

# 6강. Pandas

# 1. 시각화?

15개의 데이터가 있습니다.

	이름	그룹	소속사	성별	생년월일	키	혈액형	브랜드평판지수
0	지민	방탄소년단	빅히트	남자	1995-10-13	173.6	A	10523260
1	지드래곤	빅뱅	YG	남자	1988-08-18	177.0	A	9916947
2	강다니엘	NaN	커넥트	남자	1996-12-10	180.0	A	8273745
3	뷔	방탄소년단	빅히트	남자	1995-12-30	178.0	AB	8073501
4	화사	마마무	RBW	여자	1995-07-23	162.1	A	7650928
5	정국	방탄소년단	빅히트	남자	1997-09-01	178.0	A	5208335
6	민현	뉴이스트	플레디스	남자	1995-08-09	182.3	O	4989792
7	소연	아이들	큐브	여자	1998-08-26	NaN	B	4668615
8	진	방탄소년단	빅히트	남자	1992-12-04	179.2	O	4570308
9	하성운	핫샷	스타크루이엔티	남자	1994-03-22	167.1	A	4036489
10	태연	소녀시대	SM	여자	1989-03-09	NaN	A	3918661
11	차은우	아스트로	판타지오	남자	1997-03-30	183.0	B	3506027
12	백호	뉴이스트	플레디스	남자	1995-07-21	175.0	AB	3301654
13	JR	뉴이스트	플레디스	남자	1995-06-08	176.0	O	3274137
14	슈가	방탄소년단	빅히트	남자	1993-03-09	174.0	O	2925442



## 1. 시각화?

---

데이터가 한 눈에 보이기 때문에  
인사이트를 얻기 쉽습니다.

# 1. 시각화?

그렇다면...?

307511 row X 122

train

	SK_ID_CURR	TARGET	NAME_CONTRACT_TYPE	CODE_GENDER	FLAG_OWN_CAR	FLAG_OWN_REALTY	CNT_CHILDREN	AMT_INCOME_TOTAL
0	100002	1	Cash loans	M	N	Y	0	202500.0
1	100003	0	Cash loans	F	N	N	0	270000.0
2	100004	0	Revolving loans	M	Y	Y	0	67500.0
3	100006	0	Cash loans	F	N	Y	0	135000.0
4	100007	0	Cash loans	M	N	Y	0	121500.0
...	...	...	...	...	...	...	...	...
307506	456251	0	Cash loans	M	N	N	0	157500.0
307507	456252	0	Cash loans	F	N	Y	0	72000.0
307508	456253	0	Cash loans	F	N	Y	0	153000.0
307509	456254	1	Cash loans	F	N	Y	0	171000.0
307510	456255	0	Cash loans	F	N	N	0	157500.0

307511 rows × 122 columns



## 1. 시각화?

---

시각화는 수 많은 숫자 + 문자형으로 된  
데이터로부터  
사람이 이해하기 쉽도록  
그리고 인사이트를 얻을 수 있는 가장  
직관적인  
시각적 정보를 얻기 위함입니다.



## 2. EDA

---

# EDA

## Exploratory Data Analysis

### 탐색적 데이터 분석

수집한 데이터가 들어왔을 때, 이를 다양한 각도에서 관찰하고 이해하는 과정.

### 3. 시각화의 다양한 예시

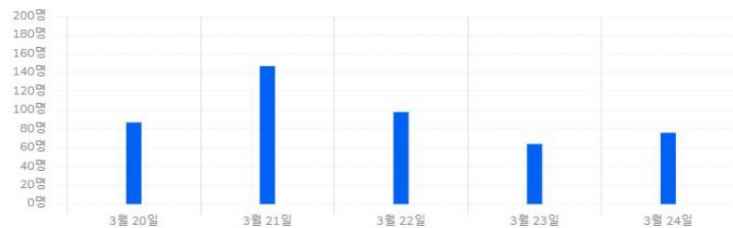


일별시세

날짜	종가	전일비	시가	고가	저가	거래량
2020.03.24	46,950	▲ 4,450	43,850	46,950	43,050	49,489,302
2020.03.23	42,500	▼ 2,900	42,600	43,550	42,400	41,701,626
2020.03.20	45,400	▲ 2,450	44,150	45,500	43,550	49,730,008
2020.03.19	42,950	▼ 2,650	46,400	46,650	42,300	56,925,513
2020.03.18	45,600	▼ 1,700	47,750	48,350	45,600	40,152,623
2020.03.17	47,300	▼ 1,600	46,900	49,650	46,700	51,218,151
2020.03.16	48,900	▼ 1,050	50,100	50,900	48,800	33,339,821
2020.03.13	49,950	▼ 850	47,450	51,600	46,850	59,462,933
2020.03.12	50,800	▼ 1,300	51,000	51,900	49,300	48,149,999
2020.03.11	52,100	▼ 2,500	54,300	54,400	52,000	45,707,281

### 3. 시각화의 다양한 예시

일일 전일대비 확진자증가 그래프

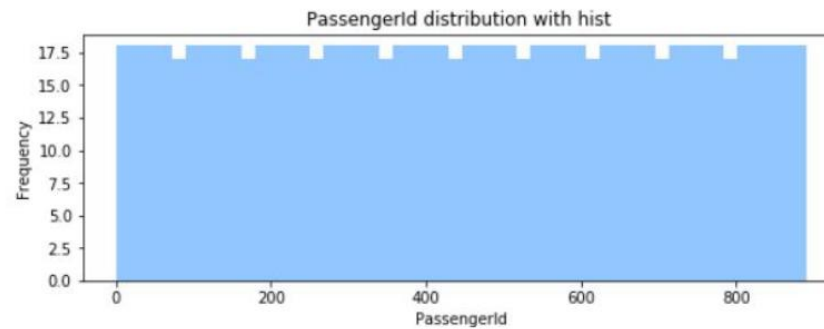
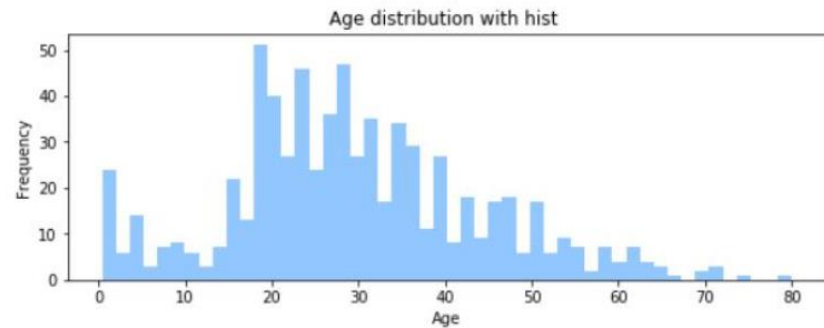
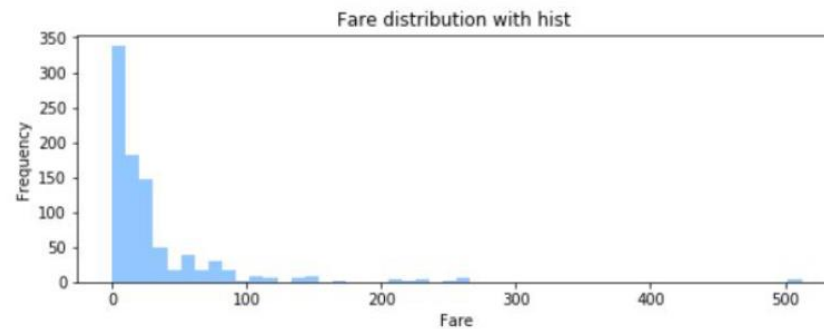


질병관리본부의 오전브리핑을 기준으로 제작되었습니다.

날짜	확진자(전일대비)
3월 24일	9037명(+76)▲
3월 23일	8961명(+64)▼
3월 22일	8799명(+98)▼
3월 21일	8799명(+147)▲
3월 20일	8652명(+87)▼



### 3. 시각화의 다양한 예시



## 4. 우리가 앞으로 활용할 시각화 그래프



## 5. 시각화는

---

### 데이터의 시각화 (Data Visualization)

숫자, 문자로 된 데이터를 시각적으로 변환  
가장 직관적인 수단입니다