K-LEAGUE 예매 분석

- 인천유나이티드 -

데이터분석가 이진우

县大

- 1. 프로젝트 개요
- 2. 프로젝트 구성
- 3. 데이터 전처리

4. 데이터 분석

5. 프로모션 제안 및 회고

프로젝트개요



K LEAGUE

K리그는 한국프로축구연맹 산하의 대한민국 프로축구리그입니다. 1983년 슈퍼리그라는 명칭으로 출범하였으며, 현재까지 AFC 챔피언스 리그 우승 횟수가 가장 많은 리그이기도 합니다.

--- 코로나 19

코로나19로 인해 피해를 입었던 관광·스포츠·공연업계의 제한이 해제되면서 관객이 조금씩 늘어나는 추세입니다. 2023년에는 더 많은 관객이 방문 할 것으로 예상하고 있습니다.

__ 월드컵

작년 월드컵의 뜨거운 열기로 K-리그에도 관심이 많아 진 지금, 적절한 전략을 통해 관객을 늘리는 방안을 검 토해보았습니다.

프로젝트 구성

비즈니스 분석	데이터 전처리	데이터 분석	제안 및 회고
• K LEAGUE	 MEMBERSHIP.CSV RESERVE_TICKET.CSV D95E264436CF42DE.XLSX SCHEDULE202212022101	• 구매 패턴 분석	프로모션 제안회고
• 인천 유나이티드	9212245.XML	• 고급좌석 이용 고객 분석	

인천 유나이티드

2003년에 창단 된 인천광역시의 K리그1 소속 프로축구단

전술력

시즌 초반 주로 하위권에 머물지만 적 극적인 여름 이적시장 활용으로 하반 기에 좋은 성적을 거둡니다.

생존왕

부족한 재정과 잦은 선수 이탈에도 승 강제 도입 이후 한 번도 강등하지 않 은 유일한 팀으로 하반기로 갈수록 팬 들의 응원이 뜨겁습니다.



데이터전처리

membership.csv

고객번호, 나이, 직업, 좋아하는 선수

reserve_ticket.csv

주문번호, 고객번호, 주문시간, 경기시간 취소유무, 티켓번호, 예매좌석 정보

d95e264436cf42de.xlsx

일자별 인천유나이티드 순위

schedule2022120221019212245.xml

경기시간, 경기정보, 관객 수

job, zipcode가 결측치이고 좋아하는 선수가 무고사 회원가입 기본 설정으로 추정됨

-> zipcode는 0, job은 '기타'로 분류

중복컬럼 제거(detail_cancel_date_time)

예매,취소 시간컬럼 병합

상대팀 정보 추가

스카이라운지구역 결측치 발견

-> seat_Attribute_5 컬럼을 'sky'로 분류

2022년 경기 순위만 남기고 제거

한글 인코딩 문제가 발생

-> 디코딩 후 인코딩을 진행하고 다시 디코딩하여 문제 해결 승부차기 진행하지 않아 home_team_tk, away_team_tk 컬럼 삭제 게임시간 컬럼을 datetime형식으로, 요일 컬럼을 int로 변경

데이터분석



7년1

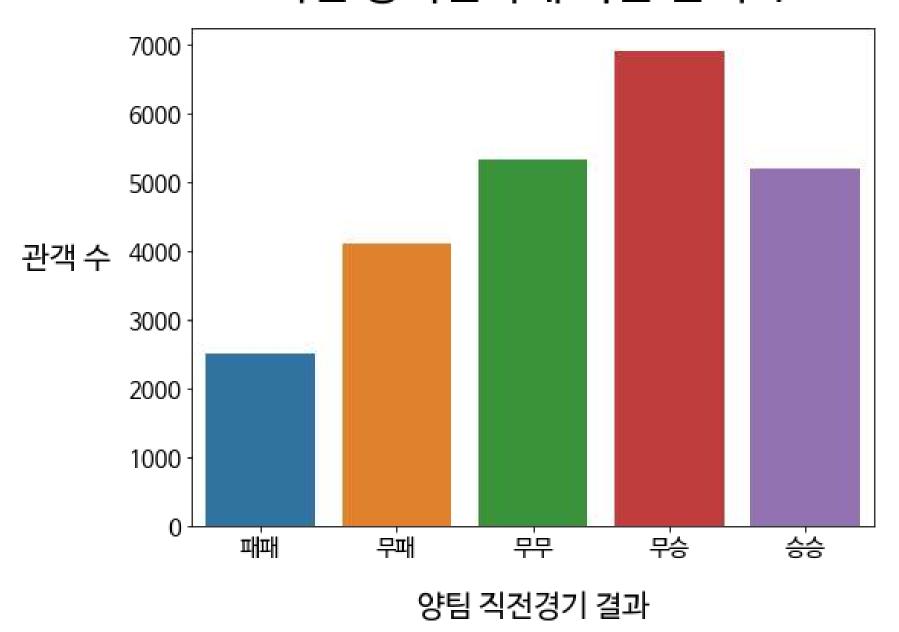
직전 경기 결과와 순위가 예매에 영향을 미친다.



직전 경기 결과

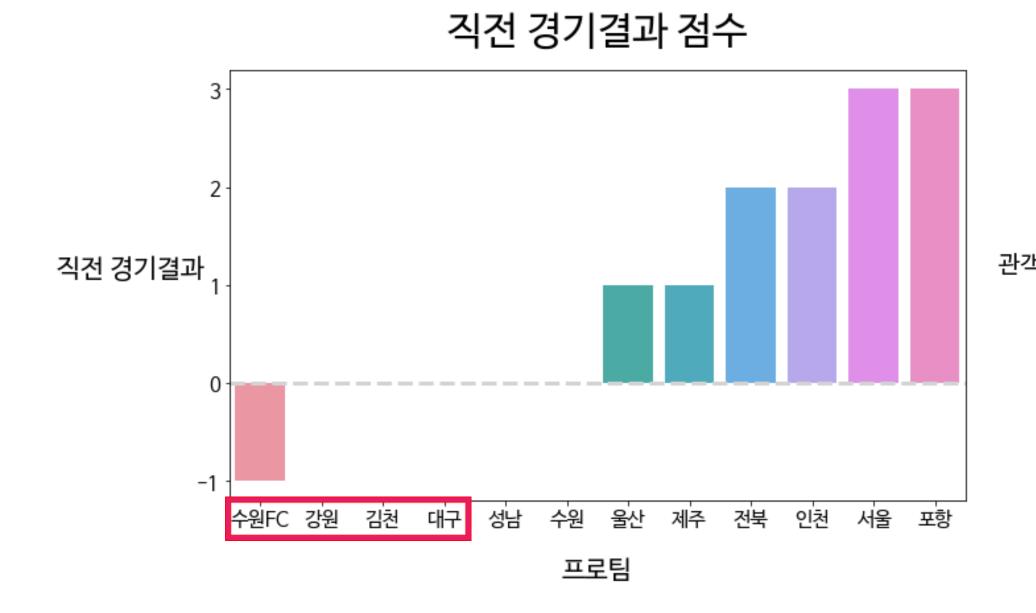
두팀 모두 직전 경기를 패배 했을 때 관중의 수가 적어지는 경향이 있다.

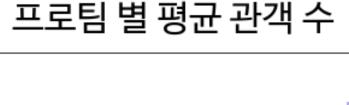
직전 경기결과에 따른 관객 수

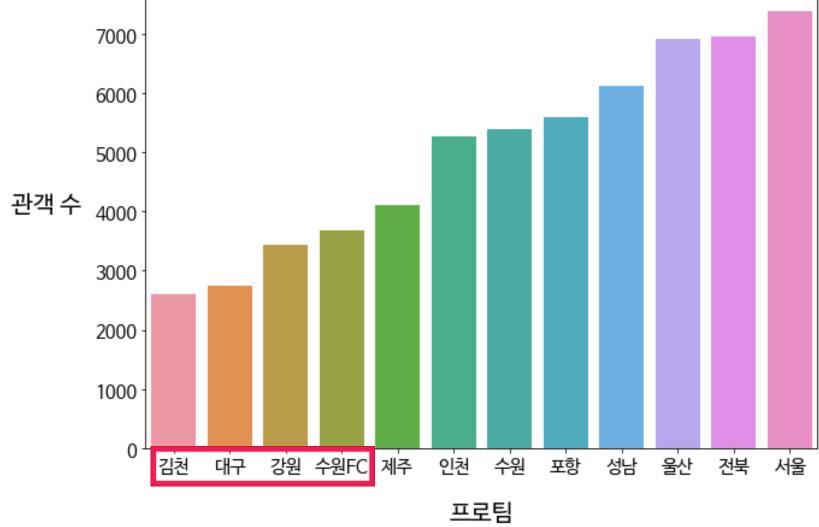


직전 경기 점수 합계

직전경기: 승리 +1 / 패배 -1 / 무승부 0 직전경기 하위권 팀의 관객 수도 하위권



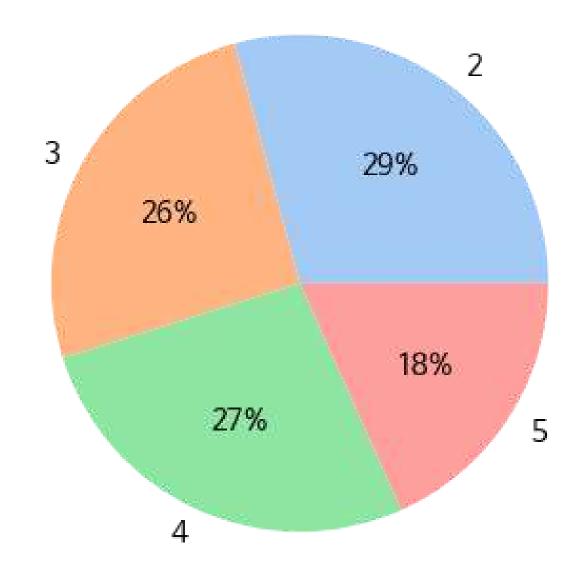




인천유나이티드 순위에 따른 관객 수

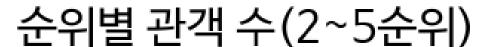
순위가 높을 때 평균 관객수가 높은 편이다.

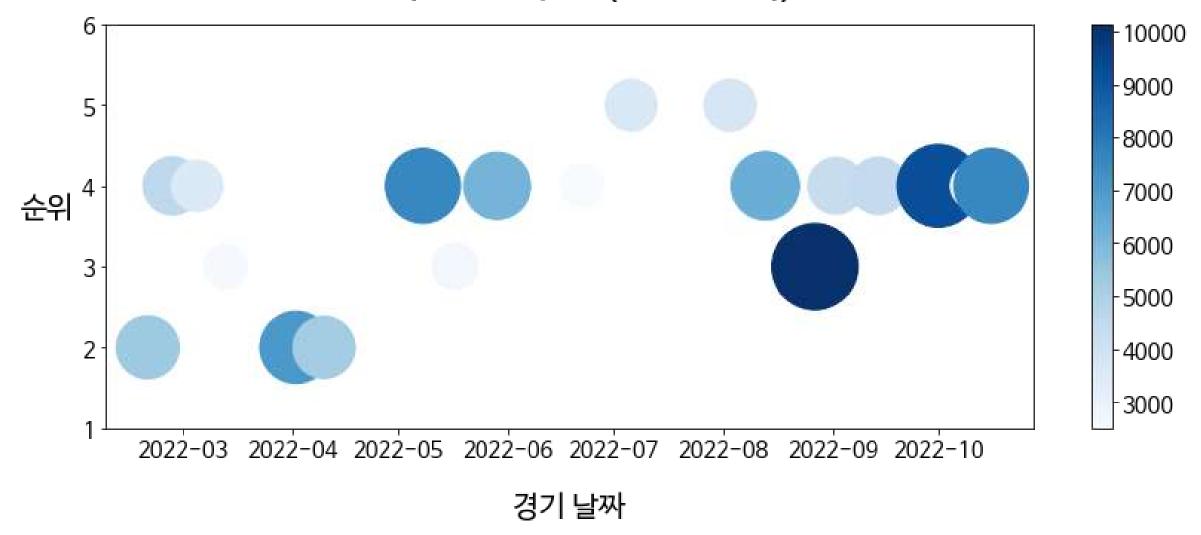
순위별 평균 관객 수 비율



인천유나이티드 순위에 따른 관객 수

구체적으로 살펴보면 높은 순위보다는 순위의 변화가 있을 때 주로 관객 수의 변화가 관찰됨 2022년의 자료만으로 순위에 따른 관객 수를 단정짓는 것은 어려워보임





가설 1 결론

직전 경기 결과와 순위가 예매에 영향을 미친다.

직전 경기 결과

- 직전 경기결과가 안좋은 팀간의 경기는 관객 수가 적을 것으로 예측할 수 있다.
- 인천이 패배한 경우 관람객 위로 및 재방문 독려를 위해 일부 고객에게 쿠폰 제공

순위

• 순위와 관객 수는 상관관계가 없는 것으로 보인다.

7 설 2

새로운 패턴을 가진 고객이 있을 것이다.



경기를 예매하는 시간은 월, 화, 수요일 14시

요일별 예매시간

- 5000

-4000

- 3000

- 2000

- 1000



주로 직업이 회사원이거나 기타인 회원

직업별 경기예매 시간

	공무원 -	12	8	8	2			1	8	22	45	55	61	41	62	396	98	88	80	51	56	75	51	50	24		
	교사 -	33	11	7	1	5	6	5	6	18	42	39	64	61	54	543	131	83	60	89	57	48	79	77	54	225	5000
	군인(소방/경찰/교정직 포함) -	1	7		1					4	24	11	19	15	7	128	43	32	12	22	20	9	22	13	15		
	금융 및 보험업 종사자 -	30		5	2				10	18	21	21	30	37	31	228	47	52	38	45	27	39	30	25	23		
	기타 -	163	42	25	13	3	16	21	47	135	253	312	355	332	457	2448	690	566	502	474	344	288	314	267	280	- A	4000
	농업 종사자(축산/수산/임업 포함) -											3	6	6	4	23	10	1	2	3			6	2			
	대학생(대학원생 포함) -	132	66	48	23	14	8	6	7	19	40	86	97	102	176	1329	328	252	203	264	178	117	137	151	106		
	무직 -	36	6	3	5		3	7	10	9	42	27	32	60	80	350	94	92	66	56	42	40	48	67	36		
7101	방송 및 언론 종사자 -	8	5	3		1		1		5	7	3	13	19	23	80	32	17	21	19	13	6	7	26	19	18-	3000
직업	법률 종사자 -	7								1	2	2	6	1	12	72	20	6	2	14	7	3	1	14	6		
	보건 및 의료종사자 -	18	3	2				3	6	12	17	40	42	39	57	383	73	48	79	76	23	50	47	41	36		
	서비스업 종사자 -	24	14	7	4	1	15	4	7	29	66	34	43	52	68	388	129	94	53	85	88	67	82	51	22	834	2000
	예술 및 디자인 종사자 -	9	3	2	1	4		1	7	1	11	20	7	8	12	79	33	19	19	27	12	19	11	18	8		2000
	자영업자 -	70	30	13	7	2		6	36	61	62	105	130	124	222	799	324	231	193	177	133	138	122	112	77		
	전업주부 -	18	4	6				7	7	26	77	74	71	65	90	858	199	159	103	108	77	106	97	118	59		
	프리랜서 -	53	32	7				14	14	21	36	84	52	59	96	353	148	94	91	86	69	59	75	64	49	8-	1000
	학생(초/중/고) -	119	83	33	20	1		2	25	56	78	96	136	162	165	1047	457	372	363	284	247	226	289	346	252		
	회사원(사무직/연구직) -	239	57	21	19	6	11	34	107	335	521	563	566	566	728	5304	1183	1006	733	691	545	464	505	440	392		
	회사원(제조/건설/기술직) -	81	28	9	3	1	15	15	51	103	176	230	253	220	230	1702	515	372	236	206	224	230	207	155	92		
		Ó	ί	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23		

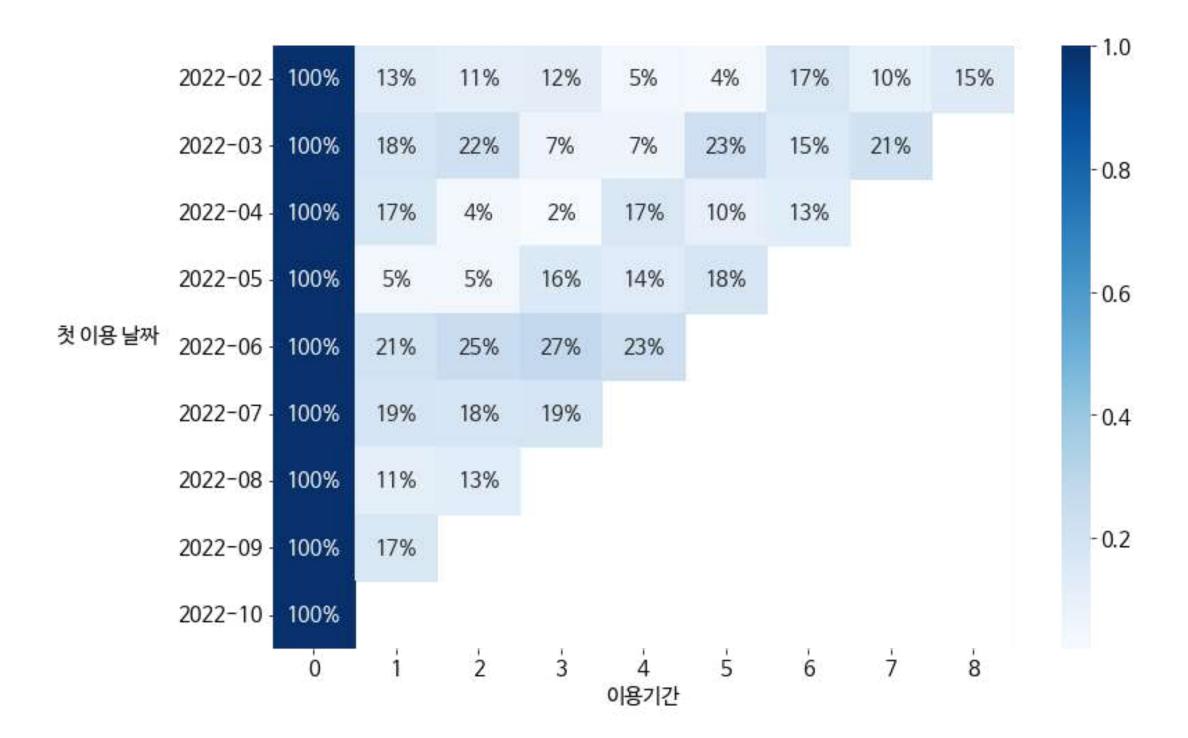
예매 시간

예매는 주로 경기 당일이나 경기 3일 후 이루어짐

경기요일별 예매요일



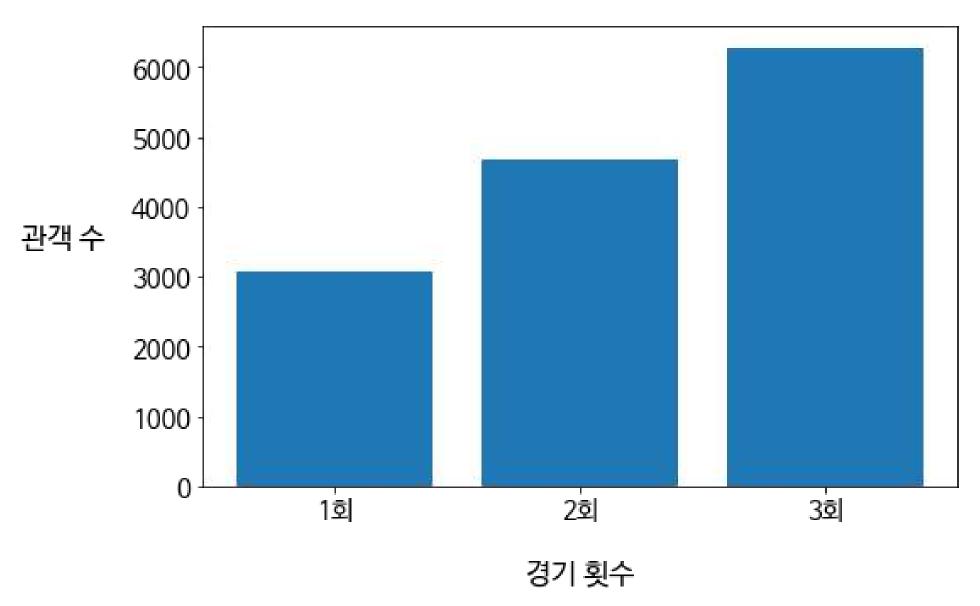
당일예매 회원 리텐션 코호트 분석 당일 예매 회원은 꾸준히 당일 예매를 이용하고 있다.



패턴2. 경기 횟수가 잦을수록 관객 수가 많다.

월 경기 횟수가 잦을수록 경기장을 이용하는 평균 관객 수가 늘어난다.

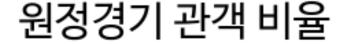
경기 횟수별 평균 관객 수

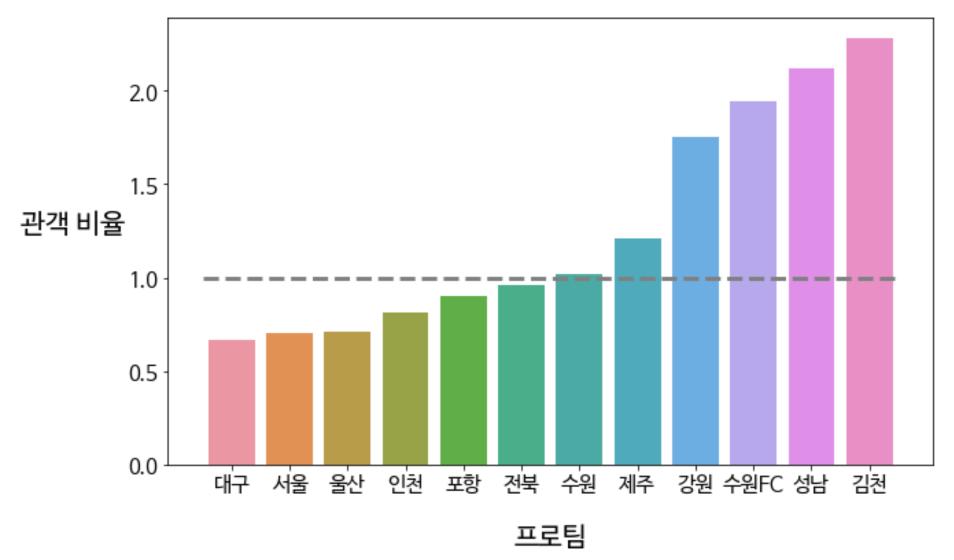


패턴3. 홈경기 관객 비율이 높다.

홈경기의 관객 수가 원정 경기의관객 수 보다 많다.

관객 비율 = 원정경기 평균관객수 / 홈경기 평균 관객 수





가설 2 결론

경기 예매 피크 타임

- 월요일, 화요일, 수요일 14시 경 많은 회사원, 기타 고객이 예매한다.
- 이는 경기 당일과 경기 3일 후에 예약하는 패턴 때문인 것으로 파악된다.
- 이 시간에 고급좌석 이용 경험이 있는 고객에게 고급좌석 얼리버드 할인 쿠폰을 지급하여 노출효과를 올린다.

월별 경기 횟수

- 월별 경기 수가 많을수록 관객 수도 늘어난다.
- 관객 수가 많을수록 좌석의 쾌적함이 떨어지므로 경기가 많은 달에 고급좌석 할인을 진행

당일예매 고객

- 당일예매 고객은 지속적으로 당일예매를 이용한다. 이는 경기 수요의 예측을 방해 하고 운영비용을 추가로 발생시킨다.
- 온라인 예약 시 할인 받을 수 있는 쿠폰을 지급하여 온라인 예매를 이용하도록 유도 한다.

홈경기 관객 비율

홈경기의 관객비율이 높다는 것은 홈팬들이 많이 온다는 뜻일 수 있지만 원정팬들의 진입장벽이 높다는 의미로도 해석될 수 있으므로 원정팬을 위한 행사도 진행할 필요가 있다.

가설3

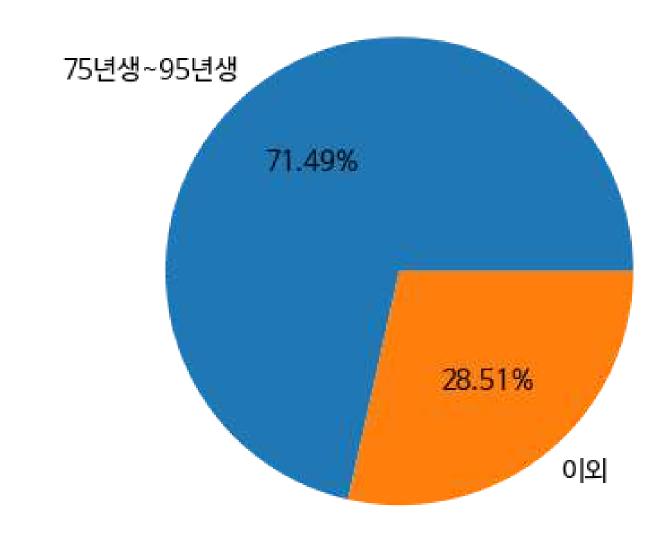
구매유도에 호응하는 고객이 있다.

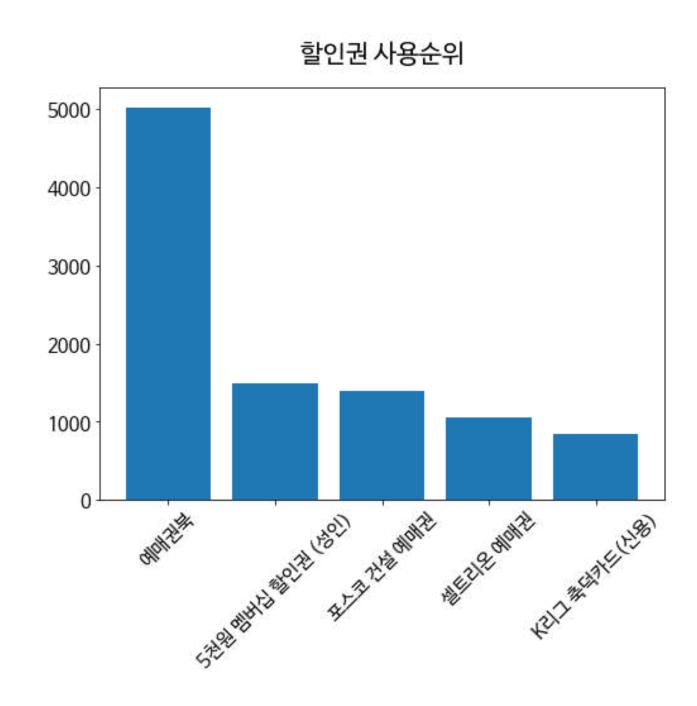


구매유도에 응하는 고객

할인권 사용 고객의 70% 이상이 75~95년생 가장 인기가 많은 할인은 예매권북

할인권 사용 나이 비율





구매유도에 응하는 고객

70년대 중반~ 80년대 중반 청소년/어린이 티켓 구매

Top4 구역 연령별 예매상품

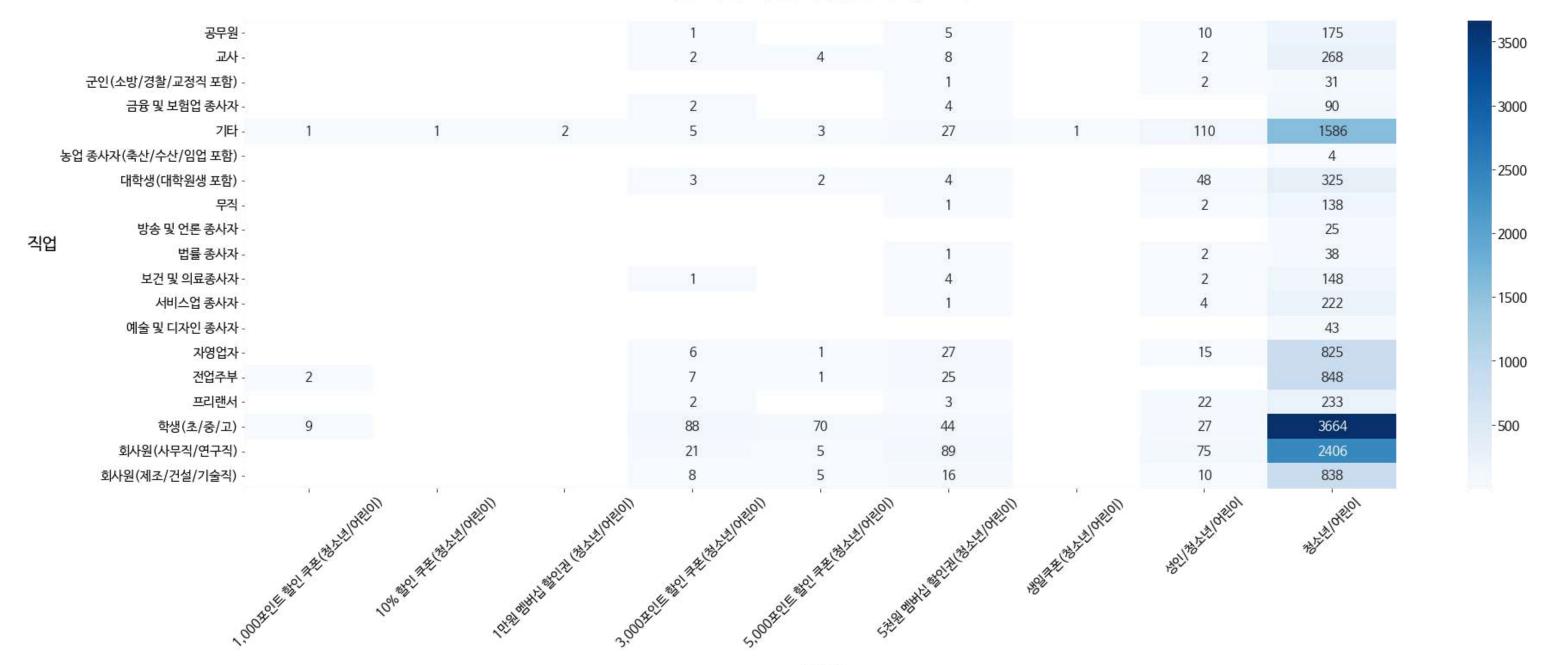
	1,000포인트 할인 쿠폰(성인)	-			1					1	1				1		1					2	2	1	1					2			1	2										
1,000포	E인트 할인 쿠폰(청소년/어린이)	-								1			1																							4								
	3,000포인트 할인 쿠폰(성인)	- 1	8	6	8	7	5	8	11	19	9	13	15	9	13	9	10	5	10	5	10	14	15	13	3	5	4	7	5	4	3	4	2	12			1							
3,000포	프인트 할인 쿠폰(청소년/어린이)	- 2	8		1		3	4	2	4	4	1	4	3			1						1	2								2	1	1	10	7	9	34	9					
	5,000포인트 할인 쿠폰(성인)	- 3	4	5	6	4	3	3	7	9	10	8	11	6	10	8	9	8	10	10	6	16	15	17	12	9	11	9	13	7	3	8	5	6	1	1		1						
5,000포	프인트 할인 쿠폰(청소년/어린이)	-	1						3	4	2		1			1			1				4										1		9		4	40	7	2				
	5천원 멤버십 할인권 (성인)	- 13	10	14	42	9	51	22	39	46	54	47	73	42	60	69	23	40	40	16	57	92	101	93	42	31	45	23	14	7	6	38	5	7	2	2	3		4		3			
5천원	원 멤버십 할인권(청소년/어린이)	-	2	2	5	2	15	9	29	32	12	10	17	9	8	5		4	1			2	1											5	6	6			12		1			
	K리그 PASS	4)												8		2		9		7	10	10	12		8			2		2	2													
직업	K리그 축덕카드(신용)	- 4	4	4	15	2	2	2	21	39	36	39	37	15	32	23	6	18	20	17	37	40	49	29	27	16	25	6										2						
	K리그 축덕카드(체크)	-		1	21	2		4	4	17	13	6	21	30	6	2	14	1	21	25	23	34	46	38	34	15	36	19	18	27	30	11	19	25	12				1					
	ㅅㅇ전 게임권	-	4	4					7	5	6	16	19	2	10	5	8	28	6	3	12	11	29	10	12	2	13	6		4	9	5	3	3	4			4	4					
	게임권	-	12	2	16	18	6	6	12	23	6	7	11	13	9	7	6	6	17	3	26	23	50	10	27	26	12	17	9	19	7	21	2	6	16	28	2	9			3			
톳	롯데백화점 30% 할인쿠폰(성인)	÷																								5																		
	문화누리카드(성인)	-	4	8		3	2		4	64	6	2	3	5	5	1	2	1		28	4	10	4	7		7	10	33	6	13	12	9	6	16		4	14				3			
	성인	111	214	196	241	321	276	322	336	546	522	488	593	540	437	517	363	417	504	462	544	587	685	750	571	584	615	679	566	452	448	255	330	303	78	47	56	44	24	3	7		4	1
	셀트리온 예매권			6	6	4	2	12	15	38	12	12	16	47	83	17	64	32	46	63	44	62	73	104	60	43	28	21	31	16	4	12					2							
	예매권북	- 23	84	54	66	77	60	1/16	164	201	ววก	122)E0	152	250	141	69	156	121	98	123	163	238	114	111	79	85	81	23	36	30	31	24	10	33	32	4	22	26	20	14	2	4	
	청소년/어린이	68	114	15	346	337	235	383	417	574	583	500	478	421	309	45	152	89	72	75	50	42	38	75	3	54	25	10	10	40	26	57	55	181	720	565	390	682	229	79	28	5	4	7
	포스코 건설 예매권	19	32	2	64	38	52	46	101	52	109	85	109	111	75	100	45	16	40	12	20	4	7	6	8	8	8	2	44		2		2	2	2	8	-		5					
		71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00	01	02	03	04	05	06	07	08	09	10	11	12	13

- 300

구매유도에 응하는 고객

청소년 / 어린이 티켓 구매자들의 쿠폰 사용률이 현저히 낮음

청소년/어린이 상품 구매고객



가설 3 결론

구매유도에 호응하는 고객이 있다.

호응고객

- 쿠폰 사용자의 70%가 75년생~95년생이다.
- 가장 인기있는 할인은 예매권북이다.

비호응고객

- 70년대초반~ 80년대 초반 청소년/어린이 티켓 구매자들은 쿠폰 사용비율이 적다.
- 미성년자 관객의 비율은 전체의 약 24%를 차지한다.
- 가족할인 혜택을 제공하여 구매호응을 유도할 수 있다.

가설4

고급좌석 예매 고객 분석



고급좌석 이용자

주로 70~90년대생 직장인이 이용

7~90년생 고급좌석 이용자

-60

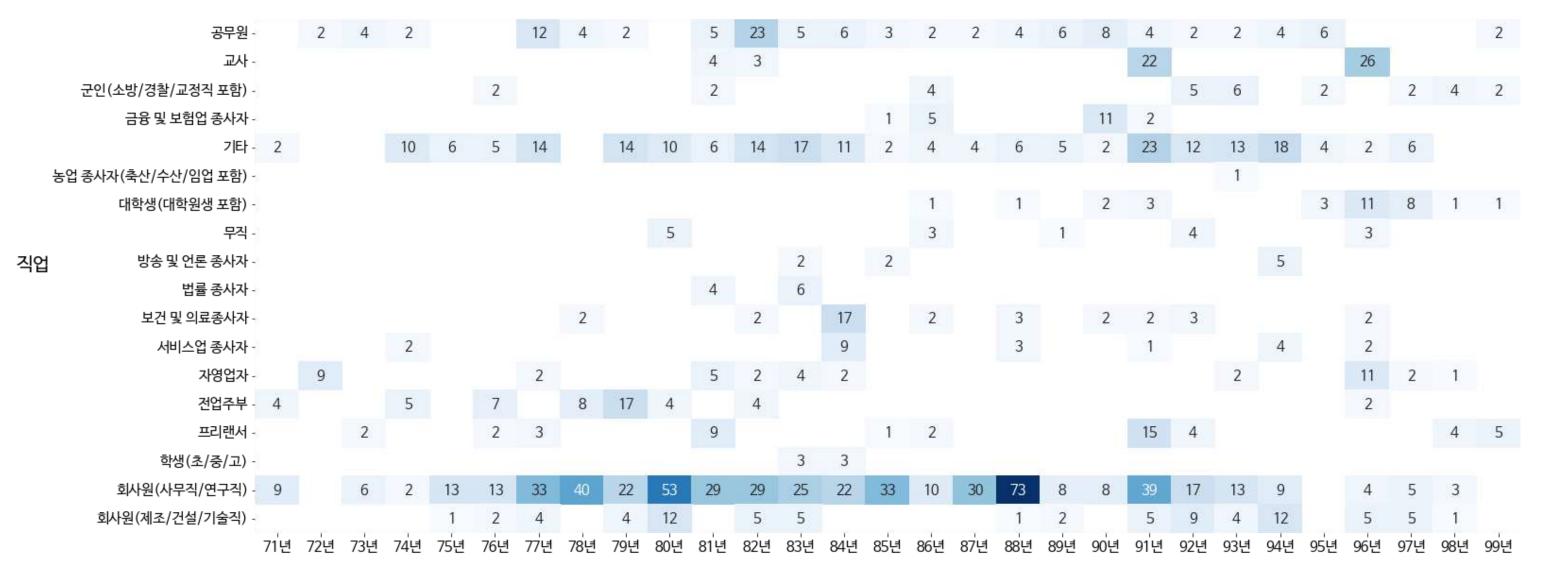
-50

- 40

- 30

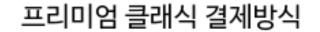
- 20

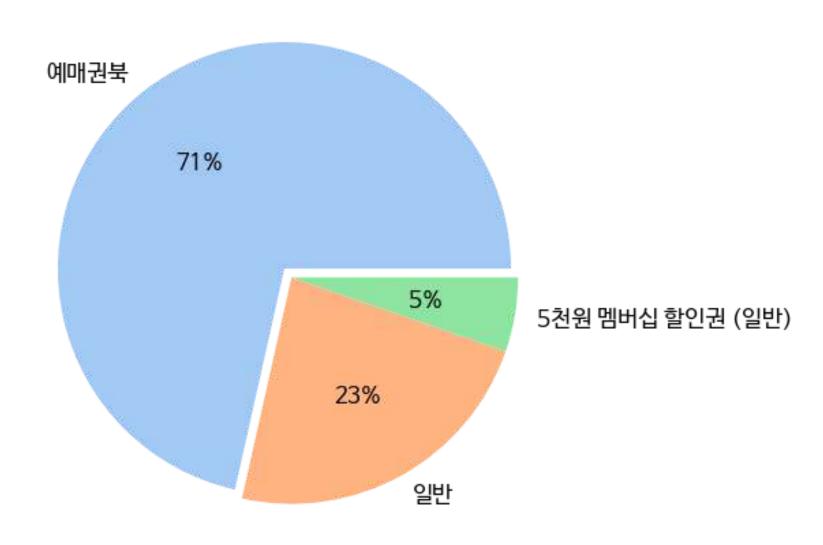
- 10



고급좌석 이용자 (프리미엄 클래식)

프리미엄 클래식 관객의 70% 이상이 예매권북을 사용한다. 해당 고객군은 청소년 / 어린이 구매 이력이 적다.

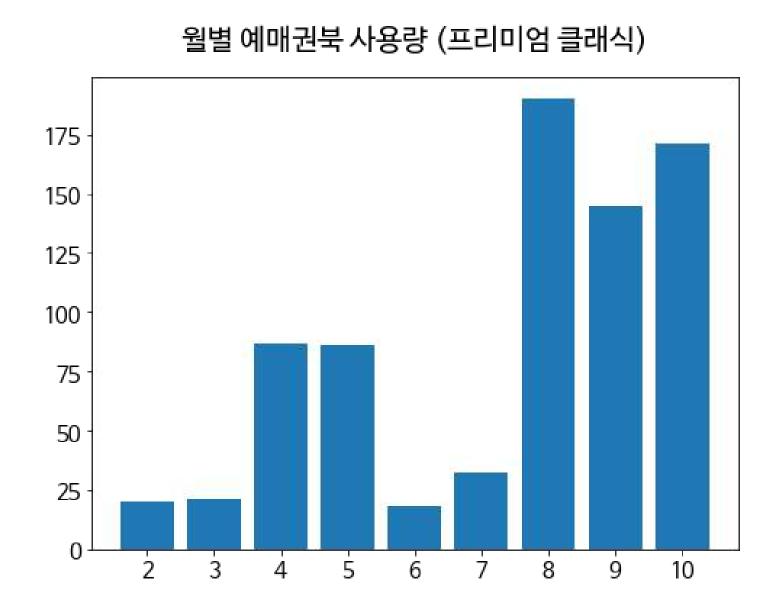


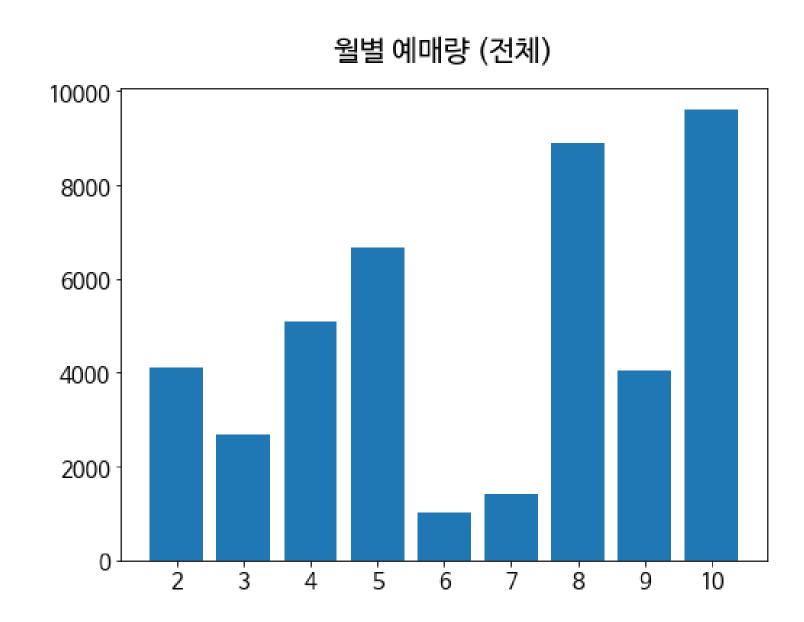


프리미엄 골드석 50% 할인쿠폰	2
3,000포인트 할인 쿠폰(일반)	9
5,000포인트 할인 쿠폰(일반)	10
5천원 멤버십 할인권 (성인)	11
K리그 축덕카드(체크)	11
K리그 축덕카드(신용)	26
청소년/어린이	26
성인	67
5천원 멤버십 할인권 (일반)	87
일반	451
예매권북	770

프리미엄 클래식 이용 추이

주로 시즌 후반에 사용하며, 전체 관객 수와 고급좌석의 수요가 비례하는 경향이 있다.





테이블석

프리미엄석에 비해 예매권북 사용이 현저히 낮고 <mark>일반티켓 이용자가 많으며</mark>, 프리미엄 클래식구역과 마찬가지로 청소년 / 어린이 구매 이력이 적다.

70~90년생 테이블석 이용자

120

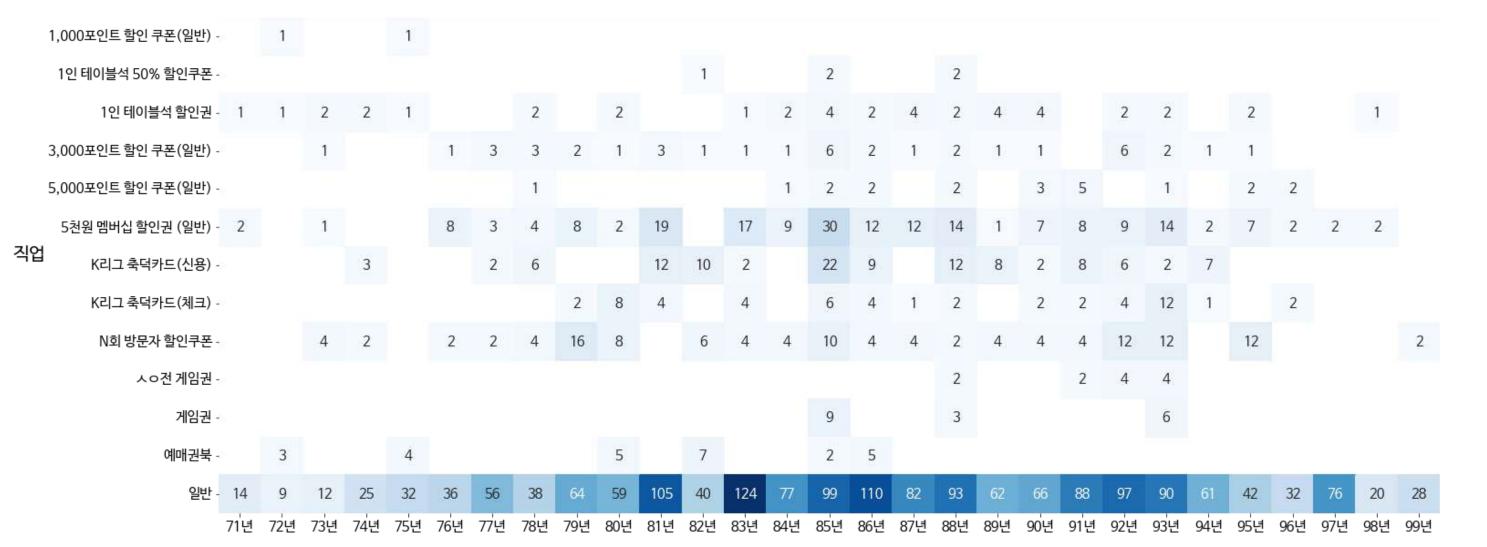
100

- 80

- 60

40

- 20



가설 4 결론

고급좌석 예매 고객 분석

프리미엄 클래식

- 주로 70년~90년대생이고, 미성년자 티켓 구매 이력이 적은 것으로 보아 자녀가 없는 고객층으로 예상된다.
- 예매권북 사용률이 높고, 하반기에 예매권북을 사용하는 경향이 있다.
- 전체 관객 수에 비례하여 고급좌석의 수요도 늘어나는 경향이 있다.
- 하반기에는 예매권북으로 인한 수요가 있으므로 상반기에 고급좌석을 할인 판매하여 예매권북 소지자도 참여할 수 있도록 진행한다.

테이블석

- 예매권북 사용률이 적고 일반티켓 수요가 있는 편이다.
- 프리미엄 클래식과 마찬가지로 자녀가 없는 고객층으로 판단된다.
- 테이블석은 부수적인 이익을 창출하기에 좋은 좌석이므로 예매권북 홍보 및 할인행사 진행 필요



프로모션 제안

- 인천 유나이티드가 패배한 경우 경기관람의 부정적인 감정을 해소하기 위해 일 부 고객에게 할인 혜택을 제공한다.
- 월, 화, 수 14시 경 70~90년대 직장인을 대상으로 마케팅을 진행한다.
- 자녀가 없는 경우 고급석 또는 테이블석을 할인을 제공한다.
- 자녀가 있는 경우 미성년 추가 할인이나 가족할인 등을 제공한다.
- 하반기에는 관객이 늘어나고 고급석의 수요가 좋은 편이므로 상반기에 할인 행사를 통해 고객의 경험을 증진한다.
- 테이블석은 예매권북 수요가 적으며 일반 예매가 주를 이루므로 즉흥적 마케팅
 을 통해 사용량을 늘린다.

프로젝트 회고

고급좌석 예매 예측모델이라는 프로젝트 목표에 도달하지 못하여 아쉽다. 하지만, 데이터를 직접 뜯어보고 들여 다 보는 것도 중요하다는 생각으로 열심히 했고 모델링을 하지 못한 것 치고는 만족한다.

생각은 많았으나 생각대로 되지 않는 부분들도 많았고 생각나는 대로 풀어가다 보니 정리하기가 쉽지 않았다. 중 간에 한 번 프로젝트를 정리하는 과정이 많은 도움이 되었다. 분석을 할 때 꼭 큰 줄기를 기억하면서 진행해야 한 다는 것을 다시 한 번 느꼈다.

또 무턱대고 데이터를 보기 시작했는데, 데이터를 보면 볼 수록 도메인 지식을 잘 알아야 한다는 생각이 많이 들었다. 도메인 지식은 데이터분석의 지도와 같다는 생각이 든다.

파이썬이 아닌 다른 기술들도 사용하고 싶었으나, 파이썬만 사용하기에도 벅찼다. 생각을 이루기 위해서는 꾸준한 공부와 새로운 기술을 연마할 필요성을 느꼈고 몇 가지 책을 구매하였지만 공부하지는 못했다. 시간을 계획적으로 사용하지 못한 것 같아 아쉽다.