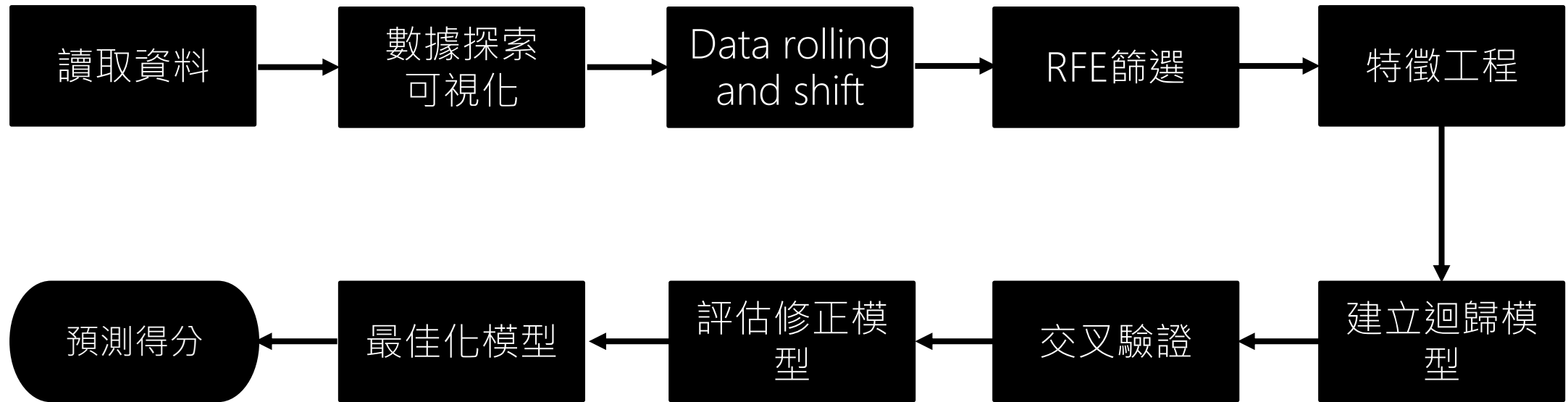


建立迴歸模型以預測賽爾提克季後賽得分

曾煥文 吳榮峻

流程圖



讀取資料

	SEASON_ID	TEAM_ID	TEAM_ABBREVIATION	TEAM_NAME	GAME_ID	GAME_DATE	MATCHUP	WL	MIN	PTS	...
0	42019	1610612738	BOS	Boston Celtics	0041900306	2020-09-27	BOS @ MIA	L	240	113	...
1	42019	1610612738	BOS	Boston Celtics	0041900305	2020-09-25	BOS vs. MIA	W	238	121	...
2	42019	1610612738	BOS	Boston Celtics	0041900304	2020-09-23	BOS @ MIA	L	239	109	...
3	42019	1610612738	BOS	Boston Celtics	0041900303	2020-09-19	BOS @ MIA	W	240	117	...
4	42019	1610612738	BOS	Boston Celtics	0041900302	2020-09-17	BOS vs. MIA	L	240	101	...
...
3490	21983	1610612738	BOS	Boston Celtics	0028300052	1983-11-05	BOS @ WAS	W	240	120	...
3491	21983	1610612738	BOS	Boston Celtics	0028300043	1983-11-04	BOS vs. IND	W	240	121	...
3492	21983	1610612738	BOS	Boston Celtics	0028300033	1983-11-02	BOS vs. MIL	W	240	119	...
3493	21983	1610612738	BOS	Boston Celtics	0028300019	1983-10-29	BOS @ CLE	W	240	108	...
3494	21983	1610612738	BOS	Boston Celtics	0028300006	1983-10-28	BOS @ DET	L	240	121	...

3495 rows x 28 columns

- 透過nba_api，讀取賽爾提克歷年的資料
- 去掉遺失值
- 定義新欄位(年度、主客場、對手)
- 將數據集分為訓練(Train)和測試(Test)

讀取資料

■ Train的資料

	SEASON_ID	TEAM_ID	TEAM_ABBREVIATION	TEAM_NAME	GAME_ID	GAME_DATE	MATCHUP	WL	MIN	PTS	...	REB	AST	STL	BLK	TOV	PF	PLUS_MINUS	year	Home	Player
17	22019	1610612738	BOS	Boston Celtics	0021901308	2020-08-13	BOS vs. WAS	L	240	90	...	51	22	8	7	18	19	-6.0	2019	True	WAS
18	22019	1610612738	BOS	Boston Celtics	0021901301	2020-08-11	BOS @ MEM	W	242	122	...	46	29	7	7	12	27	15.0	2019	False	MEM
19	22019	1610612738	BOS	Boston Celtics	0021901288	2020-08-09	BOS vs. ORL	W	266	122	...	54	32	9	7	12	19	3.0	2019	True	ORL
20	22019	1610612738	BOS	Boston Celtics	0021901279	2020-08-07	BOS @ TOR	W	240	122	...	52	27	8	2	14	22	22.0	2019	False	TOR
21	22019	1610612738	BOS	Boston Celtics	0021901267	2020-08-05	BOS vs. BKN	W	242	149	...	43	29	10	10	14	23	34.0	2019	True	BKN
...
2285	21996	1610612738	BOS	Boston Celtics	0029600066	1996-11-09	BOS @ CHI	L	238	92	...	46	14	9	2	26	24	-14.0	1996	False	CHI
2286	21996	1610612738	BOS	Boston Celtics	0029600053	1996-11-08	BOS vs. PHI	L	240	105	...	44	22	10	5	20	34	-10.0	1996	True	PHI
2287	21996	1610612738	BOS	Boston Celtics	0029600041	1996-11-06	BOS vs. IND	W	240	94	...	33	24	11	6	15	19	10.0	1996	True	IND
2288	21996	1610612738	BOS	Boston Celtics	0029600020	1996-11-02	BOS @ MIL	L	241	102	...	41	16	7	4	17	30	-22.0	1996	False	MIL
2289	21996	1610612738	BOS	Boston Celtics	0029600001	1996-11-01	BOS vs. CHI	L	240	98	...	36	20	10	2	18	33	-9.0	1996	True	CHI

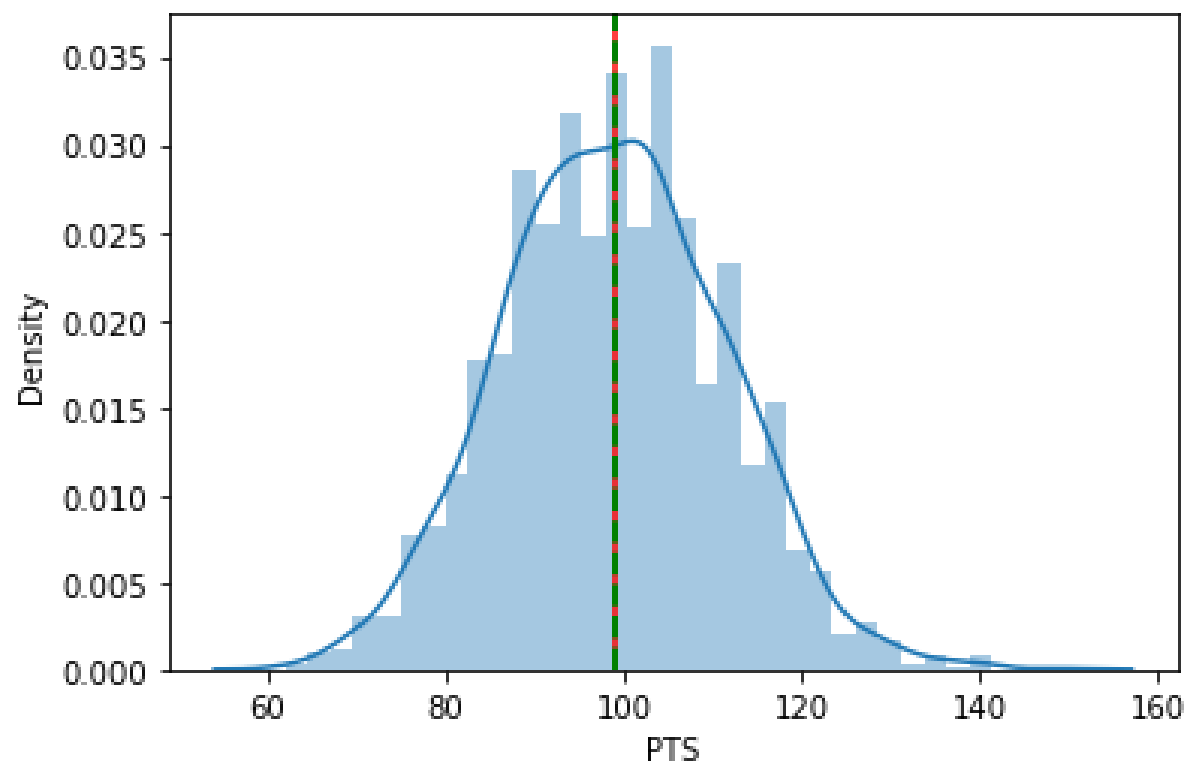
2246 rows × 31 columns

讀取資料

■ Test的資料

	SEASON_ID	TEAM_ID	TEAM_ABBREVIATION	TEAM_NAME	GAME_ID	GAME_DATE	MATCHUP	WL	MIN	PTS	...	REB	AST	STL	BLK	TOV	PF	PLUS_MINUS	year	Home	Player
0	42019	1610612738	BOS	Boston Celtics	0041900306	2020-09-27	BOS @ MIA	L	240	113	...	42	26	7	2	10	21	-12.0	2019	False	MIA
1	42019	1610612738	BOS	Boston Celtics	0041900305	2020-09-25	BOS vs. MIA	W	238	121	...	50	29	7	4	11	23	13.0	2019	True	MIA
2	42019	1610612738	BOS	Boston Celtics	0041900304	2020-09-23	BOS @ MIA	L	239	109	...	46	28	3	5	19	24	0.2	2019	False	MIA
3	42019	1610612738	BOS	Boston Celtics	0041900303	2020-09-19	BOS @ MIA	W	240	117	...	50	27	8	3	14	29	8.4	2019	False	MIA
4	42019	1610612738	BOS	Boston Celtics	0041900302	2020-09-17	BOS vs. MIA	L	240	101	...	41	19	5	3	20	22	-5.0	2019	True	MIA
5	42019	1610612738	BOS	Boston Celtics	0041900301	2020-09-15	BOS vs. MIA	L	264	114	...	37	24	9	6	10	25	-3.2	2019	True	MIA
6	42019	1610612738	BOS	Boston Celtics	0041900217	2020-09-11	BOS @ TOR	W	240	92	...	44	19	12	5	10	16	5.0	2019	False	TOR
7	42019	1610612738	BOS	Boston Celtics	0041900216	2020-09-09	BOS vs. TOR	L	290	122	...	58	32	5	4	15	25	-3.0	2019	True	TOR
8	42019	1610612738	BOS	Boston Celtics	0041900215	2020-09-07	BOS @ TOR	W	238	111	...	45	23	5	6	11	19	22.0	2019	False	TOR
9	42019	1610612738	BOS	Boston Celtics	0041900214	2020-09-05	BOS vs. TOR	L	240	93	...	45	18	6	2	14	20	-7.0	2019	True	TOR
10	42019	1610612738	BOS	Boston Celtics	0041900213	2020-09-03	BOS vs. TOR	L	239	103	...	44	21	6	7	14	16	-1.0	2019	True	TOR
11	42019	1610612738	BOS	Boston Celtics	0041900212	2020-09-01	BOS @ TOR	W	241	102	...	44	18	6	9	17	20	3.0	2019	False	TOR
12	42019	1610612738	BOS	Boston Celtics	0041900211	2020-08-30	BOS @ TOR	W	240	112	...	50	25	8	4	22	22	18.0	2019	False	TOR
13	42019	1610612738	BOS	Boston Celtics	0041900124	2020-08-23	BOS @ PHI	W	240	110	...	42	19	4	5	14	25	4.0	2019	False	PHI
14	42019	1610612738	BOS	Boston Celtics	0041900123	2020-08-21	BOS @ PHI	W	240	102	...	45	15	7	8	10	23	8.0	2019	False	PHI
15	42019	1610612738	BOS	Boston Celtics	0041900122	2020-08-19	BOS vs. PHI	W	241	128	...	43	20	4	5	6	28	27.0	2019	True	PHI
16	42019	1610612738	BOS	Boston Celtics	0041900121	2020-08-17	BOS vs. PHI	W	239	109	...	43	22	12	4	7	21	8.0	2019	True	PHI
17 rows x 31 columns																					

數據探索可視化



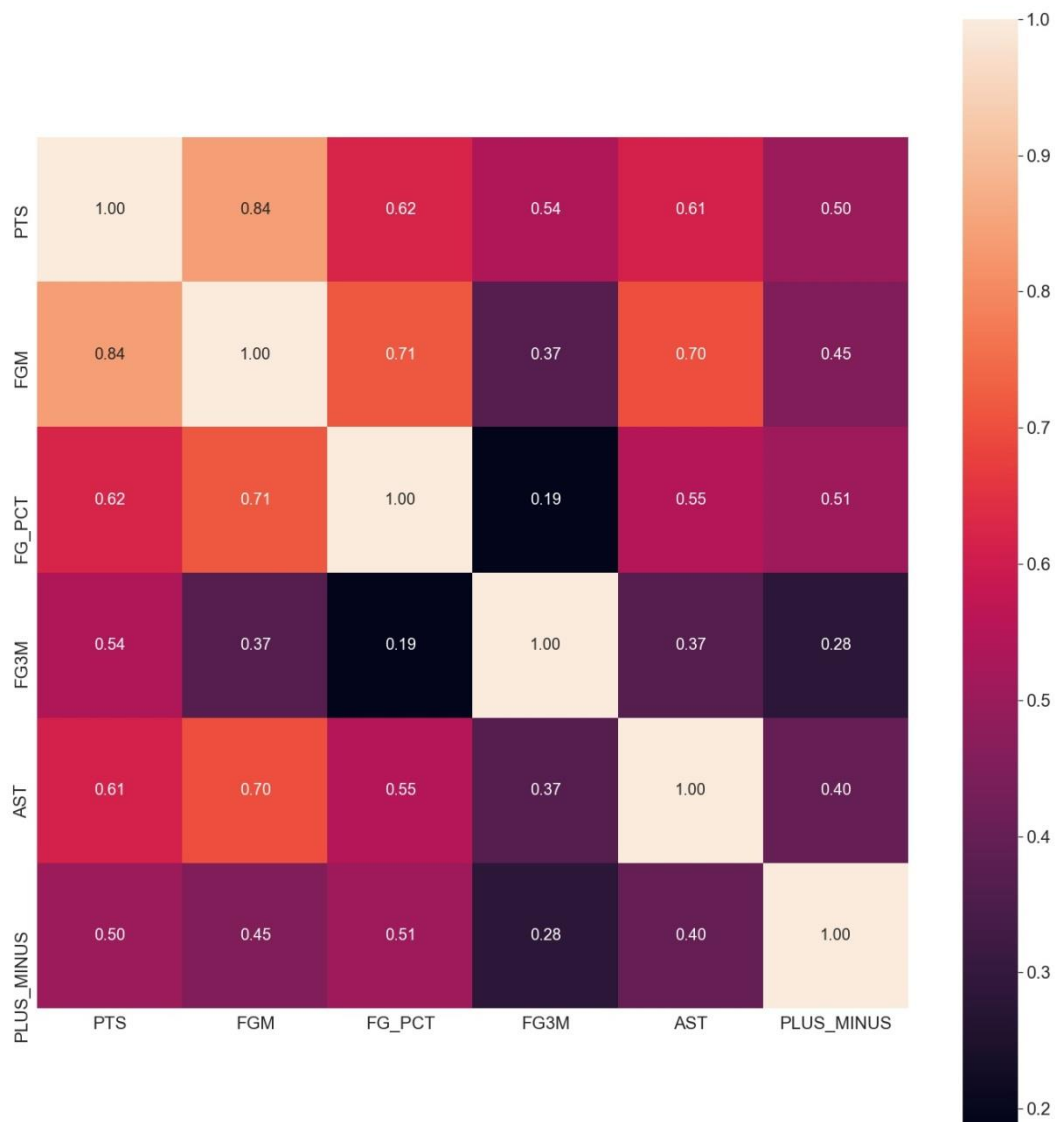
➤ 得分是常態分佈嗎？

■ **Mean: 99.0 (green line)**

■ **Median: 98.4 (red line)**

➤ 透過卡方檢定，得知在顯著水準為 0.05 時，PTS(得分)服從常態分配

數據探索可視化



➤ 得分和哪些連續性的特徵值有關係？
($|r| > 0.5$)

1. FGM(投籃進球數)

2. FG_PCT(投籃命中率)

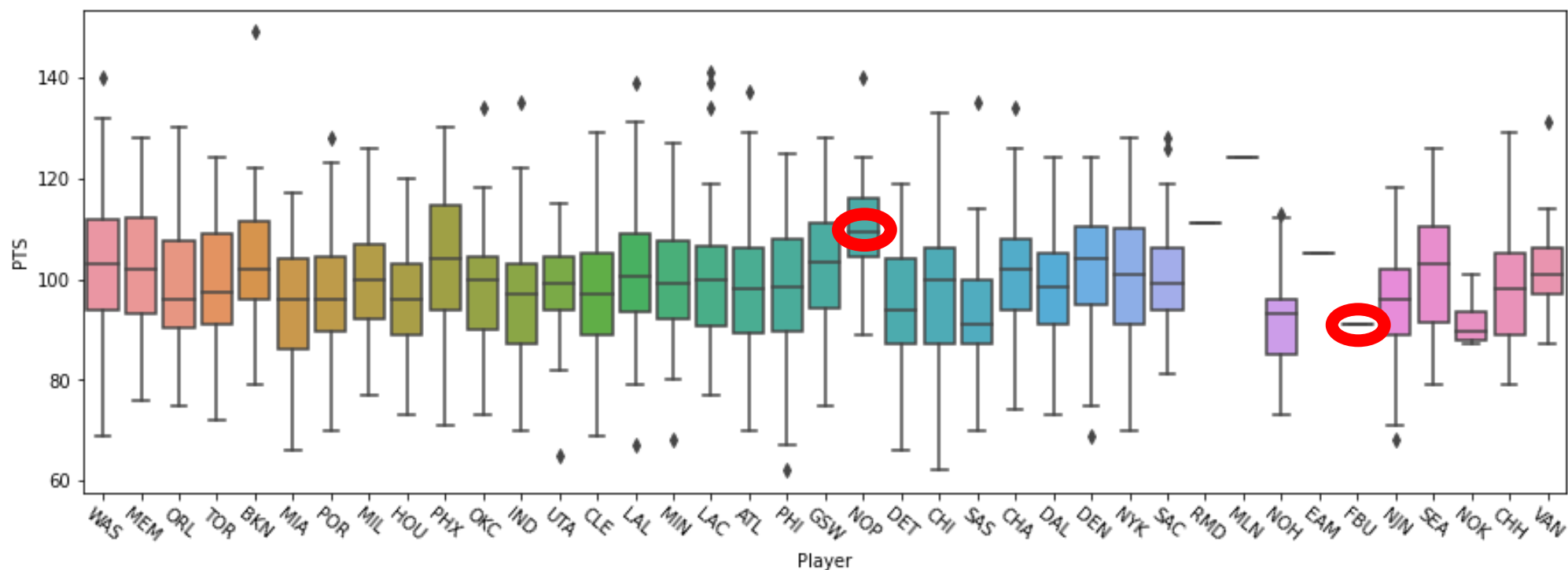
3. FG3M(三分球進球數)

4. AST(助攻)

5. PLUS_MINUS(正負分值)

數據探索可視化

➤ 得分和對戰球隊有相關嗎？



Player	PTS
ATL	98.73
BKN	103.9
CHA	101.5
CHH	99
CHI	97.89
CLE	98.3
DAL	98.35
DEN	101.8
DET	95.12
FBU	91
HOU	96.08
IND	95.96
LAC	100.6
LAL	101.1
MEM	102.4
MIA	94.96
MIL	99.98
MIN	100.2
MLN	124
NJN	95.18
NOH	91.9
NOK	91.75
NOP	110.6
NYK	99.98
OKC	98.59
ORL	98.52
PHI	97.7
PHX	102.7
POR	97.75
RMD	111
SAC	100.4
SAS	93.04
SEA	101.1
TOR	99.6
UTA	98.44
VAN	103.2
WAS	103.3

Data rolling and shift

- 對變數做data rolling and shift，變數包含PTS, FG_PCT, FGM, FGA, FG3M, FG3A, FTM, AST, PLUS_MINUS, REB, STL, BLK, TOV, PF
- 注意填補遺漏值(此處用中位數(median)補上)

```
train["MA_3_PTS"] = np.round(train["PTS"].rolling(window =3,  
                                                center = False ).mean(),2) #往前三筆含自己做平均  
  
train["Last_PTS"] = train.PTS.shift() #往前一筆  
  
train["MA_3_PTS_by_player"] = np.round(train["PTS"].rolling(window =3,  
                                                            center =False ).mean(),2) #往前三筆含自己  
  
train["Last_PTS_player"] = train.PTS.shift() #往前一筆
```

Data rolling and shift

- 做完Data rolling and shift之後，Train的資料

	SEASON_ID	TEAM_ID	TEAM_ABBREVIATION	TEAM_NAME	GAME_ID	GAME_DATE	MATCHUP	WL	MIN	PTS	MA_3_FG3A_by_player	MA_3_FTM_by_player	MA_3_AST_by_player	MA_3_PLUS_MINUS_by_player	MA_3_REB_by_player	MA_3_STL_by_player	MA_3_BLK_by_player	MA_3_TOV_by_player	MA_3_PF_by_player	Last_PTS_player	
2275	21996	1610612738	BOS	Boston Celtics	0029600226	1996-12-03	BOS @ ATL	L	239	95	...	19.33	14.67	22.00	1.33	41.33	8.33	4.33	14.33	22.33	
2276	21996	1610612738	BOS	Boston Celtics	0029600212	1996-11-30	BOS @ MIA	L	240	77	...	26.00	14.67	25.33	12.67	44.33	9.00	6.33	9.67	20.33	122.0
2277	21996	1610612738	BOS	Boston Celtics	0029600199	1996-11-29	BOS vs. HOU	L	241	94	...	27.00	16.67	23.00	-2.00	43.67	11.33	3.67	15.00	23.00	119.0
2278	21996	1610612738	BOS	Boston Celtics	0029600187	1996-11-27	BOS vs. LAL	W	240	110	...	32.33	21.33	24.33	7.00	50.67	9.33	3.33	15.00	24.33	141.0
2279	21996	1610612738	BOS	Boston Celtics	0029600169	1996-11-24	BOS vs. DAL	W	241	105	...	27.33	20.33	22.67	10.67	42.00	7.00	7.67	14.33	22.67	112.0
2280	21996	1610612738	BOS	Boston Celtics	0029600155	1996-11-22	BOS vs. SEA	L	240	91	...	33.67	26.33	24.33	-7.33	42.67	8.00	5.67	14.33	22.67	114.0
2281	21996	1610612738	BOS	Boston Celtics	0029600138	1996-11-20	BOS vs. DET	L	241	83	...	24.33	15.33	21.33	-5.33	39.33	8.67	4.33	15.67	21.00	108.0
2282	21996	1610612738	BOS	Boston Celtics	0029600116	1996-11-16	BOS @ WAS	L	242	92	...	15.67	17.33	20.33	-8.67	35.67	10.00	4.67	14.67	23.00	98.0
2283	21996	1610612738	BOS	Boston Celtics	0029600104	1996-11-15	BOS vs. DEN	L	239	80	...	29.67	21.67	18.33	4.67	47.00	4.33	7.00	14.00	23.67	109.0
2284	21996	1610612738	BOS	Boston Celtics	0029600089	1996-11-13	BOS vs. ATL	W	240	103	...	19.33	18.00	22.00	1.33	41.33	8.33	4.33	14.33	22.33	99.0
2285	21996	1610612738	BOS	Boston Celtics	0029600066	1996-11-09	BOS @ CHI	L	238	92	...	19.00	17.67	16.00	-12.00	41.67	8.67	2.33	20.00	26.67	98.0
2286	21996	1610612738	BOS	Boston Celtics	0029600053	1996-11-08	BOS vs. PHI	L	240	105	...	27.33	15.67	25.67	3.00	46.33	9.67	6.00	14.67	24.00	122.0
2287	21996	1610612738	BOS	Boston Celtics	0029600041	1996-11-06	BOS vs. IND	W	240	94	...	29.00	16.67	22.33	-5.33	45.00	10.33	5.00	16.67	22.33	112.0
2288	21996	1610612738	BOS	Boston Celtics	0029600020	1996-11-02	BOS @ MIL	L	241	102	...	27.00	25.00	18.33	-6.67	44.67	6.33	5.33	15.67	26.00	106.0
2289	21996	1610612738	BOS	Boston Celtics	0029600001	1996-11-01	BOS vs. CHI	L	240	98	...	20.67	18.00	15.00	-6.33	38.33	10.33	3.33	16.33	24.67	87.0
15 rows * 61 columns																					

Data rolling and shift

- 做完Data rolling and shift之後，Test的資料

	SEASON_ID	TEAM_ID	TEAM_ABBREVIATION	TEAM_NAME	GAME_ID	GAME_DATE	MATCHUP	WL	MIN	P	...	MA_3_FG3A_by_player	MA_3_FTM_by_player	MA_3_AST_by_player	MA_3_PLUS_MINUS_by_player	MA_3_REB_by_player	MA_3_STL_by_player	MA_3_BLK_by_player	MA_3_TOV_by_player	MA_3_PF_by_player	Last_PTS_player
5	42019	1610612738	BOS	Boston Celtics	41900301	2020-09-15	BOS vs. MIA	L	264	114	...	35.33	22.67	22.67	6.67	45.33	6.00	5.00	13.33	22.67	114.0
4	42019	1610612738	BOS	Boston Celtics	41900302	2020-09-17	BOS vs. MIA	L	240	101	...	35.33	20.00	22.67	6.67	45.33	6.00	5.00	13.33	22.67	114.0
3	42019	1610612738	BOS	Boston Celtics	41900303	2020-09-19	BOS @ MIA	W	240	117	...	32.00	22.00	23.33	0.07	42.67	7.33	4.00	14.67	25.33	101.0
2	42019	1610612738	BOS	Boston Celtics	41900304	2020-09-23	BOS @ MIA	L	239	109	...	31.33	20.67	24.67	1.20	45.67	5.33	3.67	17.67	25.00	117.0
1	42019	1610612738	BOS	Boston Celtics	41900305	2020-09-25	BOS vs. MIA	W	238	121	...	34.67	22.67	28.00	7.20	48.67	6.00	4.00	14.67	25.33	109.0
0	42019	1610612738	BOS	Boston Celtics	41900306	2020-09-27	BOS @ MIA	L	240	113	...	41.33	18.00	27.67	0.40	46.00	5.67	3.67	13.33	22.67	121.0
16	42019	1610612738	BOS	Boston Celtics	41900121	2020-08-17	BOS vs. PHI	W	239	109	...	38.33	20.00	25.67	3.00	45.00	8.67	3.33	9.33	21.67	113.0
15	42019	1610612738	BOS	Boston Celtics	41900122	2020-08-19	BOS vs. PHI	W	241	128	...	40.00	18.67	22.67	7.67	42.67	7.67	3.67	7.67	23.33	109.0
14	42019	1610612738	BOS	Boston Celtics	41900123	2020-08-21	BOS @ PHI	W	240	102	...	35.00	22.00	19.00	14.33	43.67	7.67	5.67	7.67	24.00	128.0
13	42019	1610612738	BOS	Boston Celtics	41900124	2020-08-23	BOS @ PHI	W	240	110	...	36.33	21.67	18.00	13.00	43.33	5.00	6.00	10.00	25.33	102.0
12	42019	1610612738	BOS	Boston Celtics	41900211	2020-08-30	BOS @ TOR	W	240	112	...	35.00	20.33	19.67	10.00	45.67	6.33	5.67	15.33	23.33	110.0
11	42019	1610612738	BOS	Boston Celtics	41900212	2020-09-01	BOS @ TOR	W	241	102	...	37.33	20.67	20.67	8.33	45.33	6.00	6.00	17.67	22.33	112.0
10	42019	1610612738	BOS	Boston Celtics	41900213	2020-09-03	BOS vs. TOR	L	239	103	...	35.33	18.67	21.33	6.67	46.00	6.67	6.67	17.67	19.33	102.0
9	42019	1610612738	BOS	Boston Celtics	41900214	2020-09-05	BOS vs. TOR	L	240	93	...	34.00	19.67	19.00	-1.67	44.33	6.00	6.00	15.00	18.67	103.0
8	42019	1610612738	BOS	Boston Celtics	41900215	2020-09-07	BOS @ TOR	W	238	111	...	32.67	20.00	20.67	4.67	44.67	5.67	5.00	13.00	18.33	93.0
7	42019	1610612738	BOS	Boston Celtics	41900216	2020-09-09	BOS vs. TOR	L	290	122	...	38.33	19.67	24.33	4.00	49.33	5.33	4.00	13.33	21.33	111.0
6	42019	1610612738	BOS	Boston Celtics	41900217	2020-09-11	BOS @ TOR	W	240	92	...	39.33	17.33	24.67	8.00	49.00	7.33	5.00	12.00	20.00	122.0

17 rows • 61 columns

RFE 篩選

```
import sklearn
from sklearn.feature_selection import RFE
n_features_optimal = 10

lm = LinearRegression()
lm.fit(X, y)

rfe = RFE(lm,
          n_features_to_select = n_features_optimal, step=2)
rfe = rfe.fit(X, y)
```

- RFE(recursive feature elimination)
- 目的：減少特徵數量來提升機器學習之預測準確度
- 方法：經過不斷地剔除與資料分類關係較少之特徵，來篩選特徵數目至指定數目。

RFE 篩選

- 令 $n=10$, $step=2$
- 最終得到10個變數

```
cols = np.array(cols)
cols
cols[rfe.support_]
```

```
array(['MA_3_PTS', 'MA_3_FG_PCT', 'MA_3_FGM', 'MA_3_FG3M', 'MA_3_FTM',  
      'MA_3_PTS_by_player', 'MA_3_FG_PCT_by_player',  
      'MA_3_FGM_by_player', 'MA_3_FG3M_by_player', 'MA_3_FTM_by_player'],  
      dtype='<U25')
```


特徵工程

```
size_mapping = {'W': 1, 'L': 0}

train['Last_WL'] = train['WL'].shift().map(size_mapping)
train
imr = SimpleImputer(missing_values=np.nan,
                    strategy='median')
```

- 對於有興趣的變數，但其為類別標籤，故需轉換成整數值(即為將分類變數轉為連續變數)
- 方法：有序編碼、獨熱編碼、標籤編碼
- train與test需要同步做更改

特徵工程

- 對戰球隊透過有序編碼轉為整數值

```
train["Player"].unique()
```

```
array(['WAS', 'MEM', 'ORL', 'TOR', 'BKN', 'MIA', 'POR', 'MIL', 'HOU',  
      'PHX', 'OKC', 'IND', 'UTA', 'CLE', 'LAL', 'MIN', 'LAC', 'ATL',  
      'PHI', 'GSW', 'NOP', 'DET', 'CHI', 'SAS', 'CHA', 'DAL', 'DEN',  
      'NYK', 'SAC', 'RMD', 'MLN', 'NOH', 'EAM', 'FBU', 'NJN', 'SEA',  
      'NOK', 'CHH', 'VAN'], dtype=object)
```

```
Player_mapping = {'WAS':1, 'MEM':2, 'ORL':3, 'TOR':4, 'BKN':5, 'MIA':6, 'POR':7, 'MIL':8, 'HOU':9,  
                  'PHX':10, 'OKC':11, 'IND':12, 'UTA':13, 'CLE':14, 'LAL':15, 'MIN':16, 'LAC':17, 'ATL':18,  
                  'PHI':19, 'GSW':20, 'NOP':21, 'DET':22, 'CHI':23, 'SAS':24, 'CHA':25, 'DAL':26, 'DEN':27,  
                  'NYK':28, 'SAC':29, 'RMD':30, 'MLN':31, 'NOH':32, 'EAM':33, 'FBU':34, 'NJN':35, 'SEA':36,  
                  'NOK':37, 'CHH':38, 'VAN':39}
```

```
# 將'Player' 轉換成有序編碼
```

```
train['Player'] = train['Player'].map(Player_mapping)
```

```
train
```


建立迴歸模型

➤ 預備動作：

將RFE篩選進行的結果與特徵工程後的變數，
建立一個新的DataFrame，令為totaldata。

```
parts = train[["PTS", 'GAME_ID', "Player", "Home", "Last_WL",  
              'MA_3_PTS', 'MA_3_FG_PCT', 'MA_3_FGM',  
              'MA_3_FG3M', 'MA_3_FTM', 'MA_3_PTS_by_player',  
              'MA_3_FG_PCT_by_player', 'MA_3_FGM_by_player',  
              'MA_3_FG3M_by_player', 'MA_3_FTM_by_player' ]]  
  
# 與 Dummy New_columns合併  
totaldata = pd.concat([parts], axis=1) # 左右合併  
totaldata
```

建立迴歸模型

```
y = totaldata.iloc[:,0 ]
X = totaldata.iloc[:,2: ]

#加載用於線性回歸的套件
from sklearn.linear_model import LinearRegression

#分拆數據集 for train valid
from sklearn.model_selection import train_test_split
X_train, X_valid, y_train, y_valid = train_test_split(
    X, y, test_size=0.2, random_state=42 )

#創建線性回歸
lin_mod = LinearRegression()
lin_mod.fit(X_train,y_train)
#顯示截距 / 斜率 / R2 score / 欄位名稱
print(lin_mod.intercept_)
print(lin_mod.coef_)
print(lin_mod.score(X_train,y_train))
X.columns

-29.51845822538543
[ -0.00272986  1.21259245  0.06457035  9.62653585  8.89147775
 -18.03514587 -8.81501699 -8.96403749  3.38816697  2.20355991
 -5.52578569 -3.06298422 -2.73206505]
0.601888685834486
```

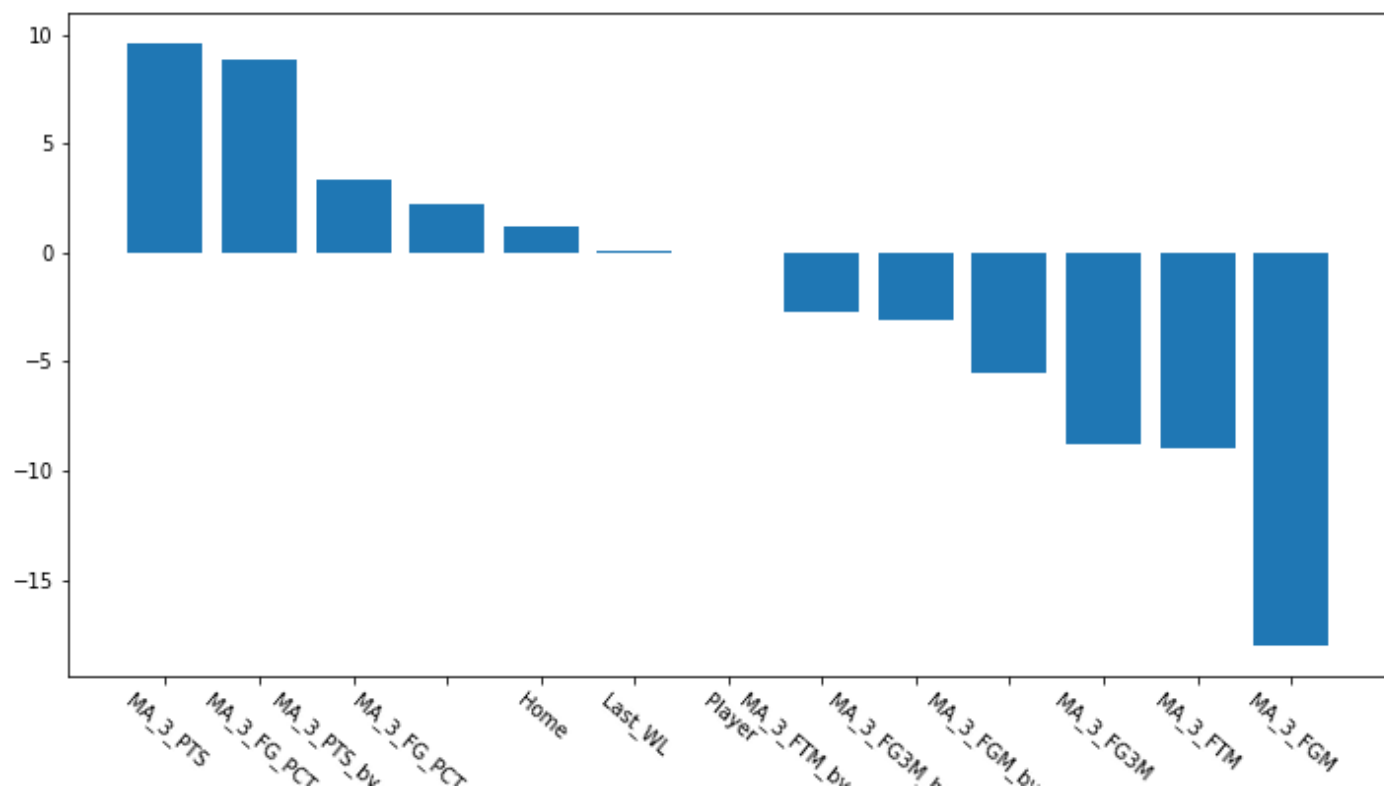
➤ 分拆數據集：

將訓練集(train set)再分割為train set與驗證集(validation set)。

➤ 建立迴歸模型：

對訓練集(train set)創建線性迴歸

建立迴歸模型



	coef
MA_3_PTS	9.626536
MA_3_FG_PCT	8.891478
MA_3_PTS_by_player	3.388167
MA_3_FG_PCT_by_player	2.203560
Home	1.212592
Last_WL	0.064570
Player	-0.002730
MA_3_FTM_by_player	-2.732065
MA_3_FG3M_by_player	-3.062984
MA_3_FGM_by_player	-5.525786
MA_3_FG3M	-8.815017
MA_3_FTM	-8.964037
MA_3_FGM	-18.035146

交叉驗證

```
model=lin_mod #reg
from sklearn.model_selection import cross_val_score
from sklearn.model_selection import TimeSeriesSplit

tscv = TimeSeriesSplit(n_splits=5)

R2 = cross_val_score(model, X_train, y_train, cv=tscv, scoring='r2')
print(R2)
print(f"\nR2: {R2.mean()} (+/- {R2.std()})")

[0.5611859  0.5955535  0.62155647 0.56951273 0.59386396]

R2: 0.5883345113672833 (+/- 0.021343354735106776)
```

➤ Cross Validation by TimeSeriesSplit :

以固定時間的間隔，來觀察時間序列的數據樣本

● n-splits : 分割數

評估修正模型

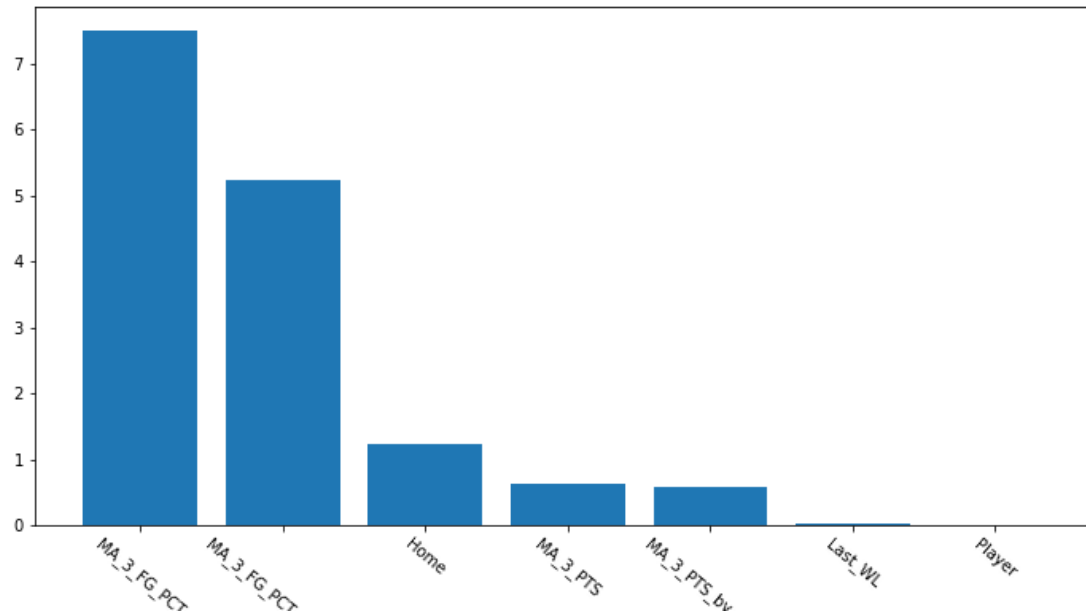
- 透過判定係數R square、R square adj、MSE(損失函數)、MAE來評估模型，並去修正此模型(變數的多寡)

MSE	53.22
RMSE	7.3
MAE	5.87
MAD	5.01
R ²	0.679
Adj R ²	0.673

最佳化模型

➤ 目前比較得到最好的為取變數

"Player", "Home", "Last_WL", "MA_3_FG_PCT", "MA_3_FG_PCT_by_player", "MA_3_PTS", "MA_3_PTS_by_player"



Name	coef
MA_3_FG_PCT	7.5
MA_3_FG_PCT_by_player	5.23
Home	1.22
MA_3_PTS	0.34
MA_3_PTS_by_player	0.58
Last_WL	0.02
Player	0.0001

最佳化模型

➤ 目前比較得到最好的為取變數

"Player" ,"Home","Last_WL","MA_3_FG_PCT","MA_3_FG_PCT_by_player","MA_3_PTS","MA_3_PTS_by_player"

MSE	53.07
RMSE	7.28
MAE	5.88
MAD	5.02
R^2	0.681
Adj R^2	0.674

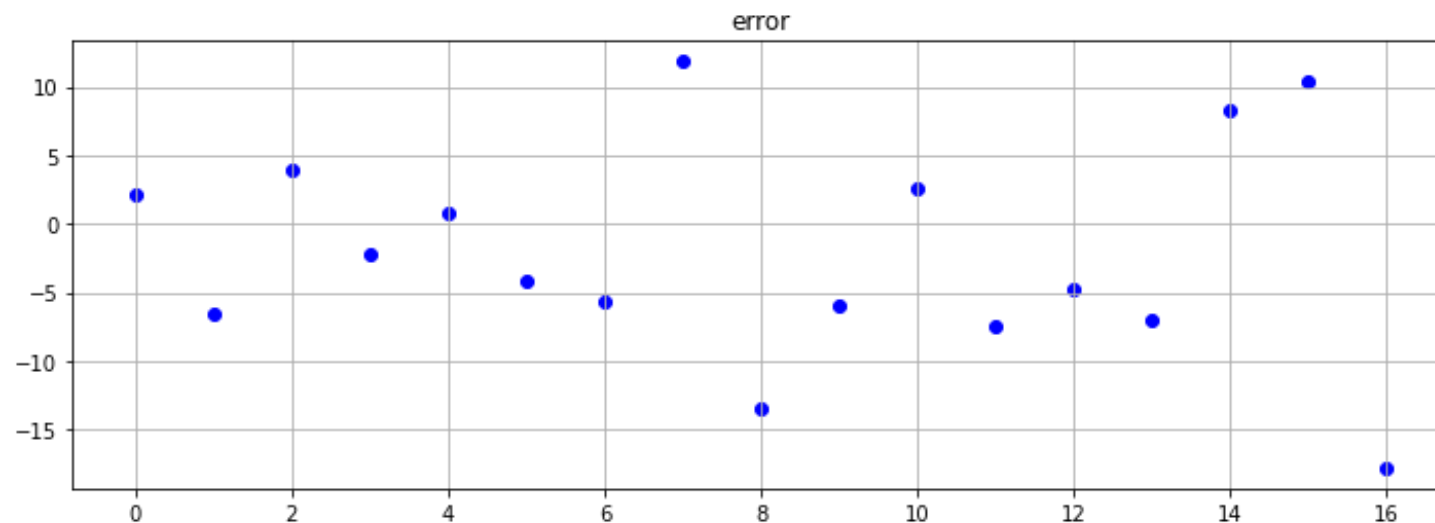
預測得分

game	answer	predict	error
0	114	112	2
1	101	108	-7
2	117	113	4
3	109	111	-2
4	121	120	1
5	113	117	-4
6	109	115	-6
7	128	116	12
8	102	115	-13
9	110	116	-6
10	112	109	3
11	102	109	-7
12	103	108	-5
13	93	100	-7
14	111	103	8
15	122	112	10
16	92	110	-18

➤ 使用最佳化模型，預測本賽季賽爾提克季後賽得分(共17場)

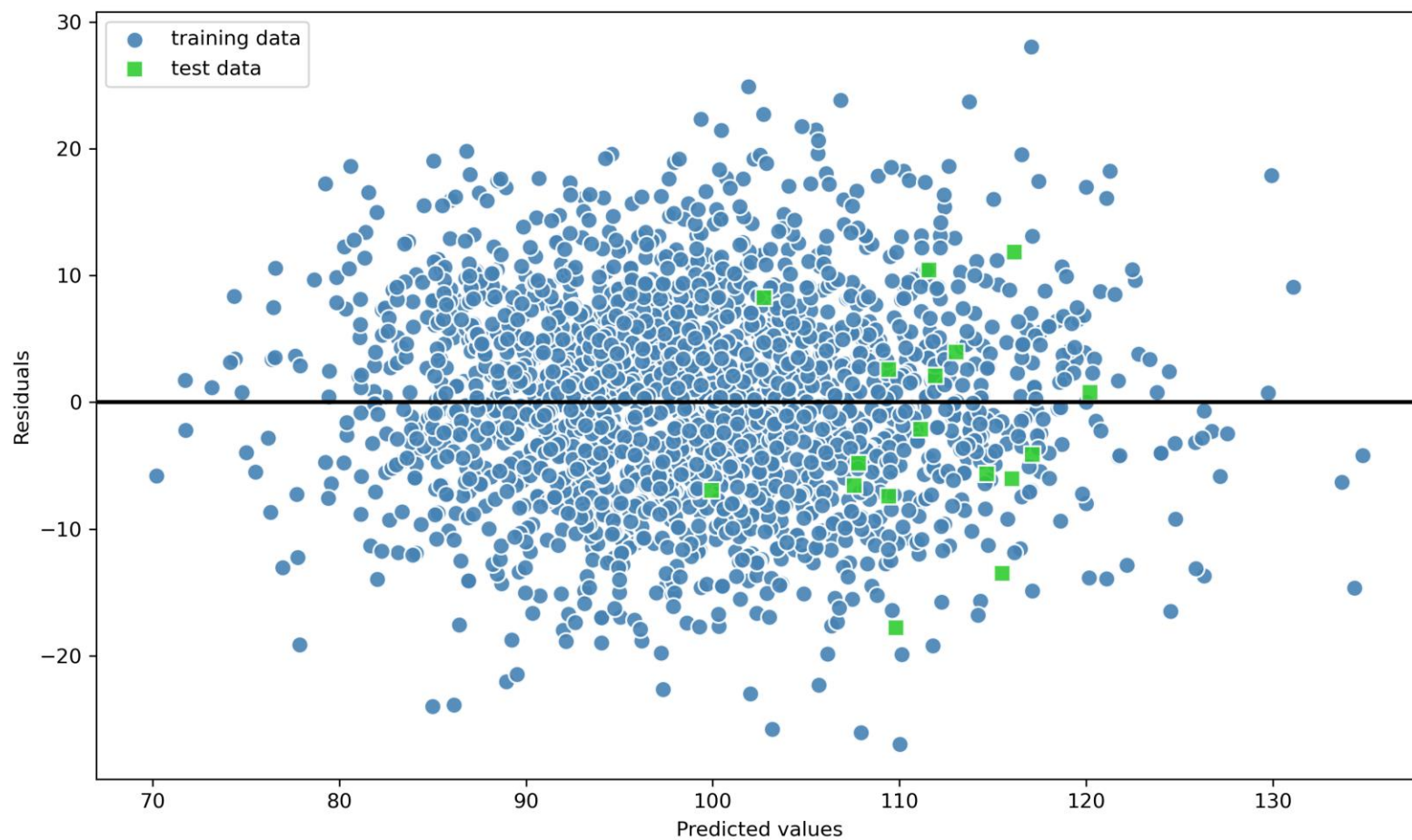
■ MSE test: 64.89

■ R^2 test: 0.281



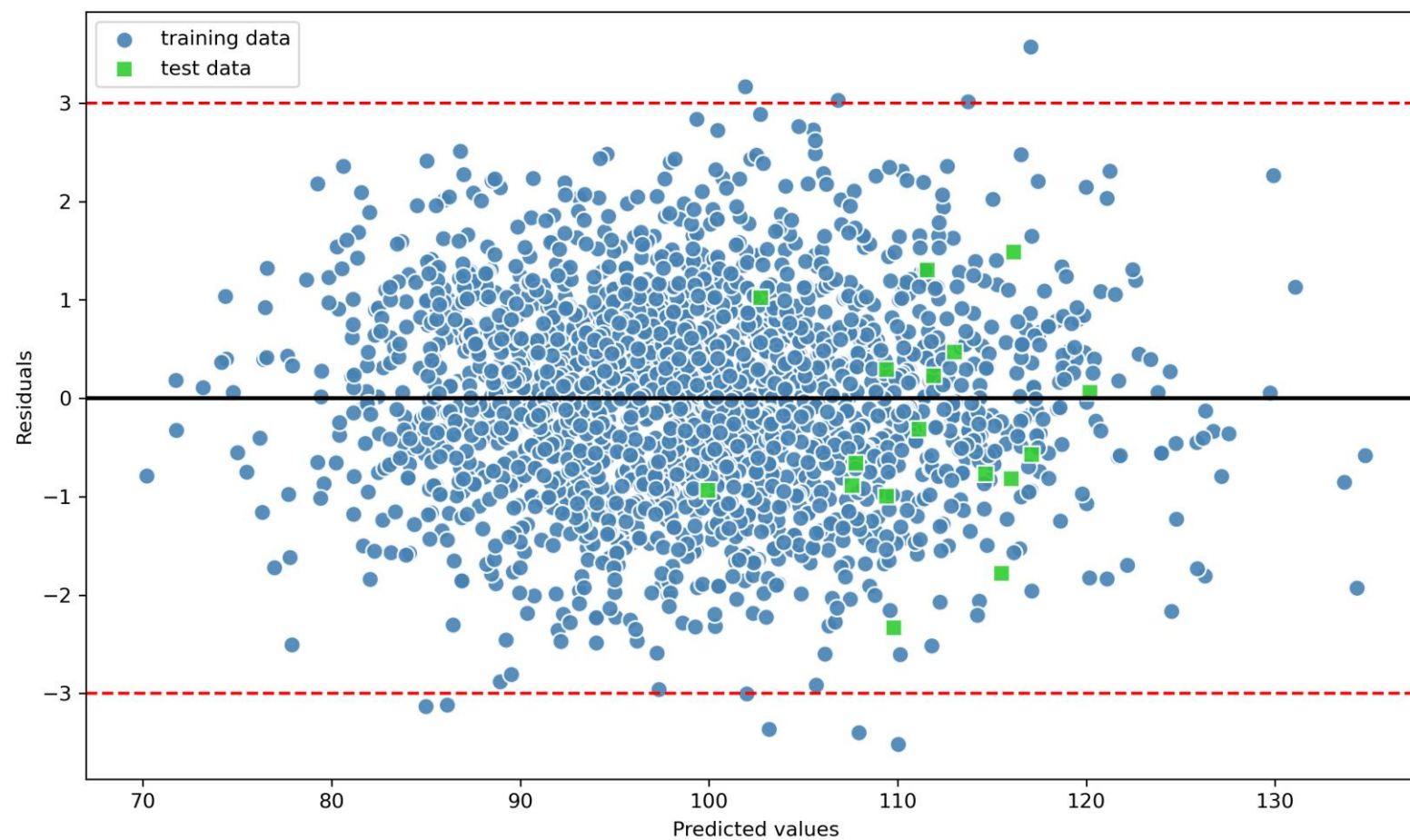
預測得分

➤ 標準化前的殘差圖(train data vs test data)



預測得分

➤ 標準化後的殘差圖(train data vs test data)



心得與未來



Black

White

Thank You

FOR WATCHING AND LISTENING