

- 도서정가제로 인한 시장상황 분석

201501489 최영진





#### 목표

도서정가제로 인한 시장 상황 빅데이터를 분석하여 향후 출판/문화 업계의 시장 변화 예측 (DNR) 해당 법안의 개정 방향과 대응 방안 추론

- 각 측의 주장 정리
- 여론 조사
- 주장에 대한 근거 빅데이터로 확인

#### 예상 분석

현행 도서정가제는 출판 생태계를 지킴과 동시에 소비 시장을 축소시킨다.

### 개정 및 대응 방안

형태가 다양해진 시장 변화에 맞추어 도서정가제를 유지하되 완화해야 한다.

#### 최종 목표

도서정가제가 유지되는 경우(=예상 분석이 일치하는 경우), 출판물 소비 시장의 확대와 성장을 위한 홍보 예측 모델 구축

### 방법

구매 결정 요인(=홍보 요인)과 각 상관 관계를 노드와 엣지로 표현하고 판매자의 투자 지표와 소비자의 인식(여론) 지표로 가중치를 계산. 가중치 변화에 따른 판매량 변화 확인.



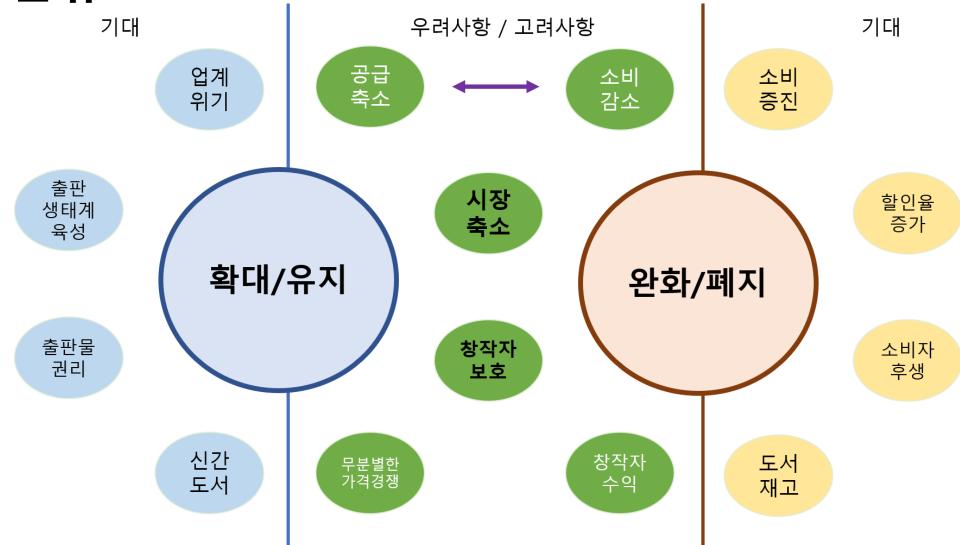
출판 업계 주장 = "출판 시장이 버티고 있는 건 현행 도서정가제 덕분"

**1** 출판 생태계가 정말 도서정가제에 달려있을까?

소비자의 주장 = "시장 축소는 높아진 가격으로 인해 소비가 줄어 발생 "

2 소비 시장의 흐름은 도서정가제에 영향을 어떻게 받았는가?

# 쟁점 분류



출판 생태계가 정말 도서정가제에 달려있을까?

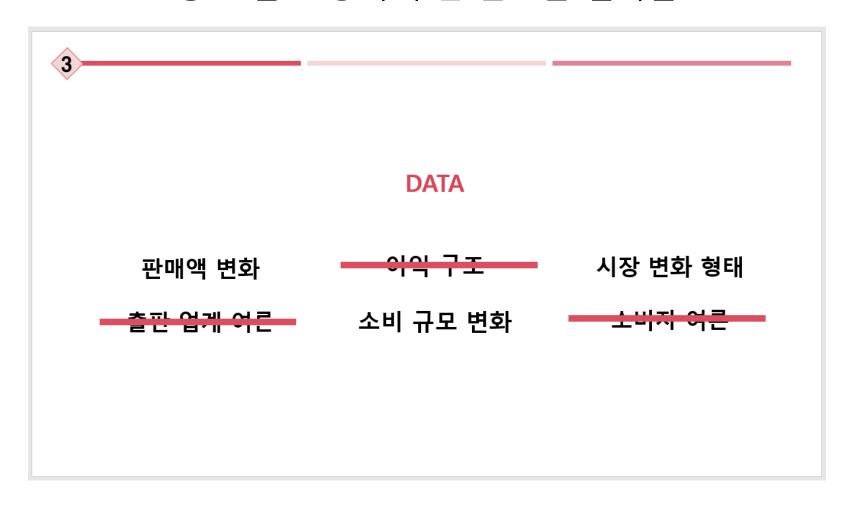
-출판업계 미래 투자액

출판업계의 판매량 변화, 시장 변화 형태에 따른 \_\_\_ 출판업계 반응 및 여론조사

소비 시장의 흐름은 도서정가제에 영향을 어떻게 받았는가?

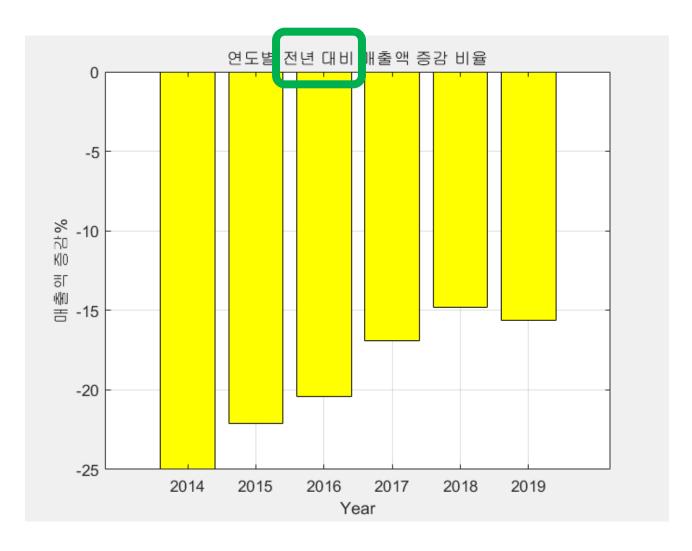
소비자 여론 조사, 시장 변화 형태에 소비자 인식에 따른 소비규모 변화 소비자 인식의 영향력 확인 소비자 여론 조사,

#### 중간 발표 당시 수집 완료한 결과물

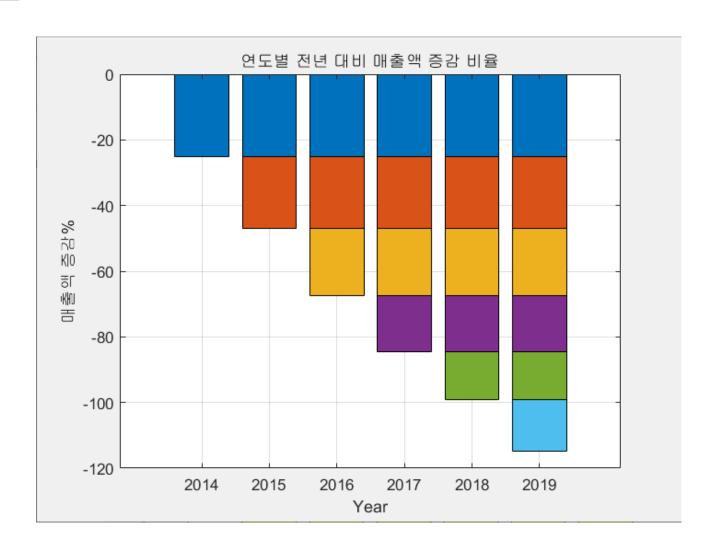




# 매출액 증감



# 매출액 감소



# 영업 이익률

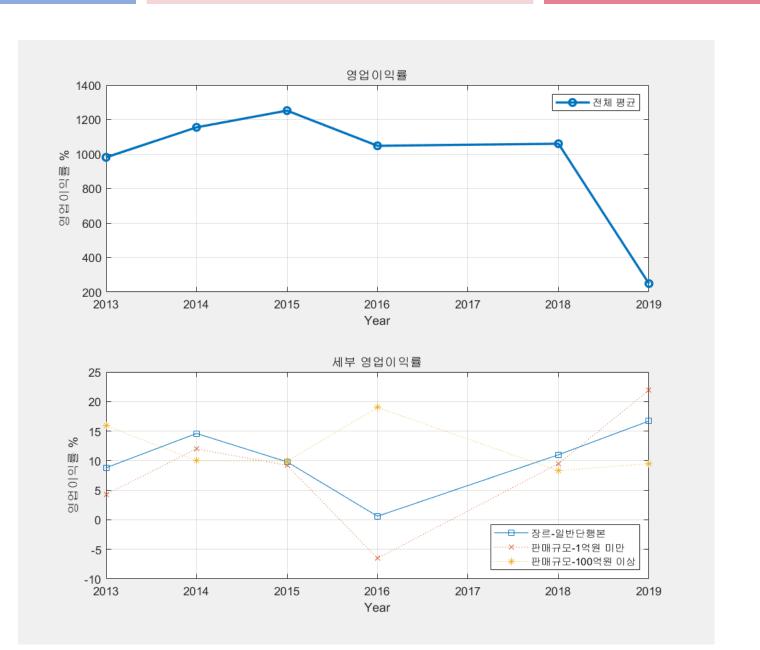
	2013	2014	2015	2016	2018	2019
전체	10.7	14.8	9.9	4.8	11.8	18.4
일반 단행본	8.8	14.6	9.8	0.6	11	16.7
학술/전문서	13.2	16.8	12.1	7.2	12.8	25.1
교과서 및 학습 참고서	14.5	12.9	1.7	8.5	18.6	12.5
아동도서	10.4	12.4	9.7	10.8	11	13.2
학습지	19	12.3	12.5	21.4	6	14.5
전집	26.3	-11.7	0.7	11.2	8.2	10.1
기타	9.6	14	6		13.4	23.1
1억원 미만	4.3	12	9.2	-6.5	9.5	21.9
1~10억원 미만	13.6	18.4	11.3	10.1	15.4	16.8
10~100억원 미만	15.8	15.9	14.8	10.6	11.9	14.9
100억원 이상	15.9	10	9.9	19	8.3	9.5
1~4인	8.2	13.5	9.1	4	23.9	36.7
5~9인	16.7	18.2	11.8	8.1	11.8	17.8
10~49인	16.3	17	11.8	-2.9	11.9	16.6
50이상	22	14.5	18.1	18	9.9	18.2
전체 사례 수	981	1155	1252	1048	1060	249

#### 평균 이익률

2017년도 제외

#### 세부 이익률

2017년도 제외



## 시장이 변화한 시기

웹소설 시장이 확대 된 계기 - 네이버 웹소설 2013년 open

하지만, 여전히 낮은 사회기여도 조사

2016년 기점으로 웹소설과 웹툰 기반

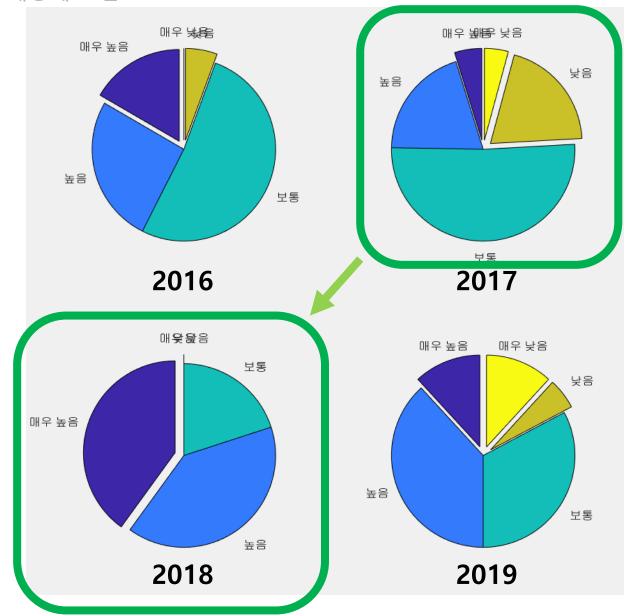
드라마, 영화화 관심도 증가 관심도가 높았던 예시 : 구르미 그린 달빛, 지즈인더트랩 등등

웹소설 및 전자책 사회기여도 상승

2015년도 기점으로 조사 대상에 포함

# 전자책

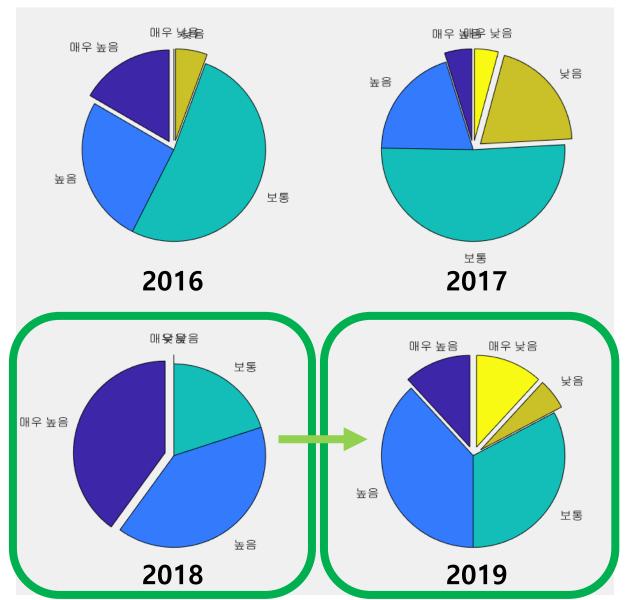
전자책 시장 가능성 - 업계의 여론



초중반에 발표된 자료이므로 실제로는 전 년도~해당 년도 상반기에 해당하는 결과 2015년도 기점으로 조사 대상에 포함

# 전자책

전자책 시장 가능성 - 업계의 여론

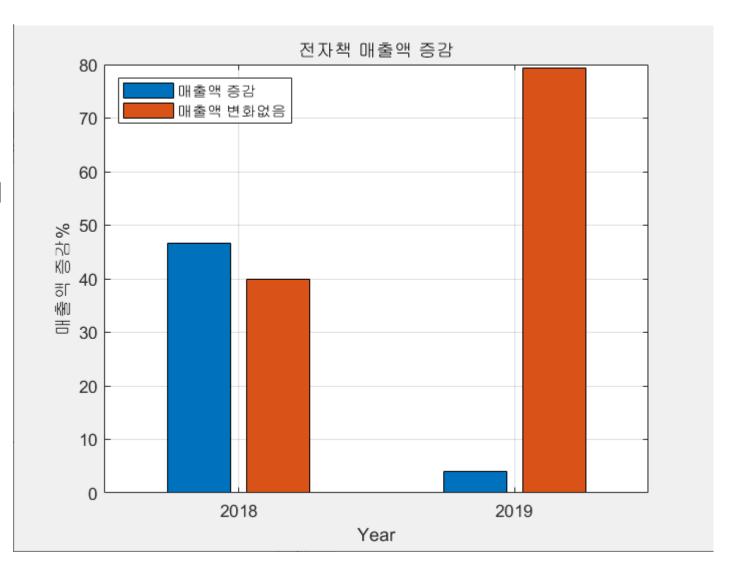


다시 등장한 부정적 여론의 이유는 무엇일까? 시장 가능성 여론에 영향을 끼친

# 매출액

꾸준히 하락하던 전반 도서 매출액과 달리

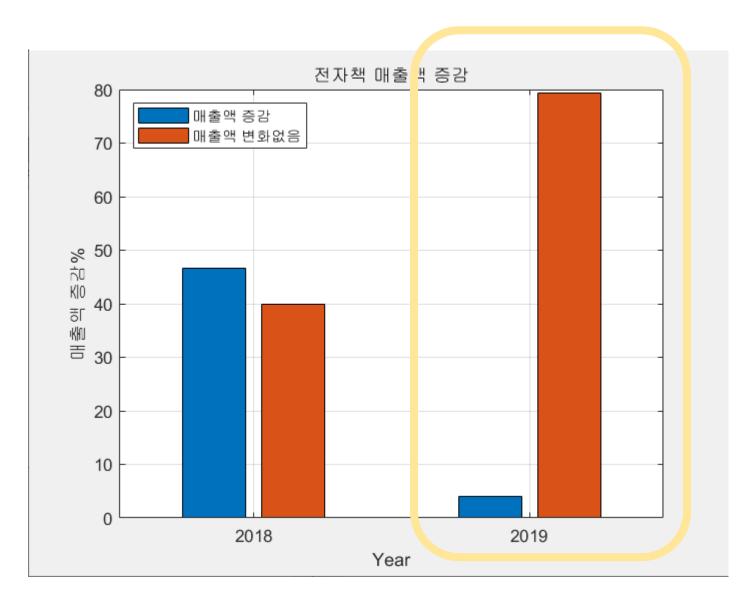
꽤나 큰 폭의 상승률의 전자책 매출액



하지만 그에 비해

2019년도에 발표된 매출액 증가율은 큰 변화 없음.

# 매출액



하지만 사실 전년도 대비 전자책 매출액은 증가된 매출액 유지

플랫폼의 다양화로 소비층 분산되었음에도

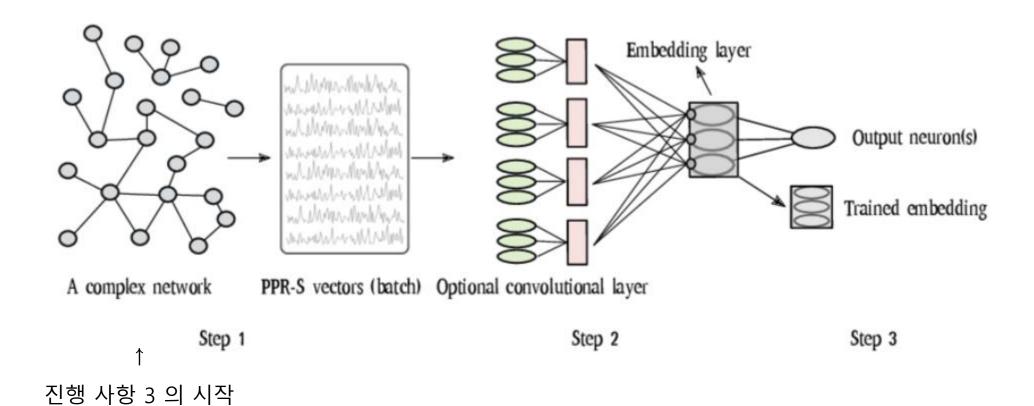
일정 매출액이 확보된 결과 출판 생태계가 정말 도서정가제에 달려있을까?

2 소비 시장의 흐름은 도서정가제에 영향을 어떻게 받았는가?

3 예상 분석이 일치했을 경우, 다음? 대응방안 추론

PageRank 알고리즘 이용한 Deep Node Ranking 시스템 구축 도전

# Deep Node Ranking



#### 한거리

웹소설협회 "웹소설 제외 도서정가 제 개악 반대"

창작스토리작가협회 "웹소설 도서정가제. 시장 죽이는것"

웹툰 산업계 도서정가제 대응 가시화, '협의체' 구성

☑ 구독하기

## 여론 신뢰도 문제











- 한국장작스토리작가협회, 웹소설 도서정가제 개정안에 입장 밝혀

포털의 시장 장악·중소기( "양대 포털 비공식 '무료 [ "도서 생태계를 적자생존

미리 수집했던 출판업계 측 발표자료

하고 연재 중인 웹소설을 도

외하기로 한 문화체육관광부의 도서정가제 개정안 초

수 수성대학교에 원소설, 원론, 원론스… [8]





웹툰산업계가 '도서정가제 대응과 웹툰 분류체계 추진협

자'. 셈0

리얼미터 발표 보도자료

데이터양 부족 → 가쟛

유통 플랫폼만이 시장을 독점하고, 중소기업은 부실의 늪에서 허덕이다 침몰하게 될 것"이라고 주장했다. 협

기반으로 해서 판매하는 소설이 아닌, 웹을 기반으로 하 는 소설"이라며 "소설이라는 형태와 도서로 만들어질 수

내 주요 업체들이 모여 별도의 협의체를 구성하게 됐다고 삐겨으 서며해다

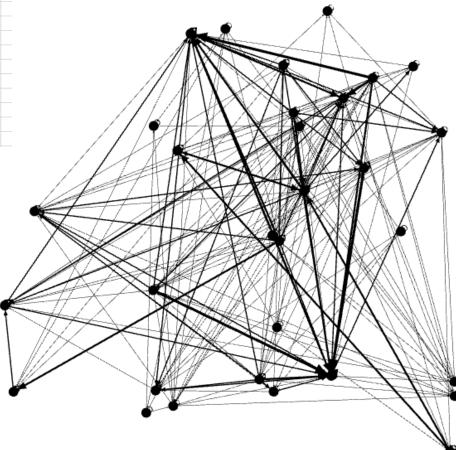
Α	ВС	
Id	Label	
0	시대흐름	
1	소장	
2	자기만족	
3	선물	
4	출판사 명성	
5	내용유용성	
6	작가 명성	
7	내용 필요성	
8	디자인	
9	책읽는 환경	
10	도서관 대출 서비:	_
11	중고도서	
12	주변인 추천	
13	영화	
14	TV시청	
15	타인평가	
16	책값	
17	서평	
18	인터넷 검색	
19	우수추천도서	
20	베스트셀러	
21	기분전환	
22	흥미	
23	재미	
24	정보획득	
25	타인과의 대화	
26	교환 반품 용이	
27	신문 잡지	
28	독서 캠페인	
29	인터넷	
30	라디오	

4	Α	В	С	D
1	id	Source	Target	Туре
2	0	0	_	Directed
3	1	1	1	Directed
4	2	1	0	Directed
5	3	4	4	Directed
6	4	5	5	Directed
7	5	7	7	Directed
8	6	7	5	Directed
9	7	7	4	Directed
10	8	5	4	Directed
11	9	9	9	Directed
12	10	5	10	Directed
13	11	4	10	Directed
14	12	9	10	Directed
15	13	0	10	Directed
16	14	11	10	Directed
17	15	10	10	Directed
18	16	5	11	Directed
19	17	4	11	Directed
20	18	9	11	Directed
21	19	0	11	Directed
22	20	11	11	Directed
23	21	5	0	Directed
24	22	4	0	Directed
25	23	9	0	Directed
26	24	0	0	Directed
27	25	5	9	Directed
28	26	4	9	Directed
29	27	9	9	Directed
30	28	5	4	Directed
31	29	4	4	Directed
32	30	5	5	Directed
33	31	24	21	Directed
34	32	23	21	Directed
25	22	22	24	Directed

4	Α	В	С	D	
384	382	10	0	Directed	Г
385	383	11	0	Directed	
386	384	27	27	Directed	
387	385	1	27	Directed	
388	386	21	27	Directed	
389	387	10	27	Directed	
390	388	11	27	Directed	
391	389	1	1	Directed	
392	390	21	1	Directed	
393	391	10	1	Directed	
394	392	11	1	Directed	
395	393	21	21	Directed	
396	394	10	21	Directed	
397	395	11	21	Directed	
398	396	10	10	Directed	
399	397	11	10	Directed	
400	308	11	11	Directed	

### 구매 결정 요인 상관관계

#### 가중치 계산 X







```
split.py
    import os
    from PyPDF2 import PdfFileReader, PdfFileWriter
    path = r'D:\학교\빅데이터 입문\도서정가제 data\분할용 pdf'
   s_path = r'D:\학교\빅데이터 입문\도서정가제 data\분할 페이지'
    for item in os.listdir(path):
        item_path = os.path.join(path, item)
       pdf = PdfFileReader(open(item_path, 'rb'))
        # pdf페이지수
       numberPages = pdf.getNumPages()
        for page in range(1, numberPages):
           pdf_writer = PdfFileWriter()
           pdf_writer.addPage(pdf.getPage(page))
           output_name = 'page{num}_{name}'.format(num=page, name=item)
           save_path = os.path.join(s_path, output_name)
           with open(save_path, 'wb') as f:
               pdf_writer.write(f)
```



#### PC > 로컬 디스크 (D:) > 학교 > 빅데이터 입문 > 도서정가제 data > 분할용 pdf

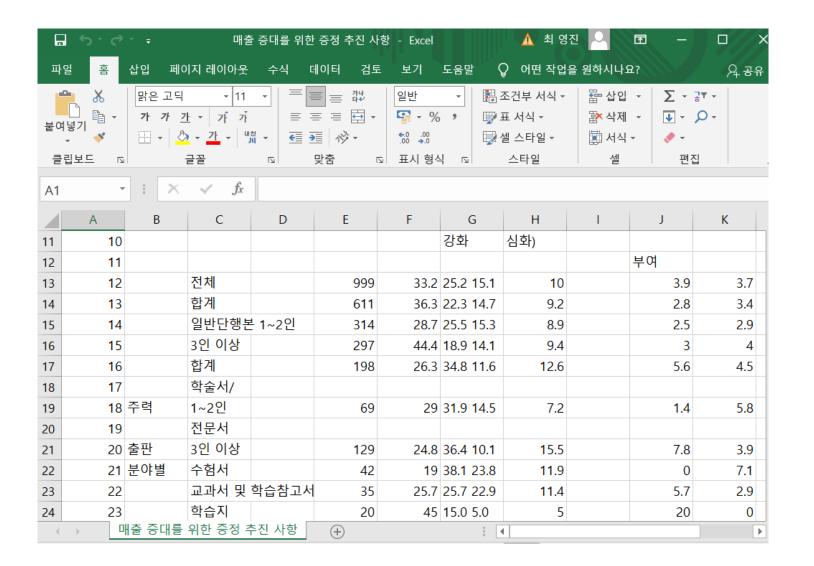
이름	수정한 날짜	유형	크기
🔈 (자료1)개정+도서정가제+바로알기(2009	2020-11-12 오후 4:50	Hpdf Document	245KB
ዬ (조사연구19-05) 출판물의 지역서점 공급	2020-11-12 오후 4:53	Hpdf Document	98,369KB
ዬ 개정 도서정가제 영향평가 및 향후 방향	2020-11-12 오후 4:55	Hpdf Document	60,820KB
ዬ 디지털 시대의 출판 저작권 보호를 위한	2020-11-12 오후 4:53	Hpdf Document	159,288KB
鴙 소비자 도서구매 결정요인 분석	2020-11-12 오후 4:54	Hpdf Document	206,238KB
🖲 웹소설산업 실태조사 보고서	2020-11-12 오후 4:54	Hpdf Document	2,385KB
ዬ 지역서점 현황조사 및 진흥정책 연구	2020-11-12 오후 4:56	Hpdf Document	5,358KB

디스크 (D:) 〉 학교 〉 빅데이터 입문 〉 도서정	3/1세 Udla / 군일 페이지	~ O	○ 분할 페이지 검색	
이름	수정한 날짜	유형	크기	^
	7070 15 07 TT 10:15	ripai Document	TIND	
ዬ page11_개정 도서정가제 영향평가 및 향	2020-12-09 오후 10:12	Hpdf Document	129KB	
ዬ page11_디지털 시대의 출판 저작권 보호	2020-12-09 오후 10:13	Hpdf Document	328KB	
ዬ page12_(조사연구19-05) 출판물의 지역	2020-12-09 오후 10:12	Hpdf Document	217KB	
🖲 page12_개정 도서정가제 영향평가 및 향	2020-12-09 오후 10:12	Hpdf Document	130KB	
ዬ page12_디지털 시대의 출판 저작권 보호	2020-12-09 오후 10:13	Hpdf Document	472KB	
ዬ page13_(조사연구19-05) 출판물의 지역	2020-12-09 오후 10:12	Hpdf Document	262KB	
ዬ page13_개정 도서정가제 영향평가 및 향	2020-12-09 오후 10:12	Hpdf Document	130KB	
ዬ page13_디지털 시대의 출판 저작권 보호	2020-12-09 오후 10:13	Hpdf Document	417KB	
ዬ page14_(조사연구19-05) 출판물의 지역	2020-12-09 오후 10:12	Hpdf Document	278KB	
ዬ page14_개정 도서정가제 영향평가 및 향	2020-12-09 오후 10:12	Hpdf Document	138KB	
ዬ page14_디지털 시대의 출판 저작권 보호	2020-12-09 오후 10:13	Hpdf Document	345KB	
ዬ page15_(조사연구19-05) 출판물의 지역	2020-12-09 오후 10:12	Hpdf Document	43KB	
ዬ page15_개정 도서정가제 영향평가 및 향	2020-12-09 오후 10:12	Hpdf Document	1KB	
ዬ page15_디지털 시대의 출판 저작권 보호	2020-12-09 오후 10:13	Hpdf Document	484KB	
ዬ page16_(조사연구19-05) 출판물의 지역	2020-12-09 오후 10:12	Hpdf Document	19KB	
🖲 page16_개정 도서정가제 영향평가 및 향	2020-12-09 오후 10:12	Hpdf Document	73KB	
ዬ page16_디지털 시대의 출판 저작권 보호	2020-12-09 오후 10:13	Hpdf Document	489KB	
ዬ page17_(조사연구19-05) 출판물의 지역	2020-12-09 오후 10:12	Hpdf Document	1KB	
🖲 page17_개정 도서정가제 영향평가 및 향	2020-12-09 오후 10:12	Hpdf Document	1KB	
ዬ page17 디지털 시대의 출판 저작궈 보호	2020-12-09 오후 10:13	Hpdf Document	641KB	~



RStudio File Edit Code View Plots Session Build Debug Profile Tools Help ■ • Addins • ata extract.R × Q cdc table × Duntitled2\* × Environment History Connections Tutorial Run 🐪 🖶 Source 🗸 🗏 1 library(magick) ■ Global Environment ▼ Data cdc\_img <- image\_read\_pdf("D:/학교/빅데이터 입문/도서정가제 data/출판 산업 density = 300, pages = 1) cdc\_table List of 1 values 6 cdc\_draw\_img <- cdc\_img %>% cdc\_draw\_img External pointer of class ' image\_scale("350") External pointer of class ' cdc\_img 9 image\_draw(cdc\_draw\_img) 11 rect(30, 80, 320, 350, border = "red", lty = "dashed", lwd = 2) 12 # dev.off() 13 14 library(tabulizer) 15 library(tidyverse) 16 17 cdc\_table <- extract\_tables("D:/학교/빅데이터 입문/도서정가제 data/출판 산업 Files Plots Packages Help Viewer 18 19 write.csv(cdc\_table,"D:/학교/빅데이터 입문/도서정가제 data/출판 산업 통합 실 20 21 RE 출판산업 설득조사 경마 출판 사업체 ··· (표 3-15) 매출 증대를 위한 중점 추진 사항(1순위) R Script \$ 21:1 (Top Level) \$ Console Terminal × Jobs ×  $\neg$ 한집 1억 원 미만 매충역 1~10억 등 미만 규모별 10~100억 원 미만 ~/@ # A tibble: 1 x 7 format width height colorspace matte filesize density <chr> <int> <int> <chr> <1q1> <int> <chr> 350 483 sRGB TRUE 0 300x300 규모병 > rect(30, 80, 320, 350, border = "red", lty = "dashed", lwd = 2) > # dev.off() > library(tabulizer) > library(tidyverse)

> totaly(trayverse)
> cdc\_table <- extract\_tables("D:/학교/빅데이터 입문/도서정가제 data/출판 산업 통합 실
태(2013~2019)/2019 출판사업 실태조사 보고서-페이지/page-123.pdf", encoding = "UTF-





**1** 출판 생태계가 정말 도서정가제에 달려있을까?

YES 도서 정가제 이전의 출판 시장은 가파른 하락세 하락세를 잠시 늦춰주는 역할

소비 시장의 흐름은 도서정가제에 영향을 어떻게 받았는가?

시행 직후 소비자 판매액 축소 소규모 출판사들은 크게 위축, 큰 출판사만 겨우 규모 유지이후로도 회복세는 보이지 않음.

오히려 소규모 전자 출판사들이 생겨나며, 도서정가제와 관련된 인식이 없었던 전자책이 눈에 띄는 성장세를 보임. → 이 오해로 인한 논란(전자책은 이미 도서정가제를 따르고 있었음.)

3 예상 분석이 일치했을 *감*수, 다음? 대응방안 추론

PageRank 인고리는 이용한
Deep Mode Ranking 지수템 구축 도전

프로젝트 최종 목표 도달 실패



# Q&A

교수님 질문 : 17page의

"드라마 인기로 전자책 매출이 일어난 것" 위험한 논리 점프 주의

데이터 시각화로 보여주지 못한 파트이기에 교수님이 하신 말씀이 옳습니다.

하지만 드라마 제목은 사실, TV매체에서 미디어 믹스를 시도한 여러 프로그램 중에 웹소설, 웹툰기반임이 (기사로 많이 나와서) 확실하게 알려졌던 유명한 드라마가 2016년 상영임을 알려주어 쉽게 언급도가 올라갔었음을 쉽게 이해시키기 위한 일종의 예시였고, 실제 논리를 도출하게 된 순서는 반대였습니다.

총 매출액 상승 -> 종이책은 상승X -> 전자책 매출액이 상승 & 전자책 시장 평가도 상승 -> 발전 가능성 & 향후 투자계획 데이터 수치 동반 상승 -> 왜지? -> (2016, 미디어, 전자책, 웹소설, ...) 등의 키워드를 조합 검색 도출한 것입니다. (뒷장 이어)

# Q8LA

아래는 전자책 매출액을 확인하던 과정에서 추출했던 표와 글 데이터의 일부입니다.

사실 매출액이나 이익률이 증가했음을 확인했던 이후에, 정확하게 데이터를 마저취합하려 했으나 연도별 '영업이익률' 계산에 필요한 모든 데이터(정확한 매출액, 지출액등)의 존재 여부가 연도별로 달라, 최종적으로 데이터로 시각화 자료에서 빼게 되었습니다. 사람들의 인식률과 연관이 있는 사회기여도는 시간 부족으로 아예 데이터 시각화 시도를할 엄두조차도 못했으므로 교수님이 지적하신 바와 같이 위험한 논리 도출이 맞습니다. 조언 감사합니다. 가르쳐 주신대로 꼼꼼하게 찾아서 노력해보겠습니다.

196

2016 출판산업 실태조사



매출 현황

#### 1. 전자책 매출액

2015년 전자책 출판사의 1곳당 평균 매출액은 약 1억 4천 8백만 원으로 나타났다. 출판사 유형별로 순수 전자책 출판사가 2억 3천 4백만 원으로 종이책 기반 출판사보다 상대적으로 전자책 매출액이 높았다. 주력 분야별로 살펴보면, 장르문학 주력 출판사가 3억 9천 7백만 원, 멀티미디어 전자출판물(아동, 어학, 교육) 주력 출판사가 1억 5천 3백만 원, 기타(웹툰 등) 분야 주력 출판사가 1억 4천 2백만 원, 일반 분야 주력 출판사는 8천 8백만 원인 것으로 집계되었다.

웹소설과 모바일앱 매출 여부 및 오디오북 제작 여부에 따른 전자책 출판사의 2015년 전자책 매출액을 살펴보면, 웹소설 매출이 있는 전자책 출판사(2억 4백만 원)가 그렇지 않은 출판사보다 상대적으로 높았고, 모바일앱 매출이 있는 출판사(1억 9천 7백만 원)가 그렇지 않은 출판사보다 상대적으로 높았으며, 오디오북을 제작하는 전자책 출판사(4억 2천 3백만 원)도 그렇지 않은 출판사보다 상대적으로 높은 것으로 나타났다.

2016년 발표(2015)

표 4-10 매출액 증감 (단위: %)

	구 분	사례수	증가함	거의 비슷함	감소함
	전 체	(364)	22.5	62.1	15.4
전자책	종이책 기반 출판사	(219)	20.1	66.7	13.2
출판사	순수 전자책 출판사	(113)	26.5	53.1	20.4
유형	기타(IT 기반 등)	(32)	25.0	62.5	12.5
	장르문학	(51)	35.3	52.9	11.8
주력 출판	일반 분야	(251)	19.1	66.1	14.7
물면 분야별	멀티미디어 전자출판물	(19)	10.5	68.4	21.1
도시크	기타(웹툰 등)	(43)	32.6	46.5	20.9
	1천만원 미만	(197)	18.8	66.5	14.7
DILEON	1-5천만원 미만	(66)	22.7	66.7	10.6
매출액 규모별	5천만원-1억원 미만	(34)	20.6	61.8	17.6
птэ	1-5억원 미만	(35)	37.1	42.9	20.0
	5억원 이상	(32)	31.3	46.9	21.9
X 1171	1-2인	(271)	20.3	62.7	17.0
종사자 규모별	3-4인	(44)	22.7	61.4	15.9
	5인 이상	(49)	34.7	59.2	6.1
웹소설	있음	(46)	41.3	45.7	13.0
매출	없음	(318)	19.8	64.5	15.7

2017년 발표(2016)

