

# 지도학습: 회귀분석(실습)

홍익 대학교  
Hyun-Sun Ryu

# 회귀분석(Regression) 실습1

# 보스턴시 주택가격 데이터

<https://www.kaggle.com/code/prasadperera/the-boston-housing-dataset/notebook>

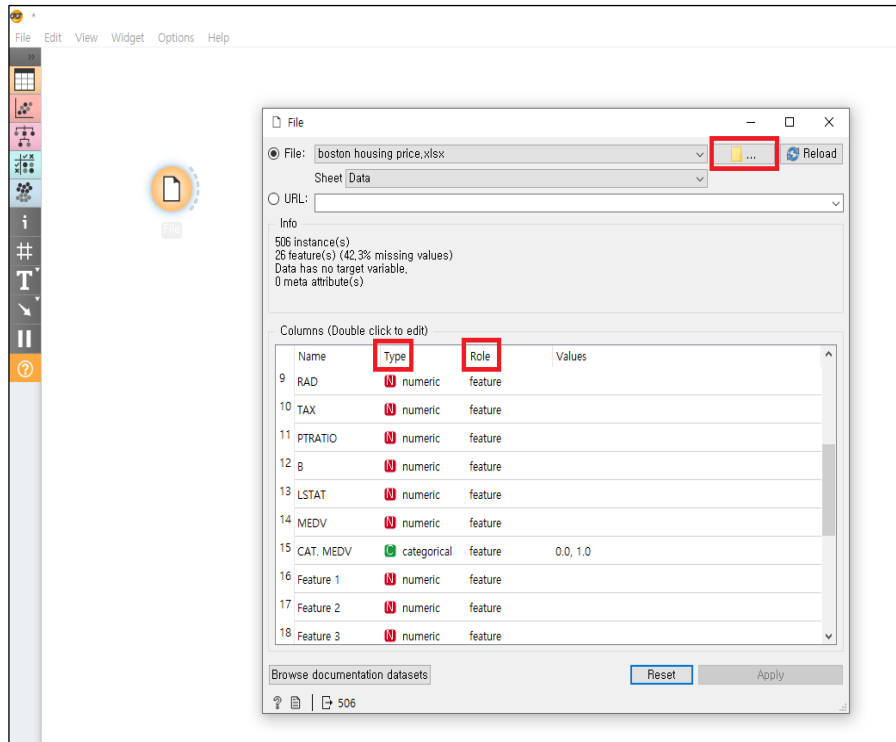


# 보스턴시 주택가격 데이터

[01] CRIM	자치시(town) 별 1인당 범죄율
[02] ZN	25,000 평방피트를 초과하는 거주지역의 비율
[03] INDUS	비소매상업지역이 점유하고 있는 토지의 비율
[04] CHAS	찰스강에 대한 더미변수(강의 경계에 위치한 경우는 1, 아니면 0)
[05] NOX	10ppm 당 농축 일산화질소
[06] RM	주택 1가구당 평균 방의 개수
[07] AGE	1940년 이전에 건축된 소유주택의 비율
[08] DIS	5개의 보스턴 직업센터까지의 접근성 지수
[09] RAD	방사형 도로까지의 접근성 지수
[10] TAX	10,000 달러 당 재산세율
[11] PTRATIO	자치시(town)별 학생/교사 비율
[12] B	$1000(Bk - 0.63)^2$ , 여기서 Bk는 자치시별 흑인의 비율을 말함.
[13] LSTAT	모집단의 하위계층의 비율(%)
[14] CAT.MDEV	MDEV가 \$30,000을 넘는지에 대한 변수 (넘는 경우 1, 아닌 경우 0)
[15] MEDV	본인 소유의 주택가격(중앙값) (단위: \$1,000)

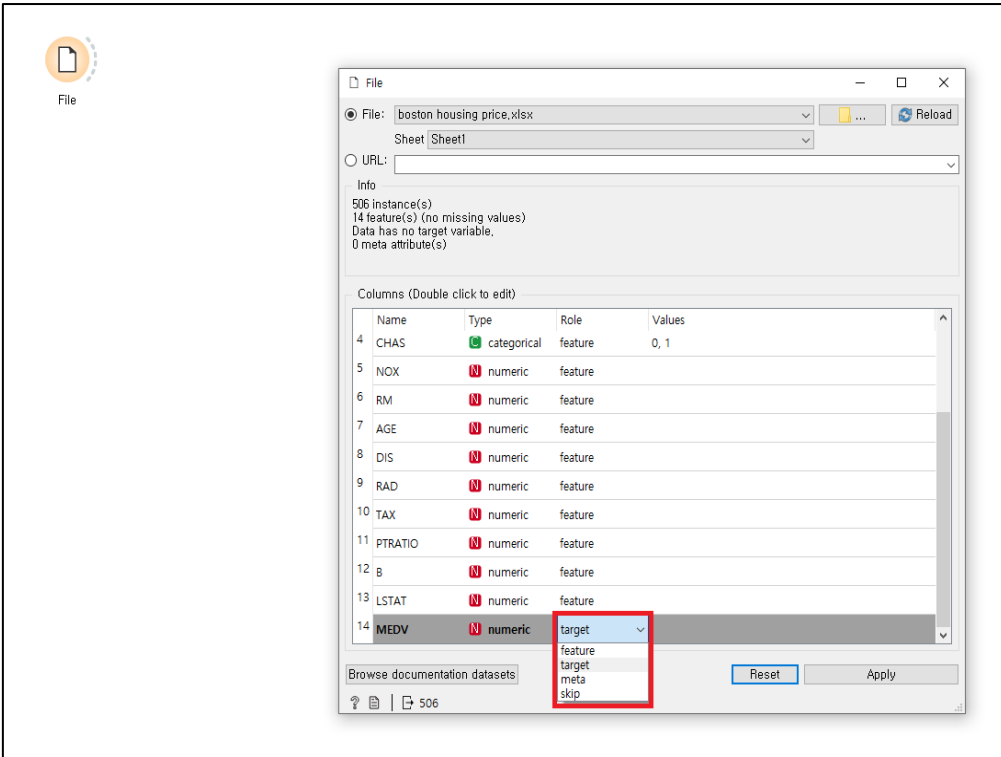
# Housing 데이터 활용

- 엑셀파일로 제공된 Housing 데이터를 활용
- 왼쪽 Data 메뉴에서 Data를 클릭하거나 드래그&드랍 데이터 위젯을 더블클릭하여 Housing 데이터를 업로드



# Housing 데이터 활용

- 데이터의 type과 role을 확인
- 집값(MEDV) 예측이 목표이므로 target 설정
- 나머지는 영향을 주는 요소이므로 feature로 설정됐는지 확인



File

File: boston housing price.xlsx

Sheet: Sheet1

URL:

Info

506 instance(s)  
14 feature(s) (no missing values)  
Data has no target variable.  
0 meta attribute(s)

Columns (Double click to edit)

	Name	Type	Role	Values
4	CHAS	categorical	feature	0, 1
5	NOX	numeric	feature	
6	RM	numeric	feature	
7	AGE	numeric	feature	
8	DIS	numeric	feature	
9	RAD	numeric	feature	
10	TAX	numeric	feature	
11	PTRATIO	numeric	feature	
12	B	numeric	feature	
13	LSTAT	numeric	feature	
14	MEDV	numeric	target	

Browse documentation datasets

506

Reset Apply

# Housing 데이터 활용

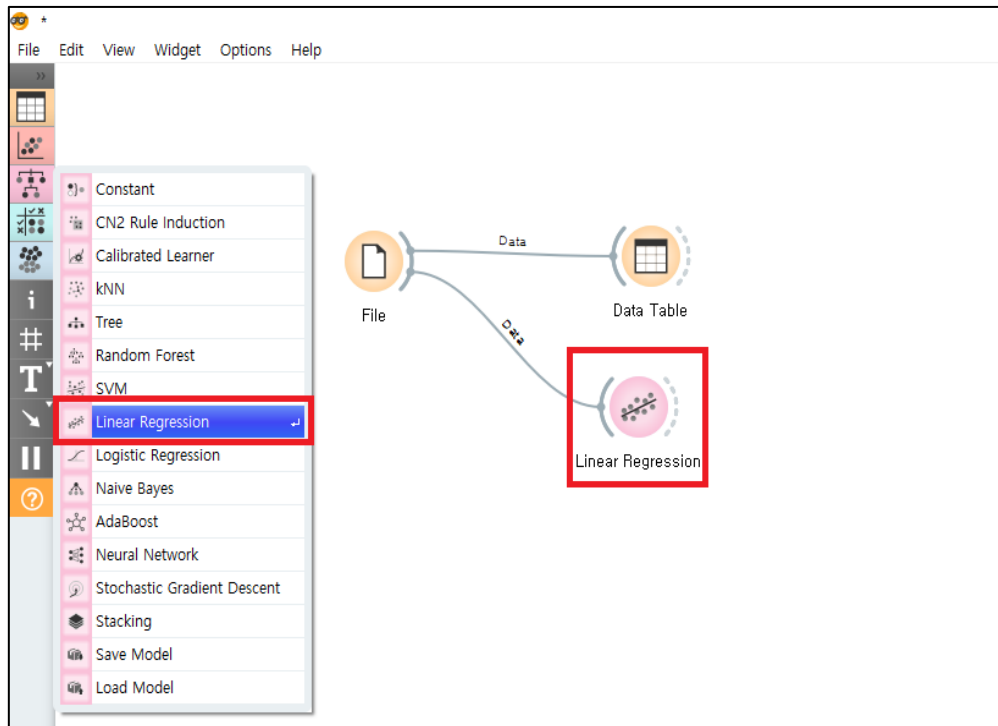
- 왼쪽 Data메뉴에서 Data table위젯을 추가하여 데이터 정렬을 확인
- 종속변수(target)는 회색 음영 처리되고 영향을 주는 독립 변수(feature)는 오른쪽에 위치함

The screenshot displays the Orange3 data mining software interface. On the left, the 'Data' menu is open, and the 'Data Table' widget is highlighted with a red box. In the center, a workflow diagram shows a 'File' widget connected to a 'Data Table' widget, both highlighted with red boxes. On the right, the 'Data Table' widget's configuration window is open, showing a data preview table. The 'MEDV' column is highlighted with a red background, indicating it is the target variable. The table contains 506 instances of housing data.

	MEDV	CRIM	ZN	INDUS	CHAS	NOX
1	24.0	0.00632	18.0	2.31	0	0.5380
2	21.6	0.02731	0.0	7.07	0	0.4690
3	34.7	0.02729	0.0	7.07	0	0.4690
4	33.4	0.03237	0.0	2.18	0	0.4580
5	36.2	0.06905	0.0	2.18	0	0.4580
6	28.7	0.02985	0.0	2.18	0	0.4580
7	22.9	0.08829	12.5	7.87	0	0.5240
8	27.1	0.14455	12.5	7.87	0	0.5240
9	16.5	0.21124	12.5	7.87	0	0.5240
10	18.9	0.17004	12.5	7.87	0	0.5240
11	15.0	0.22489	12.5	7.87	0	0.5240
12	18.9	0.11747	12.5	7.87	0	0.5240
13	21.7	0.09378	12.5	7.87	0	0.5240
14	20.4	0.62976	0.0	8.14	0	0.5380
15	18.2	0.63796	0.0	8.14	0	0.5380
16	19.9	0.62739	0.0	8.14	0	0.5380
17	23.1	1.05393	0.0	8.14	0	0.5380
18	17.5	0.78420	0.0	8.14	0	0.5380
19	20.2	0.80271	0.0	8.14	0	0.5380
20	18.2	0.72580	0.0	8.14	0	0.5380
21	13.6	1.25179	0.0	8.14	0	0.5380
22	19.6	0.85204	0.0	8.14	0	0.5380
23	15.2	1.23247	0.0	8.14	0	0.5380
24	14.5	0.98843	0.0	8.14	0	0.5380
25	15.6	0.75026	0.0	8.14	0	0.5380
26	13.9	0.84054	0.0	8.14	0	0.5380
27	16.6	0.67191	0.0	8.14	0	0.5380
28	14.8	0.95577	0.0	8.14	0	0.5380
29	18.4	0.77299	0.0	8.14	0	0.5380

# Housing 데이터 활용

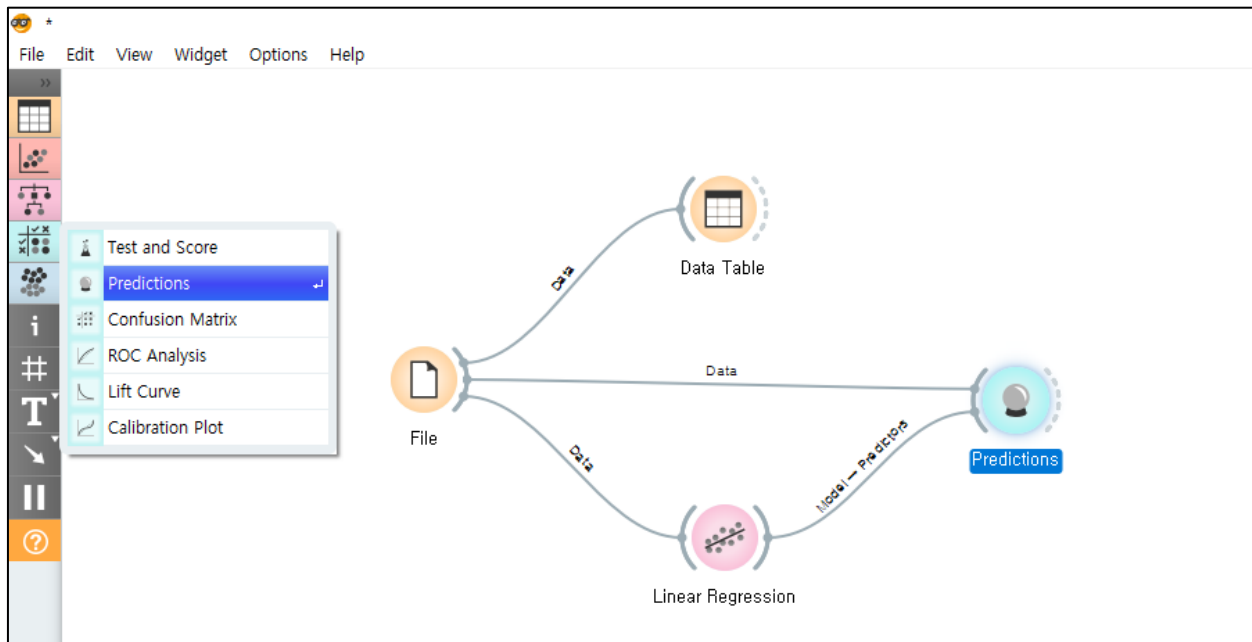
- 선형회귀를 활용한 예측을 위해 왼쪽 Prediction 메뉴에서 Linear Regression 위젯을 클릭하거나 드래그&드랍
- 추가된 선형회귀 위젯을 기본 데이터와 연결하여 훈련





# Housing 데이터 활용

- 예측을 위한 Predictions 위젯을 추가하도록 함. 그리고 선형회귀를 활용해 만든 모델을 연결하도록 함.
- 기존 데이터를 예측하여 예측 값과 실제 결과를 비교함.



# Housing 데이터 활용

- 왼쪽에 Linear Regression 열이 새롭게 생기고 MEDV는 기존에 있던 집값 데이터
- 모델을 통해 만들어진 집값과 실제 데이터의 차이점이 작을수록 선형회귀가 예측한 값이 정확

Predictions

Show probabilities for

	Linear Regression	MEDV	CRIM	ZN	INDUS	C
1	30.0	24.0	0.0632	18.0	2.31	0
2	25.0	21.6	0.0731	0.0	7.07	0
3	30.6	34.7	0.0729	0.0	7.07	0
4	28.6	33.4	0.0237	0.0	2.18	0
5	27.9	36.2	0.0905	0.0	2.18	0
6	25.3	28.7	0.0985	0.0	2.18	0
7	23.0	22.9	0.0829	12.5	7.87	0
8	19.5	27.1	0.1455	12.5	7.87	0
9	11.5	16.5	0.2124	12.5	7.87	0
10	18.9	18.9	0.1004	12.5	7.87	0
11	19.0	15.0	0.2489	12.5	7.87	0
12	21.6	18.9	0.1747	12.5	7.87	0
13	20.9	21.7	0.0378	12.5	7.87	0
14	19.6	20.4	0.6976	0.0	8.14	0
15	19.3	18.2	0.6796	0.0	8.14	0
16	19.3	19.9	0.6739	0.0	8.14	0
17	20.5	23.1	1.0393	0.0	8.14	0
18	16.9	17.5	0.7420	0.0	8.14	0
19	16.2	20.2	0.8271	0.0	8.14	0
20	18.4	18.2	0.7580	0.0	8.14	0
21	12.5	13.6	1.2179	0.0	8.14	0
22	17.7	19.6	0.8204	0.0	8.14	0
23	15.8	15.2	1.2247	0.0	8.14	0
24	13.8	14.5	0.9343	0.0	8.14	0

Model	MSE	RMSE	MAE	R2
Linear Regression	21.895	4.679	3.271	0.741

Restore Original Order

506 506

# 정확도 판별

- 대량의 데이터인 경우 정확도 판별은 위젯창의 **MSE, RMSE, MAE,  $R^2$** 을 통해 확인가능
- **MSE, RMSE, MAE**: 해당 값이 0에 가까울수록 높은 예측
- **$R^2$** : 해당 값이 1에 가까울수록 높은 예측
- MAE(평균절대오차): 예측값에서 실제 값의 차이를 평균 낸 값
- MSE(평균제곱오차): 예측값과 실제값 차이의 제곱을 평균 낸 값
- RMSE(평균제곱근편차): MSE값에 제곱근을 씌운 값
- **$R^2$** : 예측값과 실제값 평균의 차를 제곱한 것들의 합을 실제값과 실제값 평균의 차를 제곱한 것으로 나눈 값

# MSE(평균제곱오차)

- MSE(Mean Squared Error) : 예측 값과 실제 값의 차이의 제곱을 평균 낸 것
- 회귀식  $H(x)=Wx+b$ 라고 가정할 때 식으로 나타내면 다음과 같음.

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{Y}_i - Y_i)^2$$

# MSE(평균제곱오차)

- 주어진  $n$ 은 학습 데이터 세트 크기이며 실제 값과 예측 값의 차이를 제공한 평균이므로 그 크기가 작을수록, 즉 **0에 가까울수록 예측을 잘했다고 평가**
- 이 과정은 통계에  $(\text{편차})^2$ 의 평균인 분산을 구하는 과정과 비슷

# RMSE(평균 제곱근 편차)

- RMSE(Root Mean Square Deviation): MSE에 R 즉 Root가 붙은 값으로 MSE에 제곱근을 씌운 값
- **0에 가까울수록 예측을 잘했으며 의미** 있는 회귀식을 구했다고 평가

$$\sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{Y}_i - Y_i)^2}$$

# MAE(평균절대오차)

- MAE(Mean Absolute Error) :예측 값과 실제 값의 차이에 절댓값을 평균 낸 것
- MAE와 MSE는 예측 값과 실제 값의 차이를 이용하고 이를 평균을 냈다는 점에서 비슷한 양상
- MAE를 구하는 식:

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - x|$$

# 정확도 판별하기(MSE? MAE?)

- MSE라는 값이 있음에도 MAE값을 사용함에 의문이 들 수 있음.
- MAE는 실제 값과 예측 값의 차이를 절대값을 두어 그 거리를 계산하는데 MSE는 그 거리를 제곱하여 차이가 있음.
- MSE는 거리가 많이 떨어지면 떨어질수록 제곱을 통해 더 많은 페널티를 부과하여 더 많이 떨어진 것처럼 보이게 만듦
- MSE의 단점은 대부분의 데이터와 다른 특성을 보이거나 특정 속성의 값이 다른 개체들과 달리 유별난 값을 가지는 이상치(Outlier)가 존재할 때 지나치게 많은 페널티가 부여→ 이 경우 전 처리(preprocess)과정을 통해 보정이 되기도 함.
- MSE의 한계점의 이유로 MAE값을 함께 비교



# $R^2$ (알 스퀘어)

- $R^2$  (R Squared): 제곱의 형태를 가지며 결정계수(coefficient of determination)  
**모델이 얼마나 데이터를 잘 설명했는지를 의미**
- $\bar{y}$ 는  $y$ 의 평균값이며 실제 데이터에 상대적으로 얼마나 가까이 있는지 비율을 계산한 값.  
따라서 아주 근접하다면 분자, 분모가 같게되어 1에 가까워지게 됨

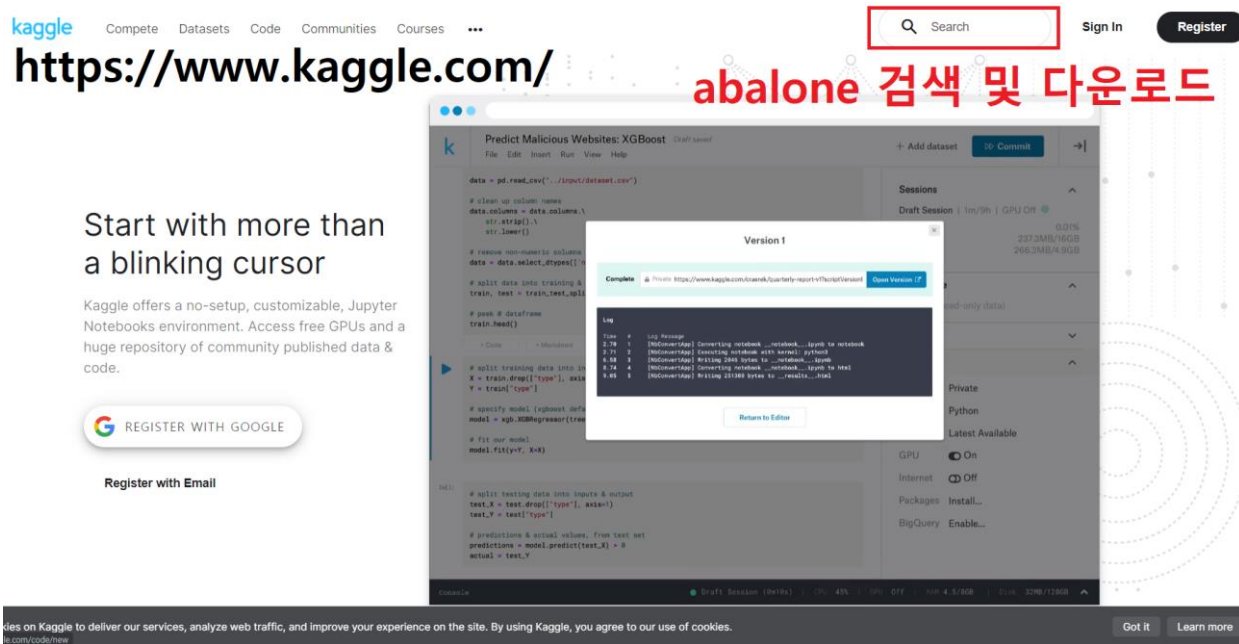
## 회귀분석(Regression) 실습2

# 전복나이 예측 모델

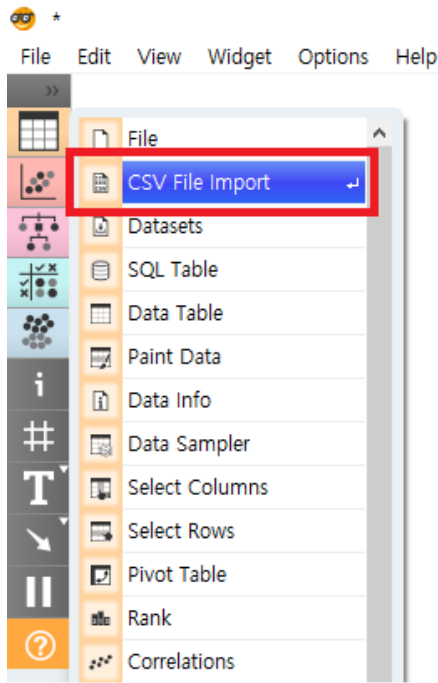
- Abalone 데이터 세트의 출처는 1994년 UCI ML Repository
- Abalone 데이터 세트는 물리적 측정에서 전복 나이를 예측하는데 활용
- 전복의 나이는 껍데기를 원뿔 모양으로 잘라 얼룩을 낸 뒤 현미경을 통해 고리의 수를 세는 방식으로 결정
- 나이는 1년 단위로 고리의 수가 1.5개 증가
- 데이터 세트에는 성별, 길이, 지름 등이 포함돼 있음.
- 이번 시간에는 오렌지3에서 제공하는 기본데이터 세트가 아닌 kaggle을 이용

## 전복나이 예측 모델

- <https://www.kaggle.com/> 에 접속하여 검색창에 abalone(전복)을 검색한 뒤 데이터셋을 다운받습니다. 많은 데이터가 공개돼 있으므로 가입하시는 것을 추천합니다



# 전복나이 예측 모델



- 엑셀의 기본 확장자인 xls 가 아닌 csv 확장자를 확인
- Csv는 각 항목이나 내용마다 쉼표로 구분하여 기록한 파일
- 데이터를 불러오기 위해 data메뉴의 csv file import위젯을 가져옴

# 전복나이 예측 모델

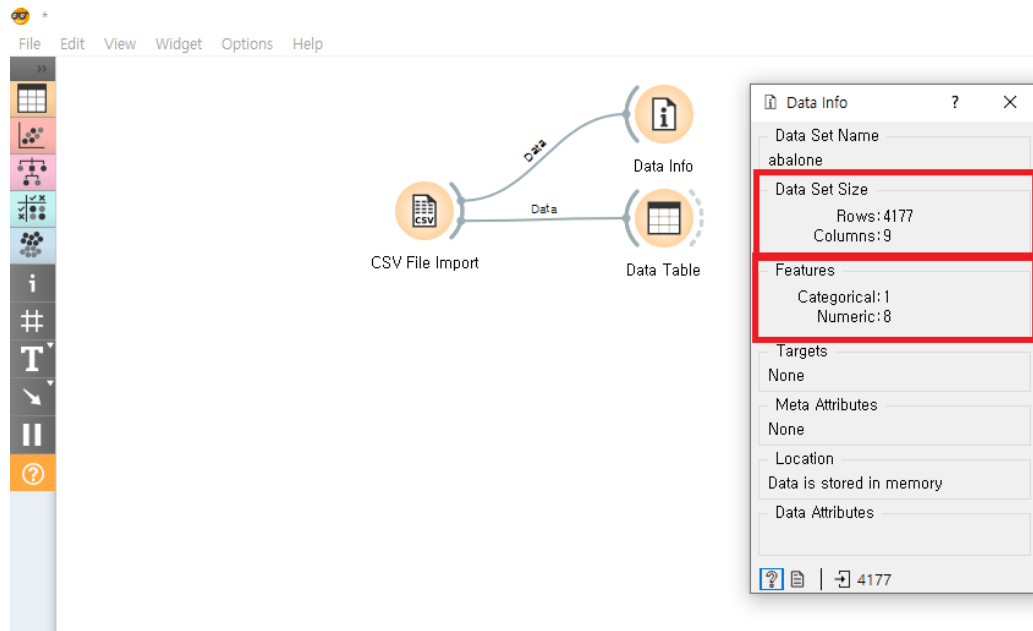
- Csv file import 위젯을 더블 클릭하여 데이터셋을 불러옴
- 각 행은 개별 전복의 instance이며 각 열은 해당 instance의 feature를 나타냄.

The screenshot displays the Orange3 data mining software interface. On the left, a vertical toolbar contains various widgets. The 'CSV File Import' widget is highlighted with a red box. A double-click action has opened the 'Import Options' dialog box on the right. This dialog allows for configuring the data import process. The 'Encoding' is set to 'Unicode (UTF-8)', the 'Cell delimiter' is 'Comma', and the 'Quote character' is an empty string. The 'Number separators' are set to 'Grouping' and 'Decimal'. The 'Column type' section shows a table with 8 columns, each corresponding to a feature in the dataset. The first row of the table is highlighted with a red box.

	1	2	3	4	5	6	7	8
1	Sex	Length	Diameter	Height	Whole weight	Shucked weight	Viscera weight	Shell weight
2	M	0.455	0.365	0.095	0.514	0.2245	0.101	0.15
3	M	0.35	0.265	0.09	0.2255	0.0995	0.0485	0.07
4	F	0.53	0.42	0.135	0.677	0.2565	0.1415	0.21
5	M	0.44	0.365	0.125	0.516	0.2155	0.114	0.155
6	I	0.33	0.255	0.08	0.205	0.0895	0.0395	0.055
7	I	0.425	0.3	0.095	0.3515	0.141	0.0775	0.12
8	F	0.53	0.415	0.15	0.7775	0.237	0.1415	0.33
9	F	0.545	0.425	0.125	0.768	0.294	0.1495	0.26
10	M	0.475	0.37	0.125	0.5095	0.2165	0.1125	0.165
11	F	0.55	0.44	0.15	0.8945	0.3145	0.151	0.32
12	F	0.525	0.38	0.14	0.6065	0.194	0.1475	0.21
13	M	0.43	0.35	0.11	0.406	0.1675	0.081	0.135
14	M	0.49	0.38	0.135	0.5415	0.2175	0.095	0.19
15	F	0.535	0.405	0.145	0.6845	0.2725	0.171	0.205
16	F	0.47	0.355	0.1	0.4755	0.1675	0.0805	0.185
17	M	0.5	0.4	0.13	0.6645	0.258	0.133	0.24
18	I	0.355	0.28	0.085	0.2905	0.095	0.0395	0.115

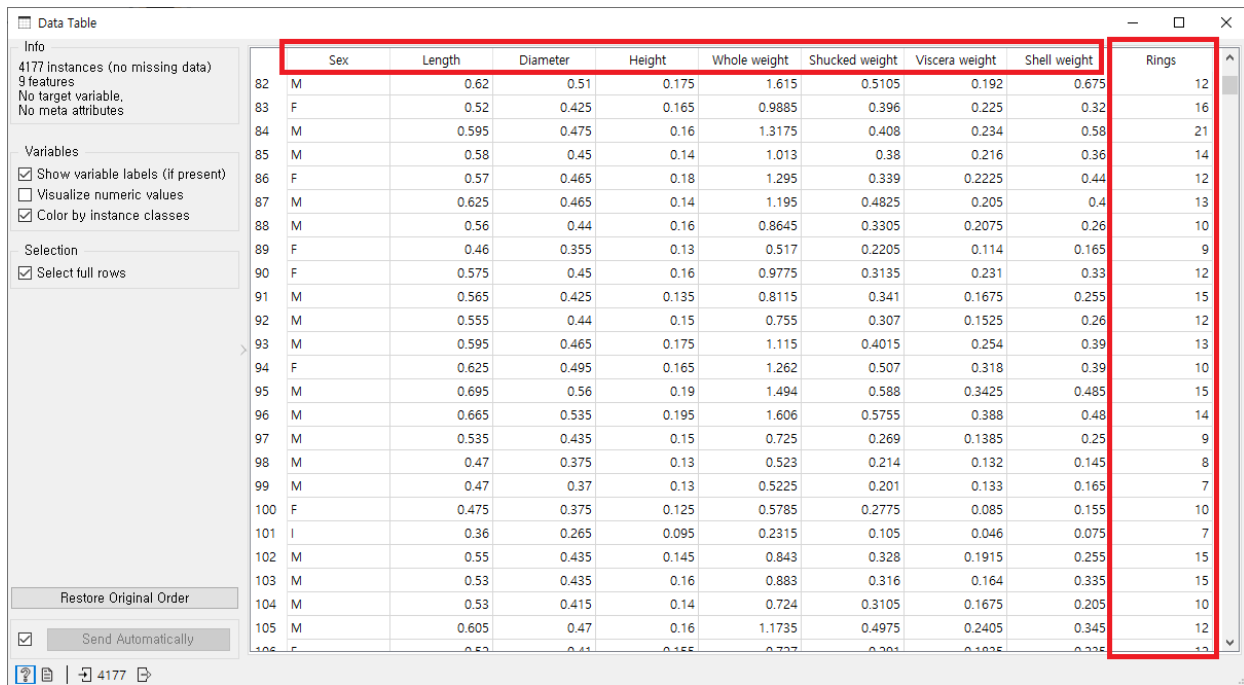
# 전복나이 예측 모델

- 먼저 데이터의 정보를 알아보기 위해 data table 위젯과 data info 위젯을 연결
- 4177개의 instance가 있으며 9개의 열이 있습니다. 9개의 열은 범주형 1개와 숫자로 된 8개의 열로 구성



# 전복나이 예측 모델

- 각 열에 대한 정보를 파악하기 위해 데이터 테이블 위젯을 더블클릭
- 전복의 특징에 관한 feature 중 우리가 파악하고자 하는 target열은 마지막에 있는 Ring임



Data Table

Info  
4177 instances (no missing data)  
9 features  
No target variable.  
No meta attributes

Variables  
☒ Show variable labels (if present)  
☐ Visualize numeric values  
☒ Color by instance classes

Selection  
☒ Select full rows

Restore Original Order

☒ Send Automatically

	Sex	Length	Diameter	Height	Whole weight	Shucked weight	Viscera weight	Shell weight	Rings
82	M	0.62	0.51	0.175	1.615	0.5105	0.192	0.675	12
83	F	0.52	0.425	0.165	0.9885	0.396	0.225	0.32	16
84	M	0.595	0.475	0.16	1.3175	0.408	0.234	0.58	21
85	M	0.58	0.45	0.14	1.013	0.38	0.216	0.36	14
86	F	0.57	0.465	0.18	1.295	0.339	0.2225	0.44	12
87	M	0.625	0.465	0.14	1.195	0.4825	0.205	0.4	13
88	M	0.56	0.44	0.16	0.8645	0.3305	0.2075	0.26	10
89	F	0.46	0.355	0.13	0.517	0.2205	0.114	0.165	9
90	F	0.575	0.45	0.16	0.9775	0.3135	0.231	0.33	12
91	M	0.565	0.425	0.135	0.8115	0.341	0.1675	0.255	15
92	M	0.555	0.44	0.15	0.755	0.307	0.1525	0.26	12
93	M	0.595	0.465	0.175	1.115	0.4015	0.254	0.39	13
94	F	0.625	0.495	0.165	1.262	0.507	0.318	0.39	10
95	M	0.695	0.56	0.19	1.494	0.588	0.3425	0.485	15
96	M	0.665	0.535	0.195	1.606	0.5755	0.388	0.48	14
97	M	0.535	0.435	0.15	0.725	0.269	0.1385	0.25	9
98	M	0.47	0.375	0.13	0.523	0.214	0.132	0.145	8
99	M	0.47	0.37	0.13	0.5225	0.201	0.133	0.165	7
100	F	0.475	0.375	0.125	0.5785	0.2775	0.085	0.155	10
101	I	0.36	0.265	0.095	0.2315	0.105	0.046	0.075	7
102	M	0.55	0.435	0.145	0.843	0.328	0.1915	0.255	15
103	M	0.53	0.435	0.16	0.883	0.316	0.164	0.335	15
104	M	0.53	0.415	0.14	0.724	0.3105	0.1675	0.205	10
105	M	0.605	0.47	0.16	1.1735	0.4975	0.2405	0.345	12
106	F	0.53	0.41	0.155	0.737	0.305	0.1635	0.235	13




# 전복나이 예측 모델

- 각 변수는 개별 instance의 특징이며 내용은 다음과 같음.

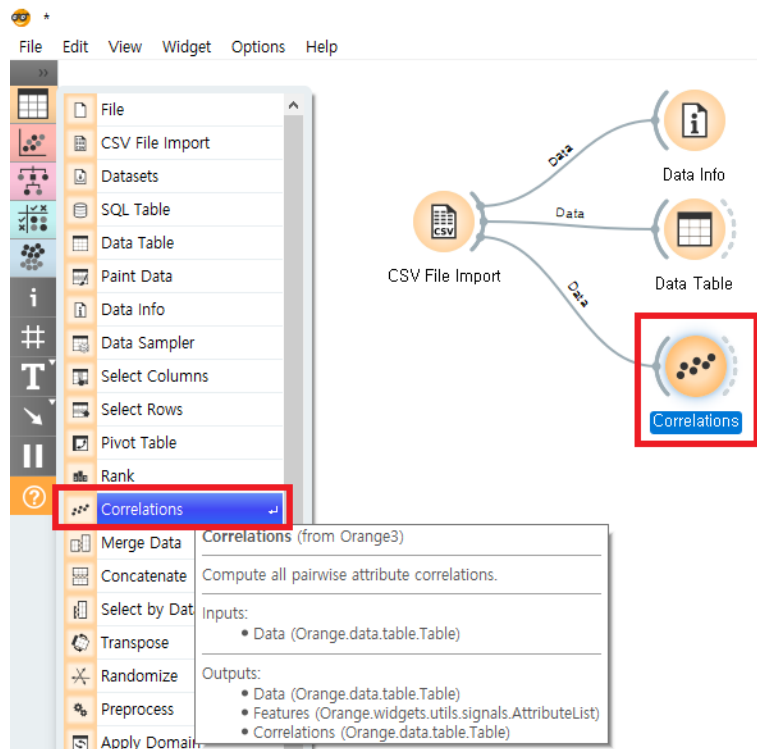
Gender	전복의 성별 Male,Female,Infant(신생아) / 범주형 타입
Length	전복의 길이 / 연속형 타입
Diameter	전복의 지름 / 연속형 타입
Height	전복의 높이 / 연속형 타입
Whole_weight	전복의 전체 무게 / 연속형 타입
Shucked_weight	전복의 무게 / 연속형 타입
Viscera_weight	전복의 내장 무게 / 연속형
Shell_weight	전복 껍데기의 무게 / 연속형
Rings(Y)	전복의 나이 / 정수형

# 전복나이 예측 모델

- Correlations 위젯은 해당 데이터 세트의 모든 features 쌍에 대한 Pearson 또는 Spearman 상관관계 점수를 계산. 이러한 방법은 단조로운 관계만 탐지

위젯	설명	입력	출력
 Correlations	모든 속성 쌍별 상관 점수를 계산한다.	Data	Data, Features, Correlations

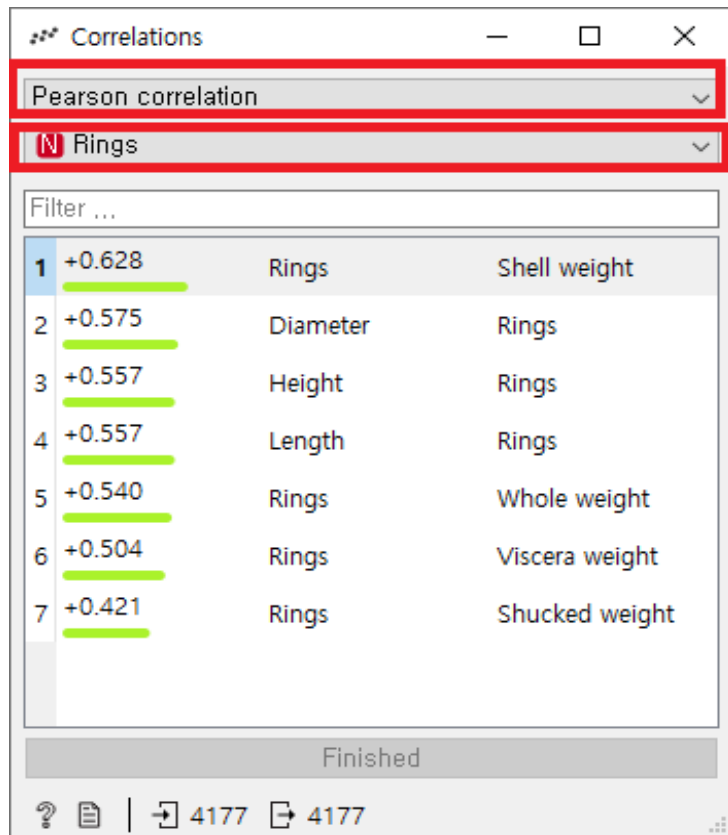
# 전복나이 예측 모델



- 두 연속형 변수 간의 상관계수를 파악하고 선형 관계 분석을 하려면 **Correlations** 위젯을 활용
- 대표적인 상관계수로는 피어슨 상관계수, 스피어만 상관계수 등

# 전복나이 예측 모델

- Correlations 위젯을 더블 클릭하여 상관계수를 확인
- 피어슨 상관계수는 두 변수 간의 선형선이 얼마나 강한지를 나타내는 지표
- 나머지 두 상관계수는 순위 상관계수로써 두 변수간의 강도를 측정하는 지표
- 파악하고 싶은 변수를 변경할 수 있음.
- Rings은 다른 모든 변수들과 양의 상관관계에 있으며 **0.42~0.63**의 범위를 가지고 있음.



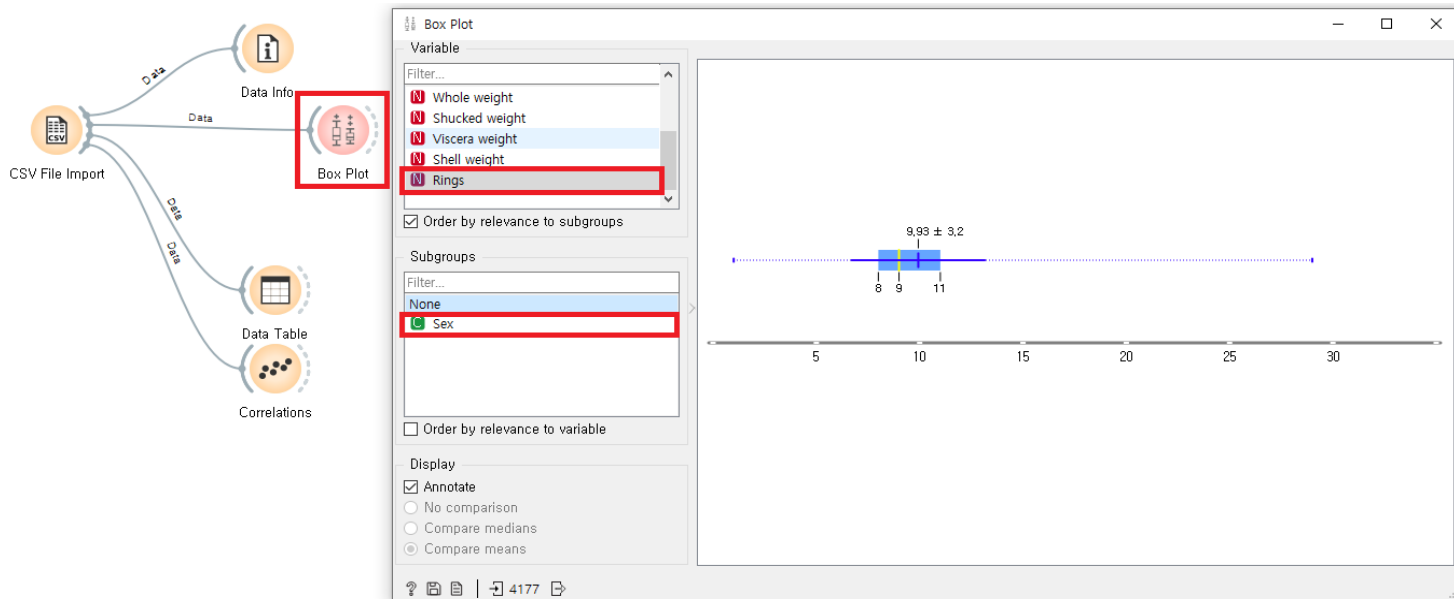
# 상관관계 크기

---

Perfect	+1	-1
Strong	+0.9	-0.9
	+0.8	-0.8
	+0.7	-0.7
Moderate	+0.6	-0.6
	+0.5	-0.5
	+0.4	-0.4
Weak	+0.3	-0.3
	+0.2	-0.2
	+0.1	-0.1
Zero	0	

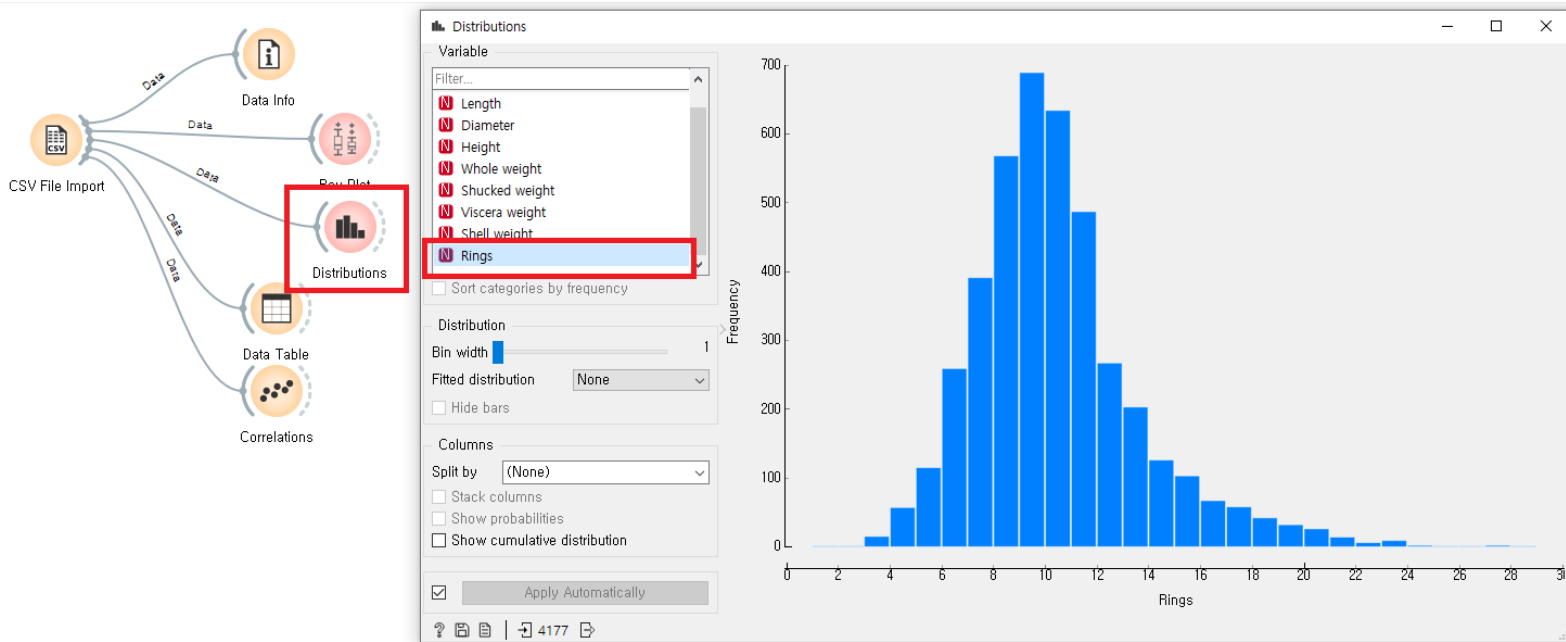
# 전복나이 예측 모델

- EDA를 위해 Box plot 위젯을 활용
- Target이 되는 Ring을 보면 현재 평균은 9.93, 중위수는 9, 제1사분위수는 8, 제3사분위수는 11임
- 성별로 나누어 보려면 Subgroups의 'sex'를 활성화



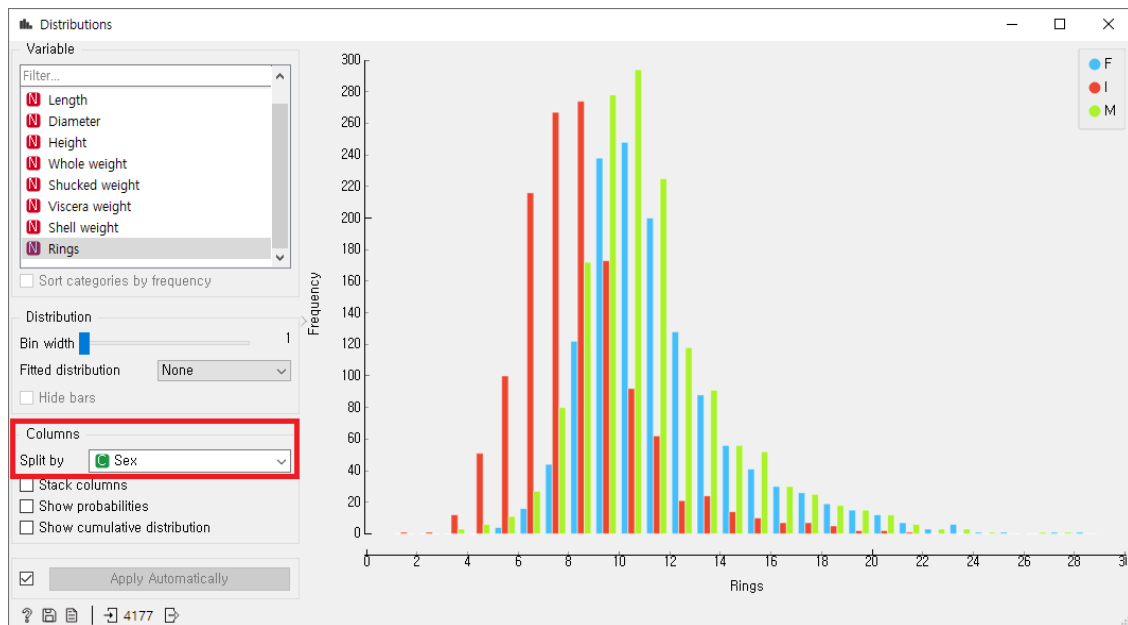
# 전복나이 예측 모델

- 데이터 분포를 시각화하기 위해 Distributions 위젯을 활용
- Distributions 위젯을 활용하여 원하는 variables을 선택하여 그 분포를 시각적으로 파악



# 전복나이 예측 모델

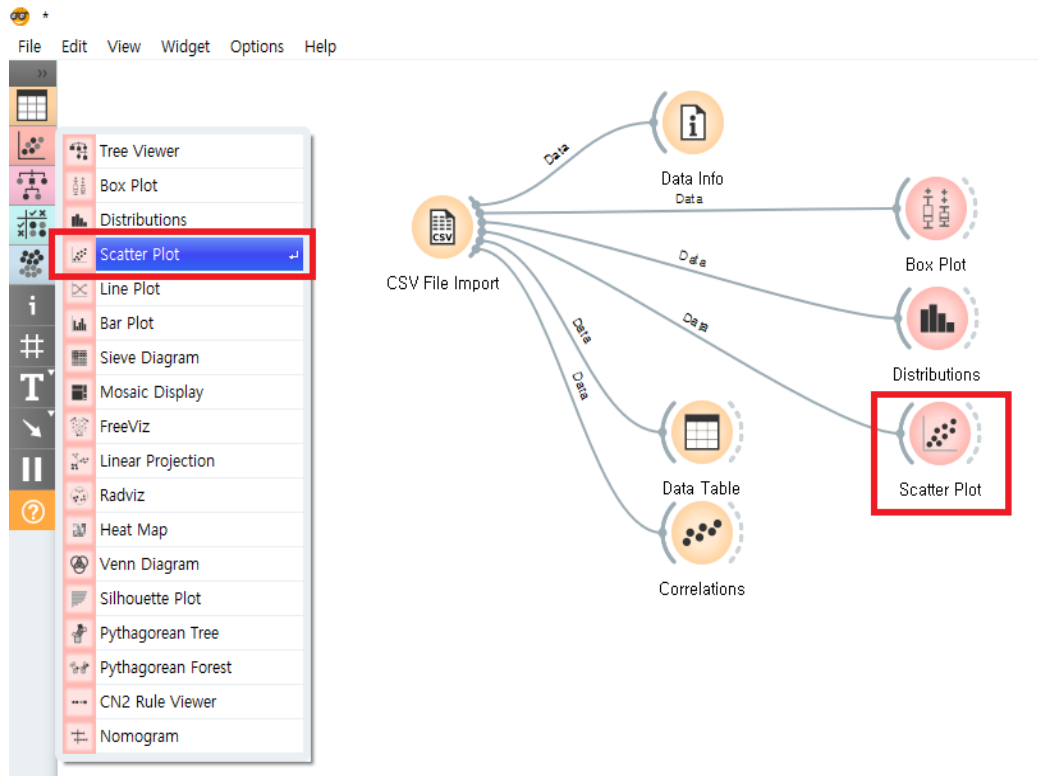
- Box plot위젯과 마찬가지로 성별을 나누어 보려면 Split by에서 성별을 활성화. 아래와 같이 전복 수컷(Male), 암컷(Female), 신생아(Infant)로 나뉘음을 확인





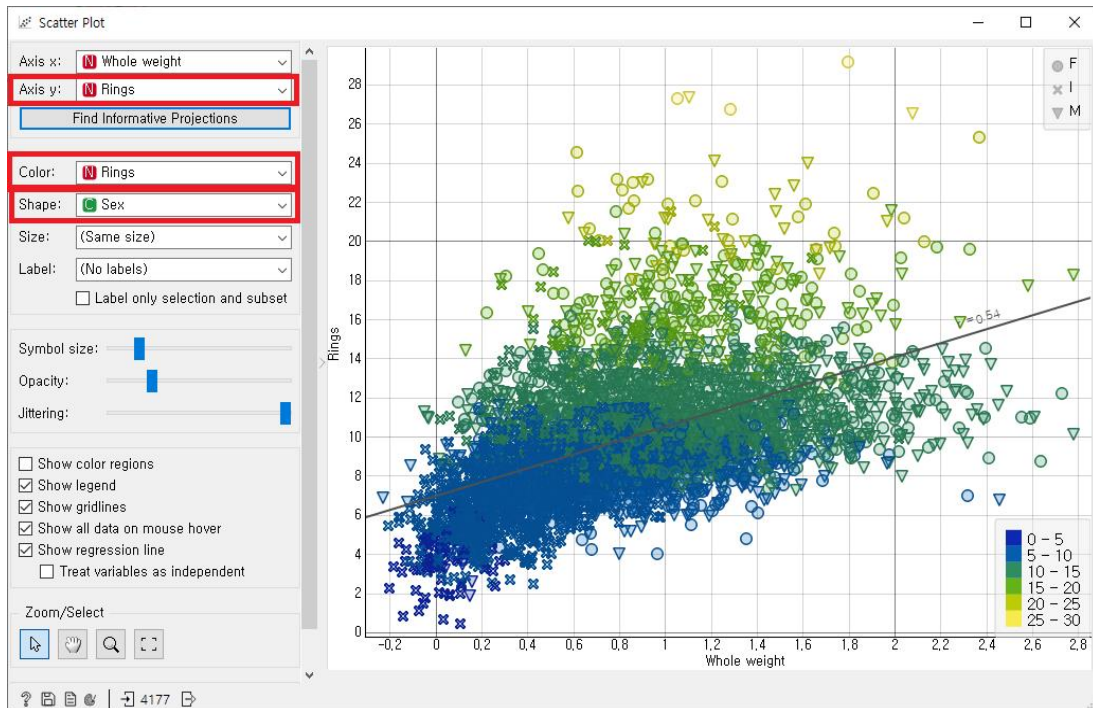
# 전복나이 예측 모델

- 각 instance의 분포를 파악하기 위해 산점도를 활용.
- Scatter plot 위젯을 추가하여 연결.

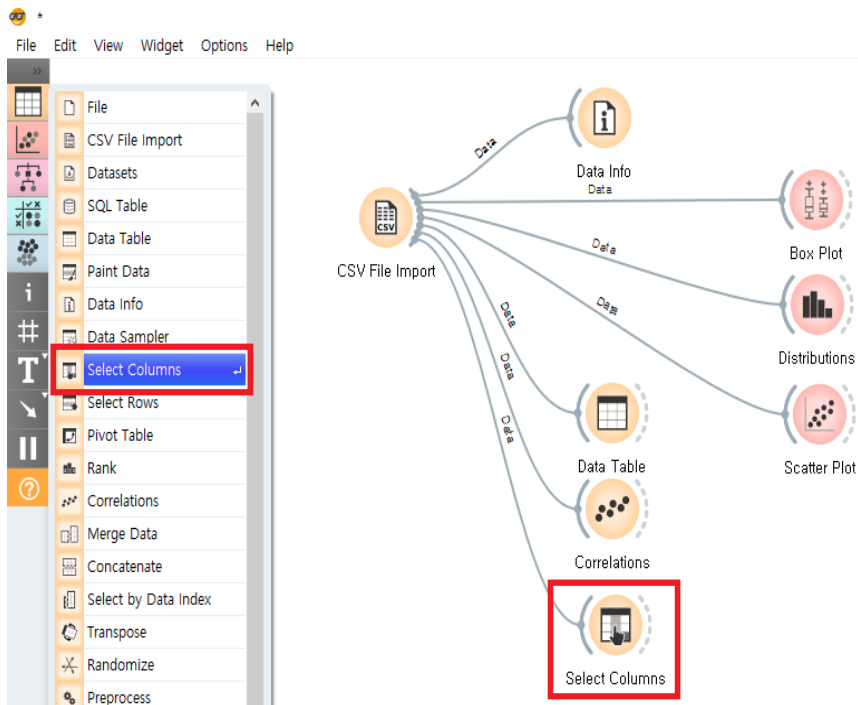


# 전복나이 예측 모델

- Y축을 Rings로 설정하여 산포도를 확인
- 5개의 ring별로 구간을 나눠 알아보기 쉽게끔 색깔을 표시
- 성별에 따라 모양을 나누어 한눈에 파악



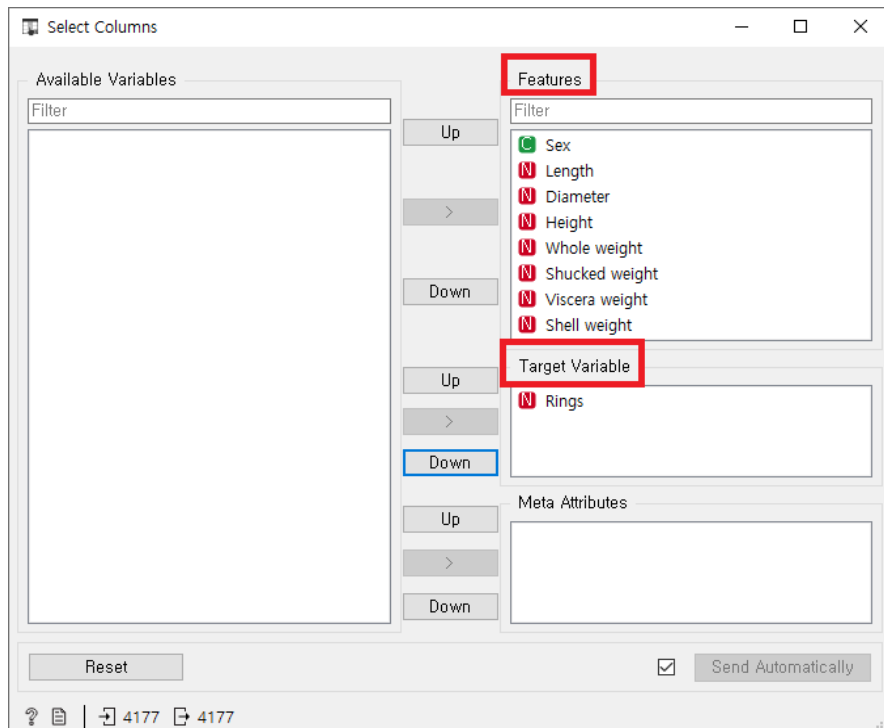
# 전복나이 예측 모델



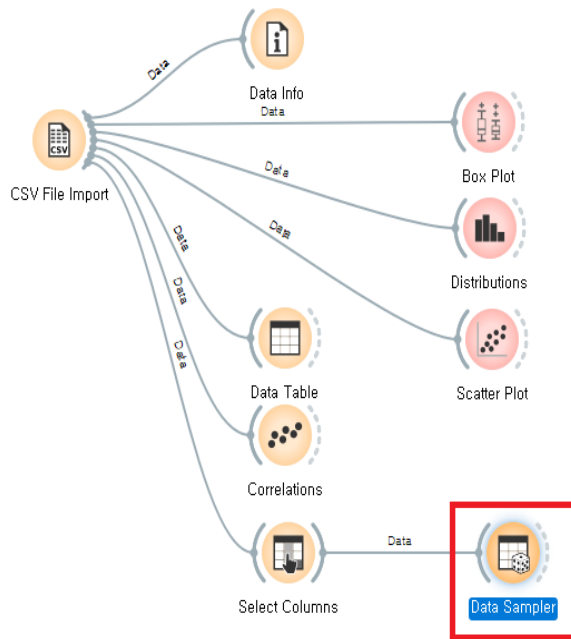
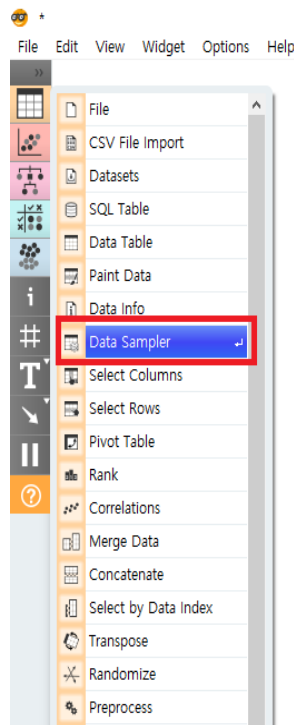
- 각 feature의 독립변수와 종속변수를 구분하기 위해 **Select Columns** 위젯을 추가
- 우리가 예측하고자 하는 Rings 변수를 **Target**으로 지정

# 전복나이 예측 모델

- Rings열을 Target variable로 지정
- 나머지 열에 대해서는 모두 Features로 지정



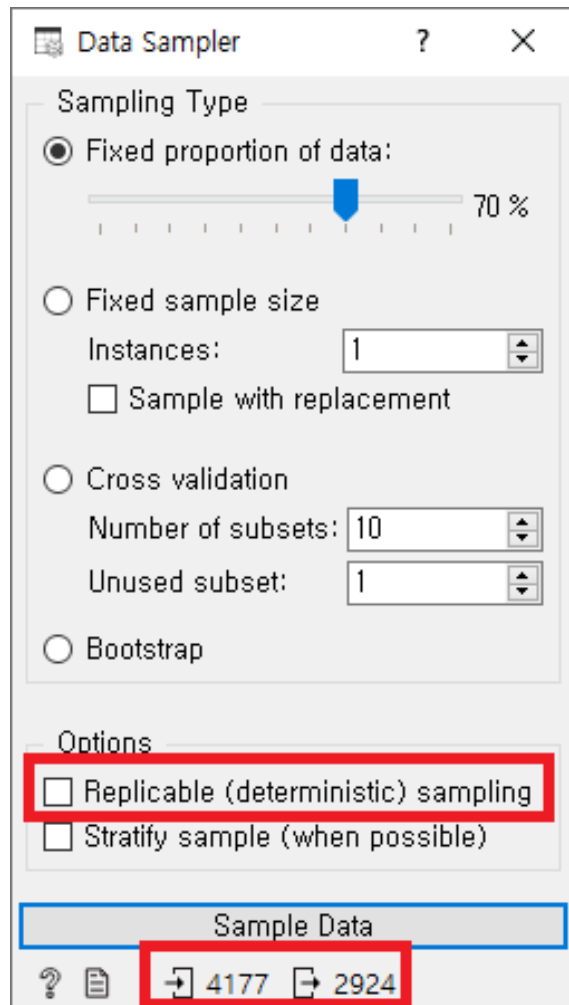
# 전복나이 예측 모델



- 훈련데이터와 예측데이터를 분리하기 위해 Data Sampler 위젯을 활용
- Data 메뉴에서 Data sampler 위젯을 클릭하거나 드래그&드랍

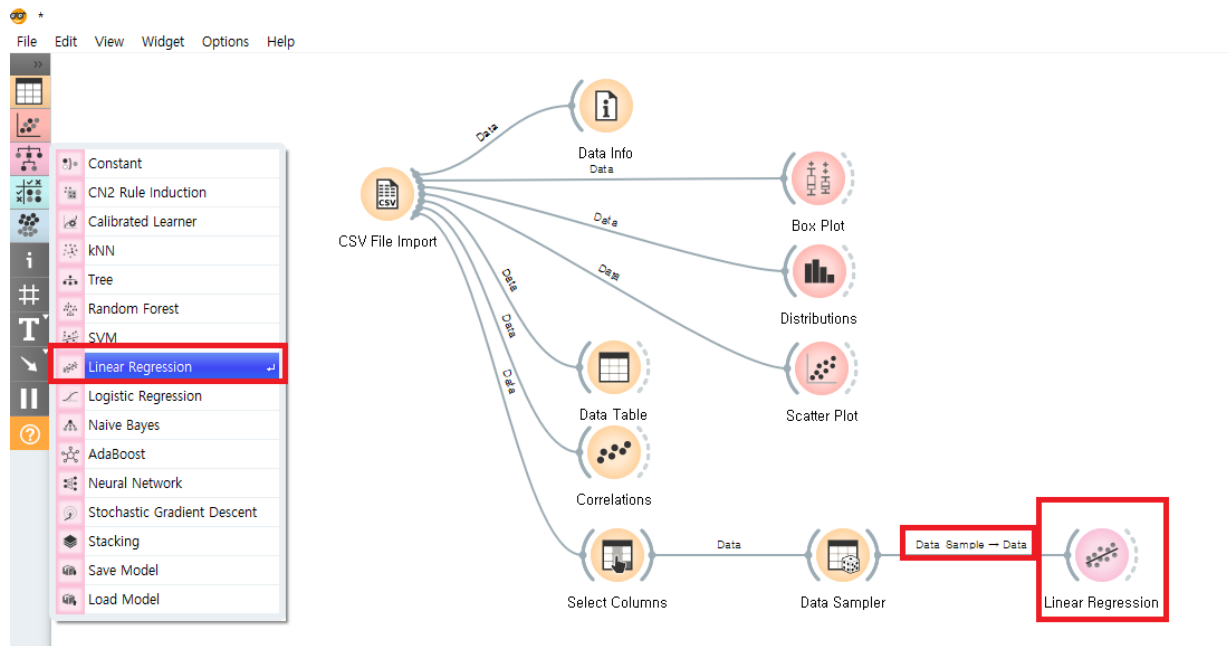
# 전복나이 예측 모델

- Sampling Type은 훈련데이터의 비율을 정하는 것임.  
현재 **70%로 설정**돼 있으며 이에 따라 하단에 4177개의 70%인 2924개가 출력됨을 확인
- Fixed sample size는 비율이 아닌 실제 개수를 입력하는 곳. 이는 4177개를 넘어설 수 없음. 단 Sample with replacement를 활성화하면 같은 것으로 대체할 수 있기 때문에 4177개를 넘길 수 있음
- 우선 70%의 훈련데이터를 설정하고 options에 replicable sampling을 해제



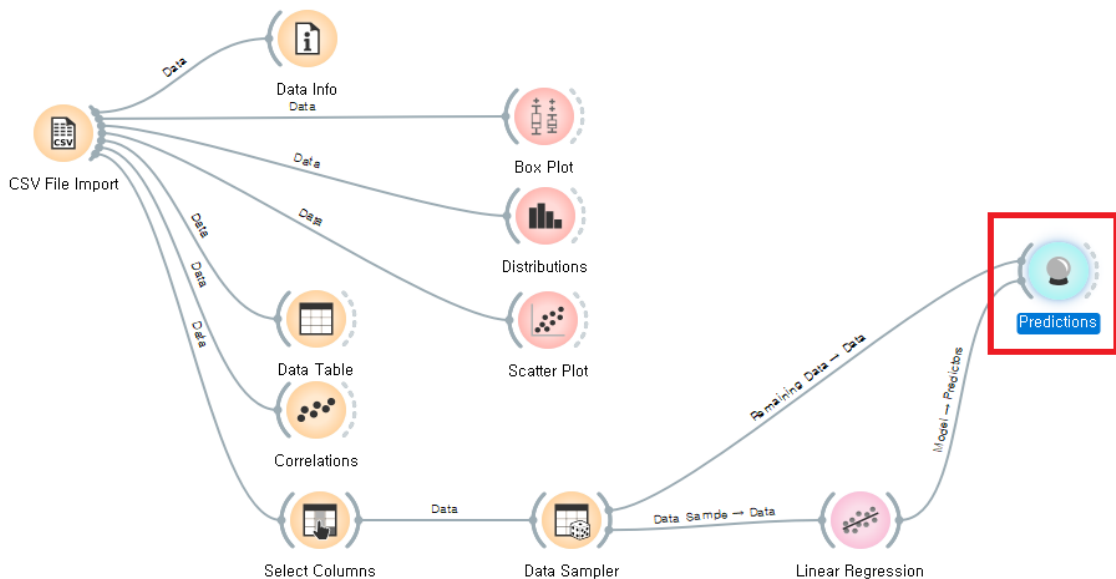
# 전복나이 예측 모델

- 선형회귀를 활용하기 위해 Linear Regression 위젯을 추가
- Data Sample 70%와 Linear Regression을 연결함으로써 모델을 생성



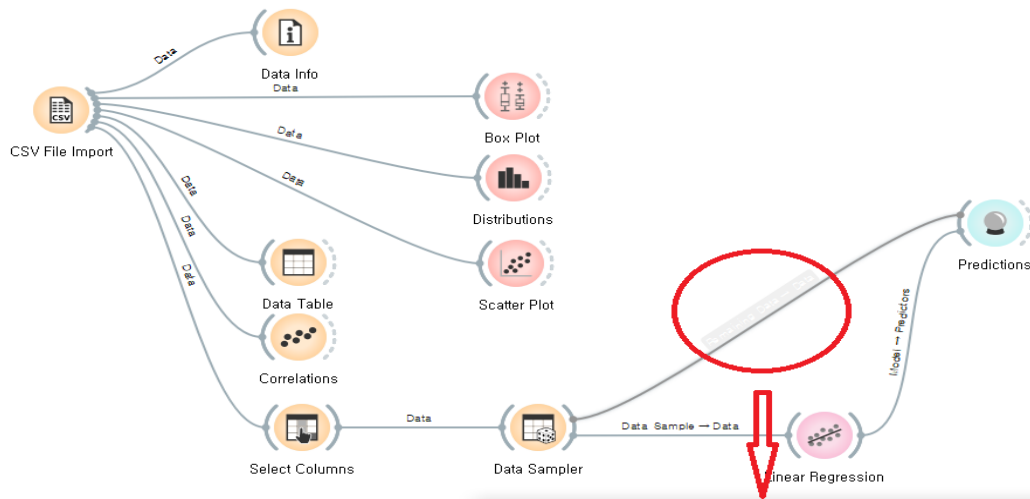
# 전복나이 예측 모델

- 예측을 위해 Predictions 위젯을 추가
- 우리는 **70%데이터를 활용**하여 선형회귀를 통해 모델 만듬
- 나머지 **30% 데이터를 테스트 데이터**로 활용하여 예측

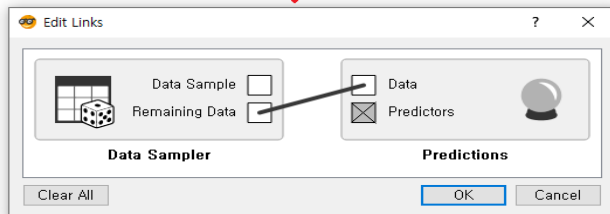




# 전복나이 예측 모델



- Data Sampler와 Predictions과의 연결이 remaining data to data가 돼야 함. 설정이 올바르게 됐는지 확인



# 전복나이 예측 모델

- Predictions 위젯을 활성화하여 30%에 해당하는 1253개의 데이터를 가지고 예측한 결과를 확인
- 추가적으로 다양한 모델을 연결하여 그 예측 값을 비교

Predictions

Show probabilities for

	Linear Regression	Rings	Sex	Length	Diameter	Height	Whole weight	Shucked weight	Viscera weight	Shell weight
172	9	15	M	0.455	0.365	0.095	0.514	0.2245	0.101	0.15
173	13	10	M	0.62	0.525	0.155	1.085	0.454	0.1965	0.35
174	13	12	F	0.585	0.475	0.185	0.8575	0.3465	0.1785	0.275
175	8	7	I	0.435	0.33	0.11	0.38	0.1515	0.0945	0.11
176	12	13	M	0.58	0.46	0.13	0.921	0.357	0.181	0.29
177	12	14	F	0.62	0.51	0.175	1.1255	0.4985	0.227	0.315
178	11	12	M	0.625	0.5	0.18	1.3705	0.645	0.303	0.3705
179	16	14	F	0.63	0.475	0.175	1.423	0.4155	0.3385	0.49
180	7	6	I	0.36	0.27	0.09	0.219	0.097	0.0405	0.065
181	13	11	F	0.725	0.57	0.205	1.6195	0.744	0.315	0.488
182	10	11	I	0.62	0.47	0.155	0.966	0.447	0.171	0.284
183	9	8	F	0.61	0.43	0.14	0.909	0.438	0.2	0.22
184	9	8	F	0.48	0.38	0.12	0.608	0.2705	0.1405	0.185
185	8	7	I	0.42	0.32	0.11	0.3625	0.174	0.0635	0.105
186	10	13	M	0.72	0.58	0.19	2.0885	0.9955	0.478	0.5305
187	10	9	M	0.58	0.47	0.165	1.041	0.54	0.166	0.279
188	29	8	F	0.455	0.355	1.13	0.594	0.332	0.116	0.1335
189	7	13	I	0.32	0.245	0.08	0.1585	0.0635	0.0325	0.05
190	11	11	M	0.53	0.41	0.155	0.7155	0.2805	0.1685	0.214
191	8	8	I	0.51	0.39	0.125	0.597	0.293	0.1265	0.1555
192	14	17	I	0.61	0.46	0.17	1.278	0.41	0.257	0.37
193	9	9	I	0.515	0.39	0.12	0.565	0.235	0.135	0.179
194	11	15	I	0.62	0.485	0.17	1.208	0.4805	0.3045	0.33
195	12	13	M	0.62	0.48	0.16	1.0765	0.412	0.253	0.3

Model MSE RMSE MAE R2

Linear Regression 5.394 2.322 1.616 0.500

Restore Original Order

1253

# 질문 있나요?

hsryu13@hongik.ac.kr

