Crawling for finance

류지승

가격

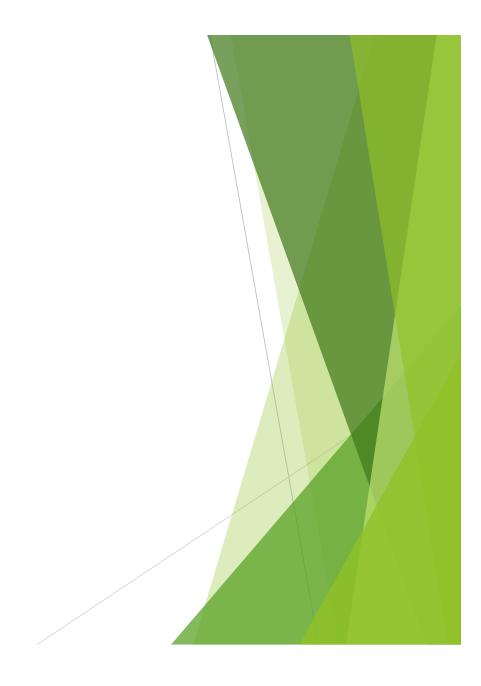# Robots.txt 이해하기

- http://egloos.zum.com/engjjang/v/10106048

- User-agent: *
  Disallow: /

- 모두 차단

- User-agent: Cowbot
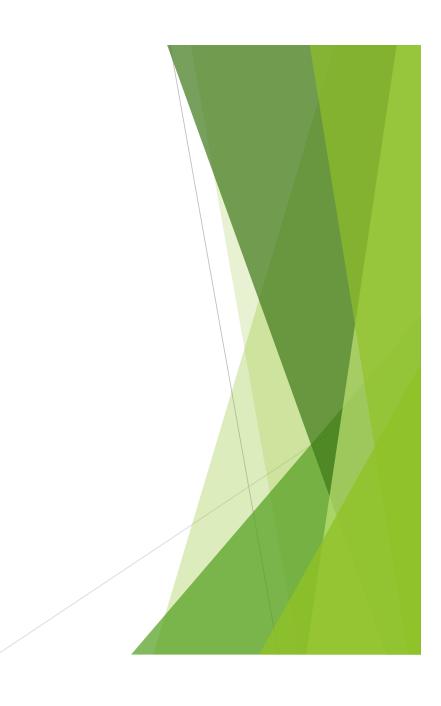 Disallow:

  User-agent: Googlebot
  Disallow:

  User-agent: *
  Disallow: /
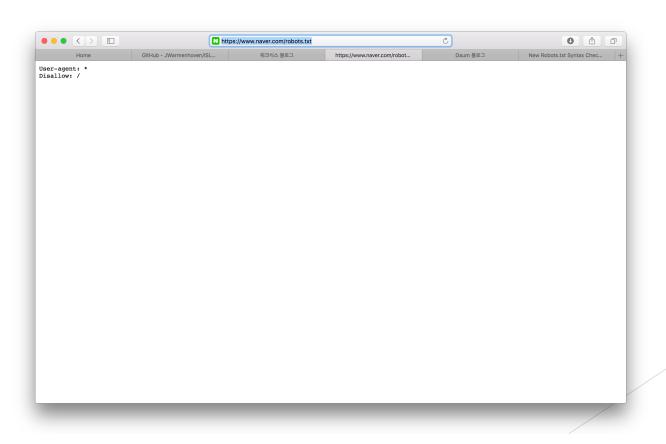
- 구글 네이버만 허용하고 다른 것은 차단
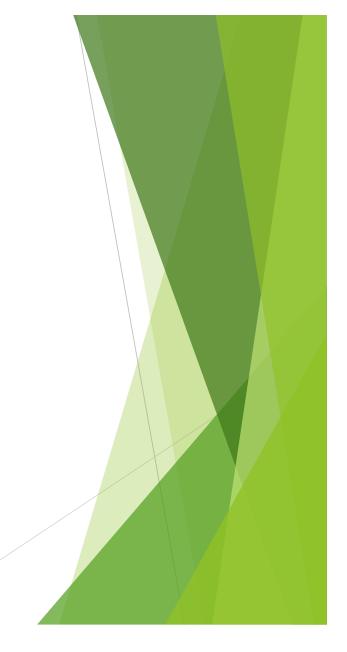
```
ch/about
ch/howsearchworks
dch
roups
ndex.html?

:

hl=*&
*&gws_rd=ssl$
hl=*&*&gws_rd=ssl
_rd=ssl$
=true$
mgres
/
references
etprefs
efault
?
/
/finance
ml?
ml/?
ml/search?
html?
html/?
html/search?
ml?
mode?
mode/?
mode/search?
sky?
sky/?
sky/search?
da?
da/?
da/search?
print_xhtml
```

www.naver.com/robots.txt
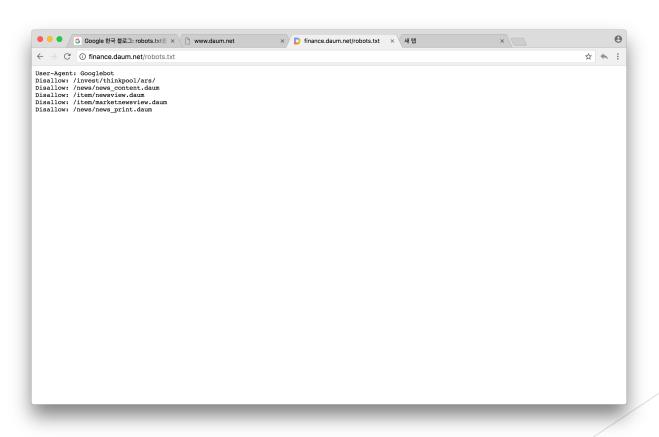
▶ https://www.naver.com/robots.txt

```
User-agent: *
Disallow: /
```
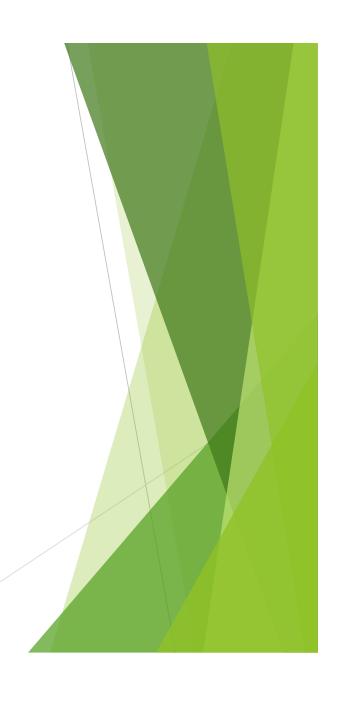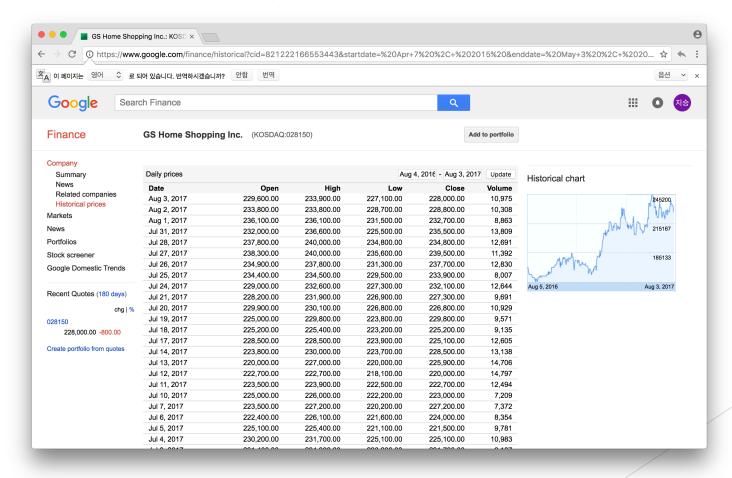
www.daum.net/robots.txt

▶ http://finance.daum.net/robots.txt

GS 홈쇼핑(Google finance)

Url 분석하기

- https://www.google.com/finance/historical?cid=821222166553443&startdate= Apr+7 %2C+ 2015 &enddate= May+3 %2C+ 2017 &num= 200

사용자 편의에 맞게 Url 수정

- library(stringr)
- startdate<-"2015/04/29"
- enddate<-"2016/03/02"
- month<-c("Jan","Feb","Mar","Apr","May","Jun","Jul","Aug","Sep","Oct","Nov","Dec")
- startyear<-as.numeric(str_sub(startdate,1,4))
- startmonth<-month[as.numeric(str_sub(startdate,6,7))]
- startday<-as.numeric(str_sub(startdate,9,10))
- endyear<-as.numeric(str_sub(enddate,1,4))
- endmonth<-month[as.numeric(str_sub(enddate,6,7))]
- endday<-as.numeric(str_sub(enddate,9,10))
- num<-200
- url<-https://www.google.com/finance/historical?cid=821222166553443&startdate=
- url1<-url%>%paste(.,startmonth,"+",startday,"%2C+",startyear,"&enddate=",endmonth,"+",endday,"%2C+",endyear,"&num=",num,sep="")

가격과 관련된 정보 가져오기 (Date,Open,High,Low,Close,Volume)
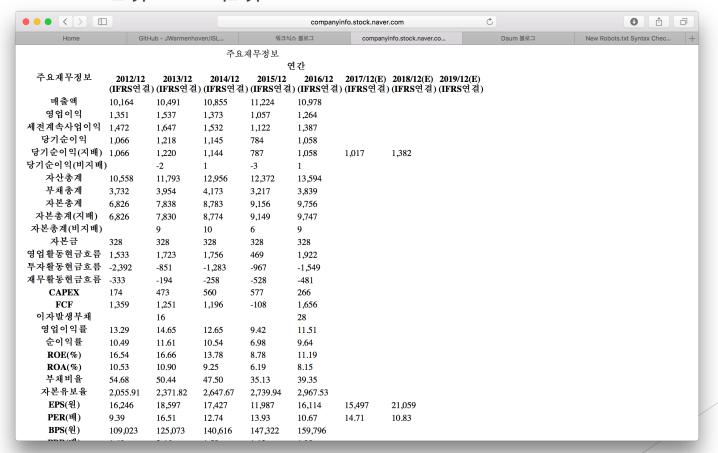
- library(rvest)

- table_url<-read_html(url1)%>%html_nodes("div#prices")%>%html_nodes(".gf-table.historical_price")

- table<-html_table(table_url)[[1]]

- for(i in 2:6) { table[,i]<-as.numeric(gsub(",","",table[,i]))}

- table

```
> table
          Date   Open    High     Low   Close Volume
1   Mar 2, 2016 188500  188500  181600  182300  17568
2  Feb 29, 2016 184100  184300  180200  184000  20797
3  Feb 26, 2016 188800  189000  183100  184700   7805
4  Feb 25, 2016 186900  189000  184000  188900  15309
5  Feb 24, 2016 187900  188000  185000  185100  12374
6  Feb 23, 2016 187500  189000  183000  187900  17572
7  Feb 22, 2016 187700  188500  184200  187700  17259
8  Feb 19, 2016 185100  188000  181800  187700  31781
9  Feb 18, 2016 185000  187600  180700  187600  36329
10 Feb 17, 2016 184300  187100  180300  186800  50020
11 Feb 16, 2016 182100  185800  182100  183600  64424
12 Feb 15, 2016 172000  178000  169000  178000  35813
13 Feb 12, 2016 173000  173000  167000  169200  33495
14 Feb 11, 2016 173400  174000  171000  174000  26086
15  Feb 5, 2016 176500  179000  173400  175000  16614
16  Feb 4, 2016 174400  178600  172900  176500  61375
17  Feb 3, 2016 164900  177500  164600  174500  88430
18  Feb 2, 2016 163000  163900  162100  163000   2887
19  Feb 1, 2016 163500  164500  162200  163100   4968
20 Jan 29, 2016 161500  164100  161500  163000  14583
```

재무제표

▶ http://companyinfo.stock.naver.com/v1/company/ajax/cF1001.aspx?cmp_cd=028150fin_typ=0&freq_typ=Y

주요재무정보

주요재무정보	2012/12 (IFRS연결)	2013/12 (IFRS연결)	2014/12 (IFRS연결)	2015/12 (IFRS연결)	2016/12 (IFRS연결)	2017/12(E) (IFRS연결)	2018/12(E) (IFRS연결)	2019/12(E) (IFRS연결)
매출액	10,164	10,491	10,855	11,224	10,978			
영업이익	1,351	1,537	1,373	1,057	1,264			
세전계속사업이익	1,472	1,647	1,532	1,122	1,387			
당기순이익	1,066	1,218	1,145	784	1,058			
당기순이익(지배)	1,066	1,220	1,144	787	1,058	1,017	1,382	
당기순이익(비지배)		-2	1	-3	1			
자산총계	10,558	11,793	12,956	12,372	13,594			
부채총계	3,732	3,954	4,173	3,217	3,839			
자본총계	6,826	7,838	8,783	9,156	9,756			
자본총계(지배)	6,826	7,830	8,774	9,149	9,747			
자본총계(비지배)		9	10	6	9			
자본금	328	328	328	328	328			
영업활동현금흐름	1,533	1,723	1,756	469	1,922			
투자활동현금흐름	-2,392	-851	-1,283	-967	-1,549			
재무활동현금흐름	-333	-194	-258	-528	-481			
CAPEX	174	473	560	577	266			
FCF	1,359	1,251	1,196	-108	1,656			
이자발생부채		16			28			
영업이익률	13.29	14.65	12.65	9.42	11.51			
순이익률	10.49	11.61	10.54	6.98	9.64			
ROE(%)	16.54	16.66	13.78	8.78	11.19			
ROA(%)	10.53	10.90	9.25	6.19	8.15			
부채비율	54.68	50.44	47.50	35.13	39.35			
자본유보율	2,055.91	2,371.82	2,647.67	2,739.94	2,967.53			
EPS(원)	16,246	18,597	17,427	11,987	16,114	15,497	21,059	
PER(배)	9.39	16.51	12.74	13.93	10.67	14.71	10.83	
BPS(원)	109,023	125,073	140,616	147,322	159,796			

```
oper<-
http://companyinfo.stock.
naver.com/v1/company/aj
ax/cF1001.aspx?cmp_cd=
code<-"028150"
oper1<-
oper%>%paste(code,"fin_typ
=0&freq_typ=Y",sep="")
```

RStudio

Source

Console ~/

```
> oper_table<-html_table(read_html(oper1)%>%html_nodes(".gHead01.all-width"),fill=T)[[1]]
> oper_table
```

주요재무정보	연간	연간	연간
1 주요재무정보	2012/12\r\n\t\t\t\t\t\t\t\t\t(IFRS연결)	2013/12\r\n\t\t\t\t\t\t\t\t\t(IFRS연결)	2014/12\r\n\t\t\t\t\t\t\t\t\t(IFRS연결)
2 매출액	10,164	10,491	10,855
3 영업이익	1,351	1,537	1,373
4 세전계속사업이익	1,472	1,647	1,532
5 당기순이익	1,066	1,218	1,145
6 당기순이익(지배)	1,066	1,220	1,144
7 당기순이익(비지배)		-2	1
8 자산총계	10,558	11,793	12,956
9 부채총계	3,732	3,954	4,173
10 자본총계	6,826	7,838	8,783
11 자본총계(지배)	6,826	7,830	8,774
12 자본총계(비지배)		9	10
13 자본금	328	328	328
14 영업활동현금흐름	1,533	1,723	1,756
15 투자활동현금흐름	-2,392	-851	-1,283
16 재무활동현금흐름	-333	-194	-258
17 CAPEX	174	473	560
18 FCF	1,359	1,251	1,196
19 이자발생부채		16	
20 영업이익률	13.29	14.65	12.65
21 순이익률	10.49	11.61	10.54
22 ROE(%)	16.54	16.66	13.78
23 ROA(%)	10.53	10.90	9.25
24 부채비율	54.68	50.44	47.50
25 자본유보율	2,055.91	2,371.82	2,647.67
26 EPS(원)	16,246	18,597	17,427
27 PER(배)	9.39	16.51	12.74
28 BPS(원)	109,023	125,073	140,616
29 PBR(배)	1.40	2.46	1.58
30 현금DPS(원)	3,000	3,500	7,700
31 현금배당수익률	1.97	1.14	3.47
32 현금배당성향(%)	17.62	17.95	42.01
33 발행주식수(보통주)	6,562,500	6,562,500	6,562,500

연간	연간	연간
1 2015/12\r\n\t\t\t\t\t\t\t\t\t(IFRS연결)	2016/12\r\n\t\t\t\t\t\t\t\t\t(IFRS연결)	2017/12(E)\r\n\t\t\t\t\t\t\t\t\t(IFRS연결)

Environment

Global En

o... 33 ob
t... 200 ol

Values

Files Plots

```
RStudio

Source                                                                                    ⬒⬚

Console ~/ ⇗                                                                             ⬓⬚

30                              7,025                              7,062
31                              3.08                               3.10
32                              0.33
33
> str(oper_table)
'data.frame':   33 obs. of  9 variables:
 $ 주요재무정보: chr  "주요재무정보" "매출액" "영업이익" "세전계속사업이익" ...
 $ 연간       : chr  "2012/12\r\n\t\t\t\t\t\t\t\t\t(IFRS연결)" "10,164" "1,351" "1,472" ...
 $ 연간       : chr  "2013/12\r\n\t\t\t\t\t\t\t\t\t(IFRS연결)" "10,491" "1,537" "1,647" ...
 $ 연간       : chr  "2014/12\r\n\t\t\t\t\t\t\t\t\t(IFRS연결)" "10,855" "1,373" "1,532" ...
 $ 연간       : chr  "2015/12\r\n\t\t\t\t\t\t\t\t\t(IFRS연결)" "11,224" "1,057" "1,122" ...
 $ 연간       : chr  "2016/12\r\n\t\t\t\t\t\t\t\t\t(IFRS연결)" "10,978" "1,264" "1,387" ...
 $ 연간       : chr  "2017/12(E)\r\n\t\t\t\t\t\t\t\t\t(IFRS연결)" "" "" "" ...
 $ 연간       : chr  "2018/12(E)\r\n\t\t\t\t\t\t\t\t\t(IFRS연결)" "" "" "" ...
 $ 연간       : chr  "2019/12(E)\r\n\t\t\t\t\t\t\t\t\t(IFRS연결)" "" "" "" ...
> head(oper_table)
       주요재무정보                          연간                          연간                          연간
1      주요재무정보 2012/12\r\n\t\t\t\t\t\t\t\t\t(IFRS연결) 2013/12\r\n\t\t\t\t\t\t\t\t\t(IFRS연결) 2014/12\r\n\t\t\t\t\t\t\t\t\t(IFRS연결)
2          매출액                       10,164                       10,491                       10,855
3          영업이익                      1,351                        1,537                        1,373
4 세전계속사업이익                      1,472                        1,647                        1,532
5          당기순이익                    1,066                        1,218                        1,145
6 당기순이익(지배)                      1,066                        1,220                        1,144
                          연간                          연간                          연간
1 2015/12\r\n\t\t\t\t\t\t\t\t\t(IFRS연결) 2016/12\r\n\t\t\t\t\t\t\t\t\t(IFRS연결) 2017/12(E)\r\n\t\t\t\t\t\t\t\t\t(IFRS연결)
2                       11,224                       10,978
3                        1,057                        1,264
4                        1,122                        1,387
5                          784                        1,058
6                          787                        1,058                         1,017
                          연간                          연간
1 2018/12(E)\r\n\t\t\t\t\t\t\t\t\t(IFRS연결) 2019/12(E)\r\n\t\t\t\t\t\t\t\t\t(IFRS연결)
2
3
4
5
6                        1,382
>
```

```
> new_table
```

	2012/12	2013/12	2014/12	2015/12	2016/12	2017/12	2018/12	2019/12
매출액	10164.00	10491.00	10855.00	11224.00	10978.00	NA	NA	NA
영업이익	1351.00	1537.00	1373.00	1057.00	1264.00	NA	NA	NA
세전계속사업이익	1472.00	1647.00	1532.00	1122.00	1387.00	NA	NA	NA
당기순이익	1066.00	1218.00	1145.00	784.00	1058.00	NA	NA	NA
당기순이익(지배)	1066.00	1220.00	1144.00	787.00	1058.00	1017.00	1382.00	NA
당기순이익(비지배)	NA	-2.00	1.00	-3.00	1.00	NA	NA	NA
자산총계	10558.00	11793.00	12956.00	12372.00	13594.00	NA	NA	NA
부채총계	3732.00	3954.00	4173.00	3217.00	3839.00	NA	NA	NA
자본총계	6826.00	7838.00	8783.00	9156.00	9756.00	NA	NA	NA
자본총계(지배)	6826.00	7830.00	8774.00	9149.00	9747.00	NA	NA	NA
자본총계(비지배)	NA	9.00	10.00	6.00	9.00	NA	NA	NA
자본금	328.00	328.00	328.00	328.00	328.00	NA	NA	NA
영업활동현금흐름	1533.00	1723.00	1756.00	469.00	1922.00	NA	NA	NA
투자활동현금흐름	-2392.00	-851.00	-1283.00	-967.00	-1549.00	NA	NA	NA
재무활동현금흐름	-333.00	-194.00	-258.00	-528.00	-481.00	NA	NA	NA
CAPEX	174.00	473.00	560.00	577.00	266.00	NA	NA	NA
FCF	1359.00	1251.00	1196.00	-108.00	1656.00	NA	NA	NA
이자발생부채	NA	16.00	NA	NA	28.00	NA	NA	NA
영업이익률	13.29	14.65	12.65	9.42	11.51	NA	NA	NA
순이익률	10.49	11.61	10.54	6.98	9.64	NA	NA	NA
ROE(%)	16.54	16.66	13.78	8.78	11.19	NA	NA	NA
ROA(%)	10.53	10.90	9.25	6.19	8.15	NA	NA	NA
부채비율	54.68	50.44	47.50	35.13	39.35	NA	NA	NA
자본유보율	2055.91	2371.82	2647.67	2739.94	2967.53	NA	NA	NA
EPS(원)	16246.00	18597.00	17427.00	11987.00	16114.00	15497.00	21059.00	NA
PER(배)	9.39	16.51	12.74	13.93	10.67	14.71	10.83	NA
BPS(원)	109023.00	125073.00	140616.00	147322.00	159796.00	NA	NA	NA
PBR(배)	1.40	2.46	1.58	1.13	1.08	NA	NA	NA
현금DPS(원)	3000.00	3500.00	7700.00	5200.00	7000.00	6900.00	7025.00	7062.0
현금배당수익률	1.97	1.14	3.47	3.11	4.07	3.03	3.08	3.1
현금배당성향(%)	17.62	17.95	42.01	41.05	40.38	0.45	0.33	NA
발행주식수(보통주)	6562500.00	6562500.00	6562500.00	6562500.00	6562500.00	NA	NA	NA

```
>
```