

โครงการปริมาณการปล่อยก๊าซเรือนกระจกเมื่อเทียบกับประชากรในประเทศไทย

วัดภูประสงค์

-เพื่อวิเคราะห์และคาดการณ์แนวโน้ม ปริมาณของการปล่อยแก๊สคาร์บอนไดออกไซด์ต่อประชากรในอนาคต ภายในปีถัดไป

ชื่อสมาชิก:

นายธีรภัทร อัครนนท์วิวัฒน์

นายชนทัต รักษ์ชัยธนกรณ์

ก้องภพ อิน้อด

แนวทางการแก้ปัญหา

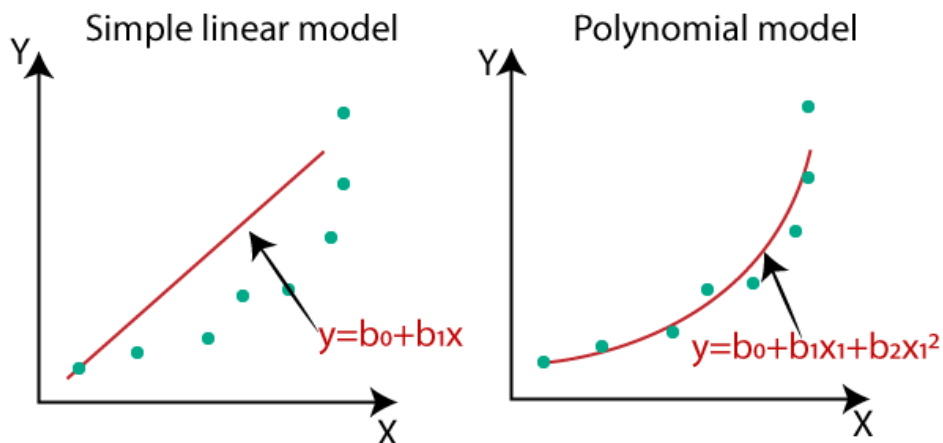
1.ค้นหาข้อมูล

หลังจากทำการค้นหา พบไฟล์จากเว็บไซต์ <https://ourworldindata.org/co2-emissions> ชื่อ `owid-co2-data.csv` ซึ่งมีทั้งข้อมูลของประชากรและปริมาณ Co2 Emission ในปีต่างๆ และตรงกับแหล่งข้อมูลของ `data.go.th` ของไทย จึงตัดสินใจใช้ข้อมูลนี้ในการเริ่มทำ polynomial regression

H2300	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
22982	Thailand	1955	THA	23711122	3.53E+10	0.189	0.008	2.373	0.322	15.73	0.1	0.067		0.132	0.006				1.454	9.965	0.583		
22983	Thailand	1956	THA	24389546	3.59E+10	0.196	0.008	2.618	0.245	10.34	0.107	0.073		0.117	0.005				1.65	12.583	0.7		
22984	Thailand	1957	THA	25096798	3.63E+10	0.2	0.008	2.922	0.304	11.61	0.116	0.08		0.143	0.006				1.85	15.506	0.843		
22985	Thailand	1958	THA	25833566	3.77E+10	0.225	0.009	3.017	0.095	3.25	0.117	0.08		0.143	0.005				2.075	18.523	0.986		
22986	Thailand	1959	THA	26600322	4.22E+10	0.251	0.009	3.321	0.304	10.07	0.125	0.079		0.147	0.005				2.326	21.844	1.133		
22987	Thailand	1960	THA	27397208	4.73E+10	0.262	0.01	3.709	0.388	11.69	0.135	0.079		0.147	0.005				2.588	25.554	1.279		
22988	Thailand	1961	THA	28224186	4.97E+10	0.4	0.014	4.141	0.431	11.62	0.147	0.083		0.121	0.004				2.988	29.694	1.4		
22989	Thailand	1962	THA	29080946	5.36E+10	0.476	0.016	5.027	0.886	21.4	0.173	0.094		0.15	0.005				3.464	34.721	1.55		
22990	Thailand	1963	THA	29966874	5.8E+10	0.491	0.016	5.562	0.535	10.64	0.186	0.096		0.15	0.005				3.954	40.283	1.7		
22991	Thailand	1964	THA	30881136	6.19E+10	0.523	0.017	7.47	1.909	34.32	0.242	0.121		0.117	0.004				4.477	47.753	1.818		
22992	Thailand	1965	THA	31822656	6.68E+10	0.618	0.019	7.44	-0.03	-0.4	0.234	0.111		0.243	0.146	0.005			5.095	55.193	1.964		
22993	Thailand	1966	THA	32789128	7.44E+10	0.734	0.022	9.465	20.25	27.22	0.289	0.127		0.26	0.194	0.006			5.829	64.659	2.158		
22994	Thailand	1967	THA	33778804	8.06E+10	0.857	0.025	12.15	2.685	28.36	0.36	0.151		0.293	0.374	0.011			6.687	76.809	2.532		
22995	Thailand	1968	THA	34791416	8.72E+10	1.17	0.034	17.076	4.926	40.54	0.491	0.196	0.321	0.348	0.01				7.857	93.884	2.88		
22996	Thailand	1969	THA	36827088	9.4E+10	1.188	0.033	14.54	-2.536	-14.85	0.406	0.155	0.258	0.392	0.011				9.045	108.424	3.272		
22997	Thailand	1970	THA	36884524	1E+11	1.297	0.035	15.353	0.813	5.59	0.416	0.153	0.229	0.462	0.013				10.343	123.777	3.734		
22998	Thailand	1971	THA	3763280	1.05E+11	1.374	0.036	19.221	3.869	25.2	0.506	0.183	0.259	0.528	0.014				11.717	142.998	4.261		
22999	Thailand	1972	THA	39058592	1.09E+11	1.668	0.043	21.805	2.584	13.45	0.558	0.199	0.242	0.414	0.011				13.385	164.804	4.676		
23000	Thailand	1973	THA	40159584	1.2E+11	1.832	0.046	24.42	2.615	11.99	0.608	0.203	0.256	0.443	0.011				15.216	189.224	5.119		
23001	Thailand	1974	THA	41252320	1.26E+11	1.937	0.047	24.211	-0.21	-0.86	0.587	0.192	0.244	0.612	0.015				17.153	213.434	5.731		
23002	Thailand	1975	THA	42326308	1.32E+11	1.966	0.046	24.372	0.161	0.66	0.576	0.185	0.223	0.63	0.015				19.119	237.806	6.361		
23003	Thailand	1976	THA	43377268	1.44E+11	2.185	0.05	28.716	4.345	17.83	0.662	0.199	0.248	0.736	0.017				21.304	266.522	7.097		
23004	Thailand	1977	THA	44405904	1.58E+11	2.478	0.056	32.065	3.349	11.66	0.722	0.203	0.252	0.63	0.014				23.782	298.587	7.727		
23005	Thailand	1978	THA	45413084	1.74E+11	2.443	0.054	34.947	2.882	8.99	0.77	0.201	0.252	0.795	0.018				26.225	333.534	8.522		
23006	Thailand	1979	THA	46401752	1.83E+11	2.5	0.054	36.586	1.64	4.69	0.789	0.2	0.25	1.539	0.033				28.726	370.12	10.091		
23007	Thailand	1980	THA	47374464	1.91E+11	2.514	0.053	39.956	3.37	9.21	0.843	0.209	0.273	1.729	0.036				31.24	410.077	11.761		

2.เขียนโปรแกรมใน python โดยใช้การ import matplotlib และ sci-kit ในการช่วยเขียนกราฟ polynomial

polynomial regression equation

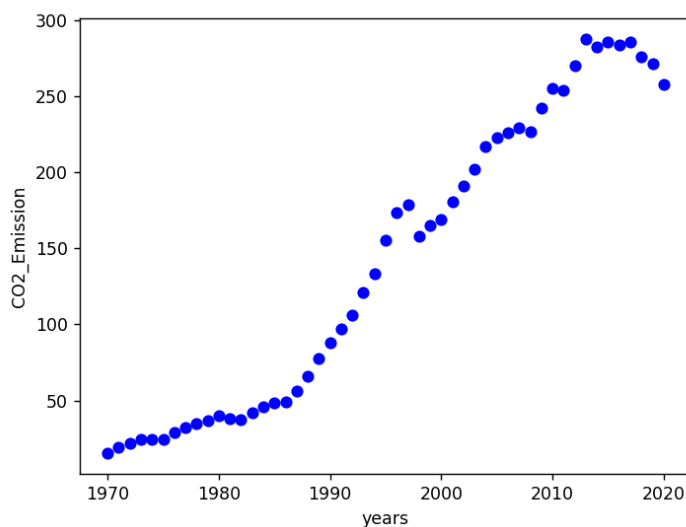


แหล่งที่มา: <https://www.javatpoint.com/machine-learning-polynomial-regression>

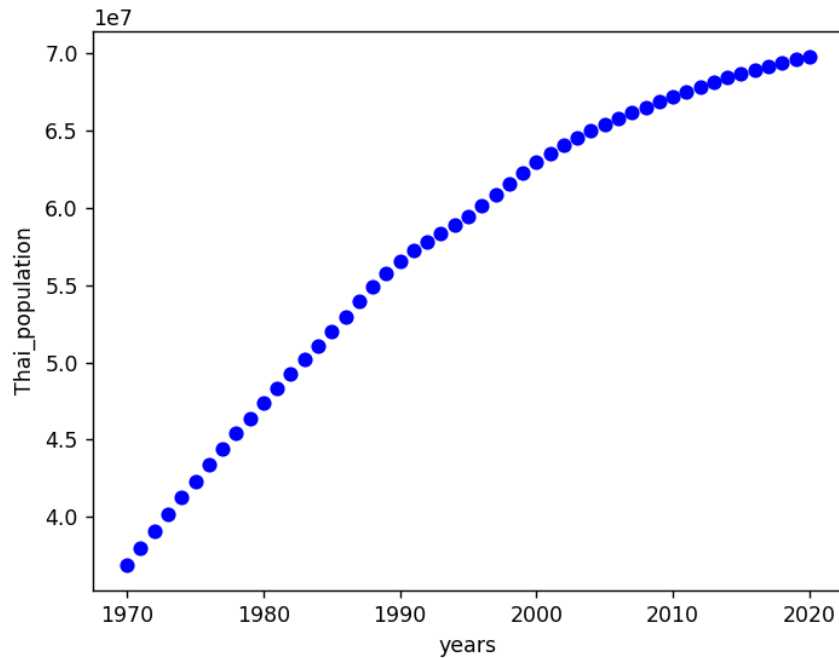
ใช้คำสั่งดังกล่าว ในการพล็อตกราฟ polynomial 1 มิติ 3 degree(โดยได้ทำการ test มาก่อนแล้ว ว่าระหว่าง polynomial 2 หรือ 3 degree นั้น แบบ 3 degree เหมาะสมที่สุด) ตามคำสั่งที่มีให้ใช้ใน numpy และ matplotlib ดังนี้

```
mymodel = np.poly1d(np.polyfit(years, th_co2_per_pop, 3))
myline = np.linspace(1970, 2020, 100)
plt.scatter(years, th_co2_per_pop)
plt.xlabel("years")
plt.ylabel("th_co2_per_pop(1ton/person/year)")
plt.plot(myline, mymodel(myline), 'red')
print(r2_score(th_co2_per_pop, mymodel(years)))
plt.show()
#for j in year:
```

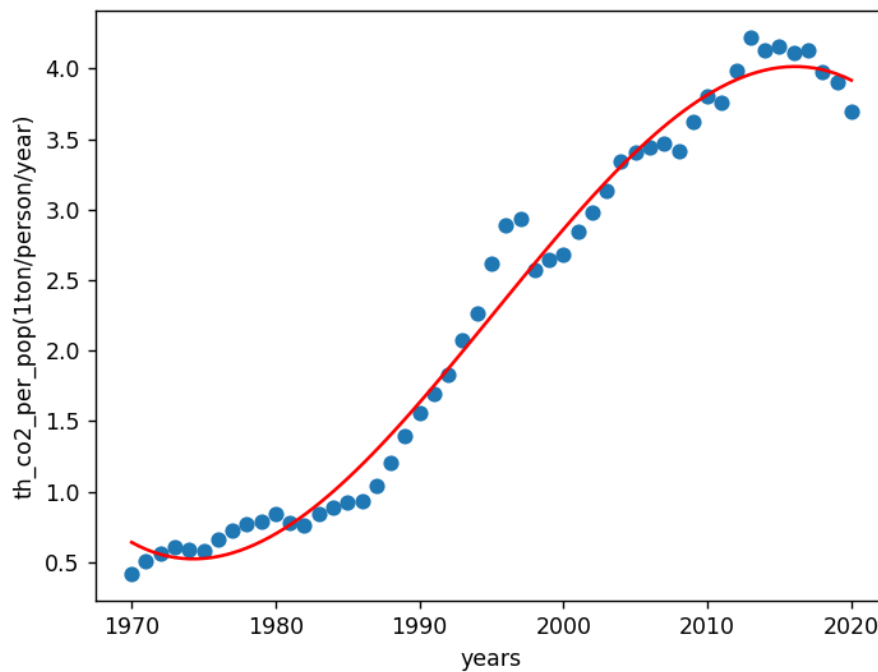
โดยตัวแปรที่ต้องการศึกษาในแกน y คือ dependent variable ในช่วงปี 1970-2020 และแกน x คือ independent variable ปริมาณแก๊สคาร์บอนไดออกไซด์ที่ปล่อยในประเทศไทย ของปี 1970-2020 พล็อตออกมาได้ดังนี้



หลังจากนั้นพล็อตกราฟเพื่อศึกษาจำนวนประชากรในช่วงปีนั้นด้วย ได้ดังนี้



หลังจากนั้น พล็อตกราฟที่เราสนใจศึกษาว่าจะนำมาวิเคราะห์ คือ กราฟความสัมพันธ์ของ ปริมาณCo2ที่ปล่อยของประเทศ/คน และปีที่ศึกษา พล็อตได้ดังนี้



R square

เนื่องจาก polynomial regression เป็นความสัมพันธ์ของ dependent variable และ independent variable แบบ non-linear ดังนั้น ความสัมพันธ์ของ variable ก็ควรจะไม่เป็น linear correlation

ใน matplotlib มี คำสั่ง `r2_score` เพื่อแสดงความสัมพันธ์ด้วย โดยมีค่า score ตั้งแต่ 0-1

$$R^2 = 1 - SS_{\text{res}} / SS_{\text{tot}}$$

Here,

- SS_{res} represents the sum of squares of the residual errors of the data model.
- SS_{tot} represents the total sum of the errors.

ลองใช้คำสั่งหาค่า r^2 ได้ดังนี้

```
C:\Users\ASUS\Pycha  
0.9839571704542494
```

Predicted Value

```
C:\Users\ASUS\PycharmProjects\pythonProject2\venv\Scripts\python.exe  
Predicted CO2 emission for 2023: 3.9553638845682144 tons/person/year  
  
Process finished with exit code 0
```

เมื่อเปรียบเทียบกับข้อมูลจริงในปี 2021 ที่ 3.90 tons/person/year พบว่า มีความคลาดเคลื่อนเล็กน้อย (ประมาณ 2%)

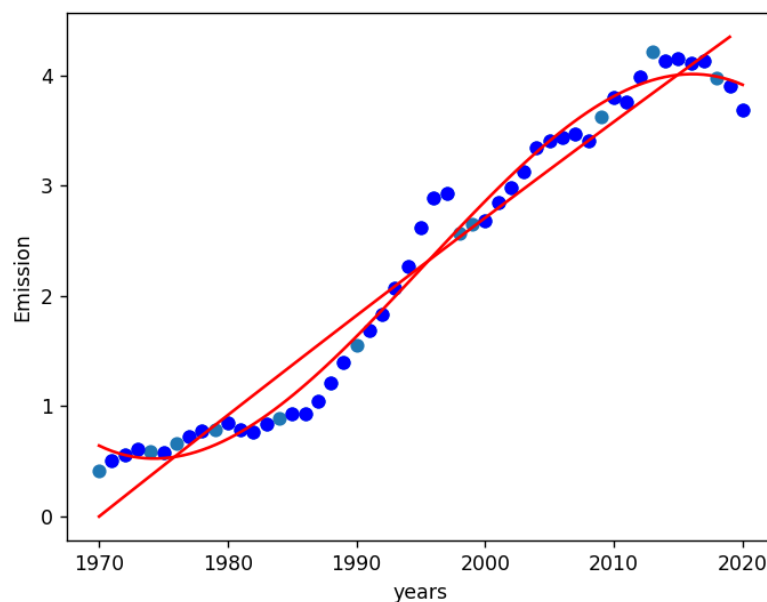
สิ่งที่我们将เพิ่ม

```
#msk = np.random.rand(len(df)) < 0.8  
#train = cdf[msk]  
#test = cdf[~msk]  
#-----
```

1. Train และ Test ข้อมูลเพื่อหาดีกรีที่อาจจะเหมาะสมกว่าเดิม โดยแบ่งอัตราส่วนเป็น Train:Test = 80:20
2. อาจจะใช้ central limit theorem ในการเปลี่ยนข้อมูลให้อยู่ในรูป normal distribution เพื่อใช้คาดการณ์ค่าในปีถัดไปที่ศึกษา
3. นำข้อมูลของประเทศที่มีนโยบายด้านการจัดการปัญหาการคาร์บอนไดออกไซด์ที่น่าจะเป็นแบบอย่างที่ดี มาเปรียบเทียบกับของไทยด้วย

Addition

Figure 1



Compare polynomial regressions of degree 2 and 3