world population - regression models

1.knowing the population data 😊

- [ ] original data

1.1 original dataframe

	country_name	country_code	1961	1962	1963	1964	1965	1
0	Aruba	ABW	2.179059	1.548572	1.389337	1.215721	1.032841	0
1	Africa Eastern and Southern	AFE	2.660180	2.732633	2.753248	2.806915	2.840787	2
2	Afghanistan	AFG	1.925952	2.014879	2.078997	2.139651	2.216007	2
3	Africa Western and Central	AFW	2.115789	2.145723	2.190827	2.211360	2.242567	2
4	Angola	AGO	1.558355	1.460738	1.410425	1.301745	1.111041	0

1.2 number of row and columns

265 * 64

- [ ] data post processing

2.1 dataframe post processing

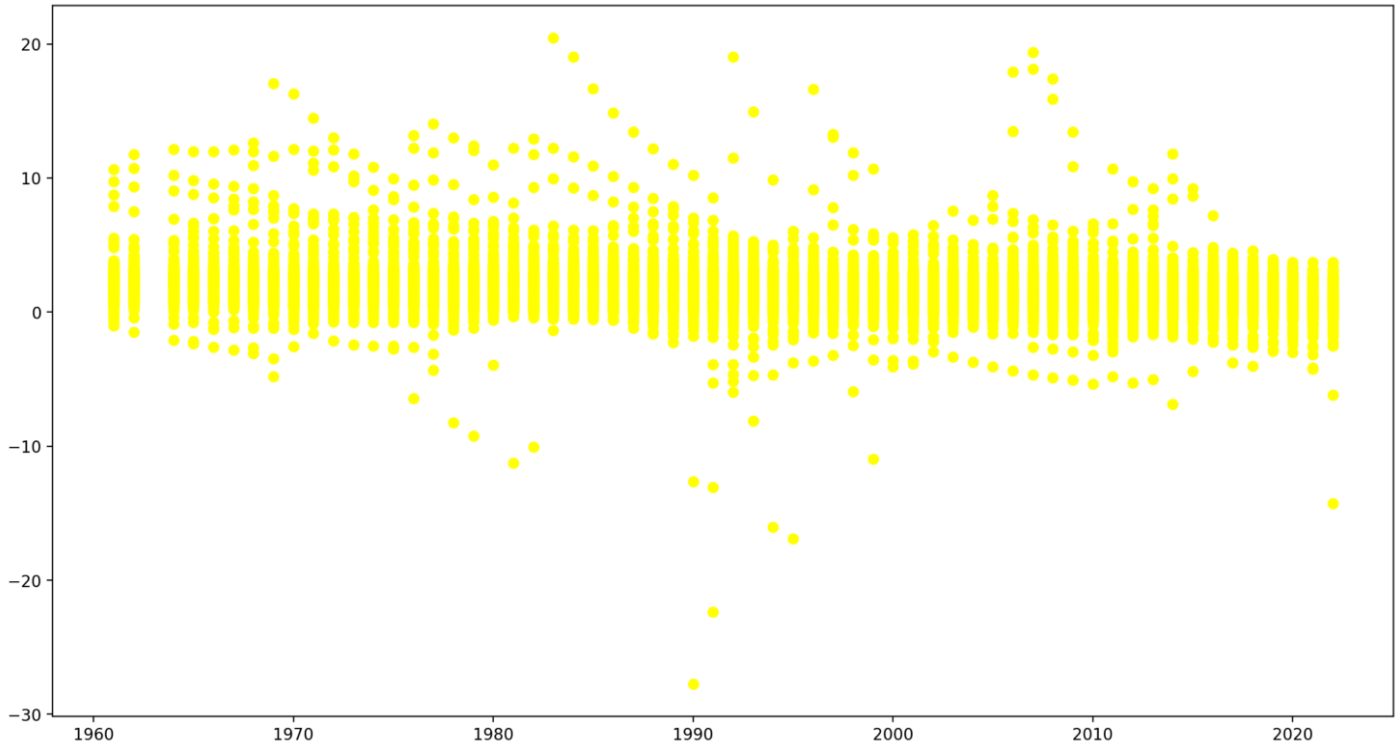
	country_name	country_code	population	year
0	Aruba	ABW	2.179059	1961
1	Africa Eastern and Southern	AFE	2.660180	1961
2	Afghanistan	AFG	1.925952	1961
3	Africa Western and Central	AFW	2.115789	1961
4	Angola	AGO	1.558355	1961

2.2 number of row and columns

```
16165 * 4
```

☐ scatterplot

3. scatterplot



2.regression models 🌞

regression models:

LinearRegression

2.1 LinearRegression

Train data(mse)*

```
1.5158338269382927
```

Test data(mse)*

```
1.1634849493913964
```

actual training vs training prediction

	train_actual	train_pred
6,835	3.022925	2.930045
12,232	3.199744	2.687633
2,893	2.691515	3.165436
3,591	1.971369	1.845584

actual training vs training prediction

	test_actual	test_pred
12,457	3.561288	1.907731
5,752	3.284728	2.977673
6,366	4.017808	3.753709
7,234	-1.120540	0.799968