

6-1-3 回帰分析regression

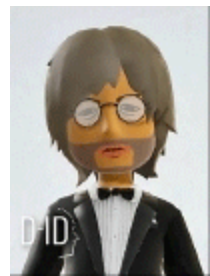
回帰分析: 複数の変数の関係を数式化してフィットする回帰直(曲)線を描画する分析手法

- ・日本の人口推移

https://strnun.fool.jp/pov-ray_strnun/population%20japan.xlsx

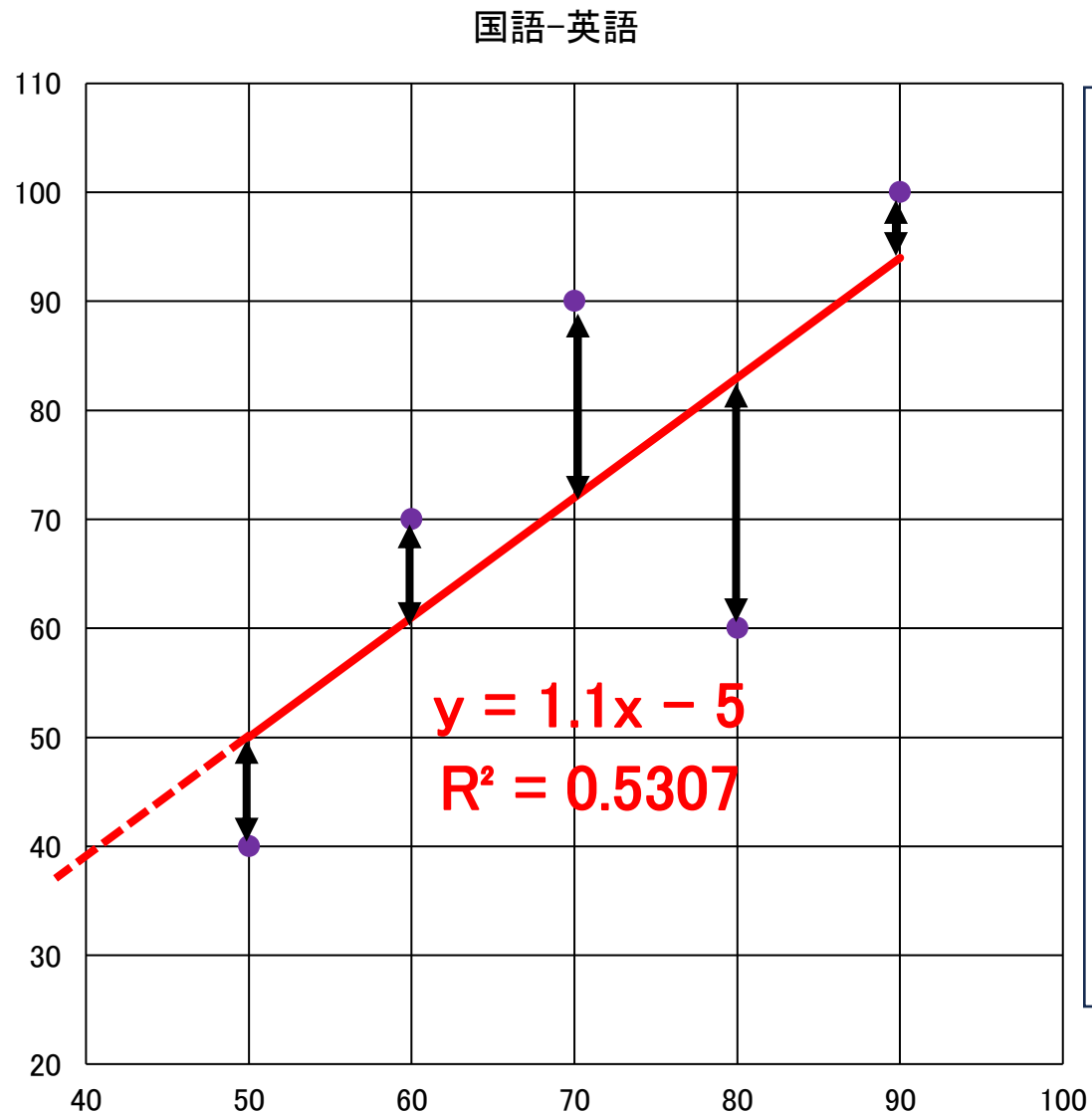
- ・日本の消費者物価指数

https://strnun.fool.jp/pov-ray_strnun/consumer_price_index.xlsx



	国語	英語
A	50	40
B	60	70
C	70	90
D	80	60
E	90	100

説明変数: 国語(x座標)
目的変数: 英語(y座標)



回帰分析

データプロット●とフィットさせる回帰式の誤差の二乗和が最小になる関数を求める

$$y = 1.1x - 5$$

この回帰式に基づくと
国語が40点の場合、英語を
39点と推測できる

$$y = 1.1 \times 40 - 5 = 39$$

	A	B	C	G	H	I	J	K	L	M	N	
1		総務省統計局発信データ https://www.stat.go.jp/data/nihon/O2.html										
2		回帰分析 過去のデータから将来の予測値をシミュレーション										
3												
4		年度	人口(万人)									
5		1920	5,596									
6		1925	5,974									
7		1930	6,445									
8		1935	6,925									
9		1940	7,193									
10		1945	7,215									
11		1950	8,411									
12		1955	9,008									
13		1960	9,430									
14		1965	9,921									
15		1970	10,467									
16		1975	11,194									
17		1980	11,706									
18		1985	12,105									
19		1990	12,361									
20		1995	12,557									
21		2000	12,693									
22		2005	12,777									
23		2010	12,806									
24		2015	12,709									
25		2020	12,577									
26		2025										
27		2030										
28		2035										
29		2040										
30		2045										
31		2050										
32		2055										
33		2060										

[解はこちら](#)

目的変数: 予測したい変数...ここでは将来の人口
 説明変数: 目的変数を説明することができる変数...年度

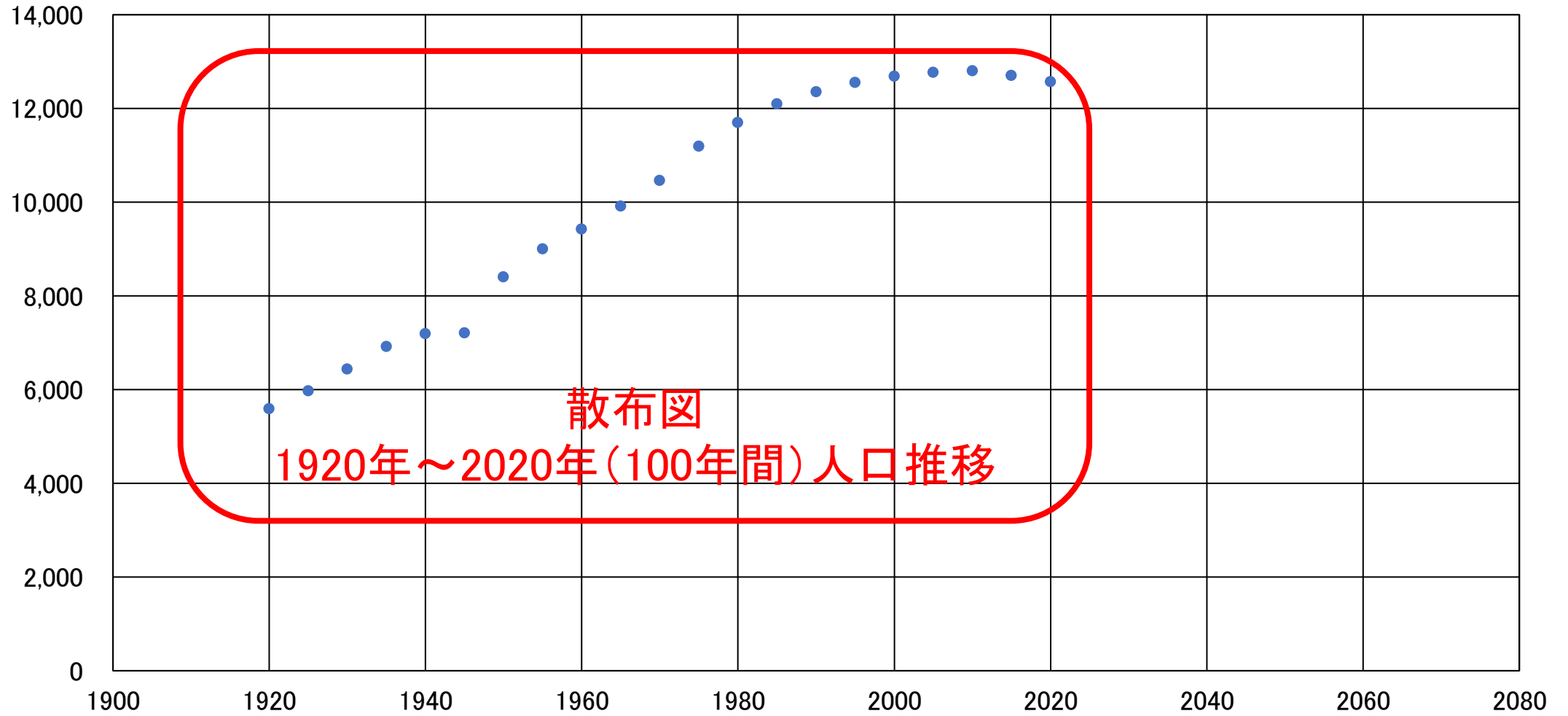
近似する数式(関数)をフィットさせ、近似直線(曲線)を描画
 確からしさ...決定係数 R^2 値(0~1)

数式(グラフ)から、説明変数(年度)における任意の値から
 目的変数(人口)を予測

注意: 多項式回帰を行うと、フィッティング精度が向上する場合があるが、適切な次数の多項式を選択する必要がある。決定係数は、回帰式がデータにどの程度適合しているかを示す指標の1つであり、それだけで妥当性を判断することはできない。



人口(万人)

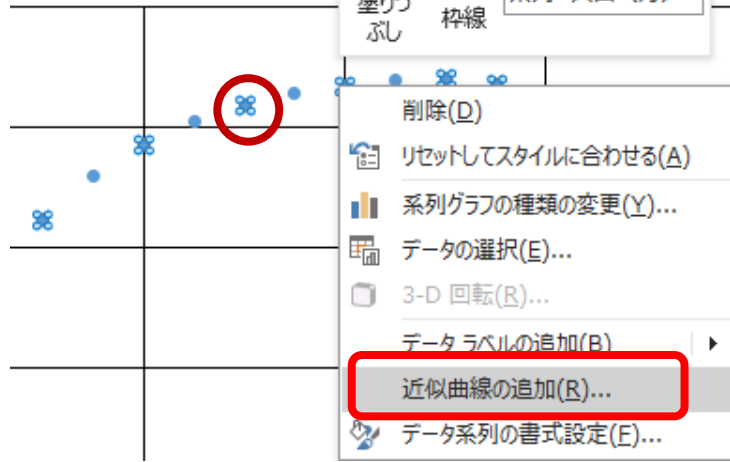


散布図

1920年~2020年(100年間)人口推移



人口(万人)



近似曲線の書式設定

近似曲線のオプション

線

- 線なし(N)
- 線(単色)(S)
- 線(グラデーション)(G)
- 自動(U)

色(C)

透明度(I)

幅(W)

一重線/多重線(C)

実線/点線(D)

線の先端(A)

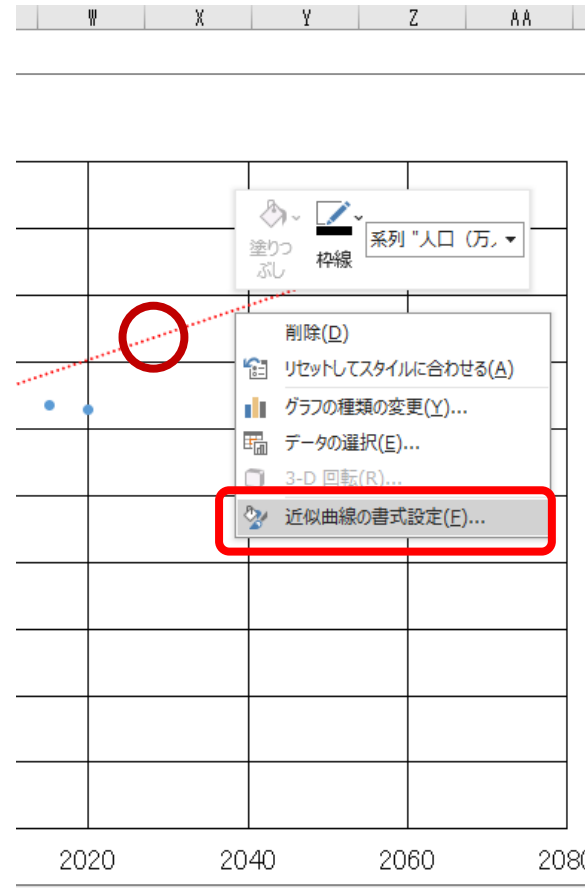
線の結合点(J)

始点矢印の種類(B)

始点矢印のサイズ(S)

終点矢印の種類(E)

終点矢印のサイズ(N)



近似曲線の書式設定

近似曲線のオプション

近似曲線のオプション

- 指数近似(X)
- 線形近似(L)
- 対数近似(O)
- 多項式近似(P) 次数(D)
- 累乗近似(W)
- 移動平均(M) 区間(E)

近似曲線名

- 自動(A) 線形 (人口)
- ユーザー設定(C)

予測

前方補外(E)

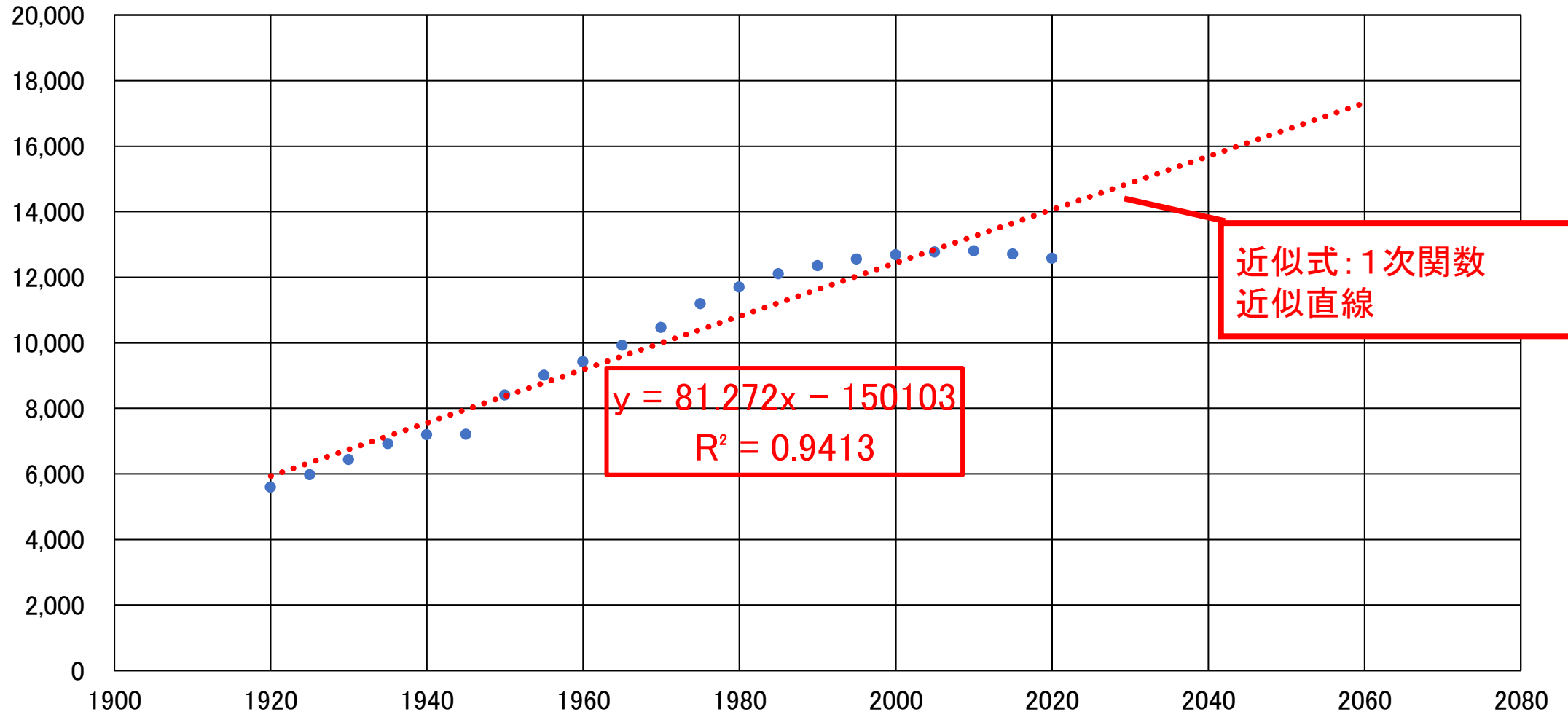
後方補外(B)

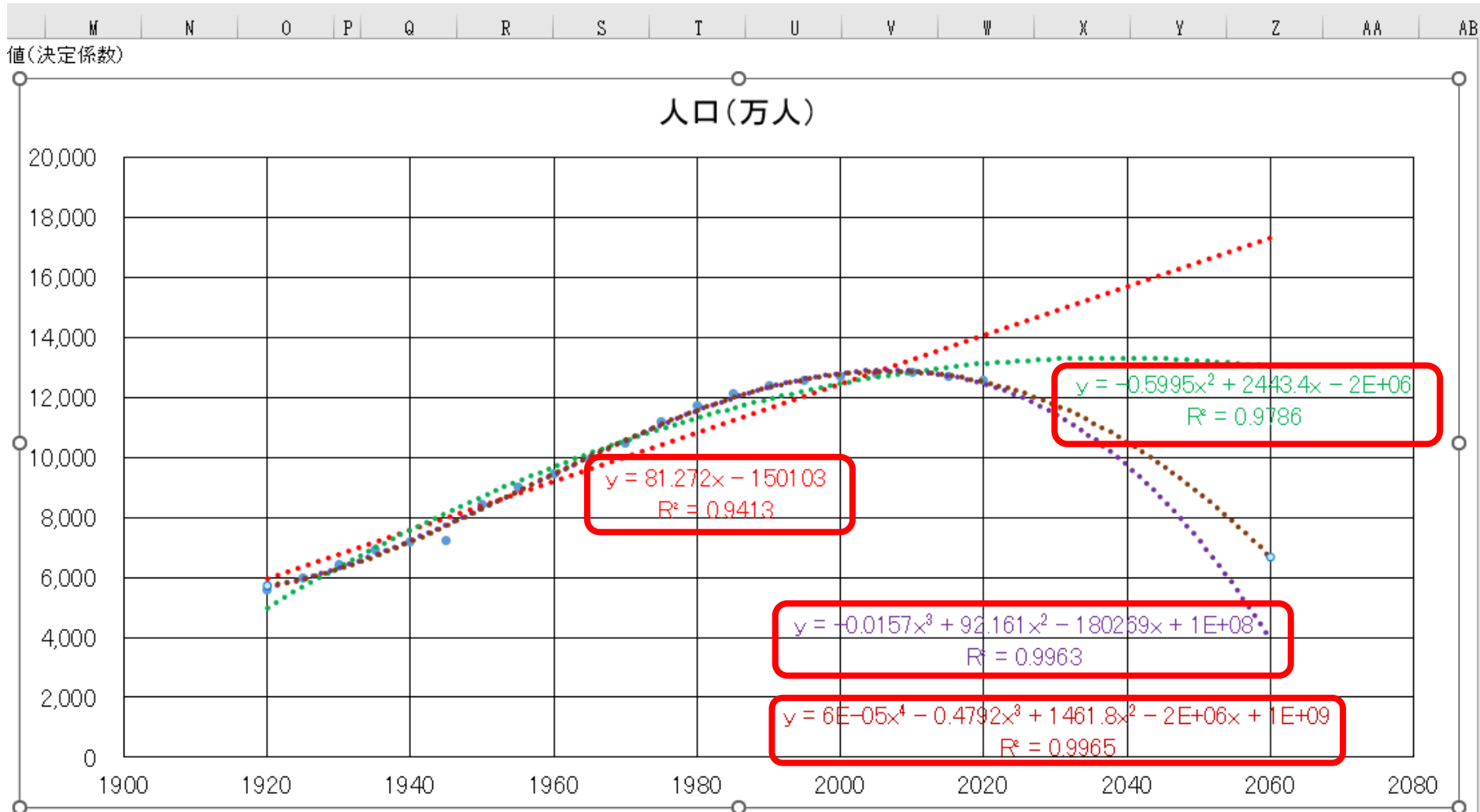
切片(S)

- グラフに数式を表示する(E)
- グラフに R-2 乗値を表示する(R)



人口(万人)





近似曲線の書式設定

近似曲線のオプション

- 指数近似(X)
- 線形近似(L)
- 対数近似(O)
- 多項式近似(P) 次数(D) 4
- 累乗近似(W)
- 移動平均(M) 区間(E) 2

近似曲線名

- 自動(A) 多項式(人口(万人))
- ユーザー設定(C)

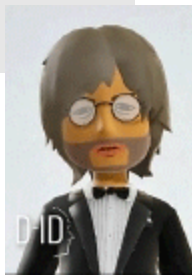
予測

前方補外(E) 0.0 区間

後方補外(B) 0.0 区間

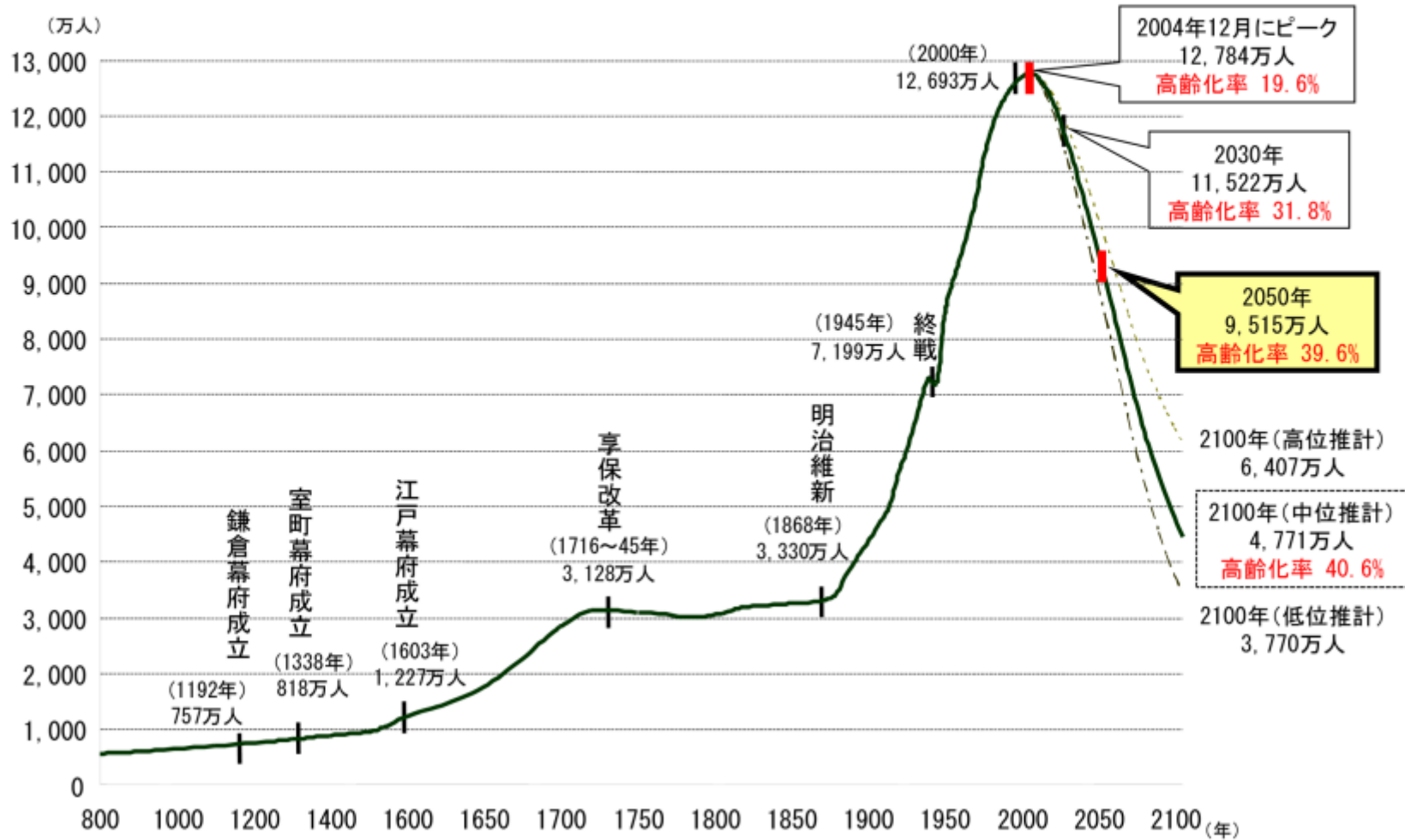
切片(S) 0.0

- グラフに数式を表示する(E)
- グラフに R-2 乗値を表示する(R)



我が国における総人口の長期的推移

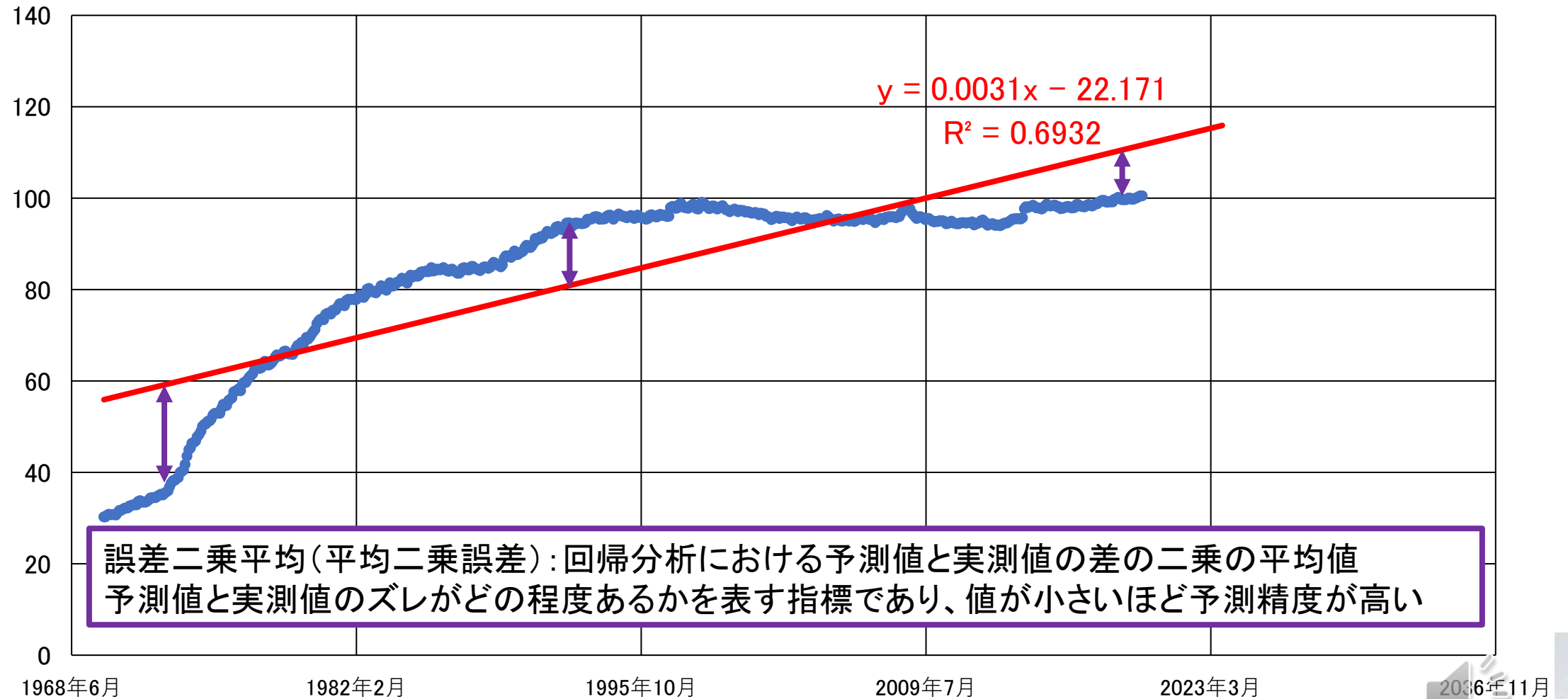
○ 我が国の総人口は、2004年をピークに、今後100年間で100年前(明治時代後半)の水準に戻っていく。この変化は、千年単位でも類を見ない、極めて急激な減少。



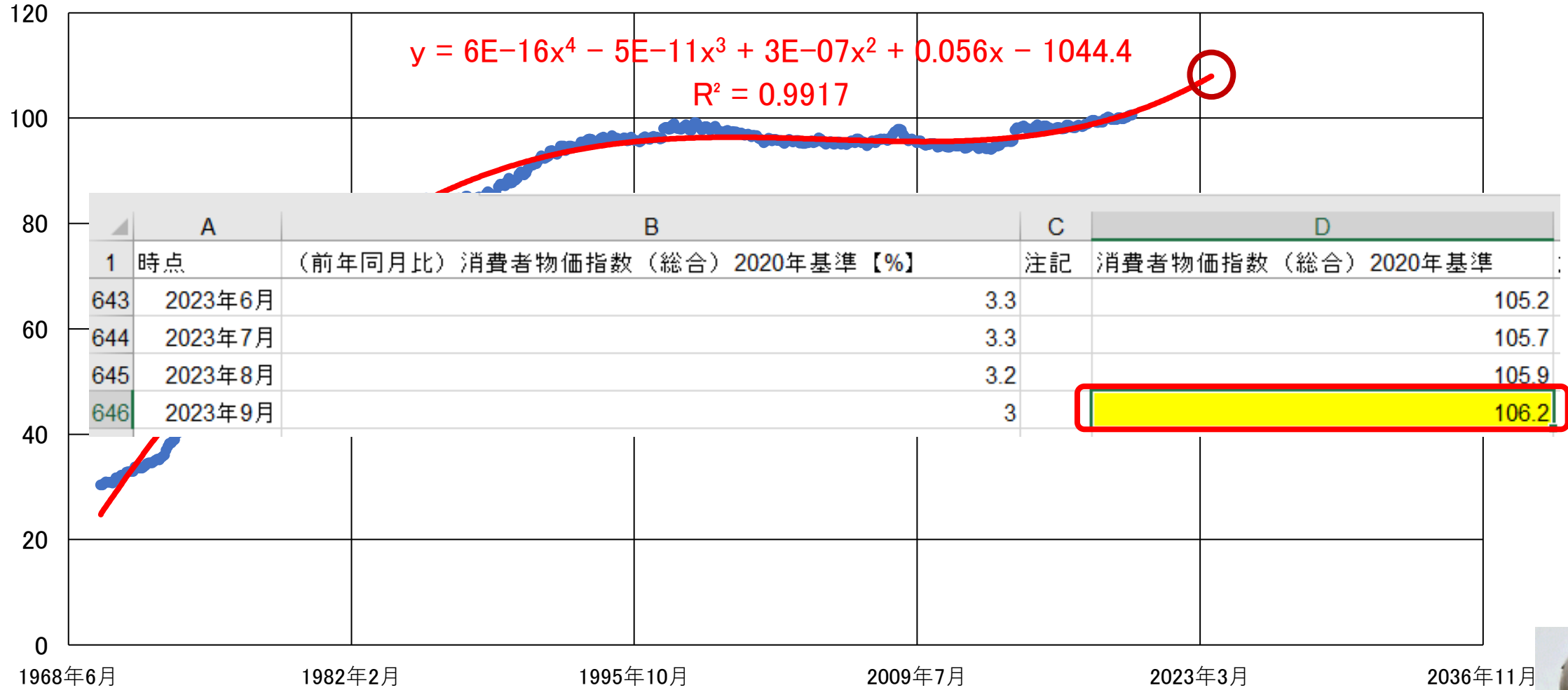
	A	B	C	D
1	時点	(前年同月比) 消費者物価指数 (総合) 2020年基準 [%]	注記	消費者物価指数 (総合) 2020年基準
2	1970年1月			30.3
3	1970年2月			30.3
4	1970年3月			30.6
5	1970年4月			30.0
	A	B	C	D
598	2019年9月	0.2		100.1
599	2019年10月	0.2		100.4
600	2019年11月	0.5		100.5
601	2019年12月	0.8		100.5
	A	B	C	D
644	2023年7月			
645	2023年8月			
646	2023年9月			

1970年1月～2019年12月のデータから2023年9月を「予測」し実測値と比較する

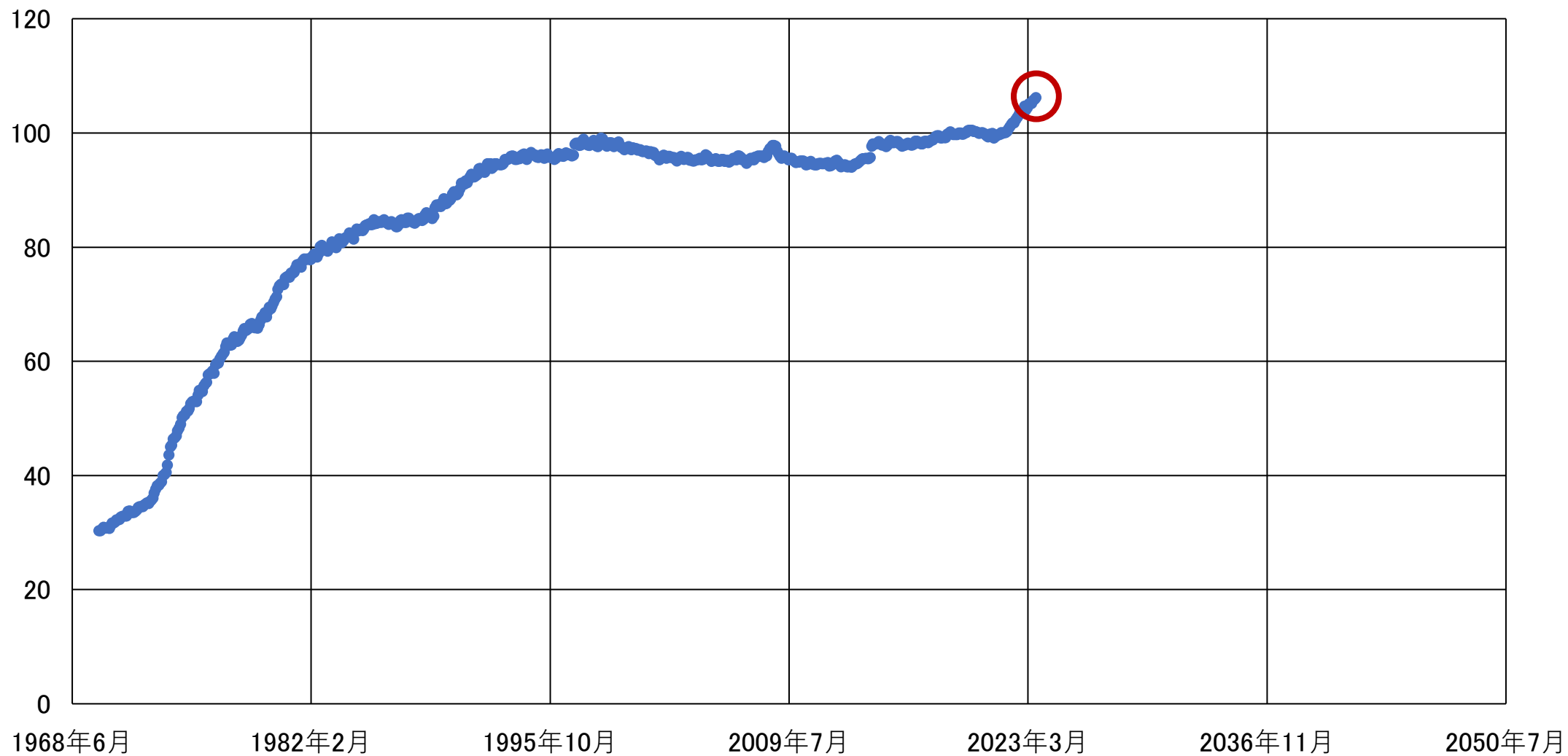
消費者物価指数(総合)_2020年基準100



消費者物価指数(総合)_2020年基準100



消費者物価指数(総合)_2020年基準100



消費者物価指数(総合)_2020年基準100

