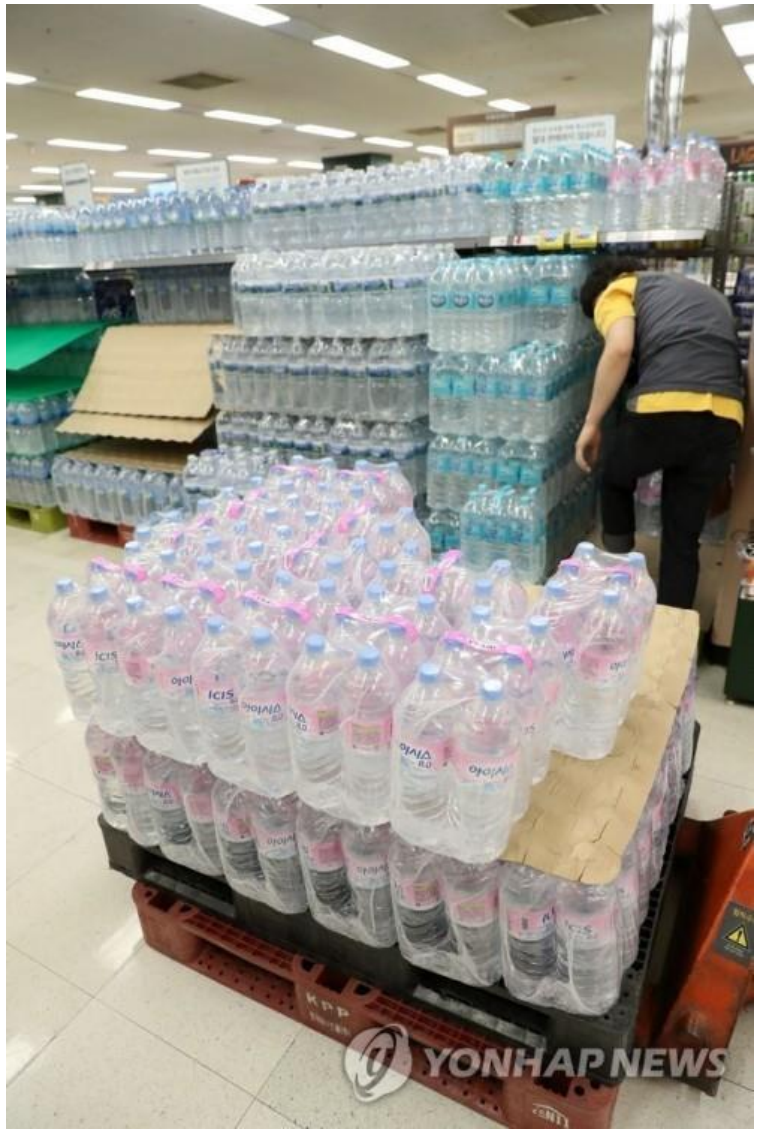
프로젝트 최종 발표

2022254019 한병엽

프로젝트 개요

- 다품종 취급하는 마트
  - 공간제약으로 인해 대량 재고 불가
  - 유통기한이 있기 때문에 적정재고의 순환이 중요
- 대상 제품은 생수
  - 자주 구매하는 상품
  - 중량이 무겁고 부피가 큼
- 과거 판매 내역 기준
  - 2016년 8주차 ~ 2020년 43주차



데이터 전처리

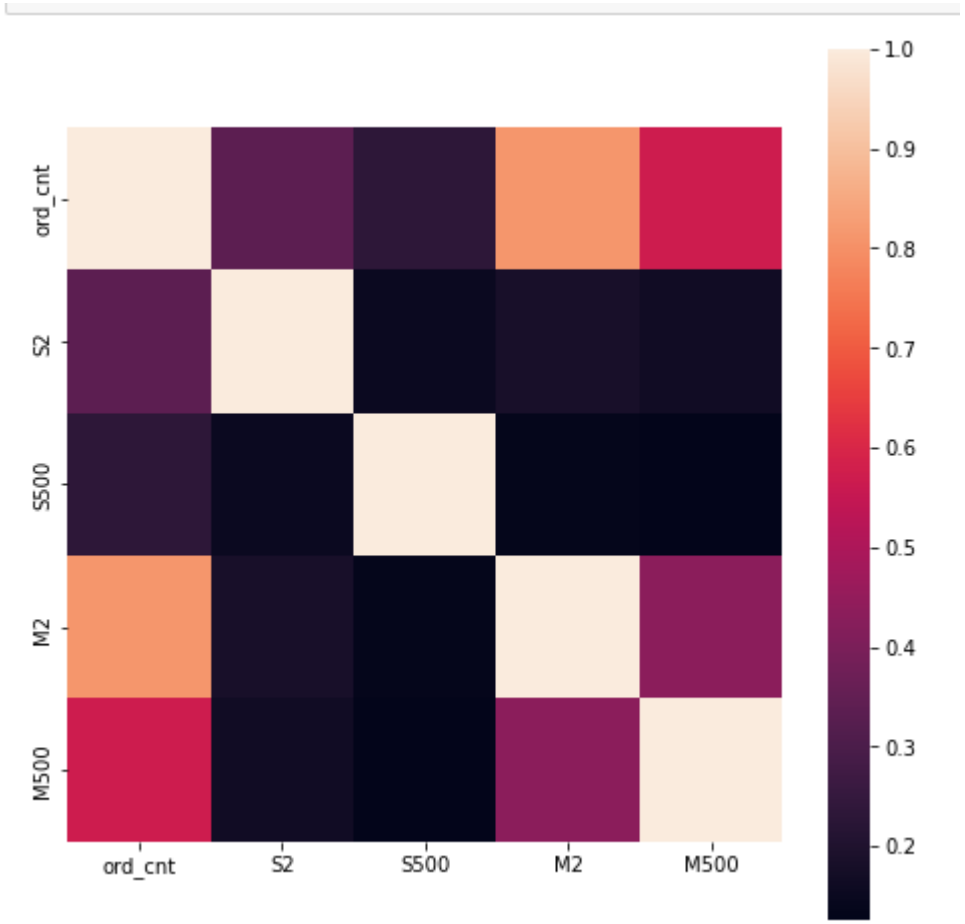
no	ono	pno	prd_type	name	sell_prc	milage	buy_ea	total_prc	total_mila	is_replace	sale_prc	api_send
4853	20160525-6DA67	835		S500	8300	83	1	8300	83	Y	0	Y
3909	20160503-BAF6A	835		S500	8300	83	5	41500	456	Y	0	Y
6304	20160711-B60D0	835	1	S500	9500	95	11	104500	1045	Y	0	Y
6168	20160629-2C5AB	835		S500	9500	95	2	19000	193	Y	0	Y
4286	20160511-D44C5	835		S500	8300	83	1	8300	83	Y	0	Y
75127	20180519-F57B5	835		S500	9500	47	2	19000	94	Y	0	Y
5712	20160613-F7B1C	835		S500	9500	95	3	28500	302	Y	0	Y
6470	20160720-56214	835		S500	9500	95	1	9500	95	Y	0	Y
6609	20160804-018BD	835		S500	9500	95	2	19000	190	Y	0	Y
6631	20160805-2A8F2	835		S500	9500	95	5	47500	493	Y	0	Y
6815	20160810-3BA9F	835		S500	9500	95	5	47500	508	Y	0	Y
6841	20160810-19EBB	835		S500	9500	95	1	9500	95	Y	0	Y
7066	20160816-FB10A	835		S500	9500	95	1	9500	95	Y	0	Y
10110	20161114-A42F2	835		S500	9500	95	1	9500	95	Y	0	Y
10124	20161116-C7E58	835		S500	9500	95	2	19000	190	Y	0	Y
10723	20161125-2073A	835		S500	9500	95	3	28500	300	Y	0	Y
12178	20170104-1B95E	835		S500	9500	95	1	9500	100	Y	0	Y
12582	20170113-63DE0	835		S500	9500	95	1	9500	100	Y	0	Y
12694	20170117-2E601	835		S500	9500	95	2	19000	0	Y	0	Y
12740	20170119-062E8	835	1	S500	9500	95	2	19000	190	Y	0	Y
14453	20170213-A0FDB	835		S500	9500	95	4	38000	341	Y	0	Y
14513	20170214-7D8C0	835		S500	9500	95	3	28500	285	Y	0	Y

ono : 주문번호(연월일 포함)
pno : 상품 고유 번호
name : 상품명
sell_prc : 판매가
milage : 적립금
buy_ea : 판매수량
total_prc : 총 판매금액
total_milage : 총 적립금
is_replace : 대체 여부
Sale_prc : 할인액

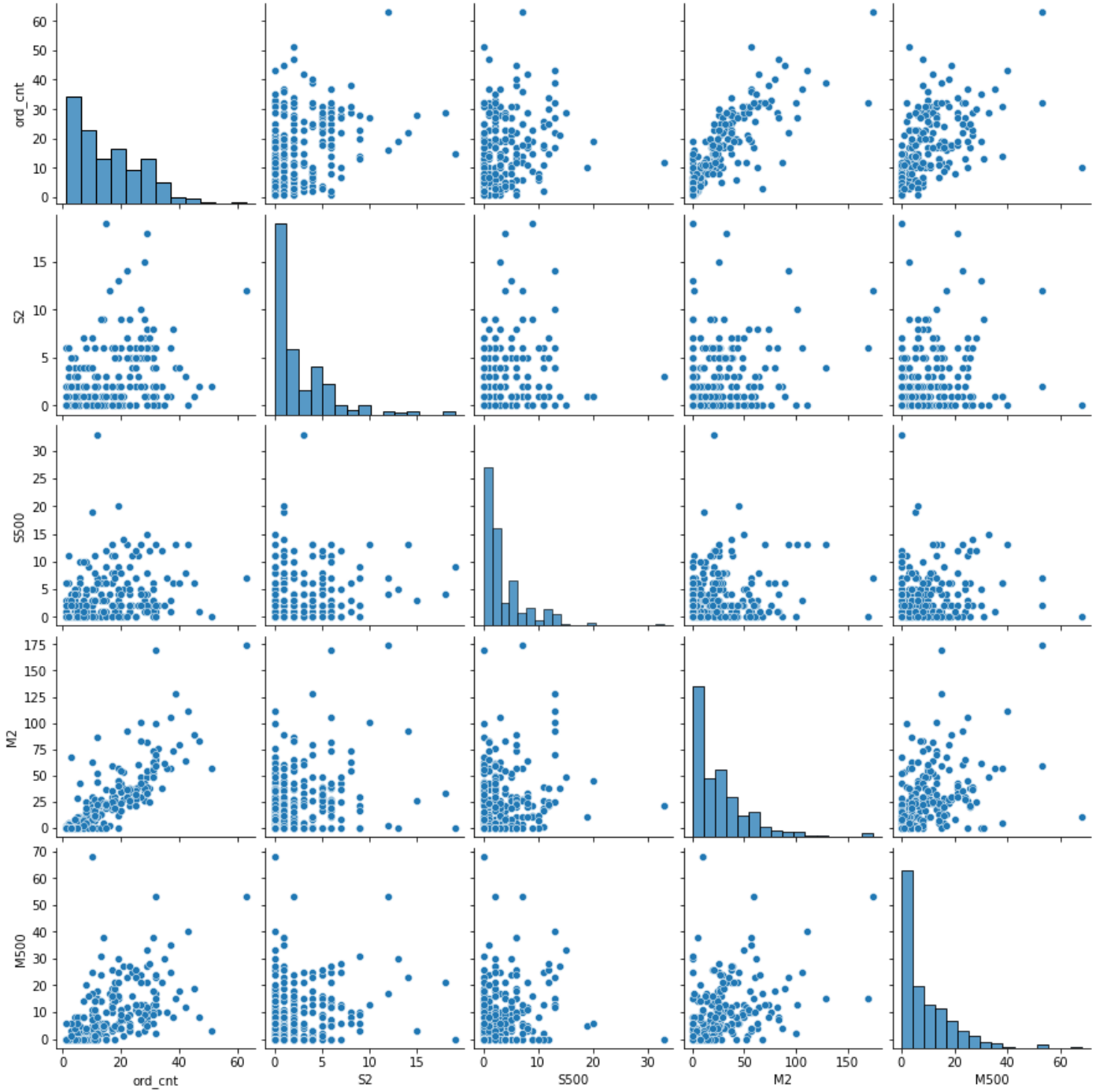
201635	32	0	0	100	2
201636	5	2	0	5	0
201637	3	0	0	1	3
201638	2	2	0	1	0
201639	7	2	0	5	14
201640	12	2	0	24	18
201641	2	0	0	3	0
201642	6	4	0	9	1
201643	1	0	0	1	0
201644	7	2	0	3	4
201645	12	0	0	44	8
201646	11	2	0	10	10
201647	13	0	3	13	6
201648	10	0	3	13	17
201649	7	2	0	7	1
201650	8	4	0	8	2
201651	3	2	0	3	0
201652	4	0	0	4	8
201653	2	3	0	1	0
201701	5	1	1	1	3
201702	5	0	1	4	4
201703	8	4	4	10	0
201704	4	3	0	4	1
201705	8	0	0	14	2
201706	5	1	0	11	2
201707	10	1	19	11	5
201708	11	0	5	22	3
201709	2	0	0	3	0
201710	8	0	0	25	2
201711	10	1	0	23	11
201712	17	1	0	59	15

Heatmap

	ord_cnt	S2	S500	M2	M500
ord_cnt	1.000000	0.336396	0.230829	0.808933	0.567009
S2	0.336396	1.000000	0.150695	0.182670	0.164580
S500	0.230829	0.150695	1.000000	0.135141	0.126263
M2	0.808933	0.182670	0.135141	1.000000	0.431276
M500	0.567009	0.164580	0.126263	0.431276	1.000000



Pairplot


ord_cnt
S2
S500
M2
M500
0
20
40
60
0
5
10
15
0
10
20
30
0
50
100
150
0
20
40
60

학습

```
import numpy as np
np.random.seed(42)
train = np.random.choice(mydata.shape[0], int(0.7 * mydata.shape[0]), replace=False)
validation = [i for i in range(mydata.shape[0]) if i not in train]
```

70% 학습 데이터 / 30% 검증 데이터

```
array([115,  15, 211, 126,   6, 170,   9, 221, 112, 220, 182, 137,  30,
       193, 113,  55,  24, 205,  86,  19, 206, 120, 141, 234,  10, 218,
       172, 109,  75,  25, 124, 185, 235,  18,  68,  60, 148, 204, 114,
        73,  82,  45,  16,  93, 186, 167,  38, 127, 183, 201, 230,  95,
       190,  84, 158,  97, 147, 184,  79, 154, 125, 119,  66, 216,  67,
       209,  69, 101, 232, 173, 108,  29, 117, 224, 150,   5,  56, 227,
       143,  96,  65, 194, 200, 140,  31,  12,  35,  28,  42, 111, 132,
       159,  51, 136, 198,  76,  41, 104, 144,  78, 139,  26, 228, 155,
       176, 164,   0,   2,  77,  46, 100, 146, 142, 199,  90,  85, 165,
       156,  98,  36, 138,  61,  22, 118, 153, 162,  33,  11, 231,  27,
       212,   4, 122,  32, 168,  62, 135, 128, 213, 177,  70, 180,  64,
        44, 152,  40, 123,  23, 175, 171,  81,  39, 197,  47,  94, 178,
        43, 145, 161,   3, 105,  53, 133, 225, 181])
```

평가

```
lin_model = LinearRegression()
my_model = lin_model.fit(X_train, y_train)
print('R2 score is ', my_model.score(X_train, y_train))
print('model coefficients:₩n', my_model.coef_, '₩nintercept: ',my_model.intercept_)
```

```
R2 score is  0.6462484197575528
model coefficients:
 [[ 2.12472304 -0.63033264 -0.60764937 -0.15318489]]
intercept:  [-1.63458559]
```

```
oos = validation
X_val = mydata.iloc[oos, :].drop(columns = ['week', 'M2'])
y_val = mydata.loc[oos, 'M2']
lin_model1 = LinearRegression()
my_model1 = lin_model1.fit(X_val, y_val)
val_pred = my_model1.predict(X_val)
val_RMSE = mean_squared_error(val_pred, y_val, squared = False)
print('the validation RMSE is ', val_RMSE)
```

```
the validation RMSE is  14.199313875011 99
```