



DATA_122

여가활동 추천을 통한 국민건강증진

스포츠여가활동 독려를 통한
스포츠산업 활성화

INDEX

01

도입

- 기획배경 및 주제선정
- 데이터 소개

02

EDA

- 여가활동 분석
- 체육활동 분석
- 스포츠산업 분석

03

정책제언

- 모델 계획

04

추천시스템

- 데이터 수집
- 데이터 전처리
- 이웃기반 협업필터링
- 추천시스템 모델링

05

정책효과 분석

- 스포츠 추천서비스 '스포츠콕'
- 이외의 기대효과

기획배경 및 주제선정

뉴스홈 | 최신기사

'체력측정' 고양시민 64%가 약골..."코로나여파"

송고시간 | 2021-12-10 08:55

코로나 '저체력' 학생 늘었다...체력평가 4·5등급
12→18% 급증(종합)

입력 2022.03.03 15:13 수정 2022.03.03 15:13

가카

뉴스 | 경제

코로나에 여가생활 절반 '뚝'...국민 5명 중 2명은 비만

입력 2022-03-15 12:07 | 업데이트 2022-03-15 12:08

코로나 거리두기 여파로 인한 체력저하
학생, 청년층의 체력적 문제가 심각하다

기획배경 및 주제선정

거리두기로 인한
프로야구, 프로스포츠 관객 감소



스포츠에 대한 관심 감소



스포츠관련 상품 소비 저하
스포츠 산업 축소

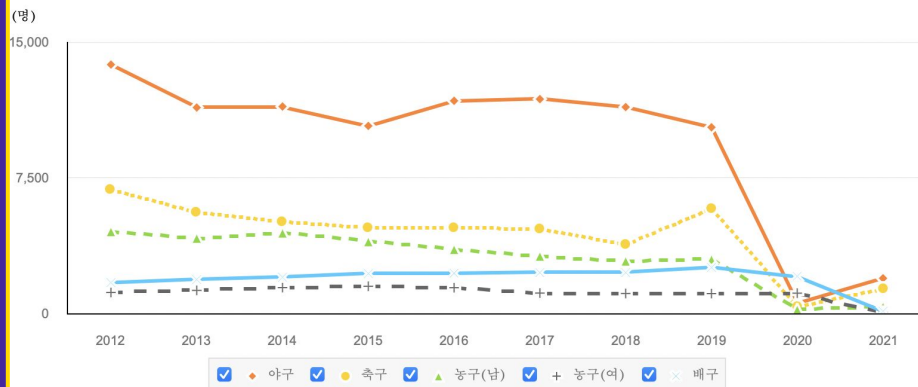
스포츠 산업의 재활성화가 필요

뉴스홈 | 최신기사

코로나 여파 무섭네...한일 프로야구 관중 24~26% 급감

송고시간 | 2022-06-23 14:40

주요 프로스포츠 경기당 평균 관중 추이



기획배경 및 주제선정



국민 건강 증진

국민 체력저하의 문제점과 원인파악
여가활동 추천을 통한 국민체력증진



스포츠산업 정상화

스포츠 여가활동의 현상황 파악
스포츠활동 장려를 통한 스포츠산업 복구 가속화

데이터 소개

- 국민여가활동조사

인구데이터, 여가생활에 대한 평가

여가활동의 목적, 경험한 여가활동의 종류

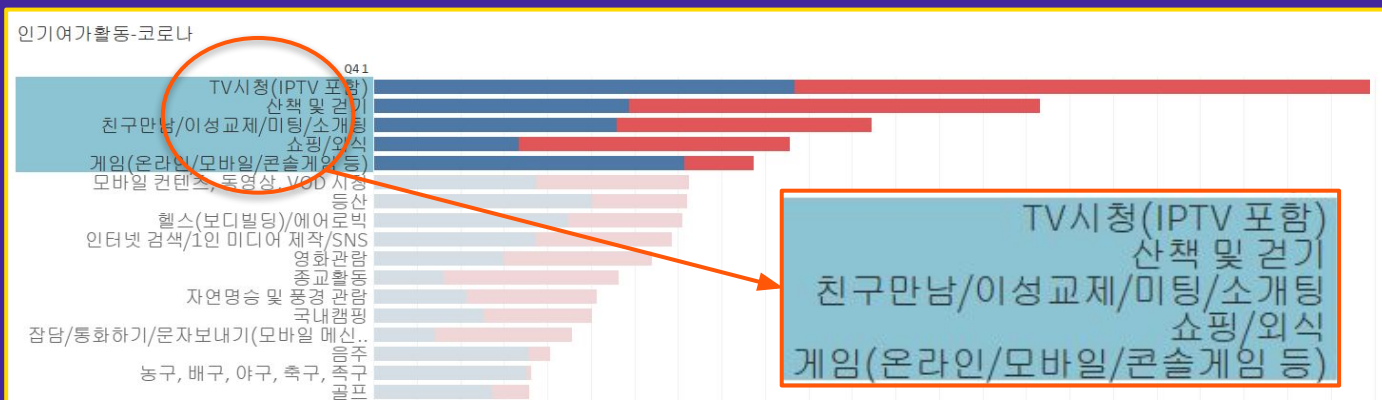
- 국민생활체육조사,스포츠산업조사

체력에 대한 인식, 체력유지방법

스포츠산업 매출액



EDA - 여가활동



코로나 거리두기 이후
스포츠 여가활동이 순위에서 많이 없어짐을 확인

EDA - 여가활동

```
1 #거리두기 이전 스포츠여가활동  
2 Countsport(df).sum()
```

```
count    9822  
dtype: int64
```



```
1 #거리두기 이후 스포츠여가활동  
2 Countsport(df).sum()
```

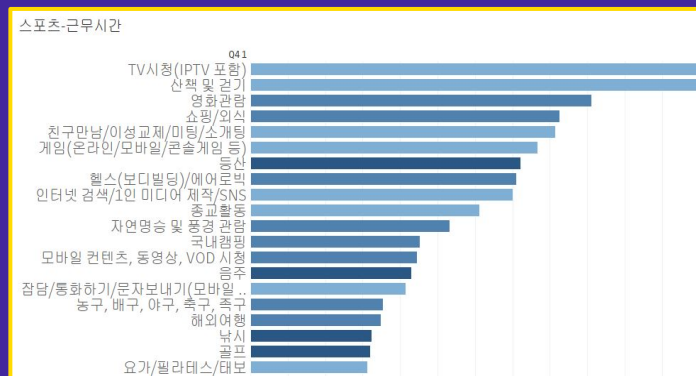
```
count    8909  
dtype: int64
```

거리두기 이후 스포츠 여가활동의 비율이
약 10%의 감소를 보임

EDA - 여가활동



결혼여부에 따른 가장 만족하는 여가활동 합계

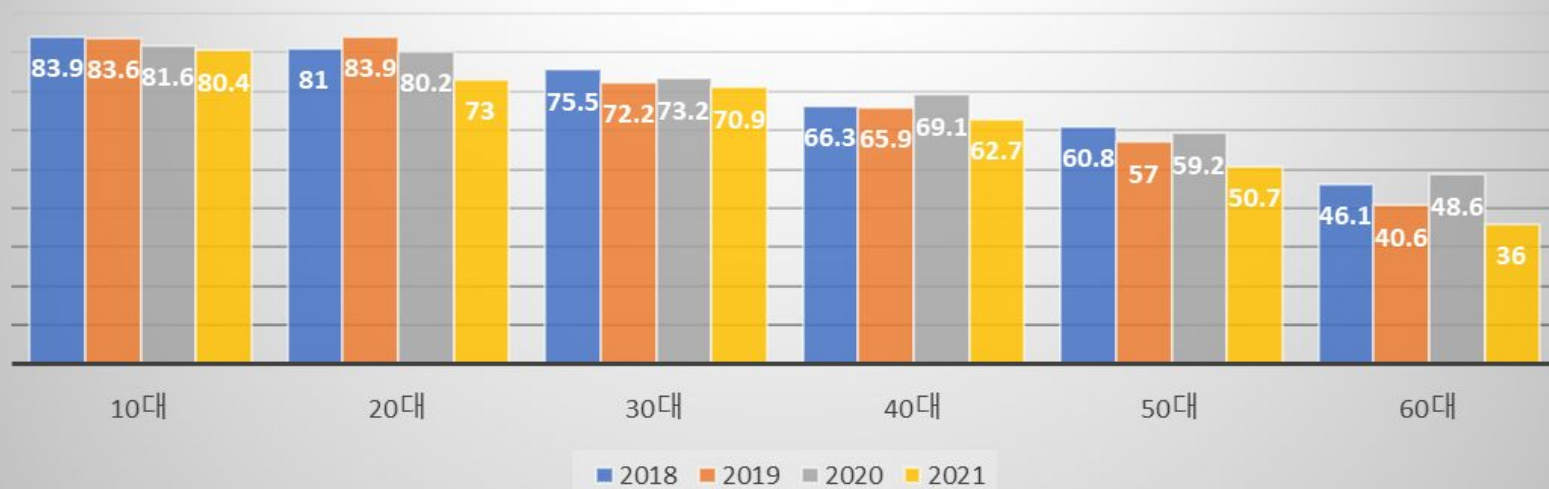


색이 진할수록 근무시간이 많은 설문자의 여가활동

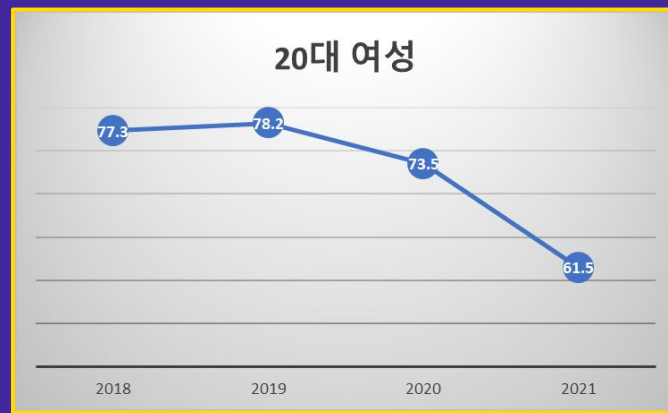
결혼여부, 근무시간, 성별, 나이가
여가활동 선택에 큰 영향을 주었다

EDA - 체력

체력에 대한 인식 - 긍정합계



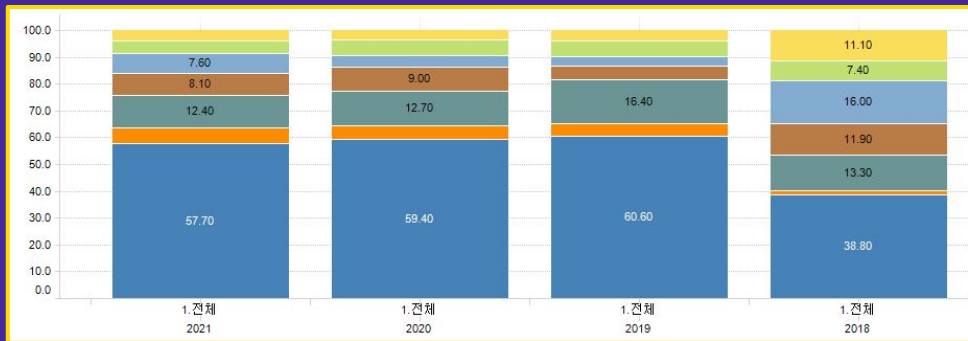
EDA - 체력



전반적인 하향 곡선

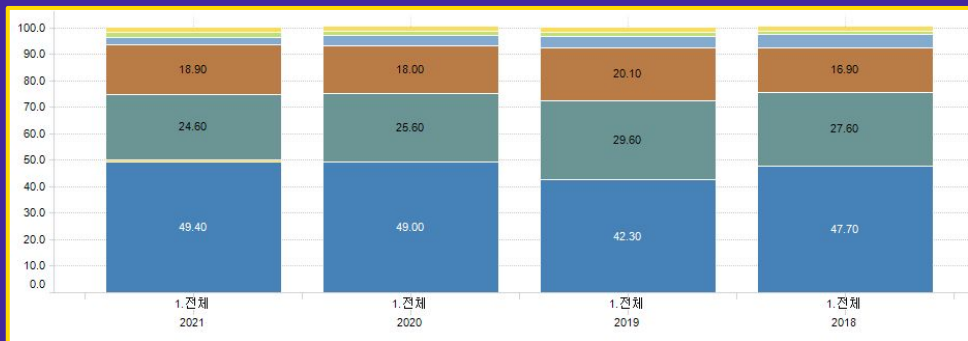
특히 10대 남성, 20대 여성의 하락폭이 뚜렷하다.

EDA - 체력



체육시설을 이용하지 않는 이유

- 1. 시간적 여유가 없어서
- 2. 체육시설에 대한 정보가 없어서
- 3. 거리상 멀어서



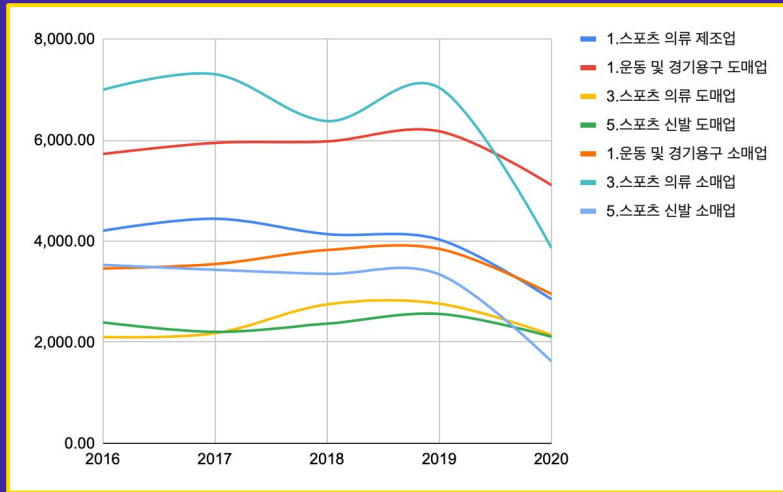
체육시설을 자주 이용하는 이유

- 1. 거리상 가까워서
- 2. 시설 이용료가 무료 또는 저렴해서
- 3. 전문적인 체육 시설을 갖추고 있...

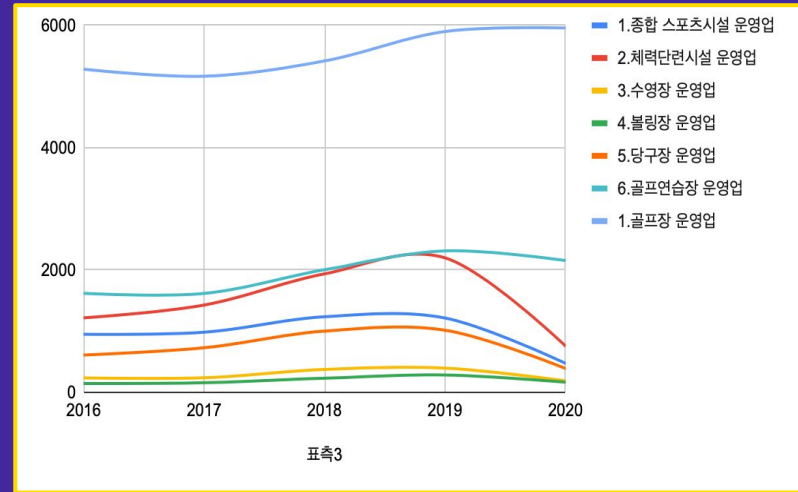
➡ 근처의 체육시설 정보를 전달하는 것이 중요

EDA - 스포츠산업

스포츠 용품



스포츠 시설



스포츠 용품, 시설 산업 모두 코로나 거리두기의 영향을 받아
매출이 떨어졌음을 알 수 있다.

EDA - 정리 및 결론

거리두기 이후 스포츠 여가활동
약 10%의 감소

체력인식 전반적인 하향 곡선

스포츠 산업,시설 매출 저하

거리두기에 익숙해진 국민들의 생활패턴이 지속된다면 악순환이 계속될 수 있다
스포츠에 대한 관심유도, 스포츠산업 활성화가 필요하다.

정책제언

스포츠에 대한 관심 유도가 필요

스포츠는 직접해보고 재미를 느꼈을 때 스포츠 산업의 활성화를 이끌 수 있다.

데이터를 이용하여 어떻게 스포츠를 유도할 수 있을까?

스포츠 추천시스템을 활용한다.



모델 계획



인구데이터(나이,성별,결혼여부,업무시간)



취향 데이터(여가활동 목적, 경험한 여가활동)



평점(여가활동에 매긴 평가)

세가지 데이터를 적절히 활용, 추천시스템 설계

추천시스템

데이터 수집

국민여가활동조사 데이터에서

추천시스템에 필요한 데이터를 추출

인구데이터

```
1 df_per = data[['DM1', 'DM2', 'DM11', 'Q45', 'Q47_N', 'DM5']]
2 df_per.head()
```

	DM1	DM2	DM11	Q45	Q47_N	DM5
20636	남성	20대	서울	하지 않았		
20637	여성	50대	서울	하였		
20638	남성	40대	서울	하였		
20639	여성	30대	서울	하지 않았		
20640	여성	60대	서울	하지 않았		

여가활동의 목적

```
1 df_etc = data[['Q3']]
2 df_etc.head()
```

	Q3
20636	개인의 즐거움을 위해
20637	마음의 안정과 휴식을 위해
20638	마음의 안정과 휴식을 위해
20639	자기만족을 위해
20640	건강을 위해

추천시스템

데이터 전처리

여가활동의 목적, 여가활동 분류 정리

카디널리티를 줄여 데이터를 응축

‘어렸을 때 해본 스포츠를 다시 하게 될 가능성이 높다’

경험의 유무가 스포츠를 시작하는데 영향이 크다고 판단, 추천시스템에 반영하였다.

여가활동 분류정리

```
1 #여가활동 컬럼이름 csv
2 q1c = pd.read_csv(path+'01_colname.csv')
3 #컬럼이름 리스트로 변경
4 temp = q1c[q1c['대분류']=='취미생활']
```

여가활동의 목적

```
[ ] 1 #카디널리티 감소를 통한 성능향상을 목적으로 4가지 카테고리로 변경
2 df['03'] = df['03'].replace(['개인의 즐거움을 위해', '스트레스 해소를 위해'], '즐거움')
3 df['03'] = df['03'].replace(['마음의 안정과 휴식을 위해', '시간을 보내기 위해'], '휴식')
4 df['03'] = df['03'].replace(['자기만족을 위해', '건강을 위해', '자기 계발을 위해'], '자기계발')
5 df['03'] = df['03'].replace(['가족과 시간을 함께하기 위해', '대인 관계·교제를 위해'], '대인관계')
```

Q1 경험한 여가활동

```
[ ] 1 col_list = data.columns
2 df_exp = data[col_list[col_list.str.contains('01')==True]]
3 df_exp.replace('미참여', 0, inplace=True)
4 df_exp.replace('참여', 1, inplace=True)
5 df_exp.head()
```

	01_1	01_2	01_3	01_4	01_5	01_6	01_7	01_8	01_9	01_10	...	01_82	01_83	01_84	01_85	01_86	01_87	01_88	01_89	01_90	01_91
20636	0	0	0	0	0	0	0	1	0	0	0	...	0	0	1	0	0	0	0	1	1
20637	1	0	0	0	1	0	1	0	0	0	...	0	0	1	1	0	0	0	0	0	1
20638	0	0	0	0	1	0	1	0	0	0	...	0	0	1	1	0	0	0	0	0	1
20639	0	0	0	0	0	0	1	1	0	0	...	0	1	0	1	0	0	0	0	0	1
20640	0	0	0	0	0	0	1	0	0	0	...	0	1	0	1	0	0	0	0	0	1

5 rows x 91 columns

추천시스템

이웃기반 협업필터링

‘비슷한 군집의 사람들이 이 여가활동을 어떻게 생각하는가’

주어진 평점을 기반으로
유저기반 협업필터링을 사용하여
평점을 예측

[] 1 df_rating

	ID	leisureID	rating
30646	202115021	취미생활중	7
30647	202115022	취미생활중	7
30648	202115023	취미생활여	7
30649	202115024	취미생활중	7
30650	202115025	취미생활중	7

	ID	202115021	202115022	202115023	202115024	202115025	202115026	202115027	202115028	구 등)	6
경투 스포츠 경기관람 (태권도, 유도, 합기도, 검도, 권투 등)	5.897895	5.898333	5.897479	5.897729	5.899326	5.899054	5.898598	5.898193		구 등)	6
경투 스포츠(태권도, 유 도, 합기도, 검도, 권투 등)	6.360615	6.363621	6.359113	6.361526	6.361058	6.357951	6.357763	6.362659		구 등)	6
골프	6.361754	6.358579	6.361881	6.35791	6.36181	6.360801	6.36177	6.358657		구 등)	6
기타	5.876157	5.874722	5.873187	5.887549	5.874864	5.872672	5.875619	5.874799			
농구, 배구, 야구, 축구, 족구	6.314411	6.314798	6.31418	6.314178	6.316611	6.316232	6.315441	6.314396			
당구/포켓볼	6.077022	6.077168	6.077148	6.076495	6.078033	6.076724	6.078323	6.077333			
댄스스포츠(탱고, 왈츠, 자이브, 맘보, 볼카, 자자 자동)	6.064592	6.064812	6.063702	6.063339	6.065725	6.063775	6.064548	6.064765			
등산	6.811081	6.811175	6.810978	6.810412	6.811128	6.811089	6.811231	6.811244			
베드민턴/줄넘기/맨손, 스트레칭 체조/올라후프	5.828386	5.829291	5.828939	5.828633	5.830665	5.830277	5.831642	5.829557			
볼링, 탁구	6.153915	6.154616	6.158109	6.154267	6.155221	6.153065	6.155216	6.154752			

추천시스템

데이터 모델링

‘군집을 구성했을 때 전체 평균보다 높은 평점을 받은
여가활동이 인기가 있을 것이다’

나이,성별 같은 인구데이터와 경험한 스포츠활동을
K-means로 군집화하여 cluster를 형성

협업필터링으로 예측된 평점을 기반으로
여가활동을 추천

```
#사용자 정보
gender = 'M'
age = 20
marriage = '미혼'
worktime = 30
purpose = '즐거움'
clu = 2
```

```
rec_list.head(10)
```

취미생활남	0.003280
취미생활중	0.000961
아이스스케이트, 아이스하키 등	0.000755
농구, 배구, 야구, 축구, 족구	0.000726
예술관람	0.000596
인라인스케이트	0.000474
온라인게임 경기관람(e-스포츠 경기 포함)	0.000379
헬스(보디빌딩)/에어로빅	0.000306
볼링, 탁구	0.000293
승마, 암벽등반, 철인삼종경기, 서바이벌	0.000192

dtype: float64

추천시스템

결론 및 한계점

추천시스템 자체의 성능은 준수하나

데이터가 적은 여가활동에 대해서는
값이 튀는 경향이 있음

이를 개선하기 위해 설문조사를 수정할 필요가 있다.

한 번 이상 참여해 본 여가활동
가장 만족하는 여가활동 1순위, 2순위...  모든 스포츠 여가활동에
대한 평점

순위 데이터 보다는 평점 데이터를 늘려
성능을 향상시킨다.

actual_rating	leisureID	pred_rating
0	여행	6.794414
1	사교활동	6.363155
2	수영	6.16656
3	예술관람	6.459399
4	스포츠 경기 직접관람- 경기장방문관람(축구, 야구, 농구, 배구 등)	6.012947

RMSE: 0.3144220922352764

여가활동 만족도 -배드민턴	문1	Q1_1	1	매우 불만족 한다
			2	
			3	
			4	보통이다
			5	
			6	
			7	매우 만족 한다
			9	해본적 없음
여가활동 만족도 -볼링	문1	Q1_2	1	매우 불만족 한다
			2	
			3	
			4	보통이다
			5	
			6	
			7	매우 만족 한다
			9	해본적 없음

정책효과 분석

스포츠 여가활동 추천서비스 '스포츠콕'

추천시스템을 통해 여가활동을 추천하고
주변시설, 프로그램정보 데이터를 제공하는

추천서비스 어플리케이션 '스포츠콕'

강남구민을 위한 개인 맞춤형 여가 추천 서비스

나에게 알맞은
여가 활동?
콕 짚어드려요!

심리학적 분석을 통한 맞춤 여가활동

- ✓ 나의 여가성향을 알려주고
최적의 여가활동 5개를 추천해 드립니다.
- ✓ 여가활동에 맞는 여가시설,
관련 영상과 제휴쿠폰을 제공해 드립니다.

★ 이용방법 : 링크 (gn.my-leisure.co.kr) 접속 또는 QR코드 스캔



정책효과 분석

스포츠 여가활동 추천서비스 '스포츠콕'

'경기도 의왕시 20대 남성, 배드민턴 추천됨'



인근 배드민턴 가능 체육시설 데이터 확인



체육시설의 정보, 프로그램 제공

경기도 의왕시 공공체육시설 및 동네체육시설 현황_20220513					
파일 수정 보기 삽입 서식 데이터 도구 확장 프로그램 도움말 모조전에 마지막					
N22 fx 100% W % .0 .00 123 기본값 (Alt... 10 B I U					
	A	B	C	D	E
1	시군명	구분	시설명	위치	주요시설
2	의왕시	체육공원	고전체육공원	오전로 18(왕곡동)	축구장, 농구장,...
3	의왕시	체육공원	부곡체육공원	부곡공원길6(월암동)	축구장(인조잔디
4	의왕시	체육공원	내산체육공원	내산공원길3(내산동)	축구장, 농구장,
5	의왕시	체육센터	의왕국민체육센터	복지로27(내산동)	수영장, 헬스장,
6	의왕시	체육센터	부곡스포츠허브	부곡시장길 77	수영장, 헬스장,
7	의왕시	체육센터	백운커뮤니티센터	포일동 487-32	수영장, 볼링장,
8	의왕시	체육센터	신곡동민체육시설	포일동 654	수영장, 볼링장,
9	의왕시	동네체육시설	내산약수 체육시설	내산동 529	축구장, 농구장,
10	의왕시	동네체육시설	장계동 체육시설	포일동 127-3	배드민턴장 8면,
11	의왕시	동네체육시설	금천마을 체육시설	이동 203-4	게이트볼장 2면(
12	의왕시	동네체육시설	백운(이동)배드민턴장	장계골길 4-12	배드민턴장 5면
13	의왕시	동네체육시설	포일배드민턴장	내산동 산48-10	배드민턴장 5면
14	의왕시	동네체육시설	하이릭 축구장	학의동 252-8	축구장 3면
15	의왕시	동네체육시설			

백운배드민턴클럽

경로 저장 대화

4.7 ★★★★★ Google 리뷰 23개
의왕시의 배드민턴장

주소: 경기도 의왕시 이동 434-4

영업시간: 영업 종료 · 수 오전 6:00에 영업 시작 ·

연락처: 031-457-5696

지역: 경기도

수정 제안하기 · 이 비즈니스의 소유주인가요?

누락된 정보 추가

클리어 추가

질문과 답변

모든 질문 보기(1개)

인기 시간대

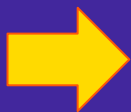
오전 11시: 일반적으로 준비는 정도가 보통인 시간대임

AM 6시 AM 8시 PM 12시 PM 3시 PM 6시 PM 8시

정책효과 분석

추가적인 기대효과

- 예측결과를 스포츠 이벤트 개최의 근거로 활용가능
- 지역 체육센터의 이용패턴을 예측하여 신규 프로그램 창설이나 시설 설치에 사용가능



스포츠에 대한 관심 증가

국민 건강증진과 스포츠산업 정상화

'AGAIN 2002'...6만4872명 붉은 함성으로 뜨거웠던 브라질전

등록 2022.06.02 22:10:00



기사내용 요약

EPL 득점왕 손흥민 vs 스타군단 브라질 맞대결에 상암별 들쭉

토트넘 내한경기에 쿠팡 '와우 멤버' 300만명 열광...유료 구독 영향력 'UP'

스 남도영 기자 | © 승인 2022.07.17 18:25



/사진=주영재출



과 브라질 축구 국가대표팀 친선
Photo1006@newsis.com

감사합니다

DATA_122

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik**

