



# 비대면 관련 산업 관심도 조사

김명환

# 목차

## 서론

- 주제선정 배경
- 산업 범주 분류

## 본론

- 데이터 수집과정 ( 파이썬 크롤링 )
- R 분석 ( 텍스트 마이닝)
- 시각화

## 결론

## 분석 배경

4차 산업혁명과 5G기술 출현으로  
비대면 산업의 규모가 커지는  
상황에서, 코로나19로 인한 거리  
두기 문화가 대중들로 하여금  
비대면 산업의 어떤 분야에  
요구를 발생시키고 있는지 확인

### 4차 산업혁명

스마트 팩토리, Ai, IoT, 로봇기술, 드론, 자율주행차, 가상현실(VR)  
등이 주도하는 차세대 산업 혁명

### 5세대 이동통신 (5G)

최대 속도가 20Gbs에 달하는 이동 통신기술  
4세대 이동통신인 LTE에 비해 속도가 20배가량 빠르고, 처리  
용량은 100배 많다.

### COVID-19

2019년 12월 중국 우한에서 처음 발생한 뒤 중국 전역과 전 세계  
로 확산된 호흡기 감염질환

# 서론

## 산업 범주

### 1. 건강/의료

디지털치료제, AI기반  
질병진단, 실시간 생체정보  
측정, 감염병 예측,  
RNA바이러스 대항백신,  
원격진료

### 2. 교육

실감형 VR기술, AI빅데이터  
기반 맞춤형 학습,  
온라인수업용 대용량  
통신기술, 상담, 원격수업,  
스마트 교육

### 3. 환경

의료폐기물 운반로봇,  
인수공통감염병 통합관리

### 4. 물류

ICT기반 물류정보 플랫폼,  
자율주행 배송로봇, 유통센터  
스마트화, 배달, 마켓컬리,  
택배(부재중)

### 5. 제조

디지털트윈, 인간증강기술,  
협동로봇기술

### 6. 문화

실감중계 서비스, 딥페이크  
탐지기술, 공연, 스포츠, 행사,  
비대면 공연

### 7. 정보/보안

화상회의 보안 확보,  
동형암호이용 동선추적시스템

### 8. 교통

자율주행차, 개인맞춤형  
라스트마일 이동 수단

### 9. 비즈니스

화상 회의, 재택근무, 면접

### 10. 금융

모바일 통장, 송금, 결제

### 11. 영상/방송

넷플릭스, 유튜브 스트리밍,  
온라인 예배, 온라인 선거유세

### 12. 거래

드라이브 스루, 홈쇼핑,  
이커머스, 키오스크, 어플  
주문

### 13. 통신

5G, 실시간기술

### 14. IT

개발, 플랫폼, 모바일, 게임,  
데이터, 스마트

[illegible]

# 데이터 수집 (파이썬 크롤링)

```

# decode할 때 [utf-8, cp949, euc-kr] 모두 사용에 가능한 종류가 많아서 어느 것을 알았음,
# 그냥 cp949가 오류가 가장 적게 발생하니 선택
# 무조건 pass하면 에러시 pass하면 안됨 에러시 수습을 위해서 try - except
try:
    result = urllib.request.urlopen(url).read().decode("cp949")
    soup = BeautifulSoup(result, 'html.parser')

    # 비록문자 = <div> 수 <div>의 번호 = 1
    result1 = soup.find_all('span', class_="class_title")
    # 인덱스 0부터 (0 ~ 768 ~ 1 ~ 23)
    artn=result1[0].get_text()
    n=artn.index(' ')
    tnum=artn[11:n].replace(' ','')
    en = int(int(tnum)/20)+1
    if pn == en+1:
        break
    print(artn)

    # 기사 주소를 가져오기 위한 코드
    result2 = soup.find_all('span', class_="art_tit")
    for j in result2:
        for k in j:
            surl.append(k.get("href"))

except:
    pass

# URL 리스트 반환
return surl

```

# 데이터 수집 (파이썬 크롤링)

```
def my_url(month):
    # my(month) 함수의 리턴값 tmpmg로 받는다.
    tmpmg = my(month)

    # 기사가 저장될 빈 리스트 생성
    param1=[]

    # 진행 함수에서 받은 리스트에서 중복없이 기사 내용을 크롤링
    for ulist in list(set(tmpmg)):
        url = urllib.request.Request(ulist)
        result = urllib.request.urlopen(url).read().decode("cp949")
        soup = BeautifulSoup(result, "html.parser")

        # 기사 내용 크롤링
        result1 = soup.find_all('div', class_='art_txt')
        for i in result1:
            param1.append(i.get_text(" ", strip=True)+'\n'+'\n')

    # csv로 저장하기 위해서 DataFrame화
    param = pd.DataFrame(param1)

    # DataFrame과 한 리스트를 csv파일로 저장
    param.to_csv('d:\data\날짜 pj\\' + url + str(month) + '.csv', encoding='utf-8')
```

# 크롤링 실행 & 결과 파일

```
In [3]: for i in range(1,7):
        mg scrll(i)
```

executed in 4h 35m 31s, finished 23:42:46 2020-07-14

```
<urllib.request.Request object at 0x00000215E22B4A48>
```

뉴스 검색결과 [총 2,211건 중 281 - 300]

```
<urllib.request.Request object at 0x00000215E25D02C8>
```

뉴스 검색결과 [총 2,211건 중 301 - 320]

유대 헌장에 주어진 조항은 "여자를 두루 보면 크게 구할 수 있습니다." 미국의 이런 군부 일에 여가 후 부업도 헌법 미국 대통령 제 62 연주가 이를 일컫는 데이다. 이런 루트스온 사형과이 사형한 지 8일 만에 8일부터 드림드 자선 활동은 그의 군사 임직을 홍보하는 피아스북 광고 약 1천 건을 쏟아냈다. 이 활동을 한다. "너무 위대한 이란 바르리스드 지도자 버그 장군에 대한 추모"라는 제목으로 삽입해 믿기지 않는 불충고로 보았기에서 이내 일과 제외비호를 한 뒤 참여자들은 드림드 판드에 후원금을 너도모 귀속받는다. 드림프 자선활동은 프랑크 지지자들에게만 "드림드 대통령의 우리 군부 다-강하게 군 다들릴 속이 막다. 어떤 또 다른 결핵과 광려하 기운에 들은 종교 종파와 악을 같이 한다고 분석한다. 드림드 지지자들 구체적으로 서남방 '독과 학교'로 의 미디어 책임자로 근무한 몇 번이나 하는 이런 "광고 공세"와 관련, "전자 헌부에서는 대통령이 성과를 낼 때마다 그가 재선하여 한다는 주장의 근거"라고 설명했다. google.org/display?google\_dtp=MC\_250x250". 하지만 미국과 이런 건 전장 우정을 능을 정도로 영웅한 사건을 정치적 이익을 위한 도구로 지킬수있다면 그건은 "공부에 기인하 시작하 무복을 둘러 그린에서 이후 곧 미래 문라임 후원 활동을 벌이는 것은 국가 안보이기 성향이지 미용사 드림드 대통령 속이 전쟁 관련 사건을 길바르게 광고로 활용하는 것은 도가 지나쳤다고 비판했다. 그는 "우리는 정부가 국가 안보 문제에 근거해 결정하는 Copyright © 연합뉴스 무단 전재 및 재배포금지 window(Query:1 document.write("<script src='//www.nate.co.kr/www/include/gnb/baner

◆ 디지털 금융 ◆ KB국민은행은 무언가를 스스로 탐색하거나 타인의 것을 있는 것을 검토는 내 손끝과 마우스를 통해 이것 그걸로 할 수에 있어 특별하고  
최초로 선보인 '리브업'이라는 알뜰폰 서비스로 디지털 금융의 혁신이란 평가를 받는다. 국민은행의 'KB나답대로적금'은 고객이 저축 방법, 저축 금액  
google.display('google\_dfp\_MCU25(x25%'); 나도 똑같을 수 있는 돈을 쉽게 관심을 '피자 반등기'로 이미지화해 누구나 쉽게 이해하고 기억할 수 있는  
돈을 가입 과정은 먼저 피자 가게에 입점해 두 가지 '피자 드'로 구성된 저축 방법(자유적립식, 정기적립식) 중 한 가지를 선택하는 것으로 시작한다. 예의  
존이나)을 자유롭게 선택한 후 여기서 '피자 드'로 구성된 무다 이 둘(급여 미리, 카드 결제 저축, 자동이체 저축, 마다나 권리비 이체, KB스카텔링 이체  
등을 할 수 있다. 마지막으로 네 가지 '피자 박스'로 구성된 브릭기(캐시박스)와다우, 피스, 고깃(여객) 및 한 가지를 선택하며 나머지 피자도 완성되며  
또 심플 가입 후 전환하면서는 다른의 피자 이름을 지을 수 있다. 타이틀 코어 중인 피자 박스에는 여섯 조각으로 나뉜 디지털 광고일이 다가올에 따  
르므로 재미를 더했다. KB국민은행과 연계한 무제한 최초로 적금의 추가 서비스로 도입된 수리나도 브랑보(피수 제외)를 제공하고 있다. 일상생활에서 그  
상품은 1인 가족 고객의 라이프 스타일에 맞춰 개발된 특화 상품이다. 여형 무담과 관련된 보험서비스를 추가 혜택으로 제공하는 등 고객 니즈를 반영한  
을 선보일 계획 가능하다. 3년에 경력직임식 기준 최고 연 2.8%까지 금리를 받을 수 있다. 무다 이 둘은 총 8개 항목으로 항목당 연 0.1%포인트의 최고  
이자, 여형 관련 무다 이 둘이 연 0.1%포인트씩 제공된다. KB가 3대 금융법인 리브(1), 리브미(2), 리브이(3)로 48마이다 회원 가입 시 각각 연 0.1%  
에 고객이 올릴 수 있는 현형경 금리도 있을 출시를 통한 다채원지 해결과 새로운 기업 이미지를 강화하고 있어 주목받고 있다. 지난해 5월 출시된 13  
단위 등이 특징이다. 고객이 많은 혜택을 지켜 지켜와 생각 할 만큼 관심을 하면 무다 금리(최고 연 1.0%)고이도로 대안고로 지원가 상해 관련 무료 브  
고통 다면 등 일상에서 쉽게 실천할 수 있는 미션을 제시하고 이를 달성하면 무다 금리를 제공해 고객의 환경 지출 실천을 증명한다. 3년에 경력직임식  
세계 올릴 것을 달성 여할하는 '달일 드와 대일 서비스'를 최근 시작했다. google.display('google\_dfp\_MCU25(x25%'); 리브업은 국민은행이 지난해  
의 실적 과정 없이 KB국민은행과 KB국민카드의 기본 알뜰폰 서비스를 사용할 수 있다. 한 번만 본인인증서를 발급받으면 나중에 추가된 기기를 바꿔도 번거로운 지  
수 있어 편리하다. (본일호 기자) (0) 매일경제 <na.co.kr> 무단전재 및 재배포 금지. window.jQuery || document.write<script src='/common/na.co



# 데이터 전처리 (R 분석)

```

1 library(KoNLP) # 한글 자연어 분석 패키지
2 library(dplyr) # 데이터 처리에 주력한 패키지
3 library(stringr) # 문자열 데이터 가공을 위해 사용하는 패키지
4
5 useSejongDic()
6
7 setwd("d:\\data\\R_pj")
8
9 for (i in 1:6){
10   txt <- readLines(paste0("d:\\data\\R_pj\\Ncr1", as.character(i), ".txt"), encoding="UTF-8")
11
12   txt <- str_replace_all(txt, "\\W", " ")
13   txt <- gsub("\\d+", "", txt)
14   txt <- gsub("\\n+", "", txt)
15   txt <- gsub("[A-Z]", "", txt)
16   txt <- gsub("[[:cntrl:]]", "", txt)
17   nouns <- extractNoun(txt)
18
19   # 추출한 명사 txt를 문자열 벡터로 변환, 단어별 빈도표 생성
20   wordcount <- table(unlist(nouns))
21
22   # 데이터 프레임으로 변환
23   re_word <- as.data.frame(wordcount, stringsAsFactors = F)
24
25   # 변수명 수정
26   re_word <- rename(re_word, word=Var1, freq=Freq)
27
28   # 두 글자 이상 단어 추출
29   re_word <- filter(re_word, nchar(word) >= 2)
30
31   # 빈도수가 많은 순으로 20개만
32   top_20 <- re_word %>%
33     arrange(desc(freq)) %>%
34     head(1000000)
35   top_20
36
37   write.csv(top_20, paste0("d:\\data\\R_pj\\nccsvacr1", as.character(i), ".csv"))
38 }

```

	A	B
1	word	20년 1월
2	가가	5
3	가건물	2
4	가게	37
5	가격	382
6	가격경쟁	2
7	가격경쟁력	4
8	가격대	1
9	가격을두고	1
10	가격지수	1
11	가격표	2
12	가격표시제	2
13	가격할인	1
14	가결	6
15	가게	6
16	가게대출	3
17	가게부	3

# 데이터 전처리 (R 분석)

```
test1 <- read.csv("d:\\data\\rpjdata\\csvacr1.csv")
test2 <- read.csv("d:\\data\\rpjdata\\csvacr12.csv")
test3 <- read.csv("d:\\data\\rpjdata\\csvacr13.csv")
test4 <- read.csv("d:\\data\\rpjdata\\csvacr14.csv")
test5 <- read.csv("d:\\data\\rpjdata\\csvacr15.csv")
test6 <- read.csv("d:\\data\\rpjdata\\csvacr16.csv")
head(test)
head(test2)

testm = merge(test1, test2, by.x='word', by.y='word', all = TRUE)
testm = merge(testm, test3, by.x='word', by.y='word', all = TRUE)
testm = merge(testm, test4, by.x='word', by.y='word', all = TRUE)
testm = merge(testm, test5, by.x='word', by.y='word', all = TRUE)
testm = merge(testm, test6, by.x='word', by.y='word', all = TRUE)
View(testm)
testm[is.na(testm)] <- 0
write.csv(testm, 'd:\\data\\rpjdata\\merge2020.csv')
```

C	D	E	F	G	H	I	J
word	X20년1월	X20년2월	X20년3월	X20년4월	X20년5월	X20년6월	CNT
몬라인	2591	4479	8773	11174	7808	7677	42502
코로나	342	2221	4566	5203	4326	3787	20445
경제	1257	1994	3134	4469	4807	4236	19897
기업	1355	1769	2944	3519	4163	4523	18273
서비스	1431	1984	3170	3850	3929	3822	18186
투자	1156	1601	2886	3364	3808	3648	16463
종목	945	1519	2775	3745	3798	3339	16121
시장	1304	1521	2313	2645	2939	3181	13903
사업	1110	1472	2151	2731	3181	3204	13849
지원	680	1294	2753	3514	2472	2761	13474
대비	524	1771	2750	2812	3262	2173	13292
시간	886	1425	2440	3379	2568	2235	12933
교육	392	1402	2593	2890	2436	2241	11954
확산	428	1965	3290	2608	1900	1637	11828
바이러스	459	1849	2676	2648	1973	1868	11473
수업	66	361	2338	4080	2833	1795	11473
학생	334	534	1869	2777	2817	2332	10663
분석	747	1070	1775	2469	2384	2169	10614
거래	783	1151	1775	2160	2340	2094	10303
고객	853	1397	1744	2010	2012	2232	10248
설명	716	1201	1905	2089	2107	2033	10051
상황	464	1102	2062	2271	2208	1704	9811
기준	537	1114	1963	2038	2215	1738	9605
중국	1203	2182	1580	1463	1556	1580	9564

merge 후 6개월 간 합 높은 순으로 정렬

# 데이터 전처리 (R 분석)

스토리	48	37	91	98	113	118	505
비난	62	89	83	92	96	82	504
시너지	37	57	70	83	117	140	504
총장	30	99	82	85	67	141	504
파이	45	32	51	120	72	184	504
인물	74	57	95	111	96	70	503
적발	50	104	82	83	96	88	503
편의성	51	42	86	95	104	125	503
단기	24	61	113	106	108	90	502
유연	30	63	111	111	100	87	502
건의	33	68	126	94	93	87	501
먼지	131	82	78	60	60	90	501
추구	64	50	70	98	114	105	501
리스크	51	39	77	129	100	104	500
설비	31	59	72	114	104	120	500
엔진	71	53	138	49	128	61	500
유안	22	48	96	121	118	95	500

비례대표	10	122	262	109	30	12	545
원격근무	4	61	155	118	90	115	543
팬데믹	0	5	134	129	181	91	540
농산물	23	65	54	165	113	115	535
박람회	39	45	68	77	161	145	535
아카데미	25	137	97	92	73	111	535
종교	8	71	210	147	32	67	535
설문조사	31	30	102	135	118	117	533
전시회	46	52	112	108	115	99	532
포럼	12	46	41	95	120	217	531
도교육	16	30	125	119	187	53	530
도서관	28	67	132	109	95	93	524
외식	22	59	94	98	128	122	523
자동화	31	41	71	105	139	135	522
친환경	33	32	76	66	136	173	516
이커머스	52	63	77	139	90	92	513
연동	49	58	74	123	99	103	506
포토	24	27	49	72	229	105	506
농가	30	43	60	154	122	96	505
편의성	51	42	86	95	104	125	503

6개월 합 500이하 제거

불필요 단어 제거  
(신문사 관련 단어, 남은 불용어)

# 데이터 전처리 (R 분석)

1	2	3	4	5	6	7	8	9	10	11	12	13	14
건강/의료	교육	환경	물류	제조	문화	정보/보안	교통	비즈니스	금융	영상/방송	거래	통신	it

15	확산	428	1965	3290	2608	1900	1637	11828	0
16	바이러스	459	1849	2676	2648	1973	1868	11473	1
17	수업	66	361	2338	4080	2833	1795	11473	2
18	학생	334	534	1869	2777	2817	2332	10663	2
19	분석	747	1070	1775	2469	2384	2169	10614	0
20	거래	783	1151	1775	2160	2340	2094	10303	12
21	고객	853	1397	1744	2010	2012	2232	10248	0
22	설명	716	1201	1905	2089	2107	2033	10051	0
23	상황	464	1102	2062	2271	2208	1704	9811	0
24	기준	537	1114	1963	2038	2215	1738	9605	0
25	중국	1203	2182	1580	1463	1556	1580	9564	0
26	상승	536	969	1655	1872	2325	2043	9400	0
27	개발	731	952	1567	1847	2039	2155	9291	14
28	디지털	682	882	1385	2047	2103	2031	9130	14
29	사진	712	993	1727	1999	1866	1813	9110	11
30	제품	867	1146	1506	1623	1812	2051	9005	5
31	운영	640	954	1658	1817	1979	1935	8983	9
32	발표	612	1072	1431	2331	2056	1421	8923	0



# 본론

## 분류결과

### 1. 건강/의료

코로나  
바이러스  
마스크  
감염증  
감염  
방역  
환자  
건강

### 2. 교육

교육  
수업  
학교  
학개  
등교  
대학  
학년

### 3. 환경

a3  
a.a.X....i...word..  
환경  
친환경

### 4. 물류

유통  
주문  
달배  
송물  
류택  
배  
쿠팡  
무르세디

### 5. 제조

제품  
상작  
제출  
수급  
공산  
반도체

### 6. 문화

행사  
활동  
스포츠  
개최  
이벤트  
영화  
문화

### 7. 정보/보안

보안  
서버  
네트워크  
사이버  
인식  
블록체인  
모니터링

### 8. 교통

자동차  
차량  
항공  
주행  
기아차  
교통  
공항

### 9. 비즈니스

운영  
비대  
론회  
의의  
업무  
언택트  
솔루션

### 10. 금융

금융  
증권  
사행  
은카드  
대출  
증권  
펀드  
자산

### 11. 영상/방송

사진  
영상  
채널  
방송  
광고  
교회  
후보  
콘텐츠

### 12. 거래

온라인  
거래  
쇼핑  
결제  
할인  
백화점  
페이

### 13. 통신

휴대폰  
실시간  
스마트폰  
전화  
텔레콤  
통신  
통화

### 14. IT

개발  
디지털  
홈페이지  
원격  
게임  
모바일  
플랫폼  
시스템

# 데이터 전처리 (R 분석)

```
a <- read.csv('d:\\data\\rpjdata\\cp12020r.csv')
```

```
b1 <- aggregate(X20년1월~X,a,sum)
```

```
b2 <- aggregate(X20년2월~X,a,sum)
```

```
br <- merge(br,b5,by='X')
```

```
br <- merge(br,b6,by='X')
```

```
write.csv(br,'d:\\data\\rpjdata\\cp12020g.csv')
```

```
> b1 <- aggregate(X20년1월~X,a,sum)
```

```
> b1
```

```
  X X20년1월
1  0    45231
2  1     4467
3  2     2521
4  3      294
5  4     1414
6  5     3527
7  6     6689
```

	B	C	D	E	F	G
	X20년1월	X20년2월	X20년3월	X20년4월	X20년5월	X20년6월
0	45231	76226	120543	139537	145922	139176
1	4467	14292	24756	23666	23390	21266
2	2521	7075	20818	27264	23762	17922
3	294	290	496	622	806	1013
4	1414	2622	3164	2888	3769	3214
5	3527	5243	8437	9558	10620	10625
6	6689	8420	13800	17655	18589	19287
7	731	927	1379	1819	2084	1638
8	873	910	1726	1802	1925	1789
9	2936	5170	10316	13074	14703	14818
10	2839	3724	6238	7758	8381	7132
11	3483	6302	12415	13368	9324	9682
12	6022	9208	14892	18555	15828	15946
13	888	1765	2817	3480	3302	3215
14	7431	9776	15691	21091	22952	23034

# 텍스트 마이닝 (R 시각화)

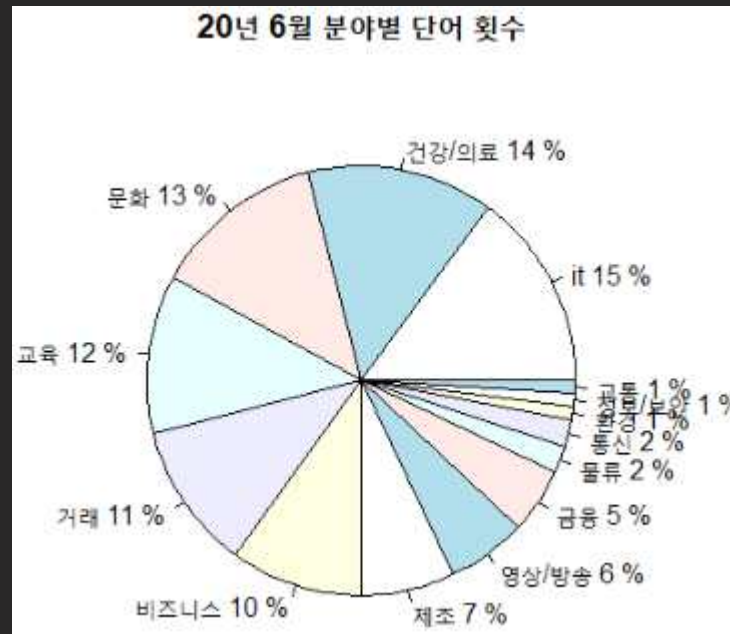
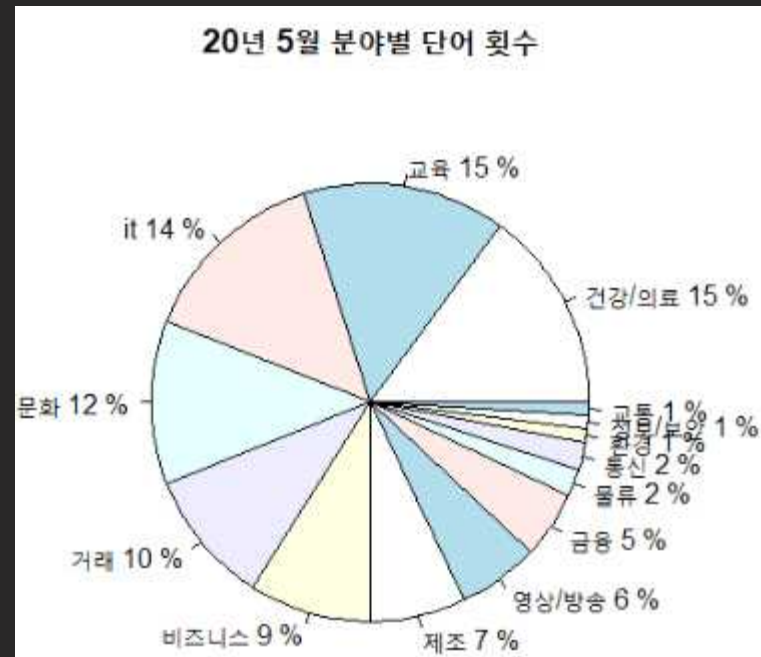
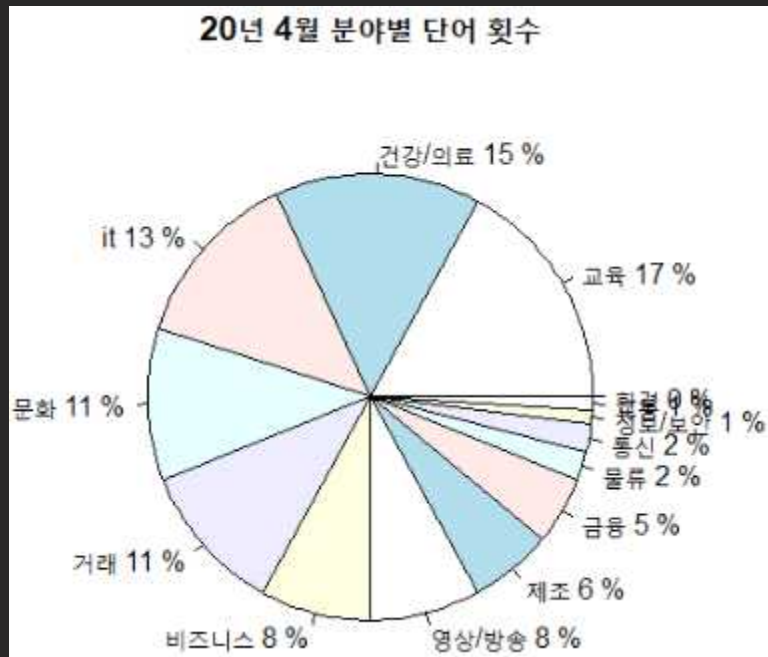
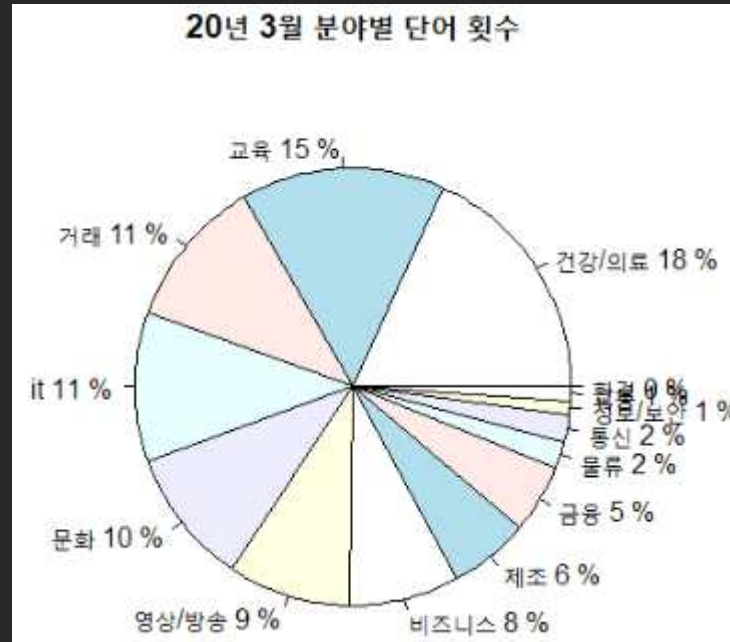
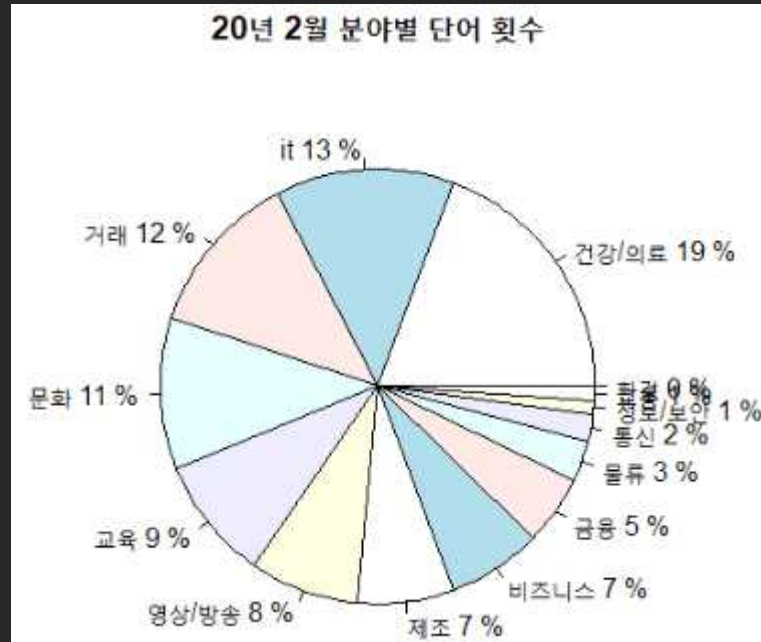
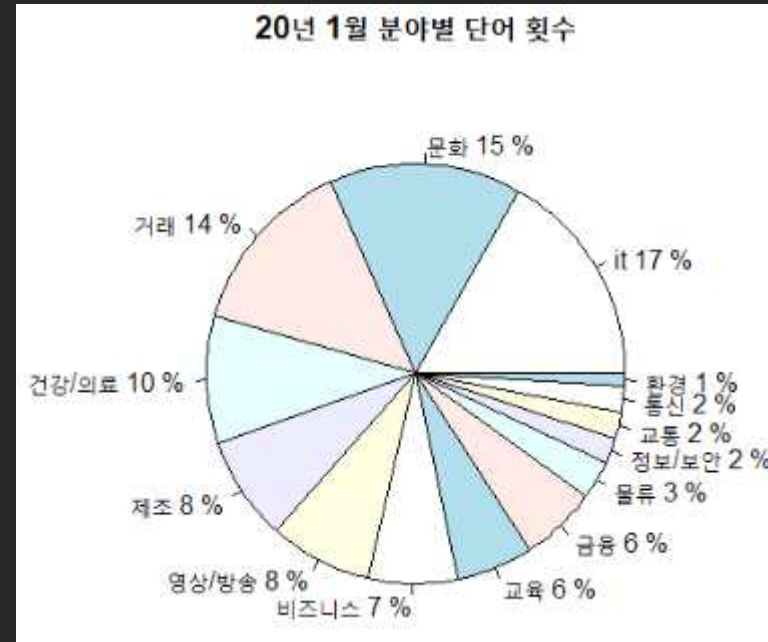
- 원형 그래프 (월별 분야 비율)

```
# 단위기간 분야별 비율  
i = 6  
data <- data.frame(  
  lb = lbls,  
  p = round(a1[i+1]/sum(a1[i+1]),2)  
)  
names(data) <- c('lb','기간')  
data  
data<- data[order(data$기간,decreasing = T),]  
pie(data$기간,labels = paste(data$lb,data$기간*100,'%'),  
  main = paste0("20년 ",i,"월 분야별 단어 횟수"))  
i
```



# 텍스트 마이닝 (R 시각화)

## - 원형 그래프 (월별 분야 비율)



	1	2	3	4	5	6	7
1월	it	문화	거래	건강	제조	영상	비즈
2월	건강	it	거래	문화	교육	영상	비즈
3월	건강	교육	it	문화	영상	영상	비즈
4월	교육	건강	it	문화	거래	비즈	영상
5월	건강	교육	it	문화	거래	비즈	제조
6월	it	건강	문화	교육	거래	비즈	제조

6달 동안 5회 이상 상위 분야  
: it, 건강, 문화, 교육, 거래, 비즈니스



# 텍스트 마이닝 (R 시각화)

- 막대 그래프 (분야별 기간 변동)

```
#ggplot/ 분야별 변동  
i = 14  
test <- data.frame(t(a1[i,-1]))  
names(test) <- "건수"  
기간 <- c('20년1월', '20년2월', '20년3월', '20년4월', '20년5월', '20년6월')  
ggplot(data = test, aes(기간, 건수)) + geom_bar(stat="identity") + ggtitle(paste('2020년 기간별 .
```

```
i = 14  
test <- data.frame(t(a1[i,-1]))  
names(test) <- "건수"  
기간 <- c('20년1월', '20년2월', '20년3월', '20년4월', '20년5월', '20년6월')  
ggplot(data = test, aes(기간, 건수)) + geom_bar(stat='identity')  
+ ggtitle(paste('2020년 기간별 <', 분야[i], '> 분야 단어언급 횟수'))  
+ theme(axis.title=element_text(size=16), plot.title = element_text(hjust=0.5, size = 20))
```

# 텍스트 마이닝 (R 시각화)

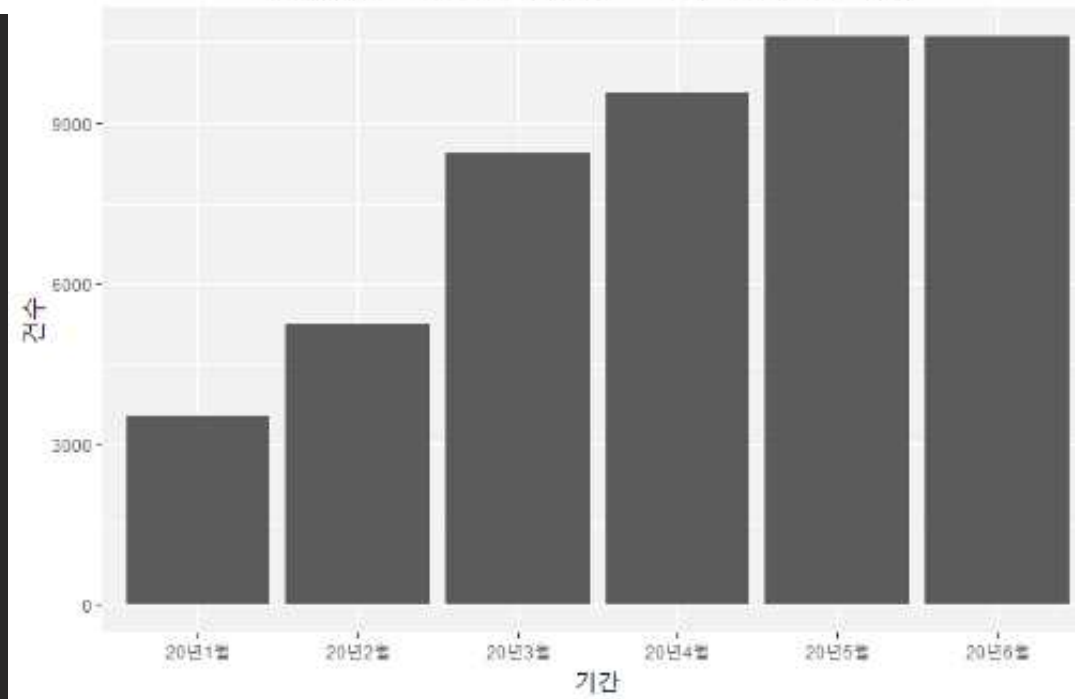
- 막대 그래프 (분야별 기간 변동)

물류 : 부재중 택배(사회적 거리두기)

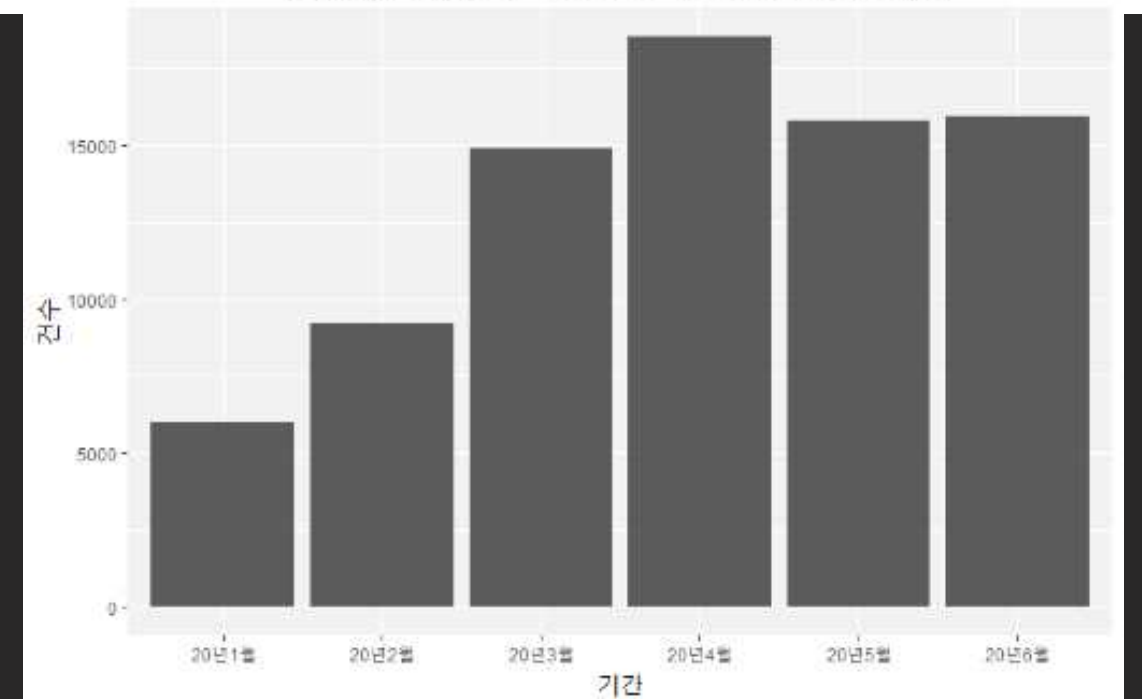
거래 : 드라이브 스루 ( 커피, 책 등)  
비 접촉 결제 (어플 결제 택시, 매장)

문화 : 방구석 콘서트, 스포츠 실시간 응원

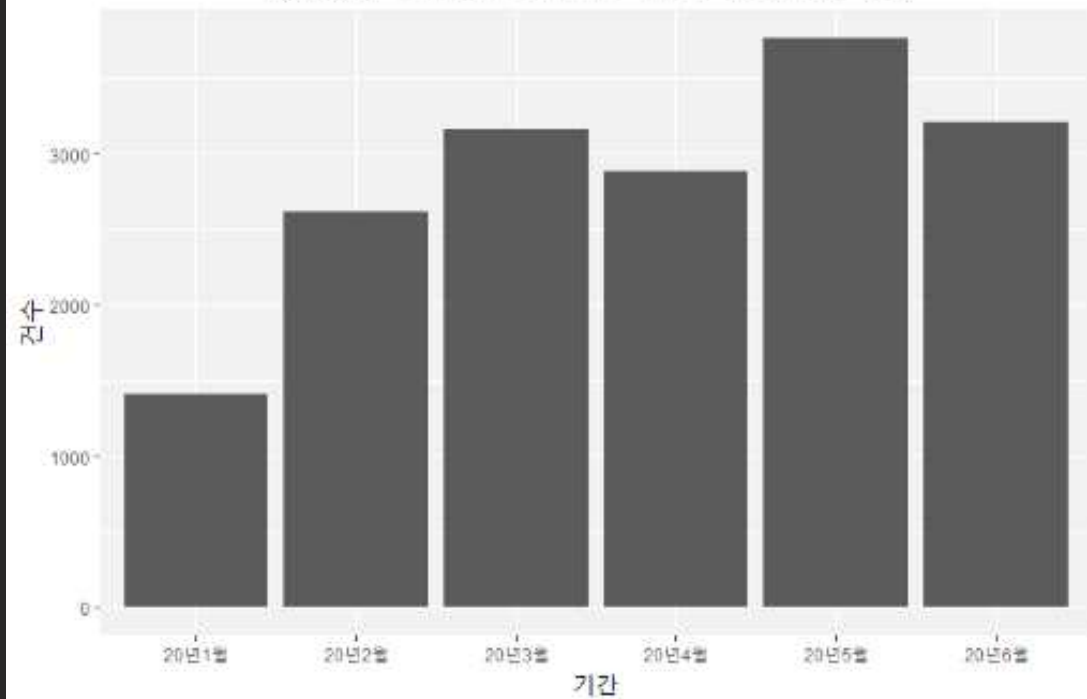
2020년 기간별 < 제조 > 분야 단어언급 횟수



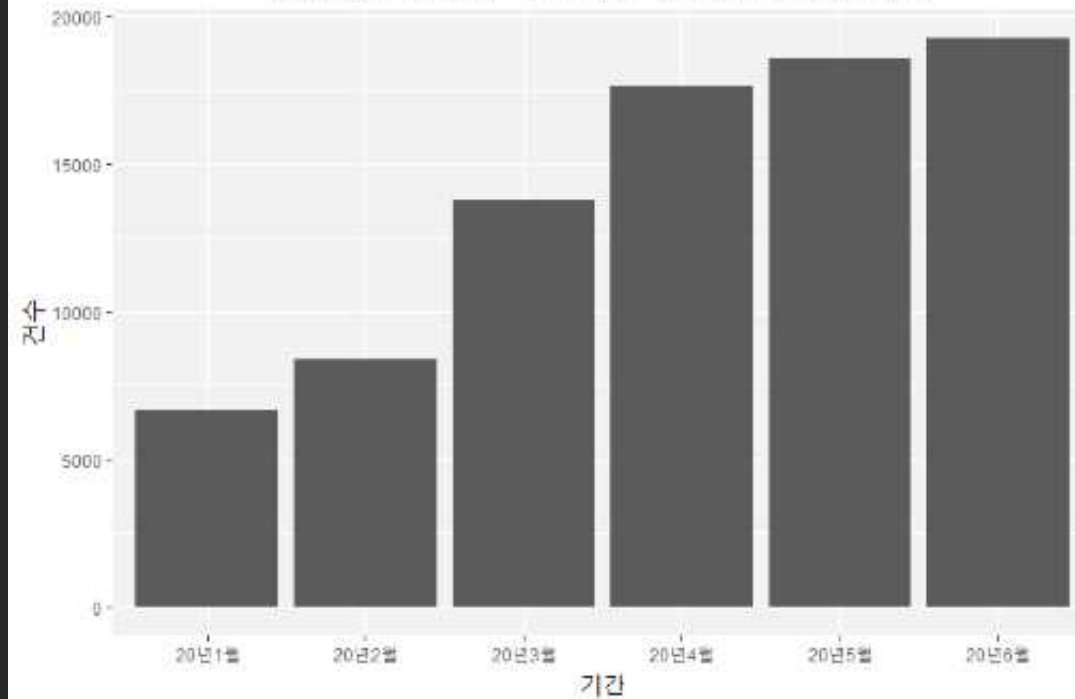
2020년 기간별 < 거래 > 분야 단어언급 횟수



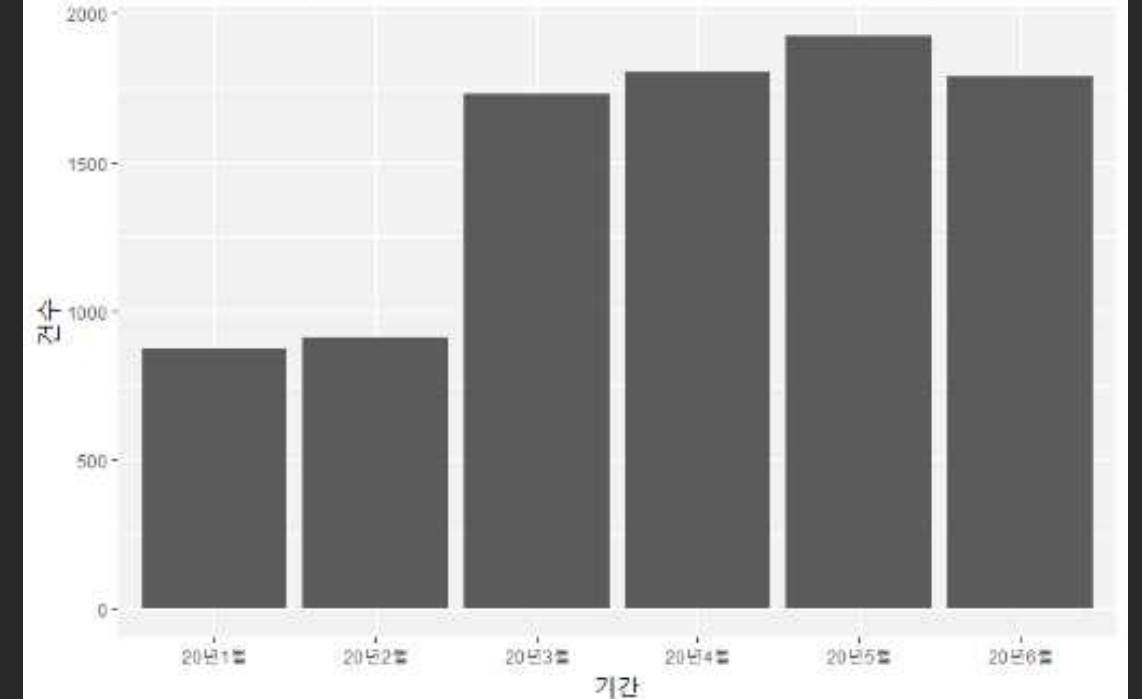
2020년 기간별 < 물류 > 분야 단어언급 횟수



2020년 기간별 < 문화 > 분야 단어언급 횟수



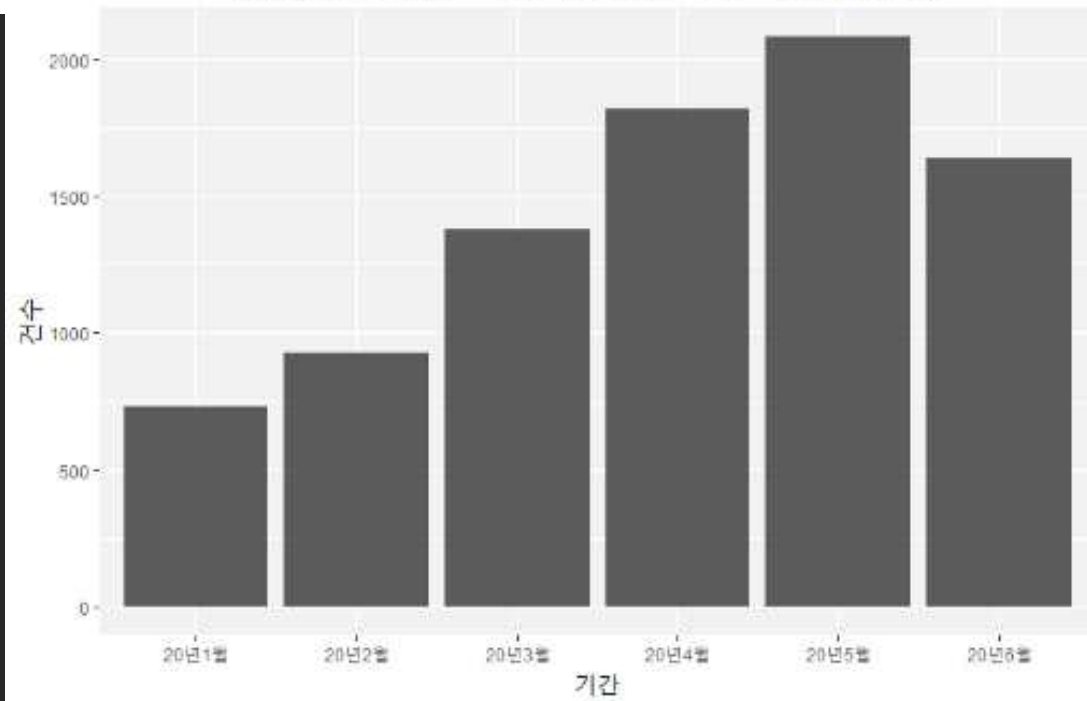
2020년 기간별 < 교통 > 분야 단어언급 횟수



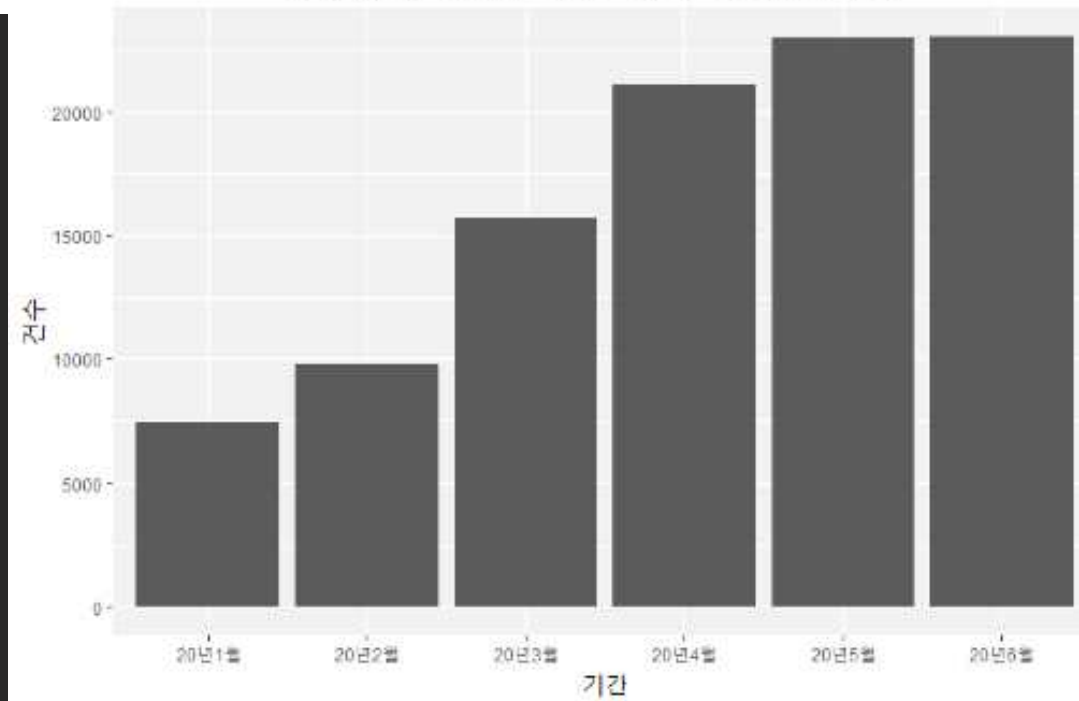
# 텍스트 마이닝 (R 시각화)

- 막대 그래프 (분야별 기간 변동)

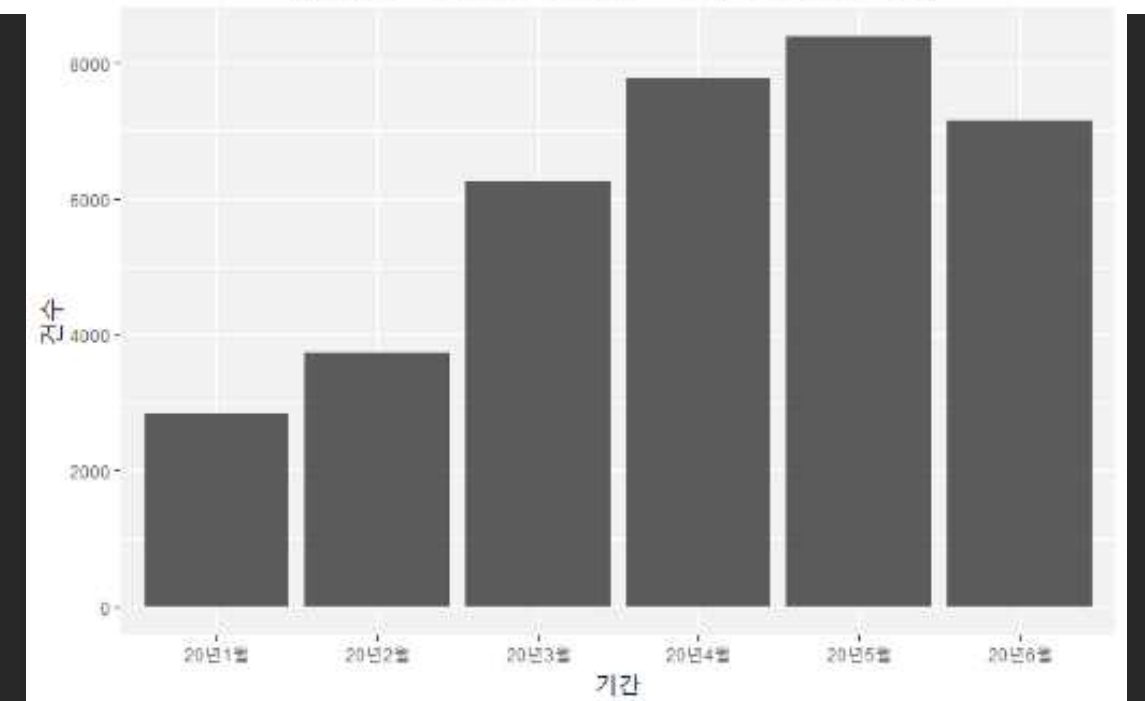
2020년 기간별 < 정보/보안 > 분야 단어언급 횟수



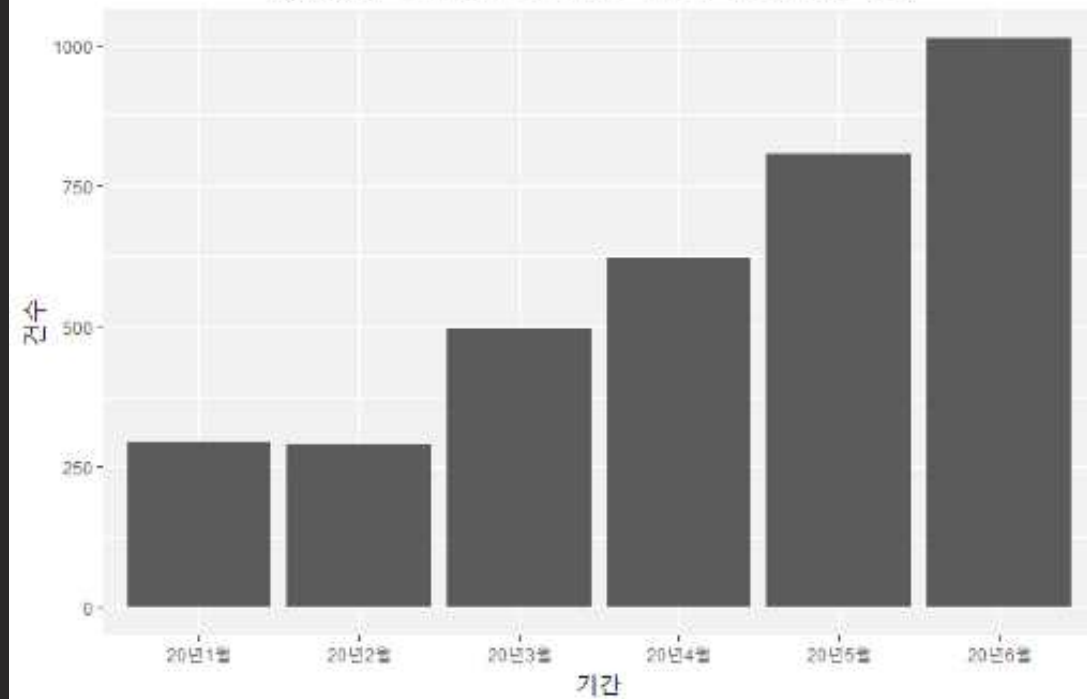
2020년 기간별 < IT > 분야 단어언급 횟수



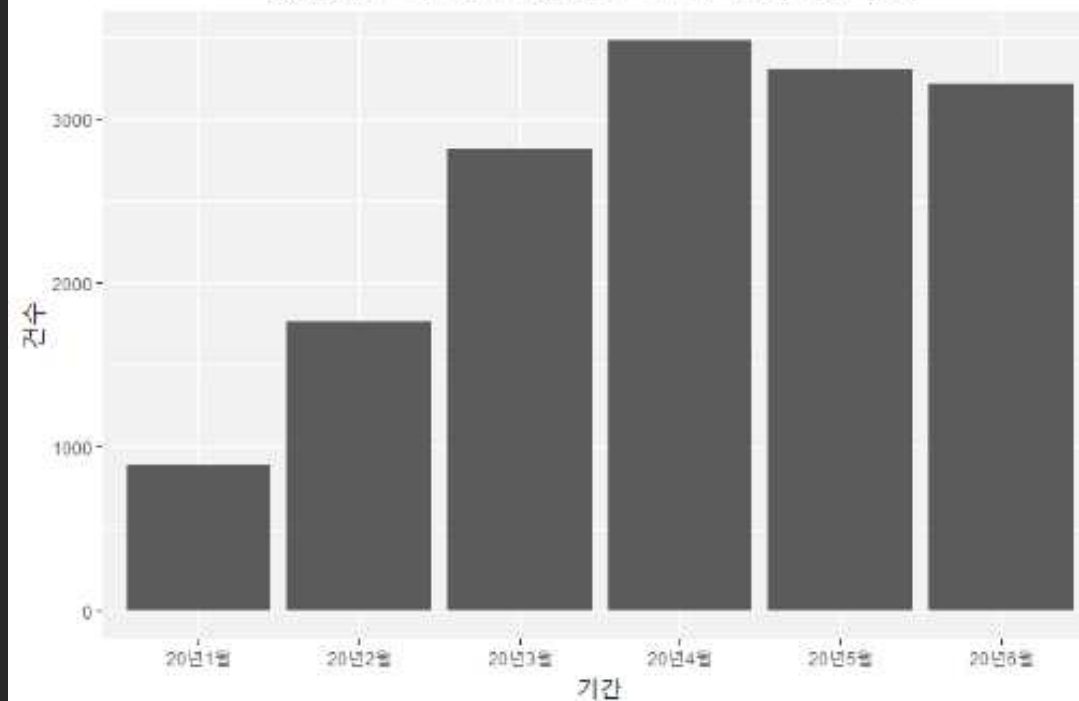
2020년 기간별 < 금융 > 분야 단어언급 횟수



2020년 기간별 < 환경 > 분야 단어언급 횟수



2020년 기간별 < 통신 > 분야 단어언급 횟수



정보/보안 : 화상 회의 보안, 어플 결제 보안

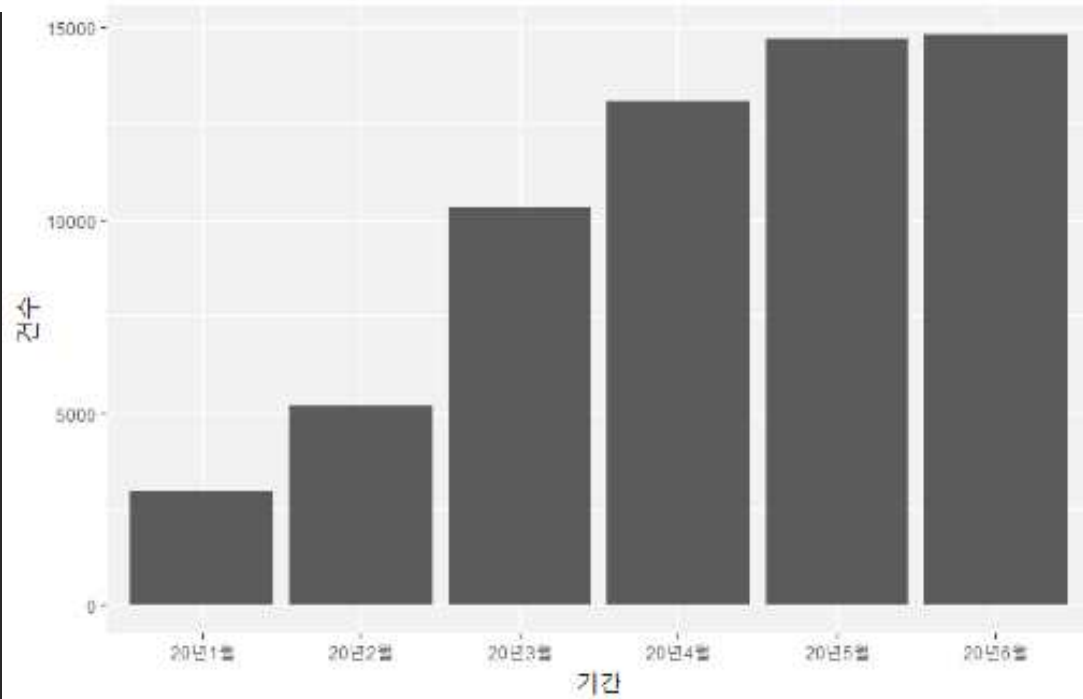
IT : 서버, 플랫폼 개발

금융 : 모바일 통장 개설, 송금

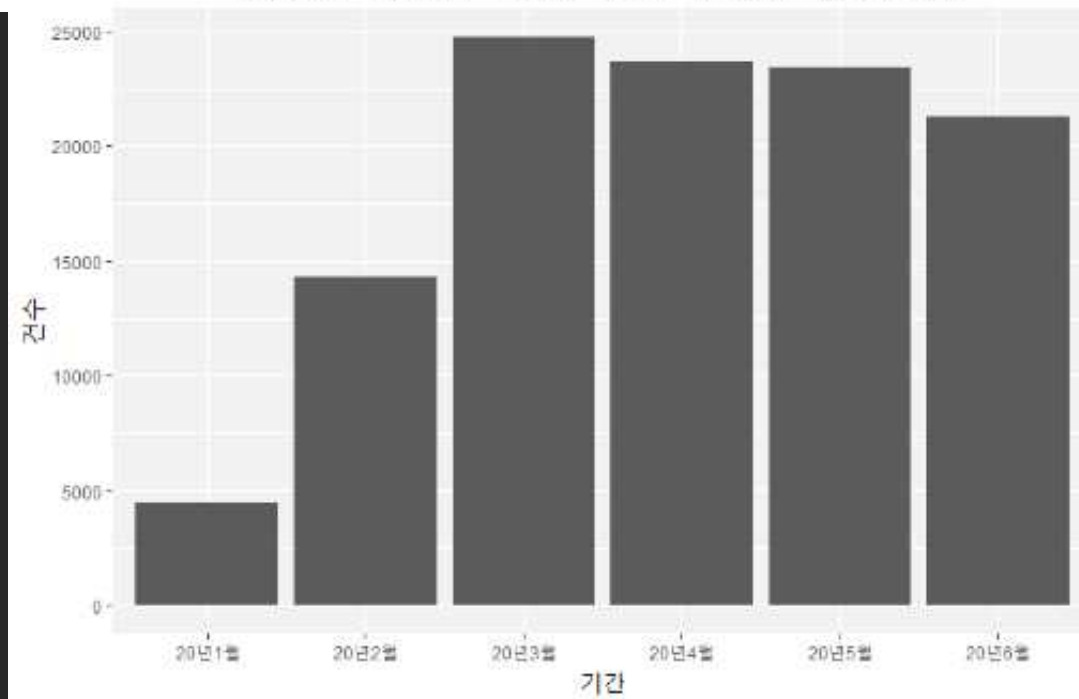
# 텍스트 마이닝 (R 시각화)

- 막대 그래프 (분야별 기간 변동)

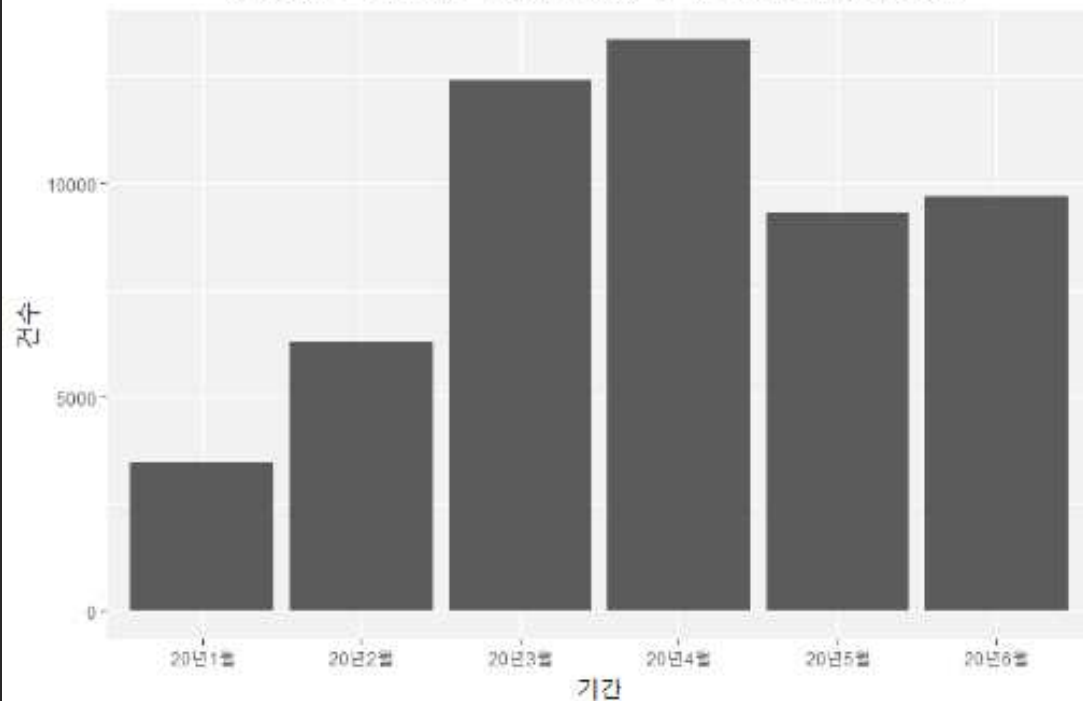
2020년 기간별 < 비즈니스 > 분야 단어언급 횟수



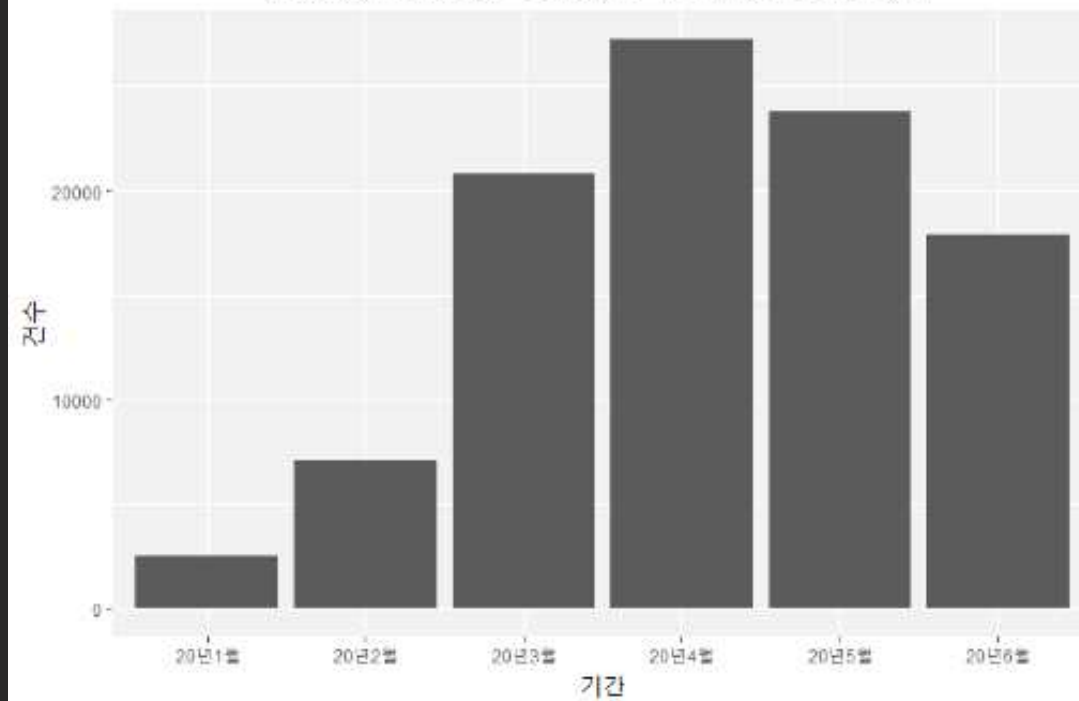
2020년 기간별 < 건강/의료 > 분야 단어언급 횟수



2020년 기간별 < 영상/방송 > 분야 단어언급 횟수



2020년 기간별 < 교육 > 분야 단어언급 횟수



비즈니스 : 화상회의, 채용(면접)

건강/의료 : 원격진료, AI기반 질병진단

영상/방송 :  
스트리밍 서비스, 선거 유세, 온라인 예배

교육 : 원격수업, 스마트 교육,  
온라인 수업용 대용량 통신기술

# 텍스트 마이닝 (R 시각화)

키워드 증가량 확인

1	2	3	4	5	6	7	8	9	10	11	12	13	14
건강/의료	교육	환경	물류	제조	문화	정보/보안	교통	비즈니스	금융	영상/방송	거래	통신	it

	A	B	C	D	E	F	G	H	I	J
1	word	X20년1월	X20년2월	X20년3월	X20년4월	X20년5월	X20년6월	cnt	X	rv1
2	예비	0	427	1131	569	100	97	2324	11	Inf
3	누릴	0	0	2	152	757	581	1492	0	Inf
4	신권지	0	549	473	44	12	17	1095	0	Inf
5	캐대믹	0	5	134	129	181	91	540	0	Inf
6	연백트	3	72	347	657	1005	1013	3097	9	338.6867
7	등록금	4	23	135	144	104	818	1028	2	153.5
8	등교	9	48	261	567	3245	1378	5506	2	151.8889
9	포스트	10	9	17	141	527	485	1189	0	47.5
10	교회	10	458	1295	397	168	435	2783	11	42.5
11	화상회의	5	24	145	273	184	168	799	9	32.6
12	항제	4	3	213	203	126	133	682	1	32.25
13	세탁근루	11	434	889	577	476	358	2745	9	31.54545
14	원격근루	4	61	155	118	90	115	543	9	27.75
15	수원	66	361	2338	4080	2833	1795	11473	2	26.19897
16	학년	28	30	373	1850	953	862	3694	2	24.46154
17	화상	17	95	392	463	577	428	1770	9	24.05882
18	민화제	9	32	21	128	178	205	573	6	21.77778
19	진료	27	105	190	370	570	542	1804	1	19.07407
20	비대	127	242	888	1533	1813	2448	7151	9	18.27559
21	포집	12	46	41	95	120	217	531	0	17.08383
22	병역	70	579	1232	1228	1875	1229	6213	1	16.55714
23	의료기기	11	29	83	136	187	184	630	1	15.72727
24	판정	29	179	480	146	481	485	1800	0	15.72414
25	연결	23	75	242	225	252	379	1196	9	15.47826
26	감염증	104	966	1953	2020	1624	1563	8230	1	14.02885
27	하구	22	27	77	208	458	310	1102	6	13.09091
28	학기	25	109	224	257	218	344	1177	2	12.76
29	메모리	11	20	107	171	184	148	639	5	12.27273
30	확진자	37	411	651	505	507	490	2601	0	12.24324
31	교육청	17	54	270	381	389	211	1312	2	11.41176
32	생중계	20	133	387	372	368	245	1520	6	11.25
33	물류센터	13	55	51	55	381	147	702	4	10.30769
34	편드대니자	37	174	332	469	477	413	1902	0	10.16216
35	코로나	342	2221	4566	5203	4326	3787	20445	1	10.0731
36	제차	12	48	81	178	149	127	595	0	9.583333
37	교직원	24	53	96	60	176	243	652	2	9.125
38	학교	168	440	1698	1666	2389	1678	8019	2	8.938095
39	금강	14	49	139	158	112	139	611	6	8.928571
40	중학교	18	26	106	338	189	153	827	2	8.5625
41	검사	79	285	687	513	823	752	3139	1	8.518987
42	선생님	8	31	41	273	192	58	589	2	8.333333

```
rv = c()
for (i in 1:nrow(a)){
  a1 <- a[i,2:7]

  rv[i] = (a[i,2:7][6]-a[i,2:7][1])/a[i,2:7][1]
}

rv1=matrix(rv,ncol=1)

rv2= cbind(a,rv1)
```

1월에 비해 6월 증가량 상위 분포 분야

비즈니스

교육

건강/의료

# 결론



4차 산업 혁명 및 5세대 이동 통신 기술의 발전으로 비대면 산업의 규모는 커지고 있다.

코로나-19 이후 행동양식이 달라진 분야에 대한 요구가 증가하고 있다.

예. 교육, 비즈니스, 공연 등



# Q&A

감사합니다.

