

EDA Methodology

23.01.10 / 8기 조찬형

CONTENTS

01. 분석이란?

분석이란 무엇인가?

분석 일반론 / 절차

02. EDA란?

탐색적 자료 분석?

EDA 일반론

03. 데이터 핸들링

데이터의 특성

구조에 따른 탐색

목적에 따른 탐색

실제 자료 EDA

04. 프로젝트 설명

EDA 프로젝트 설명

전 기수 예시 프로젝트

05. 각 조별 현황 브리핑

06. 과제 및 Q&A

0. INTRO

EDA 방법론?

여러분이 진행하게 될 EDA 프로젝트에 도움이 되고자 조직한 세션.

그러나 또한 ‘데이터 분석’이라는 큰 하나의 대주제를 분석함으로써,

데이터에 접근하는 일련의 과정에 대한 가이드라인을 제공하고자 했습니다.

0. INTRO

EDA 방법론?

여러분이 진행하게 될 EDA 프로젝트에 도움이 되고자 조직한 세션.

그러나 또한 ‘데이터 분석’이라는 큰 하나의 대주제를 분석함으로써,

데이터에 접근하는 일련의 과정에 대한 가이드라인을 제공하고자 했습니다.

그래서, ‘분석’이 뭘까요?

1. 분석이란?

분석이란 무엇인가?

‘분석’ (分析), ‘나눌 분’, ‘쪼갠 석’

- 복잡하고 많은 내용을 가진 대상을 정확히 이해하기 위해 단순한 요소로 나누어 생각함
- 정확히 이해 : 어디를 얼마나 이해해야 정확한 이해일까?

1. 분석이란?

분석이란 무엇인가?

• **정확히 이해** : 어디를 얼마나 이해해야 정확한 이해일까?



OR



1. 분석이란?

분석이란 무엇인가?

자동차를 분석한다면 : 열개나 사용 목적에 따라 분석할 수도 있고, 개별 부품 하나하나를 분석할 수도 있다.

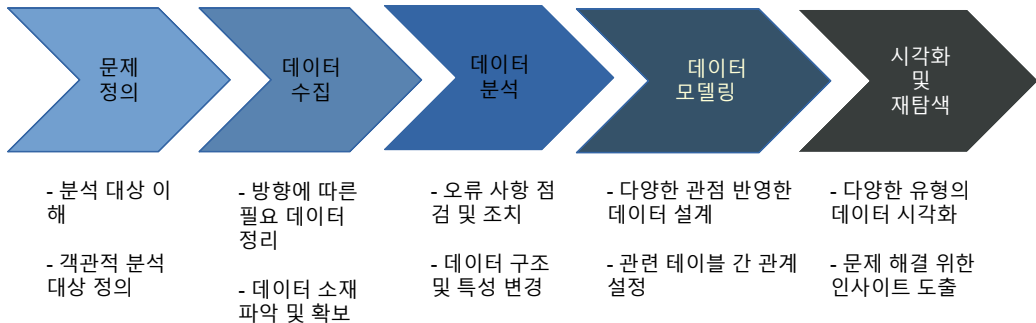
열개에 따라 분석한다면, 기능적인 부분에서의 접근을, 부품에 따라 분석한다면, 생산이나 유지/보수적인 측면에서의 접근이 가능하다.

정확한 이해는, 자신에게 필요한 것이 무엇인지 인식하는 데서 시작한다.

그럼 데이터는 어떻게 분석해야 할까?

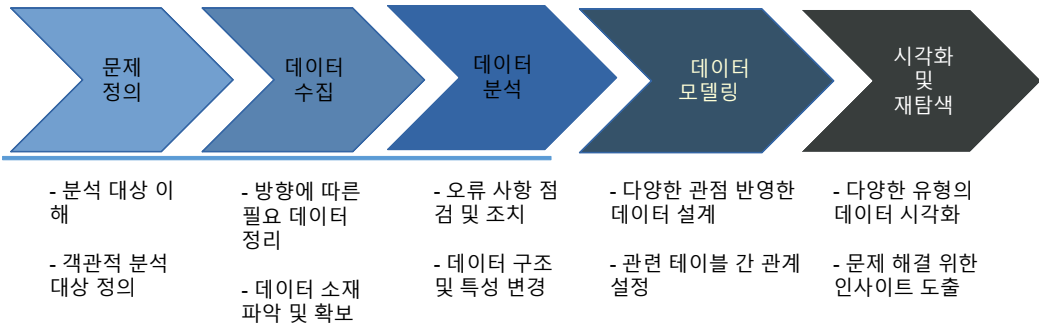
1. 분석이란?

분석 일반론 / 절차



1. 분석이란?

분석 일반론 / 절차



2. EDA란?

탐색적 자료분석?

우리의 **목적에 따라** 분석 진행 방향,
그리고 접근법이 달라진다는 것을 확인했다.

목적에 따른 통계분석 과정을
구조도로 나타내면 다음과 같다.

1-2. 통계분석과정

1) 문제발설(DDA : Descriptive Data Analysis)

- 목적 : 현재의 모습 요약, 수집 자료의 표현
- 분석도구 : 수식을 활용한 계산
 - * 추정과정을 통해 표본을 가지고 모평균을 추측
 - * 모평균 / 모분산 / 모비율을 구하여 오차와 신뢰구간 제시
 - * 신뢰구간은 얼마나 Data를 탐색했는지 알 수 있는 척도가 됨

2) 현상확인(EDA : Exploratory Data Analysis)

- 목적 : 수집된 데이터의 증상 탐색을 통한 가설도출, 가능성이 큰 x-y관계의 가설 도출
- 분석도구 : 그래프분석, 다중변수 연계 그래프 분석, 정규성 확인, 트렌드 확인
 - * 문제 파악 단계
 - * 데이터를 이용해 가설을 구축하는 단계
 - * 변수들을 이용해 주관적 Data확인
 - * 요인에 영향을 주는 변수가 제거
 - * 객관적인 Data로 바꾸는 작업
 - ⇒ 가설 : EDA과정을 통해 만들어진 객관적인 데이터

3) 가설구축 및 원인규명(CDA : Confirmatory Data Analysis)

- 목적 : 도출된 가설을 검증, p-value를 기준으로 의사결정
- 분석도구 : t-test, ANOVA, 상관분석, 카이제곱 검정, 회귀분석, 모비율검정, 비모수적 검정
 - * 확증적 분석과정
 - * p-value를 계산하여 구축된 가설 검정
 - * 과학적 의사 결정 진행

4) 대적마인(PDA : Predictive Data Analysis)

- 목적 : 관계식을 만들고, 최적 조건을 예측
- 분석도구 : Modeling 기법, Profiler최적화, Contour최적화, 시뮬레이션
 - * 가설을 통해 분석된 변수가 결과에 얼마나 영향을 주는지 파악
 - * x와 y간 모델식을 형성하는 단계
 - * 최적화 과정 진행 / 변수의 관찰과 예측이 구현되는 단계
 - * 이후 시스템을 통한 실행

2. EDA란?

탐색적 자료분석?

우리의 **목적에 따라** 분석 진행 방향,
그리고 접근법이 달라진다는 것을 확인했다.

목적에 따른 통계분석 과정을
구조도로 나타내면 다음과 같다.

EDA 방법론, 그리고 프로젝트에서는
1)에서 3)까지의 과정을 주로 다룬다.

1-2. 통계분석과정

1) 문제발생(DDA : Descriptive Data Analysis)

- 목적 : 현재의 모습 요약, 수집 자료의 표현
- 분석도구 : 수식을 활용한 계산
 - * 추정과정을 통해 표본을 가지고 모평균을 추측
 - * 모평균 / 모분산 / 모비율을 구하여 오차와 신뢰구간 제시
 - * 신뢰구간은 얼마나 Data를 탐색했는지 알 수 있는 척도가 됨

2) 현상확인(EDA : Exploratory Data Analysis)

- 목적 : 수집된 데이터의 증상 탐색을 통한 가설도출, 가능성이 큰 x-y관계의 가설 도출
- 분석도구 : 그래프분석, 다중변수 연계 그래프 분석, 정규성 확인, 토렌드 확인
 - * 문제 파악 단계
 - * 데이터를 이용해 가설을 구축하는 단계
 - * 변수들을 이용해 주관적 Data확인
 - * 요인에 영향을 주는 변수가 제거
 - * 객관적인 Data로 바꾸는 작업
 - ⇒ 가설 : EDA과정을 통해 만들어진 객관적인 데이터

3) 가설구축 및 원인규명(CDA : Conirmatory Data Analysis)

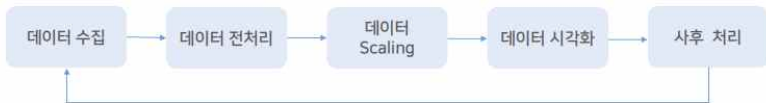
- 목적 : 도출된 가설을 검증, p-value를 기준으로 의사결정
- 분석도구 : t-test, ANOVA, 상관분석, 카이제곱 검정, 회귀분석, 모비율검정, 비모수적 검정
 - * 확증적 분석과정
 - * p-value를 계산하여 구축된 가설 검증
 - * 과학적 의사 결정 진행

4) 대략이론(PDA : Predictive Data Analysis)

- 목적 : 관계식을 만들고, 최적 조건을 예측
- 분석도구 : Modeling 기법, Profiler회직화, Contour회직화, 시뮬레이션
 - * 가설을 통해 분석된 변수가 결과에 얼마나 영향을 주는지 파악
 - * x와 y간 모델식을 형성하는 단계
 - * 회직화 과정 진행 / 변수의 관찰과 예측이 구현되는 단계
 - * 이후 시스템화를 통한 실행

2. EDA란?

EDA 일반론



1. **데이터 수집**: 데이터 수집 파이프라인 생성, 수집 필요 데이터 목록화
2. **데이터 전처리**: 결측치 처리, 이상치 탐색, 데이터 라벨링, 정규식을 통한 데이터 정리 etc..
3. **데이터 Scaling**: 데이터 범위 조율, 표준화/정규화, 데이터 양 조절(오버 샘플링/언더 샘플링)
4. **데이터 시각화**: 데이터 시각화 (모델링)
5. **사후 처리**: 결측치 처리, 이상치 탐색, FineTuning, 추후 분석 방향 결정

2. EDA란?

구조에 따른 탐색



1. **데이터 수집**: 데이터 수집 파이프라인 생성, 수집 필요 데이터 목록화
2. **데이터 전처리**: 결측치 처리, 이상치 탐색, 데이터 라벨링, 정규식을 통한 데이터 정리 etc..
3. **데이터 Scaling**: 데이터 범위 조율, 표준화/정규화, 데이터 양 조절(오버 샘플링/언더 샘플링)
4. **데이터 시각화**: 데이터 시각화 (모델링)
5. **사후 처리**: 결측치 처리, 이상치 탐색, FineTuning, 추후 분석 방향 결정

3. 데이터 핸들링

데이터의 특성

정형

우리가 가진 데이터는
어떤 특성을 가졌나?

비정형

3. 데이터 핸들링

비정형 데이터의 특성



핸들링하기 어려운 데이터.

핸들링을 위해 여러분이 차후 6개월 간의 모델링 및 사전 지식을 학습함.

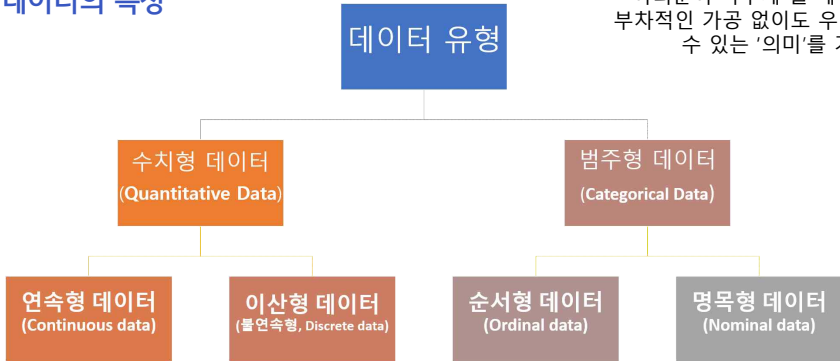
MP3, MP4, PNG, JPG 등의 파일들이 이에 속하며, 개략적으로 설명하면 자료형에 따라 서로 다른 방식의 '패턴의 분석'으로 비정형 데이터의 핸들링이 이뤄짐.

EX) 승범이와 찬형이가 경제학 시험을 보는데, 채점이 끝난 답안지를 녹화한 영상에 98 / 27이라는 점수가 적혀있었다. 이 영상은 여타 작업으로 패턴을 분석하지 않는다면 그저 영상에서 그친다. -> 자체로 의미를 창출하기 어려움

3. 데이터 핸들링

정형 데이터의 특성

여러분이 다루게 될 데이터 형태.
부차적인 가공 없이도 우리가 이해할
수 있는 '의미'를 가짐.



3. 데이터 핸들링

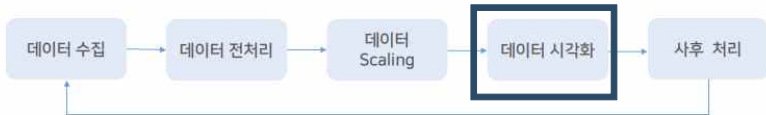
정형 데이터의 특성

- 수치형 데이터(Continuous): 1, 2, 3, 4, 5 ... 98, 99, 100
- 수치형 데이터(Discrete): 100, 393, 45, 2222, ... 12, 885
- 범주형 데이터(Binary): '남자', '여자' / 0, X / 0, 1 ...
- 범주형 데이터(MultiClass): A+, A0, A-, B+ ... , D0, D- , F

EX) 학생 시험 점수를 기록한 엑셀 문서를 보니
승범 98
찬형 27
이라고 적혀있다.
'시험 성적에서 승범이 더 높은 점수를 받았다'
는 정보를 별도의 작업 없이도 얻어낼 수 있다.

3. 데이터 핸들링

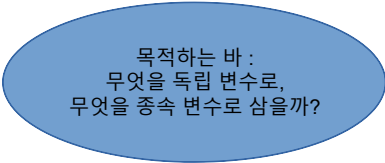
구조에 따른 탐색



1. **데이터 수집**: 데이터 수집 파이프라인 생성, 수집 필요 데이터 목록화
2. **데이터 전처리**: 결측치 처리, 이상치 탐색, 데이터 라벨링, 정규식을 통한 데이터 정리 etc..
3. **데이터 Scaling**: 데이터 범위 조율, 표준화/정규화, 데이터 양 조절(오버 샘플링/언더 샘플링)
4. **데이터 시각화**: 데이터 시각화 (모델링)
5. **사후 처리**: 결측치 처리, 이상치 탐색, FineTuning, 추후 분석 방향 결정

3. 데이터 핸들링

목적에 따른 탐색



목적하는 바 :
무엇을 독립 변수로,
무엇을 종속 변수로 삼을까?

3. 데이터 핸들링

목적에 따른 탐색

일변량

시각화

목적하는 바 :
무엇을 독립 변수로,
무엇을 종속 변수로 삼을까?

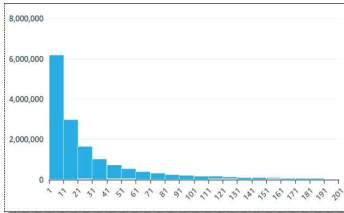
다변량

비시각화

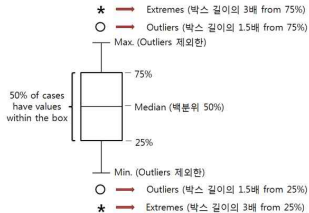
2. EDA란?

목적에 따른 탐색

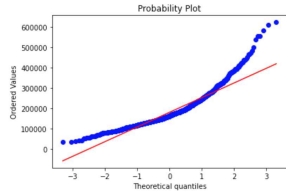
일변량/
시각화



Histogram / Bar plot



Box Plot



QQ-Plot

3. 데이터 핸들링

목적에 따른 탐색

일변량/
비시각화

Descriptive Statistics

Variable	Obs	Mean	Std.Dev.	Min	Max
price	74	6165.257	2949.496	3291	15906
mpg	74	21.297	5.786	12	41
rep78	69	3.406	.99	1	5
headroom	74	2.993	.846	.846	5
trunk	74	13.757	4.277	5	23
weight	74	3019.459	777.194	1760	4840
length	74	187.932	22.266	142	233
turn	74	39.649	4.399	31	51
displacement	74	197.297	91.837	79	425
gear_ratio	74	3.015	.456	2.19	3.89
foreign	74	.297	.46	0	1

Summary Statistics :
수치 데이터 분석에 주로 사용

2. EDA란?

목적에 따른 탐색

다변량/
비시각화

Subject ID	Age Group	Sex
GW	young	F
JA	middle	F
TJ	young	M
JMA	young	M
JMO	middle	F
JQA	old	F
AJ	old	F
MVB	young	M
WHH	old	F
JT	young	F
JKP	middle	M

Table 4.1: Sample Data for Cross-tabulation

Age Group / Sex	Female	Male	Total
young	2	3	5
middle	2	1	3
old	3	0	3
Total	7	4	11

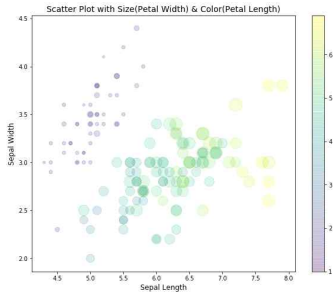
Table 4.2: Cross-tabulation of Sample Data

Cross tabulation

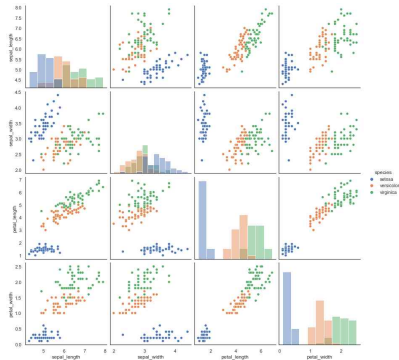
2. EDA란?

목적에 따른 탐색

다변량/
시각화



Scatter plot



범주별 시각화

3. 데이터 핸들링

전처리 / Scaling



1. **데이터 수집**: 데이터 수집 파이프라인 생성, 수집 필요 데이터 목록화
2. **데이터 전처리**: 결측치 처리, 이상치 탐색, 데이터 라벨링, 정규식을 통한 데이터 정리 etc..
3. **데이터 Scaling**: 데이터 범위 조율, 표준화/정규화, 데이터 양 조절(오버 샘플링/언더 샘플링)
4. **데이터 시각화**: 데이터 시각화 (모델링)
5. **사후 처리**: 결측치 처리, 이상치 탐색, FineTuning, 추후 분석 방향 결정

3. 데이터 핸들링

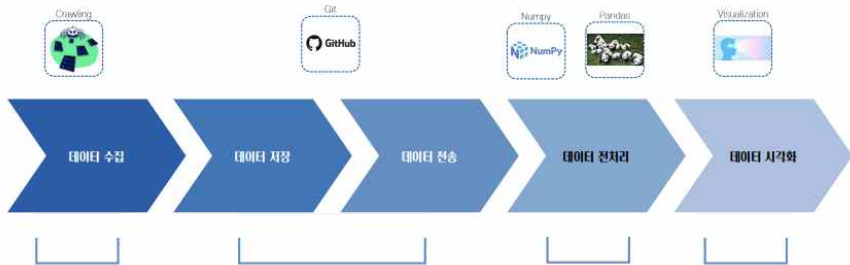
실제 자료 EDA

타이타닉 호 EDA로 이동

04. 프로젝트 설명

EDA 프로젝트 설명

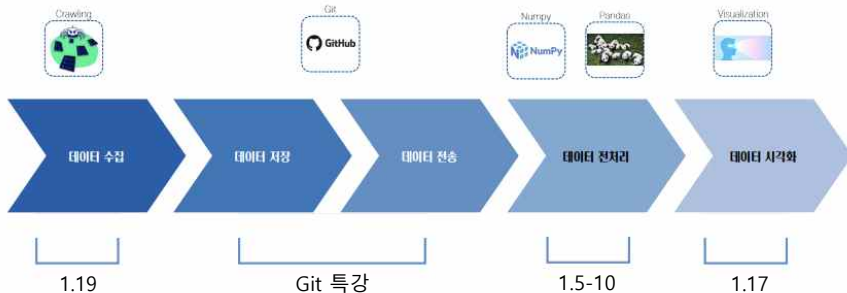
1. 기초 세션 설명



04. 프로젝트 설명

EDA 프로젝트 설명

1. 기초 세션 설명



04. 프로젝트 설명

8기 EDA 프로젝트



DSL 3조 EDA 발표
저소득계층의 신용카드 연체율 결정 요인 분석

강건우 김찬영 도승범 송규원 조찬형

05. 각 조별 현황 브리핑

1조 - 금융 데이터

Date	Close	Open	High	Low	Volume	Change
1981-05-01	123.6	123.6	123.6	123.6	3330000	0.0098
1981-05-02	123.5	123.5	123.5	123.5	2040000	-0.0008
1981-05-04	120.6	120.6	120.6	120.6	1930000	-0.0235
1981-05-06	120.7	120.7	120.7	120.7	1690000	0.0008
1981-05-07	119.3	119.3	119.3	119.3	1480000	-0.0116

지수명	종가	대비	등락률	시가	고가	저가	거래량	거래대금
코스피 (외국주포함)	-	-	-	-	-	-	398,607	6,764,112,169,245
코스피	2,289.97 ▲	25.32	+1.12	2,253.40	2,300.62	2,253.27	398,343	6,763,322,047,150
코스피 200	301.53 ▲	3.66	+1.23	296.69	303.10	296.40	121,022	5,367,994,107,791
코스피 100	2,273.23 ▲	28.48	+1.27	2,236.71	2,284.50	2,236.39	82,753	4,625,964,431,961
코스피 50	2,076.39 ▲	26.47	+1.29	2,043.38	2,086.97	2,043.20	54,177	3,705,736,671,148
코스피 200 중소형주	1,070.92 ▲	8.13	+0.76	1,053.81	1,078.20	1,051.66	38,269	742,029,675,830
코스피 200 초대형제외 ...	221.76 ▲	2.57	+1.17	217.83	222.75	217.73	103,687	4,349,224,929,891
코스피200제외 코스피지...	2,886.33 ▲	16.37	+0.57	2,859.00	2,900.80	2,857.59	243,563	1,310,036,125,812
코스피 200 커뮤니케이...	1,753.06 ▼	6.58	-0.37	1,752.78	1,766.95	1,741.99	5,256	445,103,472,356
코스피 200 건설	257.29 ▲	10.53	+4.27	244.27	258.83	243.70	11,929	345,806,479,980
코스피 200 항공업	258.78 ▲	2.03	+0.79	254.25	260.20	252.76	13,613	248,578,345,840
코스피 200 유통/소재	776.71 ▲	15.83	+2.08	757.56	777.95	753.63	2,259	177,496,289,640
코스피 200 에너지/화학	1,440.94 ▲	20.56	+1.45	1,411.34	1,448.49	1,404.31	5,394	404,542,973,710
코스피 200 정보기술	2,749.21 ▲	34.13	+1.26	2,701.22	2,766.62	2,698.98	25,452	1,730,464,117,450
코스피 200 금융	648.98 ▲	16.12	+2.55	628.42	653.70	627.98	26,764	696,597,114,529

05. 각 조별 현황 브리핑

2조 - 산업 데이터

uuid	name	type	permalink	cb_url	rank	created_at	updated_at	funding_ra	funding_rc	investor_u	investor_n	investor_ty	is_lead
524986f0-	Accel Inve	investmen	accel-inve	https://www.crunchb	#####	#####	d950d7a5-	Series A -	b08efc27-	Accel	organizati	TRL	
6556ab92-	Greylock i	investmen	greylock-i	https://www.crunchb	#####	#####	6fae3958-	Series B -	e2006571-	Greylock	organizati	TRL	
0216e06a-	Meritech C	investmen	meritech-c	https://www.crunchb	#####	#####	6fae3958-	Series B -	8d5c7e48-	Meritech C	organizati	TRL	
dadd7d86	Trinity Ver	investmen	trinity-ven	https://www.crunchb	#####	#####	bcd5a63d-	Series B -	7ca12f7a-	Trinity Ver	organization		
581c4b38-	Founders	investmen	founders-i	https://www.crunchb	#####	#####	60e6afd9-	Series A -	fb2f8884-	Founders	organizati	TRL	
aa206330-	Columbia	investmen	columbia-	https://www.crunchb	#####	#####	0c000adf-	Series A -	86aace17-	Columbia	organization		
3fa4eaa9-	Benchmar	investmen	benchmar	https://www.crunchb	#####	#####	0c000adf-	Series A -	fe2d1e8b-	Benchmar	organization		
f33089df-	Redpoint	investmen	redpoint-v	https://www.crunchb	#####	#####	7a65fabf-	Series A -	d29973c3-	Redpoint	organization		
e36e30bd-	The Kinsey	investmen	kinsey-hill	https://www.crunchb	#####	#####	7a65fabf-	Series A -	c45f8255-	The Kinsey	organization		
f33859b5-	Sevin Rose	investmen	sevin-rose	https://www.crunchb	#####	#####	18c9eabc-	Series B -	313f8a67-	Sevin Rose	organizati	FAL	
de11ea8c-	Mission V	investmen	mission-ve	https://www.crunchb	#####	#####	18c9eabc-	Series B -	c6fff75f-7	Mission V	organizati	FAL	
b759b62f-	Centennia	investmen	centennia	https://www.crunchb	#####	#####	18c9eabc-	Series B -	a6c7c343-	Centennia	organizati	TRL	
92338303-	Mission V	investmen	mission-ve	https://www.crunchb	#####	#####	1ba57957-	Series A -	c6fff75f-7	Mission V	organization		
77d02fdc-	Sevin Rose	investmen	sevin-rose	https://www.crunchb	#####	#####	1ba57957-	Series A -	313f8a67-	Sevin Rose	organization		
456263e5-	Austin Ver	investmen	austin-ver	https://www.crunchb	#####	#####	1ba57957-	Series A -	960eb982-	Austin Ver	organization		
bac91079-	Acadia W	investmen	acadia-wo	https://www.crunchb	#####	#####	35fa83eb-	Series A -	cd888a28-	Acadia W	organizati	TRL	

05. 각 조별 현황 브리핑

3조 - 산업 데이터

기준_년_코드	기준_분기_코드	상권_구분_코드	상권_구분_코드_명	상권_코드	상권_코드_명	어파트_단지_수
2022	3	A	골목상권	2110511	연희초등학교	204
2022	3	A	골목상권	2110999	윗방죽마을공원	29
2022	3	A	골목상권	2110255	상봉역 4번	397
2022	3	A	골목상권	2110510	홍연시장	295
2022	3	A	골목상권	2110766	대림역 5번	167
2022	3	A	골목상권	2110998	은곡마을(은곡마을...	47
2022	3	A	골목상권	2110127	성수1가1동주민센터	58
2022	3	A	골목상권	2110254	오거리공원	227
2022	3	A	골목상권	2110382	솔밭공원역 1번	281
2022	3	A	골목상권	2110509	연희대우아파트앞	257
2022	3	A	골목상권	2110638	신방화역 2번	315
2022	3	A	골목상권	2110765	문래동사거리	30
2022	3	A	골목상권	2110893	관악소방서(문영여...	66
2022	3	A	골목상권	2110997	대청초등학교	222
2022	3	A	골목상권	2110063	효창공원앞역 5번	259
2022	3	A	골목상권	2110126	서울숲역 1번	40
2022	3	A	골목상권	2110190	고대앞사거리	147

05. 각 조별 현황 브리핑

4조 - 환경 데이터

측정년도	측정소명	이산화질소농...	오존농도(ppm)	일산화탄소농...	아황산가스(pp...	미세먼지($\mu\text{g}/\text{m}^3$)
2023	강남구	0.036	0.015	0.7	0.004	60
2023	강남대로	0.039	0.01	0.9	0.004	63
2023	강동구	0.03	0.012	0.7	0.003	65
2023	강변북로	0.042	0.01	0.9	0.003	62
2023	강북구	0.028	0.019	0.7	0.003	64
2023	강서구	0.036	0.016	0.8	0.004	76
2023	공청대로	0.039	0.012	0.8	0.003	66
2023	관악구	0.039	0.013	0.8	0.003	66
2023	관악산	0.013	0.036	0.4	0.004	50
2023	광진구	0.022	0.012	0.9	0.002	60
2023	구로구	0.031	0.013	0.5	0.003	66
2023	금천구	0.033	0.017	0.7	0.004	59
2023	남산	0.029	0.018	0.9	0.004	60
2023	노원구	0.041	0.015	0.8	0.004	65
2023	도봉구	0.029	0.02	0.7	0.003	56
2023	도산대로	0.038	0.011	0.9	0.004	46
2023	동대문구	0.035	0.015	0.8	0.004	60

05. 각 조별 현황 브리핑

5조 - 교통 데이터

대중교통 노선명	년월	승차총승객수
지하철 1호선	201711	8633618
지하철 1호선	201712	8737235
지하철 1호선	201801	8145989
지하철 1호선	201802	7273309
지하철 1호선	201803	8692551
지하철 1호선	201804	8275767
지하철 1호선	201805	8543247
지하철 1호선	201806	7972991
지하철 1호선	201807	8150061
지하철 1호선	201808	7930624
지하철 1호선	201809	7876712
지하철 1호선	201810	8783242

사용월	호선명	지하철역	유임승차인	무임승차인	유임하차인	무임하차인	작업일자
202211	1호선	동묘앞	135459	151050	143887	154599	20221203
202211	1호선	서울역	1316422	203139	1277104	193877	20221203
202211	1호선	시청	653911	97339	662910	95891	20221203
202211	1호선	종각	924585	141959	898103	131890	20221203
202211	1호선	종로3가	485926	300364	435252	276761	20221203
202211	1호선	종로5가	436948	242398	438774	232286	20221203
202211	1호선	동대문	220069	117440	206472	116436	20221203
202211	1호선	신설동	285012	117073	275205	112432	20221203
202211	1호선	제기동	233878	266335	221879	285648	20221203
202211	1호선	청량리(서)	341039	247954	339255	251899	20221203
202211	2호선	독섬	589326	58728	624810	59614	20221203
202211	2호선	성수	994259	103533	1097054	105464	20221203
202211	2호선	건대입구	986298	90587	1058395	94453	20221203

06. 과제 및 Q&A

제출 방식

 chanspick / 9th_EDA Public
forked from DataScience-Lab-Yonsei/9th_EDA

[Code](#) [Pull requests](#) [Actions](#) [Projects](#) [Wiki](#)


 main ▾

[Go to file](#)

[Add file](#)

This branch is up to date with DataScience-Lab-Yonsei/9th_EDA:main.

[Contribute](#)

 JungGeonwooYS Merge pull request DataScience-Lab-Yon... [1](#)

 1조	Merge pull request DataScience-Lab-Yonsei...
 2조	first assignment
 3조	Merge pull request DataScience-Lab-Yonsei...
 4조	Merge pull request DataScience-Lab-Yonsei... 1 hour ago
 5조	first assignment 2 days ago

 chanspick / 9th_EDA Public

forked from DataScience-Lab-Yonsei/9th_EDA

[Pin](#)

[Watch](#) 0 ▾

[Fork](#) 20 ▾

[Star](#) 0 ▾

[Code](#) [Pull requests](#) [Actions](#) [Projects](#) [Wiki](#) [Security](#) [Insights](#) [Settings](#)

 main ▾ 9th_EDA / 1조 /

[Go to file](#)

[Add file](#) ▾

[...](#)









This branch is up to date with DataScience-Lab-Yonsei/9th_EDA:main.

[Contribute](#) ▾

[Sync fork](#) ▾

 JungGeonwooYS Merge pull request DataScience-Lab-Yonsei#1 from SeungbeomDo/main ... 2 days ago [History](#)

..

 Dataset	Add files via upload	3 days ago
 도출법	assignment tset	3 days ago
 서연우	Create test	3 days ago
 신소연	Create test	3 days ago
 유선재	Create test	3 days ago
 정주영	Create test	3 days ago
 README.md	Update README.md	3 days ago
 신소연.txt	memo	2 days ago

06. 과제 및 Q&A

오늘의 과제

오늘 세션 내용을 바탕으로, 본인 팀 데이터에 맞는 EDA를 어떻게 진행할 지,
자유 양식의 형태로 서술하여 제출해주세요.

슬랙 - 선물

1. 매 주 질문자, 답변자 한 명을 랜덤으로 뽑아, 스타벅스 아메리카노 톨 사이즈로 제공합니다!
2. 다들 활발한 슬랙 활용 바랍니다!



DATA SCIENCE LAB

발표자 조찬형

E-mail: jochanhyeong28@gmail.com