## 39644: 0 0 0  
## data\_channel\_is\_world kw\_min\_min kw\_max\_min kw\_avg\_min kw\_min\_max  
## 1: 0 0 0 0.000 0  
## 2: 0 0 0 0.000 0  
## 3: 0 0 0 0.000 0  
## 4: 0 0 0 0.000 0  
## 5: 0 0 0 0.000 0  
## ---   
## 39640: 0 -1 671 173.125 26900  
## 39641: 0 -1 616 184.000 6500  
## 39642: 0 -1 691 168.250 6200  
## 39643: 1 -1 0 -1.000 0  
## 39644: 0 -1 97 23.500 205600  
## kw\_max\_max kw\_avg\_max kw\_min\_avg kw\_max\_avg kw\_avg\_avg  
## 1: 0 0.0 0.000 0.000 0.000  
## 2: 0 0.0 0.000 0.000 0.000  
## 3: 0 0.0 0.000 0.000 0.000  
## 4: 0 0.0 0.000 0.000 0.000  
## 5: 0 0.0 0.000 0.000 0.000  
## ---   
## 39640: 843300 374962.5 2514.743 4004.343 3031.116  
## 39641: 843300 192985.7 1664.268 5470.169 3411.661  
## 39642: 843300 295850.0 1753.882 6880.687 4206.439  
## 39643: 843300 254600.0 0.000 3384.317 1777.896  
## 39644: 843300 366200.0 3035.081 3613.513 3296.909  
## self\_reference\_min\_shares self\_reference\_max\_shares  
## 1: 496 496  
## 2: 0 0  
## 3: 918 918  
## 4: 0 0  
## 5: 545 16000  
## ---   
## 39640: 11400 48000  
## 39641: 2100 2100  
## 39642: 1400 1400  
## 39643: 452 452  
## 39644: 2100 2100  
## self\_reference\_avg\_sharess weekday\_is\_monday weekday\_is\_tuesday  
## 1: 496.000 1 0  
## 2: 0.000 1 0  
## 3: 918.000 1 0  
## 4: 0.000 1 0  
## 5: 3151.158 1 0  
## ---   
## 39640: 37033.333 0 0  
## 39641: 2100.000 0 0  
## 39642: 1400.000 0 0  
## 39643: 452.000 0 0  
## 39644: 2100.000 0 0  
## weekday\_is\_wednesday weekday\_is\_thursday weekday\_is\_friday  
## 1: 0 0 0  
## 2: 0 0 0  
## 3: 0 0 0  
## 4: 0 0 0  
## 5: 0 0 0  
## ---   
## 39640: 1 0 0  
## 39641: 1 0 0  
## 39642: 1 0 0  
## 39643: 1 0 0  
## 39644: 1 0 0  
## weekday\_is\_saturday weekday\_is\_sunday is\_weekend LDA\_00  
## 1: 0 0 0 0.50033120  
## 2: 0 0 0 0.79975569  
## 3: 0 0 0 0.21779229  
## 4: 0 0 0 0.02857322  
## 5: 0 0 0 0.02863281  
## ---   
## 39640: 0 0 0 0.02503777  
## 39641: 0 0 0 0.02934870  
## 39642: 0 0 0 0.15900446  
## 39643: 0 0 0 0.04000361  
## 39644: 0 0 0 0.05000126  
## LDA\_01 LDA\_02 LDA\_03 LDA\_04 global\_subjectivity  
## 1: 0.37827893 0.04000468 0.04126265 0.04012254 0.5216171  
## 2: 0.05004668 0.05009625 0.05010067 0.05000071 0.3412458  
## 3: 0.03333446 0.03335142 0.03333354 0.68218829 0.7022222  
## 4: 0.41929964 0.49465083 0.02890472 0.02857160 0.4298497  
## 5: 0.02879355 0.02857518 0.02857168 0.88542678 0.5135021  
## ---   
## 39640: 0.02500062 0.15170116 0.02500011 0.77326035 0.4826786  
## 39641: 0.02857493 0.23186607 0.68163487 0.02857542 0.5643743  
## 39642: 0.02502466 0.02520734 0.64379353 0.14697000 0.5102958  
## 39643: 0.04000349 0.83998726 0.04000210 0.04000355 0.3585776  
## 39644: 0.79933895 0.05000041 0.05065874 0.05000064 0.5178932  
## global\_sentiment\_polarity global\_rate\_positive\_words  
## 1: 0.092561983 0.04566210  
## 2: 0.148947811 0.04313725  
## 3: 0.323333333 0.05687204  
## 4: 0.100704666 0.04143126  
## 5: 0.281003476 0.07462687  
## ---   
## 39640: 0.141964286 0.03757225  
## 39641: 0.194249311 0.03963415  
## 39642: 0.024608586 0.03393665  
## 39643: -0.008065863 0.02052786  
## 39644: 0.104891775 0.06369427  
## global\_rate\_negative\_words rate\_positive\_words rate\_negative\_words  
## 1: 0.013698630 0.7692308 0.2307692  
## 2: 0.015686275 0.7333333 0.2666667  
## 3: 0.009478673 0.8571429 0.1428571  
## 4: 0.020715631 0.6666667 0.3333333  
## 5: 0.012126866 0.8602151 0.1397849  
## ---   
## 39640: 0.014450867 0.7222222 0.2777778  
## 39641: 0.009146341 0.8125000 0.1875000  
## 39642: 0.024886878 0.5769231 0.4230769  
## 39643: 0.023460411 0.4666667 0.5333333  
## 39644: 0.012738854 0.8333333 0.1666667  
## avg\_positive\_polarity min\_positive\_polarity max\_positive\_polarity  
## 1: 0.3786364 0.10000000 0.70  
## 2: 0.2869146 0.03333333 0.70  
## 3: 0.4958333 0.10000000 1.00  
## 4: 0.3859652 0.13636364 0.80  
## 5: 0.4111274 0.03333333 1.00  
## ---   
## 39640: 0.3337912 0.10000000 0.75  
## 39641: 0.3748252 0.13636364 0.70  
## 39642: 0.3072727 0.13636364 0.50  
## 39643: 0.2368506 0.06250000 0.50  
## 39644: 0.2473377 0.10000000 0.50  
## avg\_negative\_polarity min\_negative\_polarity max\_negative\_polarity  
## 1: -0.3500000 -0.600 -0.2000000  
## 2: -0.1187500 -0.125 -0.1000000  
## 3: -0.4666667 -0.800 -0.1333333  
## 4: -0.3696970 -0.600 -0.1666667  
## 5: -0.2201923 -0.500 -0.0500000  
## ---   
## 39640: -0.2600000 -0.500 -0.1250000  
## 39641: -0.2111111 -0.400 -0.1000000  
## 39642: -0.3564394 -0.800 -0.1666667  
## 39643: -0.2052455 -0.500 -0.0125000  
## 39644: -0.2000000 -0.200 -0.2000000  
## title\_subjectivity title\_sentiment\_polarity abs\_title\_subjectivity  
## 1: 0.5000000 -0.1875000 0.00000000  
## 2: 0.0000000 0.0000000 0.50000000  
## 3: 0.0000000 0.0000000 0.50000000  
## 4: 0.0000000 0.0000000 0.50000000  
## 5: 0.4545455 0.1363636 0.04545455  
## ---   
## 39640: 0.1000000 0.0000000 0.40000000  
## 39641: 0.3000000 1.0000000 0.20000000  
## 39642: 0.4545455 0.1363636 0.04545455  
## 39643: 0.0000000 0.0000000 0.50000000  
## 39644: 0.3333333 0.2500000 0.16666667  
## abs\_title\_sentiment\_polarity shares pub\_date  
## 1: 0.1875000 593 2013-01-07  
## 2: 0.0000000 711 2013-01-07  
## 3: 0.0000000 1500 2013-01-07  
## 4: 0.0000000 1200 2013-01-07  
## 5: 0.1363636 505 2013-01-07  
## ---   
## 39640: 0.0000000 1800 2014-12-27  
## 39641: 1.0000000 1900 2014-12-27  
## 39642: 0.1363636 1900 2014-12-27  
## 39643: 0.0000000 1100 2014-12-27  
## 39644: 0.2500000 1300 2014-12-27

#Question 1  
#Unique URL  
print("Unique number of URLs")

## [1] "Unique number of URLs"

print(length(unique(news$url)), row.names=F)

## [1] 39644

#As there is no duplicates URL all the records are unique   
#Find the largest lag between publication and data collection   
cat("\nLargest lag (in days) between publication and data collection")

##   
## Largest lag (in days) between publication and data collection

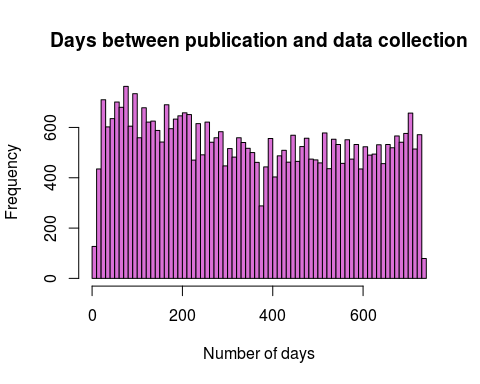
print(max(news$timedelta), row.names=F)

## [1] 731

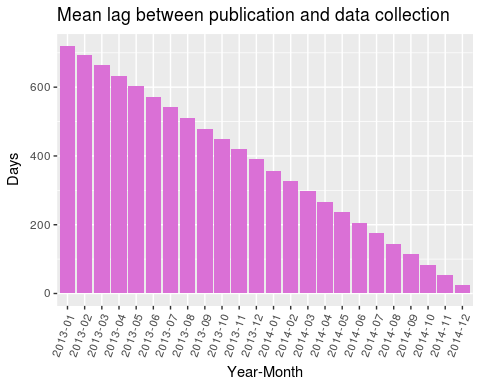
#Time Frame of publication  
summary(news$pub\_date)

## Min. 1st Qu. Median Mean 3rd Qu.   
## "2013-01-07" "2013-07-15" "2014-02-03" "2014-01-18" "2014-07-27"   
## Max.   
## "2014-12-27"

#Question 2  
#Histogram of timedelta  
hist(news$timedelta,breaks = 73,  
 main="Days between publication and data collection",  
 xlab="Number of days", col="orchid")  
suppressMessages(library(ggplot2))



#Timedelta by month  
timedelta\_bymonth = news[,.(mean\_timedelta=mean(timedelta)),  
 by=(month=format(pub\_date,"%Y-%m"))]  
ggplot(timedelta\_bymonth) +   
 geom\_bar(aes(x=month,y=mean\_timedelta),stat="identity",fill="orchid") +  
 labs(x="Year-Month",y="Days",  
 title="Mean lag between publication and data collection") +  
 theme(axis.text.x = element\_text(angle = 70, hjust = 1))



#Question 3  
#Extract topic from URL  
topics <- gsub(".\*/(.\*)/$","\\1",news$url)  
#Get topic Frequency  
topics\_freq <- as.data.frame(table(topics))  
#Topics are unique, hence maximum frequency is 1  
top20 <- topics\_freq[order(topics\_freq$Freq,decreasing = T)[1:20],]  
kable(top20, caption="\*\*Topic Frequency\*\*") #Topics are unique, hence maximum frequency is 1

**Topic Frequency**

|  |  |
| --- | --- |
| topics | Freq |
| 100000-people-join-islamic-state | 1 |
| 100-dates-of-summer | 1 |
| 100-foot-tall-six-pack-beer | 1 |
| 100-influential-books-facebook-data | 1 |
| 100-million-active-users-facebook-india | 1 |
| 100-most-shared-vines | 1 |
| 100-pound-chocolate-bar | 1 |
| 100-us-senators-interactive-data-visualization | 1 |
| 100-year-message-in-a-bottle | 1 |
| 100-year-old-moms | 1 |
| 100-year-old-time-capsule | 1 |
| 100-years-of-makeup | 1 |
| 100-year-starship-voyage | 1 |
| 101-career-tips | 1 |
| 102-year-old-woman-base-jump | 1 |
| 105-impressions | 1 |
| 105-year-old-man-high-school-diploma | 1 |
| 108-dogs-rescued | 1 |
| 10-awesome-stem-courses | 1 |
| 10-awesome-stem-jobs | 1 |

# \*\*Try topic frequency by first word and last word\*\*  
# Get firstword  
topic\_firstword <- gsub("-.\*","",topics)  
# Get lastword  
topic\_lastword <- gsub(".\*-","",topics)  
#Frequency by first word  
firstword\_freq <- as.data.frame(table(topic\_firstword))  
#Frequency by last word  
lastword\_freq <- as.data.frame(table(topic\_lastword))  
#Get top 25   
top25\_firstword <- firstword\_freq[order(firstword\_freq$Freq,  
 decreasing = T)[1:25],]  
top25\_lastword <- lastword\_freq[order(lastword\_freq$Freq,  
 decreasing = T)[1:25],]  
kable(top25\_firstword,caption="Top 25 First word of topic")

Top 25 First word of topic

|  |  |  |
| --- | --- | --- |
|  | topic\_firstword | Freq |
| 3124 | google | 957 |
| 2558 | facebook | 737 |
| 441 | apple | 723 |
| 7947 | twitter | 530 |
| 7640 | the | 338 |
| 6596 | samsung | 330 |
| 5256 | new | 283 |
| 343 | amazon | 254 |
| 4856 | microsoft | 233 |
| 7092 | social | 223 |
| 3807 | iphone | 201 |
| 5412 | obama | 197 |
| 8532 | world | 193 |
| 2926 | game | 188 |
| 5165 | nasa | 171 |
| 3751 | instagram | 162 |
| 7781 | top | 160 |
| 8639 | youtube | 155 |
| 3535 | how | 153 |
| 8586 | yahoo | 146 |
| 7259 | star | 145 |
| 7134 | sony | 143 |
| 821 | best | 137 |
| 8205 | vine | 129 |
| 88 | 3d | 114 |

kable(top25\_lastword,caption="Top 25 Last word of topic" )

Top 25 Last word of topic

|  |  |  |
| --- | --- | --- |
|  | topic\_lastword | Freq |
| 9190 | video | 469 |
| 55 | 2 | 348 |
| 66 | 2014 | 334 |
| 544 | app | 318 |
| 1850 | comic | 295 |
| 1237 | brief | 262 |
| 7201 | review | 262 |
| 8959 | twitter | 260 |
| 6387 | photos | 253 |
| 8801 | trailer | 204 |
| 286 | ad | 192 |
| 65 | 2013 | 180 |
| 564 | apps | 173 |
| 3684 | gifs | 160 |
| 8700 | tips | 143 |
| 3118 | facebook | 122 |
| 2308 | day | 112 |
| 1583 | challenge | 110 |
| 6020 | on | 108 |
| 5341 | media | 106 |
| 309 | ads | 104 |
| 3135 | facts | 102 |
| 3597 | game | 102 |
| 4413 | instagram | 100 |
| 6967 | recap | 99 |

#Question 4  
subtopics <- c("elon-musk","facebook","ebola","ipad","iphone",  
 "tornado","sharknado","taylor-swift")  
for (s in subtopics){  
 #Add a column using subtopic name and set its initial value to False  
 news[,s:=F,with=F]  
 #Find URLs that has the subtopic as substring  
 sub\_present <- grep(s,news$url)  
 #Set the subtopic column to true if it is substring of url  
 news[sub\_present,s:=T,with=F]  
}  
#Count by subtopics  
kable(news[,lapply(.SD,sum),.SDcols=subtopics],  
 caption="Count by sub-topics (substring in url)")

Count by sub-topics (substring in url)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| elon-musk | facebook | ebola | ipad | iphone | tornado | sharknado | taylor-swift |
| 37 | 1109 | 261 | 286 | 578 | 51 | 25 | 77 |

#Question 5  
#Count of subtopics by month  
sum\_by\_subtop\_month <- news[,lapply(.SD,sum),  
 .SDcols=subtopics,  
 by=(Year\_Month=format(pub\_date,"%Y-%m"))]  
kable(sum\_by\_subtop\_month,caption="Count of sub-topics by Month")

Count of sub-topics by Month

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Year\_Month | elon-musk | facebook | ebola | ipad | iphone | tornado | sharknado | taylor-swift |
| 2013-01 | 0 | 79 | 0 | 12 | 37 | 0 | 0 | 1 |
| 2013-02 | 2 | 57 | 0 | 14 | 22 | 0 | 0 | 2 |
| 2013-03 | 2 | 66 | 0 | 14 | 26 | 0 | 0 | 2 |
| 2013-04 | 1 | 80 | 0 | 10 | 32 | 1 | 0 | 1 |
| 2013-05 | 3 | 55 | 0 | 15 | 19 | 17 | 0 | 1 |
| 2013-06 | 0 | 59 | 0 | 9 | 22 | 1 | 0 | 3 |
| 2013-07 | 1 | 45 | 0 | 9 | 20 | 0 | 12 | 3 |
| 2013-08 | 6 | 70 | 0 | 8 | 33 | 0 | 1 | 1 |
| 2013-09 | 1 | 44 | 0 | 16 | 105 | 0 | 0 | 2 |
| 2013-10 | 2 | 55 | 0 | 52 | 34 | 0 | 0 | 1 |
| 2013-11 | 1 | 28 | 0 | 26 | 13 | 2 | 0 | 1 |
| 2013-12 | 0 | 37 | 0 | 8 | 13 | 0 | 0 | 2 |
| 2014-01 | 0 | 47 | 0 | 2 | 11 | 0 | 0 | 2 |
| 2014-02 | 0 | 67 | 0 | 4 | 11 | 0 | 0 | 0 |
| 2014-03 | 1 | 40 | 2 | 10 | 9 | 0 | 0 | 0 |
| 2014-04 | 2 | 47 | 3 | 3 | 9 | 17 | 3 | 1 |
| 2014-05 | 3 | 32 | 1 | 6 | 5 | 2 | 0 | 0 |
| 2014-06 | 4 | 37 | 1 | 3 | 16 | 6 | 1 | 0 |
| 2014-07 | 0 | 48 | 12 | 5 | 9 | 2 | 8 | 1 |
| 2014-08 | 2 | 25 | 34 | 4 | 16 | 1 | 0 | 8 |
| 2014-09 | 0 | 20 | 22 | 3 | 90 | 0 | 0 | 3 |
| 2014-10 | 4 | 35 | 152 | 42 | 14 | 1 | 0 | 17 |
| 2014-11 | 2 | 19 | 22 | 4 | 8 | 0 | 0 | 21 |
| 2014-12 | 0 | 17 | 12 | 7 | 4 | 1 | 0 | 4 |

#  
cols\_to\_analyze <- c("shares","num\_videos","num\_imgs",  
 "abs\_title\_subjectivity","abs\_title\_sentiment\_polarity")  
kable(news[,lapply(.SD,mean),.SDcols=cols\_to\_analyze,by=is\_weekend],  
 caption="Mean values Weekend Vs Non-Weekend")

Mean values Weekend Vs Non-Weekend

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| is\_weekend | shares | num\_videos | num\_imgs | abs\_title\_subjectivity | abs\_title\_sentiment\_polarity |
| 0 | 3318.855 | 1.275469 | 4.372439 | 0.3444084 | 0.1528588 |
| 1 | 3903.394 | 1.079962 | 5.684008 | 0.3248107 | 0.1773390 |

#Number of Shares Weekend Vs non-weekend for Entertainment News  
entertain\_shares <- news[data\_channel\_is\_entertainment > 0,lapply(.SD,mean),  
 .SDcols=cols\_to\_analyze,  
 by=is\_weekend]  
#Number of Shares Weekend Vs non-weekend for lifestyle News  
lifestyle\_shares <- news[data\_channel\_is\_lifestyle > 0,lapply(.SD,mean),  
 .SDcols=cols\_to\_analyze,  
 by=is\_weekend]  
#Number of Shares Weekend Vs non-weekend for Tech News  
tech\_shares <- news[data\_channel\_is\_tech > 0,lapply(.SD,mean),  
 .SDcols=cols\_to\_analyze,  
 by=is\_weekend]  
#Number of Shares Weekend Vs non-weekend for World News  
world\_shares <- news[data\_channel\_is\_world > 0,  
 lapply(.SD,mean),  
 .SDcols=cols\_to\_analyze,by=is\_weekend]  
kable(entertain\_shares,   
 caption="Mean values Weekend Vs Non-Weekend for Entertainment News")

Mean values Weekend Vs Non-Weekend for Entertainment News

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| is\_weekend | shares | num\_videos | num\_imgs | abs\_title\_subjectivity | abs\_title\_sentiment\_polarity |
| 0 | 2869.537 | 2.543722 | 6.088585 | 0.3230718 | 0.1709698 |
| 1 | 3647.273 | 2.560044 | 7.853712 | 0.3073490 | 0.1774945 |

kable(lifestyle\_shares,  
 caption="Mean values Weekend Vs Non-Weekend for lifestyle News")

Mean values Weekend Vs Non-Weekend for lifestyle News

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| is\_weekend | shares | num\_videos | num\_imgs | abs\_title\_subjectivity | abs\_title\_sentiment\_polarity |
| 0 | 3628.255 | 0.4633861 | 4.066198 | 0.3527943 | 0.1628971 |
| 1 | 3916.696 | 0.5255102 | 8.556122 | 0.3234864 | 0.2183741 |

kable(tech\_shares,   
 caption="Mean values Weekend Vs Non-Weekend for Tech News")

Mean values Weekend Vs Non-Weekend for Tech News

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| is\_weekend | shares | num\_videos | num\_imgs | abs\_title\_subjectivity | abs\_title\_sentiment\_polarity |
| 0 | 2974.685 | 0.4491829 | 4.400311 | 0.3482665 | 0.1300979 |
| 1 | 3753.143 | 0.4332248 | 4.673181 | 0.3197245 | 0.1774676 |

kable(world\_shares,   
 caption="Mean values Weekend Vs Non-Weekend for World News")

Mean values Weekend Vs Non-Weekend for World News

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| is\_weekend | shares | num\_videos | num\_imgs | abs\_title\_subjectivity | abs\_title\_sentiment\_polarity |
| 0 | 2229.789 | 0.5452936 | 2.824275 | 0.3608469 | 0.1271594 |
| 1 | 2679.424 | 0.5782689 | 2.955801 | 0.3632252 | 0.1315415 |