

Pràctica 1

1. La informació recollida es basa en les eleccions generals del 7 d'abril de 2019 del Principat d'Andorra.

La idea principal és recollir els diferents titulars publicats pels medis digitals andorrans durant la campanya electoral del 2019. A efecte de determinar si existeix una desigualtat de publicitat en el conjunt de partits polítics. S'ha tractat aquest tema donat que és un tema d'actualitat i de summa importància pel principat.

Convé subratllar que la societat, per comoditat, s'informa cada vegada més a través dels medis digitals. Per consegüent, s'han triat els medis digitals més influents en la societat andorrana. A continuació se n'exposa el llistat:

- www.forum.ad
- www.altaveudigital.com
- www.ara.ad

Finalment, la pràctica ha estat realitzada íntegrament i de forma individual per *Gil Santolària Rossell*.

2. Preferències polítiques dels principals medis digitals en la campanya electoral andorrana del 2019.
3. S'han extret les dades dels diferents titulars publicats pels medis citats anteriorment en el període electoral. A continuació, s'ha dut a terme una cerca de les mencions de cada partit, ja sigui sota el seu nom/acrònim o bé citant els diferents caps de llista que el formaven. Finalment, s'ha realitzat un recompte de les mencions per a cada un dels partits.

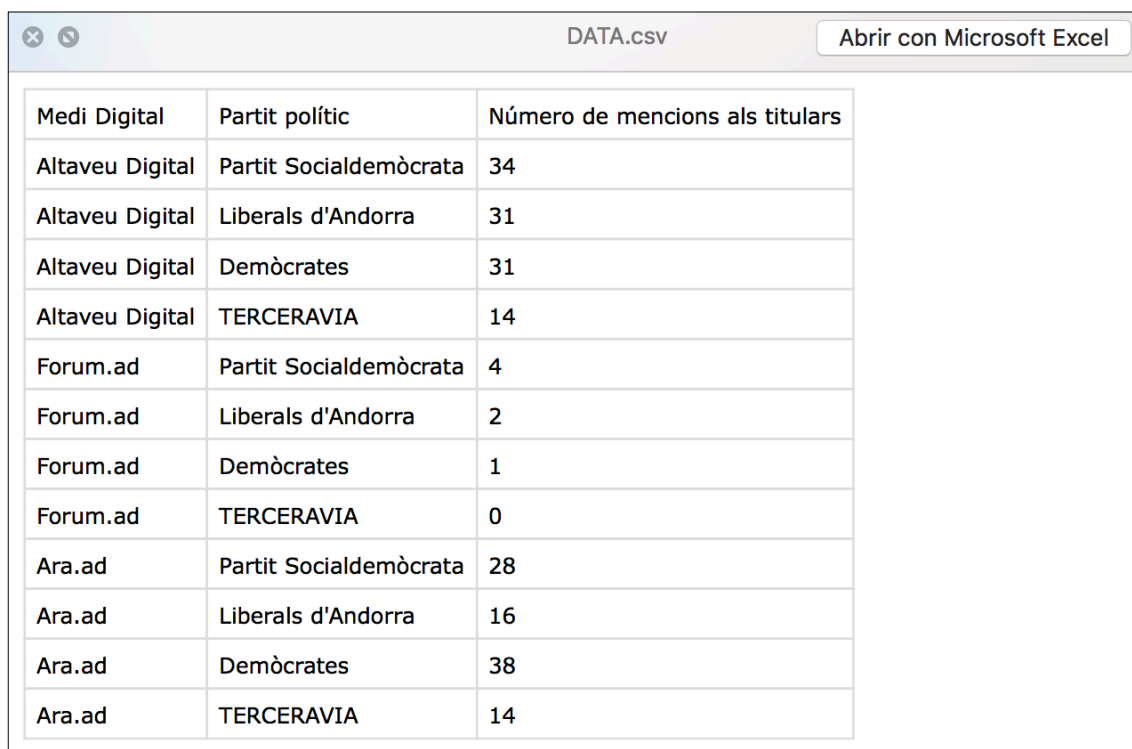
4.



Figura 1. Resultat de les eleccions generals 2019

5. El dataset està format per tres columnes, la primera fent referència medi digital d'on s'ha extret la informació, la segona al partit polític en qüestió i la tercera columna correspon al recompte de mencions en els titulars emesos durant la campanya electoral andorrana (*Període: del 01/03/19 al 05/04/19*). S'ha recollit la informació necessària mitjançant un codi Python que utilitza la llibreria "BeautifulSoup" per facilitar el scraping web dels diferents medis digitals seleccionats.

6. Es volen agrair els diferents medis digitals per dur a terme la feina de recerca i publicació i, també, per difondre tota l'actualitat andorrana. La informació i els resultats de les eleccions generals del 2019 del govern d'Andorra es poden consultar en el següent enllaç: www.eleccions.ad
7. Primerament s'intenta donar a conèixer un tema d'actualitat en una societat andorrana fracturada per l'anterior govern, on les polítiques socials cada dia agafen més força. En segon lloc, a través del dataset, es pretén esbrinar si els medis andorrans tenen preferències polítiques i si aquests tenen una influència directa en la decisió de vot del ciutadà. Finalment, cal dir que la recerca s'ha centrat en les eleccions actuals donat que alguns dels medis o bé no existien fa quatre anys o bé no tenien la suficient repercussió en la societat.
8. Com a llicència pel data set resultant s'ha escollit:
Released Under CC0: Public Domain License.
Donat que se'n renuncia tots els drets sobre l'obra pel que fa a la propietat intel·lectual, incloent-hi els drets afins, fins on és possible amb la llei vigent aplicable.
Per acabar, s'ha decidit la selecció d'aquesta llicència en concret ja que el dataset resultant parteix d'un scraping web i no d'un dataset previ. Aleshores, no es disposa de cap dret en la informació resultant presentada.
9. Dataset en format CSV:



Medi Digital	Partit polític	Número de mencions als titulars
Altaveu Digital	Partit Socialdemòcrata	34
Altaveu Digital	Liberals d'Andorra	31
Altaveu Digital	Demòcrates	31
Altaveu Digital	TERCERAVIA	14
Forum.ad	Partit Socialdemòcrata	4
Forum.ad	Liberals d'Andorra	2
Forum.ad	Demòcrates	1
Forum.ad	TERCERAVIA	0
Ara.ad	Partit Socialdemòcrata	28
Ara.ad	Liberals d'Andorra	16
Ara.ad	Demòcrates	38
Ara.ad	TERCERAVIA	14

Figura 2. Dataset obtingut.

10. Codi Python:

```

#!/usr/bin/env python3
# -*- coding: utf-8 -*-
"""
Created on Tue Apr  9 20:22:27 2019

@author: Gil
File: Scraping_1.py
"""

from urllib.request import urlopen
from bs4 import BeautifulSoup
import ssl
import csv

ctx = ssl.create_default_context()
ctx.check_hostname = False
ctx.verify_mode = ssl.CERT_NONE

i=8
llista=list()

PS_count=0
Li_count=0
DA_count=0
TV_count=0

while i>1:
    url = 'https://www.altaveu.com/categoria/eleccions-generals/pagina/'
    html = urlopen(url, context=ctx).read()
    soup = BeautifulSoup(html, "html.parser")

    tag1 = soup('h2')

    for tag in tag1:
        llista.append((tag.contents[0]))

    i=i-1

for x in llista:
    PS=x.find('PS')
    PS_1=x.find('López')
    PS_2=x.find("d'Acord")
    PS_3=x.find("socialdemòcrates")
    if PS != -1 or PS_1 !=-1 or PS_2 !=-1:
        PS_count=PS_count+1
        PS=-1
        PS_1=-1
        PS_3=-1

```

```

Li=x.find('Liberals')
Li_1=x.find('Gallardo')
Li_2=x.find("L'A")

if Li != -1 or Li_1 !=-1 or PS_2 !=-1 or Li_2 !=-1:
    Li_count=Li_count+1
    Li =-1
    Li_1 =-1
    PS_2=-1
    Li_2 =-1

DA=x.find('DA')
DA_1=x.find('Demòcrates')
DA_2=x.find('Espot')

if DA !=-1 or DA_1 !=-1 or DA_2 !=-1:
    DA_count = DA_count+1
    DA=-1
    DA_1=-1
    DA_2=-1

TV=x.find('TERCERAVIA')
TV_1=x.find('Pintat')
TV_2=x.find('Terceravia')

if TV !=-1 or TV_1 !=-1 or TV_2 !=-1 :
    TV_count=TV_count+1
    TV=-1
    TV_1=-1
    TV_2=-1

print ('Resultats Altaveu')
print('PS:',PS_count)
print('Liberals:',Li_count)
print('Demòcrates:',DA_count)
print('TERCERAVIA:',TV_count)

row = ["Medi Digital", "Partit polític", "Número de mencions als titula
row1 = ["Altaveu Digital", "Partit Socialdemòcrata", PS_count]
row2= ["Altaveu Digital", "Liberals d'Andorra", Li_count]
row3= ["Altaveu Digital", "Demòcrates", DA_count]
row4= ["Altaveu Digital", "TERCERAVIA", TV_count]

with open ('DATA.csv', 'w') as csvFile:
    writer = csv.writer(csvFile)
    writer.writerows(row)

```

```
csvFile.close()

with open('DATA.csv', 'a') as csvFile:
    writer = csv.writer(csvFile)
    writer.writerow(row1)
    writer.writerow(row2)
    writer.writerow(row3)
    writer.writerow(row4)

csvFile.close()
```

```

#!/usr/bin/env python3
# -*- coding: utf-8 -*-
"""
Created on Wed Apr 10 01:15:05 2019

@author: Gil
File: Scraping_2.py
"""

from urllib.request import urlopen
from bs4 import BeautifulSoup
import ssl
import csv

ctx = ssl.create_default_context()
ctx.check_hostname = False
ctx.verify_mode = ssl.CERT_NONE

llista=list()
PS_count=0
Li_count=0
DA_count=0
TV_count=0

url = 'https://forum.ad/category/politica/page/2'
html = urlopen(url, context=ctx).read()
soup = BeautifulSoup(html, "html.parser")

tags = soup('h2')

for tag in tags:
    llista.append(tag.string)

for x in llista:
    PS=x.find('PS')
    PS_1=x.find('López')
    PS_2=x.find("d'Acord")
    PS_3=x.find("socialdemòcrates")
    if PS != -1 or PS_1 !=-1 or PS_2 !=-1:
        PS_count=PS_count+1
        PS=-1
        PS_1=-1
        PS_3=-1

    Li=x.find('Liberals')
    Li_1=x.find('Gallardo')

```

```

Li_2=x.find("L'A")

if Li != -1 or Li_1 !=-1 or PS_2 !=-1 or Li_2 !=-1:
    Li_count=Li_count+1
    Li =-1
    Li_1 =-1
    PS_2=-1
    Li_2 =-1

DA=x.find('DA')
DA_1=x.find('Demòcrates')
DA_2=x.find('Espot')

if DA !=-1 or DA_1 !=-1 or DA_2 !=-1:
    DA_count = DA_count+1
    DA=-1
    DA_1=-1
    DA_2=-1

TV=x.find('TERCERAVIA')
TV_1=x.find('Pintat')
TV_2=x.find('Terceravia')

if TV !=-1 or TV_1 !=-1 or TV_2 !=-1 :
    TV_count=TV_count+1
    TV=-1
    TV_1=-1
    TV_2=-1

row1 = ["Forum.ad", "Partit Socialdemòcrata", PS_count]
row2= ["Forum.ad", "Liberals d'Andorra", Li_count]
row3= ["Forum.ad", "Demòcrates", DA_count]
row4= ["Forum.ad", "TERCERAVIA", TV_count]

with open('DATA.csv', 'a') as csvFile:
    writer = csv.writer(csvFile)
    writer.writerow(row1)
    writer.writerow(row2)
    writer.writerow(row3)
    writer.writerow(row4)

csvFile.close()

```



```

#!/usr/bin/env python3
# -*- coding: utf-8 -*-
"""
Created on Wed Apr 10 09:38:46 2019

@author: Gil
File: Scraping_4.py
"""

from urllib.request import urlopen
from bs4 import BeautifulSoup
import ssl
import csv

ctx = ssl.create_default_context()
ctx.check_hostname = False
ctx.verify_mode = ssl.CERT_NONE

llista=list()
pag=10
i=1

PS_count=0
Li_count=0
DA_count=0
TV_count=0

while pag>2:
    url = 'https://www.ara.ad/arxiu/politica/5001000/?page='+str(pag)
    html = urlopen(url, context=ctx).read()
    soup = BeautifulSoup(html, "html.parser")

    tags = soup('a')

    for tag in tags:
        if i >= 34 and i<= 92 :
            if tag.string != None:
                llista.append(tag.string)
            i=i+1

    i=1
    pag=pag-1

for x in llista:
    PS=x.find('PS')

```

```

PS_1=x.find('López')
PS_2=x.find("d'Acord")
PS_3=x.find("socialdemòcrates")
if PS != -1 or PS_1 !=-1 or PS_2 !=-1:
    PS_count=PS_count+1
    PS=-1
    PS_1=-1
    PS_3=-1

Li=x.find('Liberals')
Li_1=x.find('Gallardo')
Li_2=x.find("L'A")

if Li != -1 or Li_1 !=-1 or PS_2 !=-1 or Li_2 !=-1:
    Li_count=Li_count+1
    Li =-1
    Li_1 =-1
    PS_2=-1
    Li_2 =-1

DA=x.find('DA')
DA_1=x.find('Demòcrates')
DA_2=x.find('Espot')

if DA !=-1 or DA_1 !=-1 or DA_2 !=-1:
    DA_count = DA_count+1
    DA=-1
    DA_1=-1
    DA_2=-1

TV=x.find('TERCERAVIA')
TV_1=x.find('Pintat')
TV_2=x.find('Terceravia')

if TV !=-1 or TV_1 !=-1 or TV_2 !=-1 :
    TV_count=TV_count+1
    TV=-1
    TV_1=-1
    TV_2=-1

row1 = ["Ara.ad", "Partit Socialdemòcrata", PS_count]
row2= ["Ara.ad", "Liberals d'Andorra", Li_count]
row3= ["Ara.ad", "Demòcrates", DA_count]
row4= ["Ara.ad", "TERCERAVIA", TV_count]

with open('DATA.csv', 'a') as csvFile:
    writer = csv.writer(csvFile)

```

```
writer.writerow(row1)
writer.writerow(row2)
writer.writerow(row3)
writer.writerow(row4)

csvFile.close()
```