

연도별 정보취약계층(장애인,노인,다문화) 이용자수 합계 변화 추이

- 1) 데이터 추출
- 2) 데이터 분석
- 3) 데이터 시각화 – 꺾은선 그래프
- 4) 결과 분석

1) 데이터 추출

```
jupyter Untitled Last Checkpoint: 46 seconds ago

File Edit View Run Kernel Settings Help Trusted

[1]: import pandas as pd

# CSV 파일 읽기
df = pd.read_csv('통합데이터.csv')

[2]: # desired_columns 정의
# 다문화봉사대상자수(2022), 장애인봉사대상자수(2020), 노인봉사대상자수(2020), 다문화봉사대상자수(2020) 이 네 개 컬럼에 컬럼치가 있음.
desired_columns = ['장애인_봉사대상자수_2023', '노인_봉사대상자수_2023', '다문화_봉사대상자수_2023',
'지식정보취약계층_서비스_이용자수_장애인_2023', '지식정보취약계층_서비스_이용자수_노인_2023', '지식정보취약계층_서비스_이용자수_다문화_2023',
'장애인_봉사대상자수_2022', '노인_봉사대상자수_2022', '다문화_봉사대상자수_2022',
'지식정보취약계층_서비스_이용자수_장애인_2022', '지식정보취약계층_서비스_이용자수_노인_2022', '지식정보취약계층_서비스_이용자수_다문화_2022',
'장애인_봉사대상자수', '노인_봉사대상자수', '다문화_봉사대상자수',
'지식정보취약계층_서비스_이용자수_장애인', '지식정보취약계층_서비스_이용자수_노인', '지식정보취약계층_서비스_이용자수_다문화',
'장애인_봉사대상자수_2020', '노인_봉사대상자수_2020', '다문화_봉사대상자수_2020',
'지식정보취약계층_서비스_이용자수_장애인_2020', '지식정보취약계층_서비스_이용자수_노인_2020', '지식정보취약계층_서비스_이용자수_다문화_2020']

# 새로운 데이터프레임 설정
new_df = df[desired_columns]

# 새로운 데이터프레임 출력
print(new_df)
```

- (1) '통합 데이터'에서 분석에 필요한 기본 데이터들 2020-2023 장애인, 노인, 다문화 봉사대상자수와 서비스 이용자수를 먼저 추출해 보았다.
- (2) 2019 데이터는 도서관코드가 일치하지 않아 통합 X, 4개년만 취합하였다.

*테이블명은 2020-2023이지만, 실제 2019-2022의 데이터를 담고 있다.

	장애인_봉사대상자수_2023	노인_봉사대상자수_2023	다문화_봉사대상자수_2023	\
0	22375	79897	3944	2808
1	22489	73543	3944	3944
2	9552	24660	1327	1327
3	28611	67924	2161	2161
4	17621	72235	2639	2639
...
1115	5834	27704	1178	1178
1116	16319	47812	3884	3884
1117	31196	89334	8289	8289
1118	12348	55962	2814	2814
1119	9552	24660	1327	1327

	지식정보취약계층_서비스_이용자수_장애인_2023	지식정보취약계층_서비스_이용자수_노인_2023	\
0	150	600	600
1	52	2422	2422
2	5	1735	1735
3	1330	39364	39364
4	15	3882	3882
...
1115	36535	4899	4899
1116	59	48	48
1117	458	557	557
1118	2	151	151
1119	0	0	0

	지식정보취약계층_서비스_이용자수_다문화_2023	장애인_봉사대상자수_2022	노인_봉사대상자수_2022	\
0	0	22395	67236	67236
1	156	21868	69884	69884
2	0	5634	24837	24837
3	645	28832	64348	64348
4	0	17525	68483	68483
...
1115	14764	5929	27689	27689
1116	7	16818	44682	44682
1117	18918	38163	82476	82476
1118	19	12311	53822	53822
1119	0	5634	24837	24837

	다문화_봉사대상자수_2022	지식정보취약계층_서비스_이용자수_장애인_2022	...	다문화_봉사대상자수	\
0	NaN	58	...	1838	1838
1	NaN	104	...	2354	2354
2	NaN	4	...	1282	1282
3	NaN	531	...	1966	1966
4	0	0	...	2580	2580
...
1115	NaN	2905	...	1135	1135
1116	NaN	54	...	3338	3338
1117	NaN	559	...	7254	7254
1118	NaN	1	...	1827	1827
1119	NaN	26	...	1282	1282

	지식정보취약계층_서비스_이용자수_장애인	지식정보취약계층_서비스_이용자수_노인	지식정보취약계층_서비스_이용자수_다문화	\
0	0	80	0	0
1	63	2620	31	31
2	52	1889	...	184
3	452	5422	...	78
4	0	289	...	12
...
1115	3288	2750	...	4857
1116	14	21	...	8
1117	258	287	...	423
1118	0	380	...	18
1119	0	49	...	0

	장애인_봉사대상자수_2020	노인_봉사대상자수_2020	다문화_봉사대상자수_2020	\
0	NaN	NaN	NaN	NaN
1	NaN	NaN	NaN	NaN
2	NaN	NaN	NaN	NaN
3	NaN	NaN	NaN	NaN
4	NaN	NaN	NaN	NaN
...
1115	NaN	NaN	NaN	NaN
1116	NaN	NaN	NaN	NaN
1117	NaN	NaN	NaN	NaN
1118	NaN	NaN	NaN	NaN
1119	NaN	NaN	NaN	NaN

	지식정보취약계층_서비스_이용자수_장애인_2020	지식정보취약계층_서비스_이용자수_노인_2020	\
0	0	35	35
1	184	4368	4368
2	182	431	431
3	1291	11838	11838
4	0	487	487
...
1115	1452	2646	2646
1116	7	38	38
1117	638	592	592
1118	0	358	358
1119	16	1872	1872

	지식정보취약계층_서비스_이용자수_다문화_2020	
0	0	0
1	52	2422
2	0	1735
3	134	39364
4	8	3882
...
1115	1460	4899
1116	12	48
1117	898	557
1118	13	151
1119	156	0

[1120 rows x 24 columns]

```
[3]: # 데이터프레임 슬라이싱
new_df = df[df.columns].copy()

# '2023 합계' 칼럼 생성
# 결측치는 특정 변수로 대체함.
new_df['2023 합계'] = new_df['지식정보취약계층_서비스_이용자수_장애인_2023'] + new_df['지식정보취약계층_서비스_이용자수_노인_2023'] + new_df['지식정보취약계층_서비스_이용자수_다문화_2023']

# 각 칼럼의 합 출력
print(new_df['지식정보취약계층_서비스_이용자수_장애인_2023'].sum())
print(new_df['지식정보취약계층_서비스_이용자수_노인_2023'].sum())
print(new_df['지식정보취약계층_서비스_이용자수_다문화_2023'].sum())

# 결과 출력
print(new_df['2023 합계'].sum())
```

426253
2462610
388237
3277100

[]: # 따라서, 2023년 다문화, 노인, 장애인 이용자수의 합계는 3277100이다.

```
[5]: # 데이터프레임 슬라이싱
new_df = df[df.columns].copy()

# '2022 합계' 칼럼 생성
# 결측치는 특정 변수로 대체함.
new_df['2022 합계'] = new_df['지식정보취약계층_서비스_이용자수_장애인_2022'] + new_df['지식정보취약계층_서비스_이용자수_노인_2022'] + new_df['지식정보취약계층_서비스_이용자수_다문화_2022']

# 각 칼럼의 합 출력
print(new_df['지식정보취약계층_서비스_이용자수_장애인_2022'].sum())
print(new_df['지식정보취약계층_서비스_이용자수_노인_2022'].sum())
print(new_df['지식정보취약계층_서비스_이용자수_다문화_2022'].sum())

# 결과 출력
print(new_df['2022 합계'].sum())
```

473876
2247728
409410
3131014

[]: # 따라서, 2022년 다문화, 노인, 장애인 이용자수의 합계는 3131014이다.

- 1차적으로 추출한 데이터를 기반으로 필요한 칼럼들을 이용하여 원하는 결과를 도출하였다.

- 각 계층의 이용자수 합계를 먼저 구한 뒤, 세 개의 결과값을 합치는 방식으로 진행.

→ 왼쪽의 사진은 순서대로 2023-2022 연도별 정보취약계층(장애인,노인,다문화) 이용자수 합계를 구하는 과정.

```
[6]: # 데이터프레임 슬라이싱
new_df = df[df.columns].copy()

# '2021' 합계' 칼럼 생성
# 결측치는 특정 변수로 대체함.
new_df['2021 합계'] = new_df['지식정보취약계층_서비스_이용자수_장애인'] + new_df['지식정보취약계층_서비스_이용자수_노인'] + new_df['지식정보취약계층_서비스_이용자수_다문화']

# 각 칼럼의 합 출력
print(new_df['지식정보취약계층_서비스_이용자수_장애인'].sum())
print(new_df['지식정보취약계층_서비스_이용자수_노인'].sum())
print(new_df['지식정보취약계층_서비스_이용자수_다문화'].sum())

# 결과 출력
print(new_df['2021 합계'].sum())
```

379950
1608176
272968
2261094

[]: # 따라서, 2021년 다문화, 노인, 장애인 이용자수의 합계는 2261094이다.

```
[7]: # 데이터프레임 슬라이싱
new_df = df[df.columns].copy()

# '2021' 합계' 칼럼 생성
# 결측치는 특정 변수로 대체함.
new_df['2020 합계'] = new_df['지식정보취약계층_서비스_이용자수_장애인_2020'] + new_df['지식정보취약계층_서비스_이용자수_노인_2020'] + new_df['지식정보취약계층_서비스_이용자수_다문화_2020']

# 각 칼럼의 합 출력
print(new_df['지식정보취약계층_서비스_이용자수_장애인_2020'].sum())
print(new_df['지식정보취약계층_서비스_이용자수_노인_2020'].sum())
print(new_df['지식정보취약계층_서비스_이용자수_다문화_2020'].sum())

# 결과 출력
print(new_df['2020 합계'].sum())
```

530672
2880700
545141
3956513

[]: # 따라서, 2020년 다문화, 노인, 장애인 이용자수의 합계는 3956513이다.

→ 왼쪽의 사진은 순서대로 2022-2020 연도별 정보취약계층(장애인,노인,다문화) 이용자수 합계를 구하는 과정.

2) 데이터 분석

[2023년 정보취약계층 이용자수 합계] = 2022지표
장애인426253 / 노인2462610 / 다문화388237
총합 **3277100**

[2022년 정보취약계층 이용자수 합계] = 2021지표
장애인473876 / 노인2247728 / 다문화409410
총합**3131014**

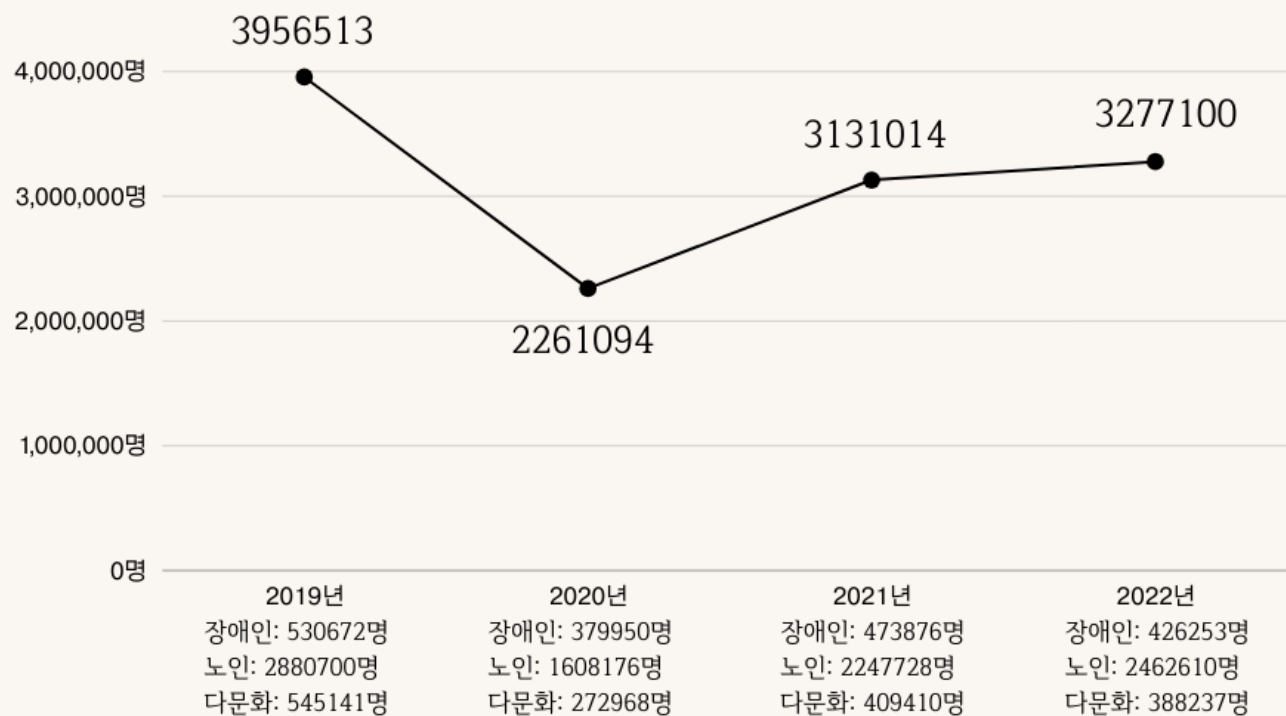
[2021년 정보취약계층 이용자수 합계] = 2020지표
장애인379950 / 노인1608176 / 다문화272968
총합**2261094**

[2020년 정보취약계층 이용자수 합계] = 2019 지표
장애인530672 / 노인2880700 / 다문화545141
총합**3956513**

3) 데이터 시각화

*테이블명은 2020-2023이지만, 실제 2019-2022의 데이터를 담고 있다.

연도에 따른 정보취약계층(장애인, 노인, 다문화)
이용자수 변화 추이



가로축 : 연도(년)
세로축 : 이용자수(명)

4) 결과 분석

[분석]

- 2019년에 가장 정보취약계층 이용자수가 많았고, 2020년 급감했다가 2021년부터 다시 상승하는 추세이지만 2019년만큼의 이용자수 확보는 이루어지지 못하고 있다.
- ↳ 2020년 코로나 팬데믹이 시작된 것이 이용자수 급감의 원인이라고 추정해볼 수 있다.
- ↳ 특히, 노인 인구가 급격하게 감소한 것이 근거가 된다. (특히 심리적/실질적으로 질병에 취약한 계층)

*

[코로나19 사회적경제 여파] ④ '파산위기, 감염위험'...위기의 돌봄서비스

(<https://www.eroun.net/news/articleView.html?idxno=11051>)

코로나로 노인요양서비스 이용 줄어...노인요양보험 2년 연속 흑자

(<http://www.bokjinews.com/news/articleView.html?idxno=92247>)

-> 위 두개의 기사 내용에서 알 수 있듯이, 코로나 팬데믹은 정보취약계층, 그중에서도 노인 계층의 각종 서비스 이용률을 확연히 낮추었다.

연도별 정보취약계층(장애인,노인,다문화) 봉사대상자수 대비 이용자수 변화 추이

- 1) 데이터 추출
- 2) 데이터 분석
- 3) 데이터 시각화 – 꺾은선 그래프
- 4) 결과 분석

1) 데이터 추출

```
[69]: #PART 2 - 정보취약계층 봉사대상자수 대비 이용자수 구하기

# [2023]
# 각 계층별 봉사대상자수와 이용자수의 합을 구함
봉사대상자수_장애인 = df['장애인_봉사대상자수_2023'].sum()
봉사대상자수_노인 = df['노인_봉사대상자수_2023'].sum()
봉사대상자수_다문화 = df['다문화_봉사대상자수_2023'].sum()

이용자수_장애인 = df['지식정보취약계층_서비스_이용자수_장애인_2023'].sum()
이용자수_노인 = df['지식정보취약계층_서비스_이용자수_노인_2023'].sum()
이용자수_다문화 = df['지식정보취약계층_서비스_이용자수_다문화_2023'].sum()

# 3개 계층의 봉사대상자수와 이용자수의 총합 계산
봉사대상자수_총합 = 봉사대상자수_장애인 + 봉사대상자수_노인 + 봉사대상자수_다문화
이용자수_총합 = 이용자수_장애인 + 이용자수_노인 + 이용자수_다문화

# 비율 계산
비율 = 이용자수_총합 / 봉사대상자수_총합

# 결과 출력
print(f"봉사대상자수 총합: {봉사대상자수_총합}")
print(f"이용자수 총합: {이용자수_총합}")
print(f"비율: {비율:.2%}")
# 소수점 이하 2자리까지 출력하고 그 값을 백분율로 표시

봉사대상자수 총합: 95375388
이용자수 총합: 3277100
비율: 3.44%
```

- 1차적으로 추출한 데이터를 기반으로 필요한 칼럼들을 이용하여 원하는 결과를 도출하였다.

- 연도별로 세 계층의 봉사대상자수 합계 구하기 -> 연도별로 세 계층의 이용자수를 합계 구하기 -> 연도별로 정보취약계층 이용자수를 정보취약계층 봉사대상자수로 나누어 비율 계산

- 왼쪽의 사진은 2023 정보취약계층(장애인,노인,다문화) 봉사대상자수 대비 이용자수 합계를 구하는 과정.

```
[70]: # [2022]
# 각 계층별 봉사대상자수와 이용자수의 합을 구함
봉사대상자수_장애인 = df['장애인_봉사대상자수_2022'].sum()
봉사대상자수_노인 = df['노인_봉사대상자수_2022'].sum()
봉사대상자수_다문화 = df['다문화_봉사대상자수_2022'].sum()

이용자수_장애인 = df['지식정보취약계층_서비스_이용자수_장애인_2022'].sum()
이용자수_노인 = df['지식정보취약계층_서비스_이용자수_노인_2022'].sum()
이용자수_다문화 = df['지식정보취약계층_서비스_이용자수_다문화_2022'].sum()

# 3개 계층의 봉사대상자수와 이용자수의 총합 계산
봉사대상자수_총합 = 봉사대상자수_장애인 + 봉사대상자수_노인 + 봉사대상자수_다문화
이용자수_총합 = 이용자수_장애인 + 이용자수_노인 + 이용자수_다문화

# 비율 계산
비율 = 이용자수_총합 / 봉사대상자수_총합

# 결과 출력
print(f"봉사대상자수 총합: {봉사대상자수_총합}")
print(f"이용자수 총합: {이용자수_총합}")
print(f"비율: {비율:.2%}")
# 소수점 이하 2자리까지 출력하고 그 값을 백분율로 표시

봉사대상자수 총합: 88086789.0
이용자수 총합: 3131014
비율: 3.55%
```

```
[71]: # [2021]
# 각 계층별 봉사대상자수와 이용자수의 합을 구함
봉사대상자수_장애인 = df['장애인_봉사대상자수'].sum()
봉사대상자수_노인 = df['노인_봉사대상자수'].sum()
봉사대상자수_다문화 = df['다문화_봉사대상자수'].sum()

이용자수_장애인 = df['지식정보취약계층_서비스_이용자수_장애인'].sum()
이용자수_노인 = df['지식정보취약계층_서비스_이용자수_노인'].sum()
이용자수_다문화 = df['지식정보취약계층_서비스_이용자수_다문화'].sum()

# 3개 계층의 봉사대상자수와 이용자수의 총합 계산
봉사대상자수_총합 = 봉사대상자수_장애인 + 봉사대상자수_노인 + 봉사대상자수_다문화
이용자수_총합 = 이용자수_장애인 + 이용자수_노인 + 이용자수_다문화

# 비율 계산
비율 = 이용자수_총합 / 봉사대상자수_총합

# 결과 출력
print(f"봉사대상자수 총합: {봉사대상자수_총합}")
print(f"이용자수 총합: {이용자수_총합}")
print(f"비율: {비율:.2%}")
# 소수점 이하 2자리까지 출력하고 그 값을 백분율로 표시

봉사대상자수 총합: 83908659
이용자수 총합: 2261094
비율: 2.69%
```

- 위 사진은 순서대로 2022-2021 정보취약계층(장애인,노인,다문화) 봉사대상자수 대비 이용자수 합계를 구하는 과정.

[73]:

```
# [2020]
# 각 계층별 봉사대상자수와 이용자수의 합을 구함
봉사대상자수_장애인 = df['장애인_봉사대상자수_2020'].sum()
봉사대상자수_노인 = df['노인_봉사대상자수_2020'].sum()
봉사대상자수_다문화 = df['다문화_봉사대상자수_2020'].sum()

이용자수_장애인 = df['지식정보취약계층_서비스_이용자수_장애인_2020'].sum()
이용자수_노인 = df['지식정보취약계층_서비스_이용자수_노인_2020'].sum()
이용자수_다문화 = df['지식정보취약계층_서비스_이용자수_다문화_2020'].sum()

# 3개 계층의 봉사대상자수와 이용자수의 총합 계산
봉사대상자수_총합 = 봉사대상자수_장애인 + 봉사대상자수_노인 + 봉사대상자수_다문화
이용자수_총합 = 이용자수_장애인 + 이용자수_노인 + 이용자수_다문화

# 비율 계산
비율 = 이용자수_총합 / 봉사대상자수_총합

# 비율 계산(2020 봉사대상자수 데이터가 없기 때문에 총합이 0이 나올경우 1로 취급.)
# 봉사대상자수가 없는데, 이용자가 있다는 점을 봉사대상자수가 전부 이용자라는 걸로 해석.
# 봉사대상자수 데이터 찾지 못하면 아예 비교하지 않는 걸 고려.
if 봉사대상자수_총합 == 0:
    비율 = 1.0 # 봉사대상자수 총합이 0인 경우 100%로 처리
else:
    비율 = 이용자수_총합 / 봉사대상자수_총합

# 결과 출력
print(f"봉사대상자수 총합: {봉사대상자수_총합}")
print(f"이용자수 총합: {이용자수_총합}")
print(f"비율: {비율:.2%}")
# 소수점 이하 2자리까지 출력하고 그 값을 백분율로 표시
```

```
봉사대상자수 총합: 0.0
이용자수 총합: 3956513
비율: 100.00%
```

```
C:\Users\erika\AppData\Local\Temp\ipykernel_40536\2277332585.py:16: RuntimeWarning: divide by zero encountered in scalar divide
    비율 = 이용자수_총합 / 봉사대상자수_총합
```

[]:

: 왼쪽 사진은 2020 정보취약계층(장애인,노인,다문화) 봉사대상자수 대비 이용자수 합계를 구하는 과정.

2) 데이터 분석

[특이 사항]

- (1) 2022 다문화 봉사대상자수 칼럼의 데이터가 결손되어 있음. (통합데이터 뿐 아니라 원데이터 내에서도 비어있음.)
- (2) 2020의 경우 세 계층의 봉사대상자수 칼럼의 데이터가 모두 결손되어 있음.. (통합데이터 뿐 아니라 원데이터 내에서도 비어있음.)
 - 이에 따라, 2020 지표를 계산할 때, 봉사대상자수가 없는데, 이용자수가 있다는 점을 봉사대상자수가 전부 이용자라는 걸로 해석하여 봉사대상자수의 총합을 1로 취급하여 계산하는 등의 시도를 하였음.
 - 그러나, 정확한 추이가 나오지 않을 것이라는 점을 우려하여 2020, 2022칼럼은 제외하고 시각화하기로 결정함.

2) 데이터 분석

[2023] = 2022

봉사대상자수 총합: 95375388이용자수 총합: 3277100비율: 3.44%

[2022] = 2021

봉사대상자수 총합: 88086789.0이용자수 총합: 3131014비율: 3.55%

[2021] = 2020

봉사대상자수 총합: 83908659이용자수 총합: 2261094비율: 2.69%

[2020] = 2019

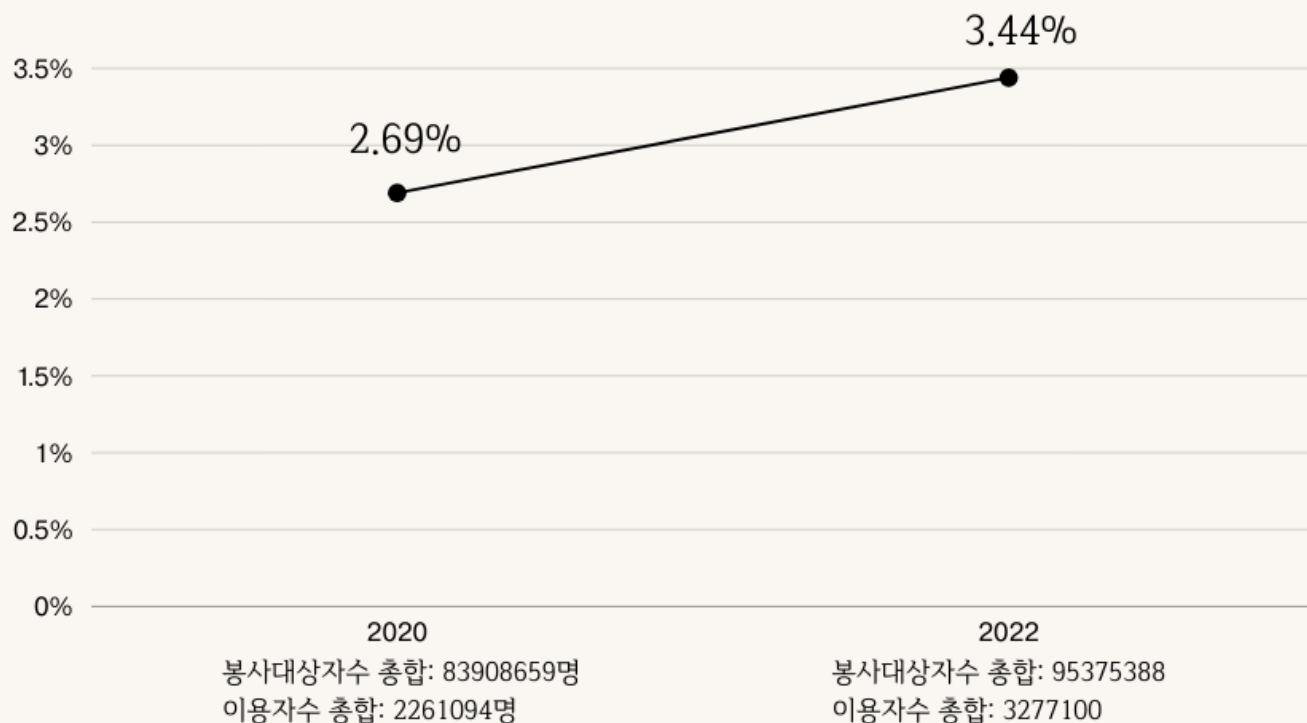
봉사대상자수 총합: 0.0 이용자수 총합: 3956513비율: 100.00

→ 이 중, 2023/2021 결괏값만 이용하여 시각화

3) 데이터 시각화

*테이블명은 2020-2023이지만, 실제 2019-2022의 데이터를 담고 있다.

연도에 따른 정보취약계층(장애인, 노인, 다문화)
봉사대상자수 대비 이용자수 변화 추이



가로축 : 연도(년)

세로축 : 정보취약계층 봉사대상자수 대비 이용자수(%)

4) 결과 분석

[분석]

- 2019년, 2021년 정보취약계층 봉사대상자수 원데이터에 결손이 있어 2개년만 비교.
- 2020년 코로나 펜데믹이 시작되었을 때와 비교하였을 때, 봉사대상자수 대비 이용자수는 증가하였음을 알 수 있다.
 - ↳ 2022년의 봉사대상자수가 더 많다는 것을 감안하면, 위 비율이 오른 것은 긍정적인 지표이다.

[분석 의의]

- : 도서관이 정보취약계층의 정보접근성에 얼마나 도움이 되고 있는지를 알아볼 수 있는 지표
- 봉사대상자수 대비 이용자수가 증가한 것은 정보접근성이 좋아진 정보취약계층의 비율 및 인원이 늘어났다는 것임으로, 정보접근성이 좋아졌음을 유추해 볼 수 있다.
 - BUT, 비율 자체만을 놓고 봤을 때 2.69%, 3.44%는 턱없이 낮은 수치라고 분석되며 앞으로 도서관이 좀 더 정보취약계층 이용자의 이용률을 늘려야 함을 보여준다.