

Regresi Linear Menganalisa Perokok Umur 15 Tahun Keatas Tahun 2021 Dengan Prediktor Tahun 2020

Muhammad wigig purbandanu

Informatika,Fakultas Teknik, Universitas Muhammadiyah Semarang

E-Mail: danupurban@gmail.com

Abstract

smoking is a very big health problem, with the number of smokers continuing to increase every year. Therefore, it is important to know the prevalence rate of smoking in people aged over 15 years by province. In this article, we will analyze the percentage of smoking in the population aged more than 15 years by province in Indonesia. We will use the linear regression method to analyze the relationship between the percentage of smoking in the population aged more than 15 years in each province in 2021 and the predictor variables in 2020 Through this analysis, we hope to provide useful recommendations for reducing smoking rates in Indonesia. The method used is the correlation method which is useful for determining the relationship or connection between two or more variables. Correlation can be used to find out how close the relationship between two variables is, which is indicated by the correlation value ranging from -1 to 1. The value 1 indicates a perfect positive correlation, which means that the two variables always move in the same direction. A value of -1 indicates a perfect negative correlation, which means that the two variables always move in opposite directions. A value of 0 indicates no correlation between the two variables. The correlation analysis method can be used in various fields, such as economics, social, health sciences, and others. This indicates that the higher a person's education level, the lower their likelihood of smoking. Based on the results of our linear regression analysis, it can be concluded that there is a significant relationship between the percentage of smokers in the population aged over 15 years in each province in 2021 and predictor variable in 2020. The variable that has the most influence on the percentage of smoking is the level of education. This shows that the higher a person's educational level, the lower their likelihood of smoking.

Keywords: analyze, linear regression, smoker, age, predictor.

Abstrak

merokok menjadi masalah kesehatan yang sangat besar, dengan jumlah perokok yang terus meningkat setiap tahunnya. Oleh karena itu, penting untuk mengetahui tingkat prevalensi merokok pada penduduk umur lebih dari 15 tahun menurut provinsi. Dalam artikel ini, kami akan menganalisis presentase merokok pada penduduk umur lebih dari 15 tahun menurut provinsi di Indonesia kami akan menggunakan metode regresi linear untuk menganalisis hubungan antara persentase merokok pada penduduk umur lebih dari 15 tahun di setiap provinsi pada tahun 2021 dengan variabel prediktor tahun 2020. Melalui analisis ini, kami berharap dapat memberikan rekomendasi yang berguna untuk mengurangi tingkat merokok di Indonesia. metode yang digunakan adalah metode korelasi yang berguna untuk menentukan hubungan atau kaitan antara dua atau lebih variabel. Korelasi dapat digunakan untuk mengetahui seberapa erat hubungan antara dua variabel, yang ditunjukkan dengan nilai korelasi yang berkisar antara -1 sampai 1. Nilai 1 menunjukkan korelasi positif yang sempurna, yang berarti bahwa kedua variabel selalu bergerak ke arah yang sama. Nilai -1 menunjukkan korelasi negatif yang sempurna, yang berarti bahwa kedua variabel selalu bergerak ke arah yang berlawanan. Nilai 0 menunjukkan tidak ada korelasi antara kedua variabel. Metode analisis korelasi dapat digunakan dalam berbagai bidang, seperti ekonomi, sosial, ilmu kesehatan, dan lain-lain. Hal ini menunjukkan bahwa semakin tinggi tingkat pendidikan seseorang, semakin rendah kemungkinan mereka untuk merokok. Berdasarkan hasil analisis regresi linear yang kami lakukan, dapat disimpulkan bahwa terdapat hubungan yang signifikan antara persentase merokok pada penduduk umur lebih dari 15 tahun di setiap provinsi pada tahun 2021 dengan variabel prediktor tahun 2020. Variabel yang paling berpengaruh terhadap persentase merokok adalah tingkat pendidikan. Hal ini menunjukkan bahwa semakin tinggi tingkat pendidikan seseorang, semakin rendah kemungkinan mereka untuk merokok.

Kata kunci : menganalisa, regresi linear, perokok, umur, prediktor.

1. Pendahuluan

Merokok merupakan masalah kesehatan yang sangat besar di dunia. Menurut data dari Organisasi Kesehatan Dunia (WHO), setiap tahunnya sekitar 7 juta orang meninggal karena komplikasi yang

disebabkan oleh merokok. Di Indonesia sendiri, merokok menjadi masalah kesehatan yang sangat besar, dengan jumlah perokok yang terus meningkat setiap tahunnya.

Oleh karena itu, penting untuk mengetahui tingkat prevalensi merokok pada penduduk umur lebih dari 15 tahun menurut provinsi. Dalam artikel ini, kami akan menganalisis presentase merokok pada penduduk umur lebih dari 15 tahun menurut provinsi di Indonesia

kami akan menggunakan metode regresi linear untuk menganalisis hubungan antara persentase merokok pada penduduk umur lebih dari 15 tahun di setiap provinsi pada tahun 2021 dengan variabel prediktor tahun 2020 seperti pendapatan per kapita, tingkat pendidikan, dan tingkat urbanisasi. Melalui analisis ini, kami berharap dapat memberikan rekomendasi yang berguna untuk mengurangi tingkat merokok di Indonesia.

2. Metode Penelitian



Figure 1. Research Methodology

2.1 Sumber Data

data yang digunakan pada penelitian ini adalah data jumlah merokok pada penduduk lebih dari 15 tahun. Adapun data yang digunakan terdiri dari elemen provinsi dari tahun pertahun yang di peroleh dari <https://www.bps.go.id/indicator/30/1435/1/persentase-merokok-pada-penduduk-umur-15-tahun-menurut-provinsi.html>

2.2 Metode yang digunakan

metode yang digunakan adalah metode analisis kolerasi yang berguna untuk menentukan hubungan atau kaitan antara dua atau lebih variabel. Korelasi dapat digunakan untuk mengetahui seberapa erat hubungan antara dua variabel, yang ditunjukkan dengan nilai korelasi yang berkisar antara -1 sampai 1. Nilai 1 menunjukkan korelasi positif yang sempurna, yang berarti bahwa kedua variabel selalu bergerak ke arah yang sama. Nilai -1 menunjukkan korelasi negatif yang sempurna, yang berarti bahwa kedua variabel selalu bergerak ke arah yang berlawanan. Nilai 0 menunjukkan tidak ada korelasi antara kedua variabel. Metode analisis korelasi dapat digunakan dalam berbagai bidang, seperti ekonomi, sosial, ilmu kesehatan, dan lain-lain.

Algoritma

- a. Ambil data tentang persentase merokok pada penduduk umur lebih dari 15 tahun di setiap provinsi pada tahun 2021.
- b. Ambil data tentang variabel prediktor tahun 2020 yang mungkin berpengaruh terhadap persentase merokok.
- c. Gunakan metode regresi linear untuk menganalisis hubungan antara variabel prediktor tahun 2020 dengan persentase merokok pada penduduk umur lebih dari 15 tahun di setiap provinsi pada tahun 2021.
- d. Hitunglah nilai kolerasi (r) antara data tahun 2020 dan tahun 2021 untuk mengetahui tingkat keterkaitan antara kedua variabel tersebut.
- e. Hitunglah nilai b dengan rumus $b = (n * \text{Sigma } xy - \text{Sigma } x * \text{Sigma } y) / (n * \text{Sigma } xx - (\text{Sigma } x)^2)$
- f. Hitung nilai a dengan rumus $a = \bar{Y} - b * \bar{X}$
- g. Buat tabel persamaan regresi linier menggunakan nilai Y yang dicari dengan cara $Y = a + (b * x)$
- h. Buat model regresi linear yang membandingkan hubungan antara variabel prediktor dan persentase merokok.
- i. Buat kesimpulan berdasarkan hasil analisis.

3. Hasil dan Pembahasan

3.1. Menggunakan excel

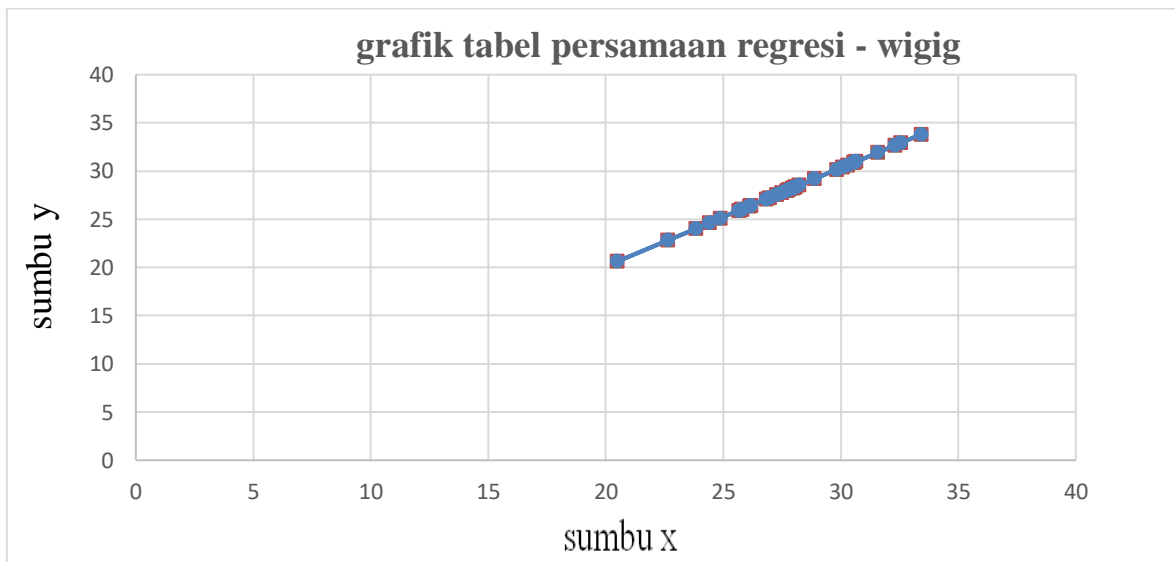
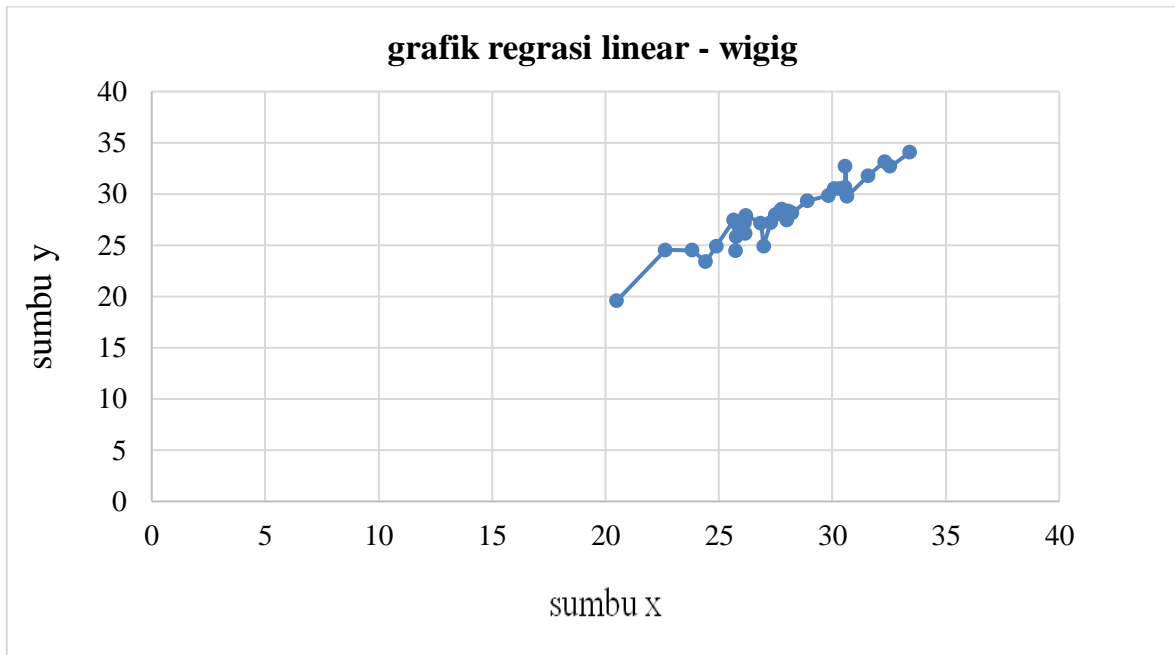
Persentase Merokok Pada Penduduk Umur ≥ 15 Tahun Menurut Provinsi (Persen)

Provinsi	2020(X)	2021(Y)	XY	X ²	(Y-Ybar) ²	(Y-a-bX) ²
BALI	20.5	19.58	401.39	420.25	70.975669	1.095306351
DI YOGYAKARTA	22.64	24.54	555.5856	512.5696	12.004187	2.992998043
KALIMANTAN SELATAN	23.83	24.51	584.0733	567.8689	12.212969	0.236092128
KALIMANTAN TIMUR	24.42	23.37	570.6954	596.3364	21.480499	1.577719565
SULAWESI SELATAN	24.89	24.91	620.0099	619.5121	9.5772045	0.038261499
KALIMANTAN UTARA	25.66	27.46	704.6236	658.4356	0.2967045	2.461063255
DKI JAKARTA	25.75	24.44	629.33	663.0625	12.707128	2.38099627
SULAWESI TENGGARA	25.77	25.85	666.1545	664.0929	4.6427574	0.023547932
PAPUA BARAT	25.8	27.07	698.406	665.64	0.8736751	1.073167942
NUSA TENGGARA TIMUR	26.14	27.22	711.5308	683.2996	0.6157633	0.703992182
KEP. RIAU	26.16	26.17	684.6072	684.3456	3.3661457	0.053528937
MALUKU	26.18	27.9	730.422	685.3924	0.0109633	2.185167339
SULAWESI BARAT	26.85	27.17	729.5145	720.9225	0.6967339	0.004178642
PAPUA	26.97	24.91	671.8227	727.3809	9.5772045	5.372156897
SUMATERA UTARA	27.28	27.24	743.1072	744.1984	0.5847751	0.09246386
KALIMANTAN BARAT	27.49	27.93	767.7957	755.7001	0.005581	0.029467889
JAWA TENGAH	27.7	28.24	782.248	767.29	0.0553633	0.071504355
JAWA TIMUR	27.78	28.53	792.5634	771.7284	0.2759339	0.226367034
SULAWESI UTARA	27.95	27.87	778.9665	781.2025	0.0181457	0.127925998
JAMBI	28.01	27.47	769.4347	784.5601	0.2859104	0.670571565
RIAU	28.06	28.34	795.2204	787.3636	0.1124221	1.03173E-08
ACEH	28.06	28.3	794.098	787.3636	0.0871986	0.001591884
KEP. BANGKA BELITUNG	28.23	28.16	794.9568	796.9329	0.0241163	0.124853617
KALIMANTAN TENGAH	28.89	29.33	847.3437	834.6321	1.7564045	0.020525654
MALUKU UTARA	29.83	29.84	890.1272	889.8289	3.3683045	0.093511816
SUMATERA BARAT	30.08	30.5	917.44	904.8064	6.2264927	0.009827264
GORONTALO	30.3	30.5	924.15	918.09	6.2264927	0.015707462
SUMATERA SELATAN	30.56	30.65	936.664	933.9136	6.997581	0.057889636
NUSA TENGGARA BARAT	30.58	32.71	1000.272	935.1364	22.139793	3.236371188
SULAWESI TENGAH	30.64	29.77	912.1528	938.8096	3.1162633	1.445345401
BANTEN	31.58	31.76	1002.981	997.2964	14.102234	0.029340256
BENGKULU	32.31	33.17	1071.723	1043.936	26.680263	0.243941546
JAWA BARAT	32.55	32.68	1063.734	1059.503	21.858375	0.058063271
LAMPUNG	33.43	34.07	1138.96	1117.565	36.787793	0.063095862

Jumlah	942.87	952.16	26682.1	26418.97	309.74705	26.81654255
Rata - Rata	27.73147	28.00471	784.7677	777.0284	9.1102073	0.78872184

Nilai a	Nilai b	Dt ²	D ²	r
-0.289205806	1.020282	309.747	26.81654	0.955732

TABEL PERSAMAAN REGRESI	
2020(X)	Y
20.5	20.62657
22.64	22.80997
23.83	24.02411
24.42	24.62607
24.89	25.10561
25.66	25.89122
25.75	25.98305
25.77	26.00345
25.8	26.03406
26.14	26.38096
26.16	26.40136
26.18	26.42177
26.85	27.10536
26.97	27.22779
27.28	27.54408
27.49	27.75834
27.7	27.9726
27.78	28.05422
27.95	28.22767
28.01	28.28888
28.06	28.3399
28.06	28.3399
28.23	28.51335
28.89	29.18673
29.83	30.1458
30.08	30.40087
30.3	30.62533
30.56	30.8906
30.58	30.91101
30.64	30.97223
31.58	31.93129
32.31	32.6761
32.55	32.92096
33.43	33.81881



SUMMARY OUTPUT

<i>Regression Statistics</i>	
Multiple R	0.955732379
R Square	0.91342438
Adjusted R Square	0.910718892
Standard Error	0.91543266
Observations	34

ANOVA

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	1	282.9305	282.9305045	337.6190695	1.45467E-18
Residual	32	26.816543	0.838016955		
Total	33	309.74705			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95,0%</i>	<i>Upper 95,0%</i>
Intercept	-0.289205806	1.5478364	-0.186845203	0.852961584	3.442045368	2.863634	-3.44205	2.863634
2020(X)	1.02028169	0.0555273	18.37441345	1.45467E-18	0.907176268	1.133387	0.907176	1.133387

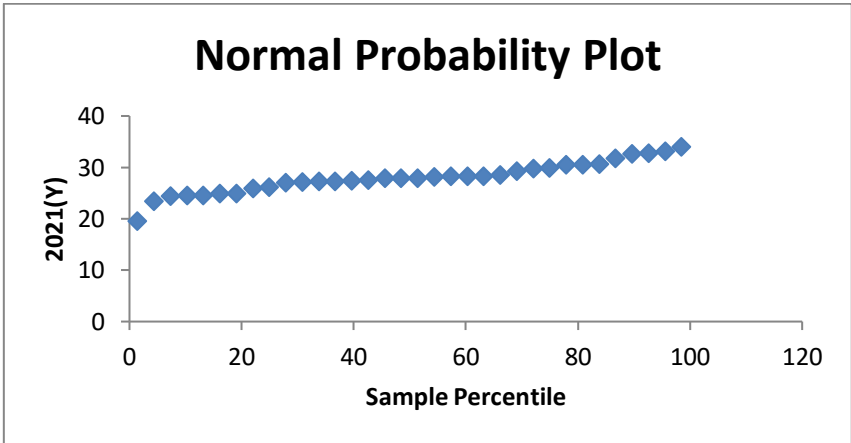
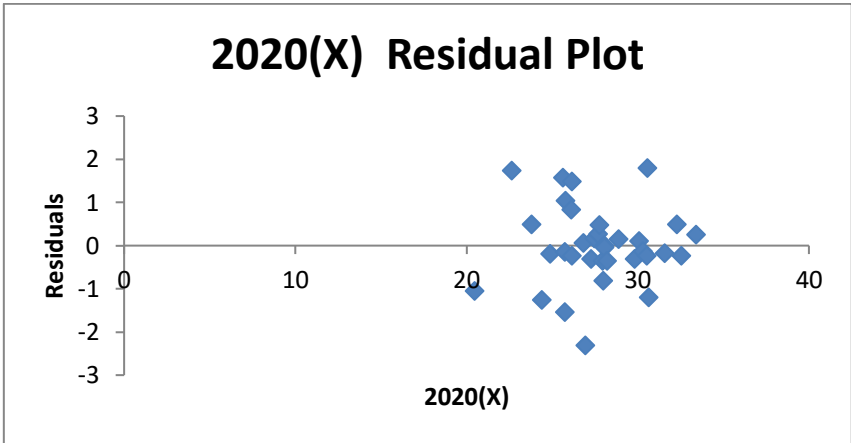
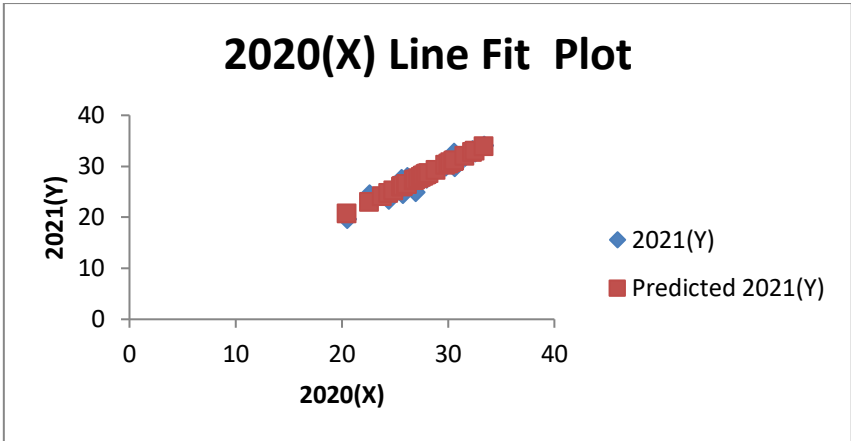
RESIDUAL OUTPUT

<i>Observation</i>	<i>Predicted 2021(Y)</i>	<i>Residuals</i>	<i>Standard Residuals</i>
1	20.62656885	-1.0465688	-1.160976372
2	22.80997166	1.7300283	1.919149444
3	24.02410688	0.4858931	0.539009391
4	24.62607307	-1.2560731	-1.393382923
5	25.10560547	-0.1956055	-0.216988425
6	25.89122237	1.5687776	1.740271333
7	25.98304772	-1.5430477	-1.711728713
8	26.00345335	-0.1534534	-0.17022838
9	26.03406181	1.0359382	1.149183611
10	26.38095758	0.8390424	0.930763826
11	26.40136321	-0.2313632	-0.256655093
12	26.42176885	1.4782312	1.639826606
13	27.10535758	0.0646424	0.07170892
14	27.22779138	-2.3177914	-2.571164849
15	27.54407871	-0.3040787	-0.337319609
16	27.75833786	0.1716621	0.190427688
17	27.97259702	0.267403	0.296634613
18	28.05421955	0.4757804	0.527791229
19	28.22766744	-0.3576674	-0.396766489
20	28.28888434	-0.8188843	-0.908402132
21	28.33989843	0.0001016	0.000112678
22	28.33989843	-0.0398984	-0.044259993
23	28.51334631	-0.3533463	-0.391972991
24	29.18673223	0.1432678	0.158929341
25	30.14579702	-0.305797	-0.33922576
26	30.40086744	0.0991326	0.109969411
27	30.62532941	-0.1253294	-0.139030019
28	30.89060265	-0.2406027	-0.266904556
29	30.91100829	1.7989917	1.995651676
30	30.97222519	-1.2022252	-1.33364856

31	31.93128998	-0.17129	-0.190014843
32	32.67609561	0.4939044	0.547896422
33	32.92096322	-0.2409632	-0.267304536
34	33.8188111	0.2511889	0.278648055

PROBABILITY OUTPUT

<i>Percentile</i>	<i>2021(Y)</i>
1.470588235	19.58
4.411764706	23.37
7.352941176	24.44
10.29411765	24.51
13.23529412	24.54
16.17647059	24.91
19.11764706	24.91
22.05882353	25.85
25	26.17
27.94117647	27.07
30.88235294	27.17
33.82352941	27.22
36.76470588	27.24
39.70588235	27.46
42.64705882	27.47
45.58823529	27.87
48.52941176	27.9
51.47058824	27.93
54.41176471	28.16
57.35294118	28.24
60.29411765	28.3
63.23529412	28.34
66.17647059	28.53
69.11764706	29.33
72.05882353	29.77
75	29.84
77.94117647	30.5
80.88235294	30.5
83.82352941	30.65
86.76470588	31.76
89.70588235	32.68
92.64705882	32.71
95.58823529	33.17
98.52941176	34.07

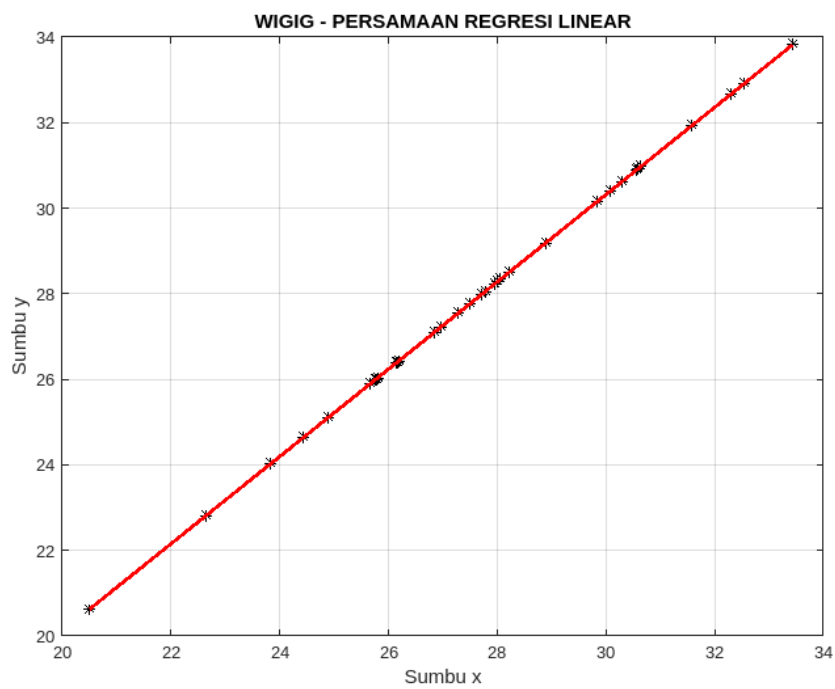
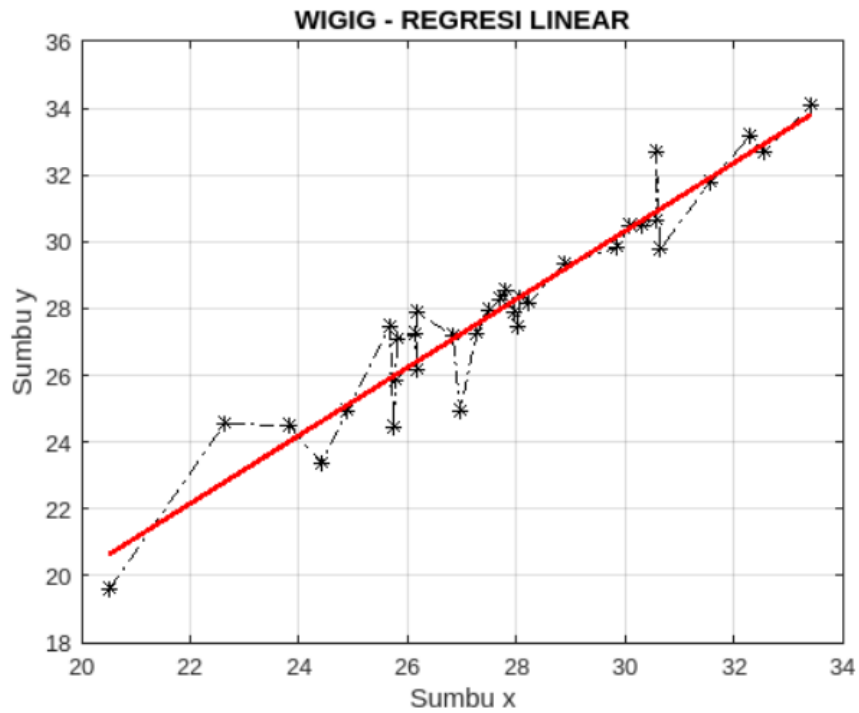


3.2. menggunakan matlab

Command Window					
=====					
TABEL REGRESI LINEAR					
=====					
X	Y	XY	X ²	(Y-ybar) ²	(Y-a-b*X) ²
20.5	19.58	401.39	420.25	70.976	1.0953
22.64	24.54	555.59	512.57	12.004	2.993
23.83	24.51	584.07	567.87	12.213	0.23609
24.42	23.37	570.7	596.34	21.48	1.5777
24.89	24.91	620.01	619.51	9.5772	0.038261
25.66	27.46	704.62	658.44	0.2967	2.4611
25.75	24.44	629.33	663.06	12.707	2.381
25.77	25.85	666.15	664.09	4.6428	0.023548
25.8	27.07	698.41	665.64	0.87368	1.0732
26.14	27.22	711.53	683.3	0.61576	0.70399
26.16	26.17	684.61	684.35	3.3661	0.053529
26.18	27.9	730.42	685.39	0.010963	2.1852
26.85	27.17	729.51	720.92	0.69673	0.0041786
26.97	24.91	671.82	727.38	9.5772	5.3722
27.28	27.24	743.11	744.2	0.58478	0.092464
27.49	27.93	767.8	755.7	0.005581	0.029468
27.7	28.24	782.25	767.29	0.055363	0.071504
27.78	28.53	792.56	771.73	0.27593	0.22637
27.95	27.87	778.97	781.2	0.018146	0.12793
28.01	27.47	769.43	784.56	0.28591	0.67057
28.06	28.34	795.22	787.36	0.11242	1.0317e-08
28.06	28.3	794.1	787.36	0.087199	0.0015919
28.23	28.16	794.96	796.93	0.024116	0.12485
28.89	29.33	847.34	834.63	1.7564	0.020526
29.83	29.84	890.13	889.83	3.3683	0.093512
30.08	30.5	917.44	904.81	6.2265	0.0098273
30.3	30.5	924.15	918.09	6.2265	0.015707
30.56	30.65	936.66	933.91	6.9976	0.05789
30.58	32.71	1000.3	935.14	22.14	3.2364
30.64	29.77	912.15	938.81	3.1163	1.4453
31.58	31.76	1003	997.3	14.102	0.02934
32.31	33.17	1071.7	1043.9	26.68	0.24394
32.55	32.68	1063.7	1059.5	21.858	0.058063
33.43	34.07	1139	1117.6	36.788	0.063096

Command Window					
=====					
JUMLAH					
=====					
X	Y	XY	X ²	(Y-ybar) ²	(Y-a-b*X) ²
942.87	952.16	26682	26419	309.75	26.817
=====					
RATA - RATA					
=====					
X	Y	XY	X ²	(Y-ybar) ²	(Y-a-b*X) ²
27.731	28.005	784.77	777.03	9.1102	0.78872
=====					
DATA LAINNYA					
=====					
a	b	DT ²	D ²	r	
-0.28921	1.0203	309.75	26.817	0.95573	

Command Window	
=====	
TABEL PERSAMAAN REGRESI LINEAR	
=====	
X	Y
20.5	20.627
22.64	22.81
23.83	24.024
24.42	24.626
24.89	25.106
25.66	25.891
25.75	25.983
25.77	26.003
25.8	26.034
26.14	26.381
26.16	26.401
26.18	26.422
26.85	27.105
26.97	27.228
27.28	27.544
27.49	27.758
27.7	27.973
27.78	28.054
27.95	28.228
28.01	28.289
28.06	28.34
28.06	28.34
28.23	28.513
28.89	29.187
29.83	30.146
30.08	30.401
30.3	30.625
30.56	30.891
30.58	30.911
30.64	30.972
31.58	31.931
32.31	32.676
32.55	32.921
33.43	33.819



4. Kesimpulan dan saran

4.1. Kesimpulan

Kesimpulan yang dapat diambil yaitu regresi linear digunakan untuk menganalisis hubungan antara variabel dependen dan independen. Dalam konteks perokok 15 tahun ke atas pada tahun 2021 dengan prediktor tahun 2020, regresi linear dapat digunakan untuk menentukan seberapa besar pengaruh tahun 2020 terhadap jumlah perokok 15 tahun ke atas pada tahun 2021. Hasil dari analisis ini dapat digunakan untuk memprediksi jumlah perokok di masa depan dan untuk mengambil tindakan preventif untuk menurunkan jumlah perokok.

4.2. Saran

Rekomendasi yang dapat diberikan berdasarkan hasil analisis ini adalah perlu adanya upaya untuk meningkatkan tingkat pendidikan masyarakat, khususnya di provinsi-provinsi dengan tingkat merokok yang tinggi. Selain itu, perlu juga adanya kebijakan yang lebih ketat dalam mengontrol penjualan rokok dan promosi rokok serta perlu adanya perhatian lebih dari pemerintah dalam memberikan edukasi tentang bahaya merokok bagi kesehatan bagi masyarakat.

5. Daftar pusaka

- [1] <https://www.who.int/indonesia>
- [2] <https://www.ugm.ac.id/id/berita/17409-jumlah-perokok-indonesia-di-atas-15-tahun-tinggi>
- [3] <https://www.bps.go.id/indicator/30/1435/1/persentase-merokok-pada-penduduk-umur-15-tahun-menurut-provinsi.html>