LAPORAN TUGAS BESAR IF2214 – Probabilitas dan Statistika

Dipersiapkan oleh:

Muhammad Ariefudin Satria Dharma 119140149

Fahri Novaldi 119140205

Jaya Megelar Cakrawarty 119140227



DAFTAR ISI

BAB I		3
PENAKS	SIRAN INTERVAL PARAMETER RATA - RATA POPULASI	3
1.1.	Latar Belakang	3
1.2.	Tabal Data Sampel	3
1.3.	Nilai Variansi Populasi	5
1.4.	Diagram Batang Data Sampel	5
1.5.	Nilai Rata-Rata, Variansi Sampel, Modus, Median, Kuartil Atas, Kuartil Bawah	6
1.6.	Penaksiran Interval.	7
1.6.1.	Penaksiran Interval Sampel dengan Tingkat Kepercayaan 95%	7
1.6.2.	Penaksiran Interval Sampel dengan Tingkat Kepercayaan 99%	7
1.7.	Kesimpulan	8
BAB II		9
PENGU	IIAN HIPOTESIS STATISTIK SATU ARAH DAN DUA ARAH	9
2.1	Latar Belakang	9
2.2	Sampel Data	9
2.3	Diagram Sampel data dan Populasi	10
2.4	Perhitungan Nilai Variansi Sampel dan Populasi	11
2.5	Perhitungan Nilai Rata-rata Sampel dan Populasi	11
2.6	Pengujian Hipotesis	12
2.6.1	Pengujian Hipotesis Satu Arah	12
2.6.2	Pengujian Hipotesis Dua Arah	13
2.7	Kesimpulan	13
BAB III		15
PENGHI SEDERH	TUNGAN KOEFISIEN KORELASI DAN PEMODELAN ANALISIS REGRESI L HANA	INIER 15
3.1	Latar Belakang	15
3.2	Data Sampel dan Populasi	
3.3	Koefisien Korelasi Data	16
3.4	Kesimpulan dari Nilai Koefisien Korelasi Data	17
3.5	Permodelan Regresi Linier	
3.6	Prediksi Regresi Linier	
3.7	Uji T	19
DAFTAI	R PUSTAKA	
I AMDID	DANI	22

BABI

PENAKSIRAN INTERVAL PARAMETER RATA - RATA POPULASI

1.1. Latar Belakang

Vitamin C adalah unsur penting yang harus dimasukkan dalam pola makan harian kita. Konsumsi sumber makanan kaya Vitamin C membantu meningkatkan imun tubuh, yang diperlukan tubuh untuk melawan penyakit. Menambah asupan makanan tinggi Vitamin C atau menggunakannya secara topikal dapat membantu memerangi masalah kulit dan memiliki kulit yang tampak lebih muda. Vitamin C kaya akan antioksidan dan bisa membantu mengurangi risiko penyakit kronis, termasuk penyakit jantung. Ketika menyebut Vitamin C, banyak orang langsung mengacu pada buah jeruk. Buah-buahan sitrus, seperti jeruk, memang merupakan salah satu sumber terbaik Vitamin C. Namun, selain jeruk, ada sejumlah buah dan sayur tinggi Vitamin C yang juga bisa dikonsumsi sehari-hari.

Kandungan nutrisi pada buah tentu memiliki nilai yang berbeda, tidak semua buah mengandung banyak Vitamin C. Dalam hal ini penerapan materi Probabilitas dan Statistika mengenai Penaksiran Interval akan dilakukan terhadap kandungan Vitamin C yang terdapat pada buah dengan mengambil 30 sampel buah yang berbeda.

1.2. Tabal Data Sampel

Data pada tabel merupakan nilai kandungan Vitamin C pada buah yang berbeda, data yang didapatkan menunjukkan setiap buat memiliki kandungan Vitamin C yang bervariasi, data kandungan Vitamin C ini didapatkan dari web nilaigizi.

No.	Nama Buah	Kandungan Vitamin C (mg)
1	Alpukat	13
2	Anggur	3
3	Apel	5
4	Belimbing	35
5	Buah Naga Merah	1
6	Cempedak	15
7	Duku	9

8	Durian	53
9	Jambu Air	5
10	Jambu Biji	87
11	Jambu Monyet	197
12	Jeruk Garut	31
13	Kedondong	32
14	Kelapa Muda	4
15	Kesemek	11
16	Lemon	50
17	Mangga	12
18	Mangga Golek	65
19	Manggis	5
20	Markisa	10
21	Matoa	54
22	Nanas	22
23	Pepaya	78
24	Piasang Ambon	9
25	Pisang Kepok	9
26	Rambutan	58
27	Salak	2
28	Semangka	6
29	Sirsak	20
30	Srikaya	30

Tabel 1. Data Sampel Vitamin C pada Buah

1.3. Nilai Variansi Populasi

Beradasarkan perhitungan menggunakan exel dengan menggunakan fungsi "=VAR.S" didapatkan nilai varansi sampel sebesar 1569,689 = 1570. Untuk nilai variansi populasi, kami melakukan klaim bahwa nilainya akan lebih besar dibandingkan dengan nilai variansi sampel yang telah didapatkan. Maka dari itu, nilai variansi populasi yang kami gunakan adalah 1580.

1.4. Diagram Batang Data Sampel

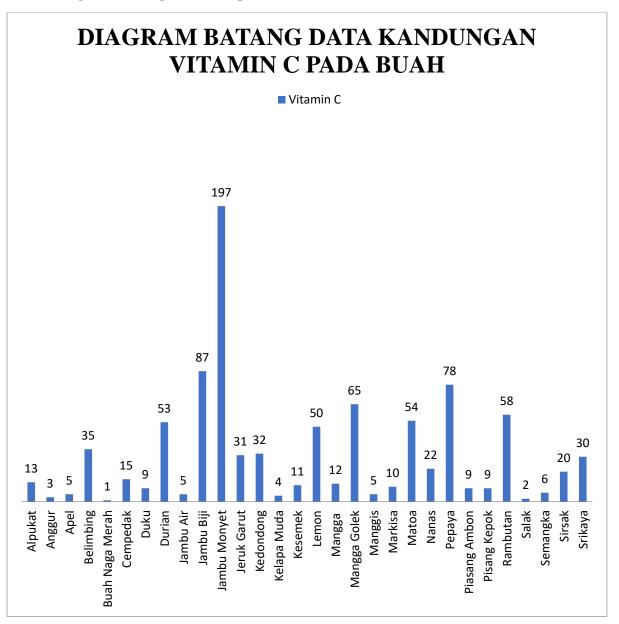


Diagram 1. Diagram Batang Sampel data

1.5. Nilai Rata-Rata, Variansi Sampel, Modus, Median, Kuartil Atas, Kuartil Bawah

Dalam mencari nilai rata rata dari data Kandungan Vitamin C digunakan fungsi pada Excel, yaitu dengan menggunakan fungsi "=AVERAGE". Pada penggunaan fungsi ini didapatkan nilai rata rata sebesar 31, 033 = 31

$$\bar{x} = \frac{\sum_{i=1}^{n} x^{i}}{n} = \frac{13+3+5+35+1+15+9+...+30}{30} = \frac{931}{30} = 31,033 = 31$$

Untuk mencari nilai variansi sampel dari Kandungan Vitamin C pada buah digunakan fungsi pada Exel, yaitu dengan menggunakan fungsi " = VAR.S". Dari penggunaan fungsi tersebut didapatkan nilai variansi sampel sebesar 1569,689 = 1570

$$\sigma^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n - 1}$$

$$\sigma^2 = \frac{(13 - 31,033)^2 + (3 - 31,033)^2 + \dots + (30 - 31,033)^2}{30 - 1} = 1569,689 = 1570$$

Untuk mencari modus digunakan fungsi pada Exel, yaitu dengan menggunakan fungsi "=MODE". Dari penggunaan fungsi ini didapatkan nilai modus sebesar 5.

Dalam mencari nilai rata rata dari data Kandungan Vitamin C digunakan fungsi pada Exel, yaitu dengan menggunakan fungsi "=MEDIAN". Pada penggunaan fungsi ini didapatkan nilai rata rata sebesar 14

$$Me = \frac{1}{2} \left(X_{\frac{n}{2}} + X_{\left(\frac{n}{2}+1\right)} \right) = \frac{1}{2} \left(\frac{30}{2} + \left(\frac{30}{2} + 1 \right) \right) = \frac{1}{2} (15 + 16) = 15,5$$

Data ke
$$15.5 = \frac{1}{2}(13+15) = 14$$

Pada pencarian nilai maksimum dan minimum, digunakan fungsi excel "=MAX" untuk mencari nilai maksimum dan "=MIN" untuk mencari nilai minimum. Hasil yang didapatkan dalam pencarian menggunakan fungsi ini adalah nilai maksimum sebesar 197 dan nilai minimum sebesar 1.

Berdasarkan perhitungan yang telah dilakukan dari data yang telah didapatkan, besar nilai quartil bawah sebesar 6

$$Q1 = x\left(\frac{n+2}{4}\right)$$
 untuk n genap dan $n+1$ tidak habis dibagi 4

$$Q1 = x\left(\frac{30+2}{4}\right) = x(8) \ data \ ke \ 8 = 6$$

Dan untuk besar nilai kuartil atas dari data yang telah dikumpulkan, didapatkan nilai quartil atas sebesar 50

$$Q3 = x\left(\frac{3n+2}{4}\right) untuk \ n \ genap \ dan \ n+1 \ tidak \ habis \ dibagi \ 4$$

$$Q3 = x\left(\frac{3.30 + 2}{4}\right) = x(23) \ data \ ke \ 23 = 50$$

1.6. Penaksiran Interval

Variansi Sampel:

$$\sigma^2 = 1569.689$$

Standar Deviasi Populasi:

$$\sigma = \sqrt{1569,689} = 39,619$$

$$P\left(\bar{x} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

1.6.1. Penaksiran Interval Sampel dengan Tingkat Kepercayaan 95%

Nilai taksiran μ menggunakan rerata sampel yang bernilai 31. Sedangkan untuk nilai α dapat dihasilkan dari persamaan 1 – α = 95% sehingga α bernilai 0,05. Nilai z dengan luas 0,025 berada di bagian kanan dan 0,975 di bagian kiri adalah $\frac{Z_{0,05}}{2} = Z_{0,025} = 1,96$ sehingga selang kepercayaan 95% yang yang dicari dapat menggunakan persamaan :

$$\left(\bar{x} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right)$$

$$\left(31 - (1,96)\frac{39,619}{\sqrt{30}} < \mu < 31 + (1,96)\frac{39,619}{\sqrt{30}}\right)$$

Persamaan tersebut akan menghasilkan interval:

$$16,822 < \mu < 45,177$$

1.6.2. Penaksiran Interval Sampel dengan Tingkat Kepercayaan 99%

Nilai taksiran μ menggunakan nilai rerata sampel bernilai 31 Sedangkan untuk nilai α dapat dihasilkan dari persamaan $1-\alpha=99\%$ sehingga ∞ bernilai 0,01. Nilai z yang memiliki luas 0.005 berada di bagian kanan dan 0.995 di bagian kiri adalah $\frac{Z_{0,01}}{2}=Z_{0,005}=2,575$ sehingga selang kepercayaan 99% yang dicari dapat menggunakan persamaan :

$$\left(\bar{x} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right)$$

$$\left(31 - (2,575) \frac{39,619}{\sqrt{30}} < \mu < 31 + (2,575) \frac{39,619}{\sqrt{30}}\right)$$

Persamaan tersebut akan menghasilkan interval:

$$12,373 < \mu < 49,626$$

1.7. Kesimpulan

Berdasarkan hasil pencarian pada selang interval dengan tingkat kepercayaan yang berbeda (95% dan 99%), dapat disimpulkan bahwa semakin tinggi tingkat kepercayaan maka interval yang dihasilkan semakin besar. Selain itu, rerata sampel dan variansi juga berpengaruh terhadap penaksiran suatu interval. Pada kasus ini, nilai kandungan Vitamin C yang bervariasi menyebabkan interval yang cukup luas.

BAB II

PENGUJIAN HIPOTESIS STATISTIK SATU ARAH DAN DUA ARAH

2.1 Latar Belakang

Tekanan darah adalah tekanan dari sirkulasi darah terhadap dinding syaraf, hal ini disebabkan oleh jantung yang memompa darah. Tekanan darah dikatakan normal jika memiliki nilai kurang dari 120. Angka tekanan darah yang tidak normal dapat berpengaruh pada kesehatan seseorang dengan menimbulkan banyak penyakit. Oleh karena itu kami berniat untuk mengunakan data tekanan darah orang diumur 30-50 tahun dan melakukan uji hipotesis pada data tersebut.

2.2 Sampel Data

Data pada tabel berikut ini merupakan sampel data tekanan darah (*resting blood pressure in mmHg*) pada rentang umur 30-50 tahun. sampel ini di dapatkan dari sebuah forum penyedia dataset untuk kebutuhan riset yaitu Kaggle. dataset yang disediakan berisi 303 sampel data namun, untuk memudahkan pengolahan data disarankan untuk menggunakan hanya 30 sampel secara acak.

		T		1
Sampel	Umur	Tekanan Darah	$(x_i - \overline{x})$	$(x_i - \overline{x})^2$
1		$(mmHg)(x_i)$		
A	50	120	-4.86207	23.63971
В	50	129	4.137931	17.12247
С	50	140	15.13793	229.157
D	50	144	19.13793	366.2604
Е	49	130	5.137931	26.39834
F	48	130	5.137931	26.39834
G	47	110	-14.8621	220.8811
Н	47	112	-12.8621	165.4328
Ι	46	142	17.13793	293.7087
J	45	110	-14.8621	220.8811
K	44	140	15.13793	229.157
L	44	108	-16.8621	284.3294
M	44	120	-4.86207	23.63971
N	44	112	-12.8621	165.4328
О	44	120	-4.86207	23.63971
P	42	136	11.13793	124.0535
Q	42	148	23.13793	535.3639
R	41	112	-12.8621	165.4328
S	41	130	5.137931	26.39834
T	41	112	-12.8621	165.4328
U	41	110	-14.8621	220.8811
V	40	110	-14.8621	220.8811

W	38	138	13.13793	172.6052
X	38	138	13.13793	172.6052
Y	35	138	13.13793	172.6052
Z	35	120	-4.86207	23.63971
AA	35	126	1.137931	1.294887
AB	34	118	-6.86207	47.08799
AC	34	118	-6.86207	47.08799
RATA-				
RATA	42.72414	124.8621		
VARIANSI	26.27833	157.5517		

Tabel 2. Sampel data Tekanan darah rentang 30 – 50 tahun

2.3 Diagram Sampel data dan Populasi

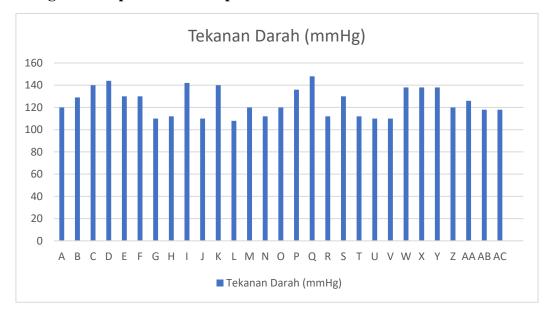


Diagram 2. Diagram Batang Sampel data

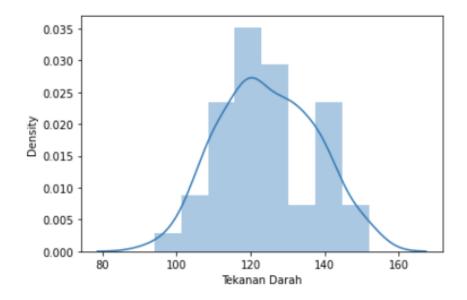


Diagram 3. Diagram Batang Populasi data dengan densitas sampel

2.4 Perhitungan Nilai Variansi Sampel dan Populasi

Untuk mendapatkan nilai variansi data sampel acak data tekanan darah yang diambil dari populasi, persamaan matematis yang akan digunakan untuk mencari variansi sample adalah sebagai berikut

$$S^{2} = \frac{\sum_{i=1}^{n} (x_{i} - \overline{x})^{2}}{n-1}$$

Dengan bantuan fungsi "VAR.S" dari excel, Sehingga didapatkan nilai $S^2 = 157.5517$ sebagai variansi sampel data tekanan darah.

Kemudian Untuk mendapatkan nilai variansi populasi, persamaan matematis yang akan digunakan untuk mencari variansi sample adalah sebagai berikut

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \mu)^2}{N}$$

Berbeda dengan variansi sampel, perhitungan variansi populasi sangatlah panjang. Penyebabnya adalah jumlah populasi yang digunakan sebanyak 94 data, kami menggunakan bantuan bahasa pemrograman python untuk memudahkan pengolahan data variansi populasi karena sangat tidak efisien jika dimuat langsung pada dokumen ataupun spreadsheet. Maka diperoleh lah nilai variansi populasi $\sigma^2 = 161.252001$

2.5 Perhitungan Nilai Rata-rata Sampel dan Populasi

Untuk mendapatkan nilai rata-rata sampel acak data tekanan darah yang diambil dari populasi, persamaan matematis yang akan digunakan untuk mencari nilai rata-rata sample adalah

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Dengan bantuan fungsi "AVERAGE" dari excel, Sehingga didapatkan nilai $\bar{x}=124.862$ sebagai variansi sampel data tekanan darah.

Kemudian Untuk mendapatkan nilai rata-rata populasi, persamaan matematis yang akan digunakan untuk mencari variansi sample adalah sebagai berikut

$$\mu = \frac{\sum_{i=1}^{n} x_i}{N}$$

Berbeda dengan variansi sampel, perhitungan nilai rata-rata populasi sangatlah panjang. Penyebabnya adalah jumlah populasi yang digunakan sebanyak 94 data, kami menggunakan bantuan bahasa pemrograman python untuk memudahkan pengolahan data variansi populasi karena sangat tidak efisien jika dimuat langsung pada dokumen ataupun spreadsheet. Maka diperoleh lah nilai variansi populasi $\mu=124.606383$.

2.6 Pengujian Hipotesis

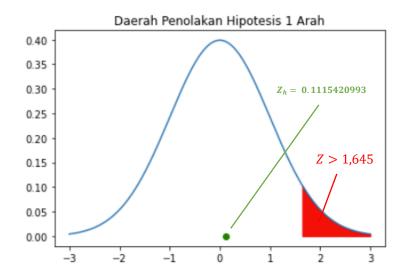
2.6.1 Pengujian Hipotesis Satu Arah

Dari nilai rata-rata sampel yang diperoleh maka dapat diketahui rumusan hipotesis statistiknya untuk mengecek apakah data rata-rata laju pertumbuhan penduduk yang ada benar-benar valid yaitu :

$$H_0: \mu = 124.606383$$

$$H_1: \mu > 124.606383$$

Dengan tingkat kepercayaan 95% atau signifikansi yaitu $\alpha=5\%$. Nilai kritisnya adalah $Z_{0.05}=1,645$ yang didapat dari tabel distribusi normal. Dengan sampel berukuran n=30, simpangan baku sampel $S=\sqrt{S^2}=\sqrt{157.5517}=12.55196$ dan rata-rata sampel $\bar{x}=124.862$, maka pengujian ini dapat kita cari dengan cara sebagai berikut



$$Z_h = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{124.862 - 124.606383}{12.55196/\sqrt{30}} = 0.1115420993$$

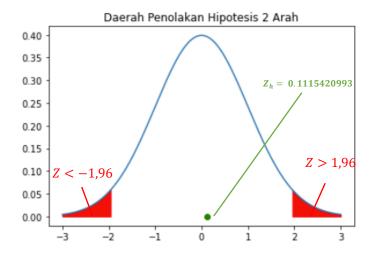
2.6.2 Pengujian Hipotesis Dua Arah

Dari nilai rata-rata sampel yang diperoleh maka dapat diketahui rumusan hipotesis statistiknya untuk mengecek apakah data rata-rata laju pertumbuhan penduduk yang ada benar-benar valid yaitu :

$$H_0: \mu = 124.606383$$

$$H_1: \mu \neq 124.606383$$

Dengan tingkat kepercayaan 95% atau signifikansi yaitu $\alpha=5\%$. Maka Nilai kritis yang akan diuji dari 2 arah sebesar $\alpha_{/2}=2.5\%$, Nilai kritisnya adalah $Z_{0.025}=1,96$ yang didapat dari tabel distribusi normal. Dengan sampel berukuran n=30, simpangan baku sampel $S=\sqrt{S^2}=\sqrt{157.5517}=12.55196$ dan rata-rata sampel $\bar{x}=124.862$, maka pengujian ini dapat kita cari dengan cara sebagai berikut



$$Z_h = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{124.862 - 124.606383}{12.55196/\sqrt{30}} = 0.1115420993$$

2.7 Kesimpulan

Berdasarkan hasil pengujian hipotesis yang dilakukan secara satu arah, nilai $Z_h=0.1115420993$ berada di luar daerah penolakan hipotesis, sehingga dapat disimpulkan hipotesis H_0 adalah benar dan menyebabkan H_1 mengalami penolakan. Begitu juga dengan pengujian hipotesis dua arah, walaupun daerah penolakannya berubah nilai $Z_h=0.1115420993$ masih berada di luar daerah penolakan hipotesis. Sehingga H_1 masih mengalami penolakan.

Kesimpulan yang dapat kita tarik adalah pengambilan sampel dan perhitungan rata-rata tekanan darah tersebut adalah valid dan sudah teruji secara 1 arah maupun 2 arah.

BAB III

PENGHITUNGAN KOEFISIEN KORELASI DAN PEMODELAN ANALISIS REGRESI LINIER SEDERHANA

3.1 Latar Belakang

Kebahagiaan bagi setiap penduduk negara merupakan hal yang sangat penting untuk suatu negara, oleh karena itu United Nations Sustainable Development Solutions Network (SDSN) membuat laporan nilai kebahagiaan negara-negara di seluruh dunia yang diantaranya berisikan *Happiness Score* dan GDP per capita yang setelah kami perhatikan negara dengan GDP yang rendah maka happiness score nya pun ikut rendah begitu juga sebaliknya. Oleh karena itu kami ingin membuat model menggunakan metode regresi linier sederhana untuk memprediksi memprediksi *Happiness Score* yang dimiliki Suatu negara berdasarkan GDP per capita nya.

3.2 Data Sampel dan Populasi

Index	Country or region	GDP per capita (X)	Happiness Score (Y)	<i>X</i> ²	<i>Y</i> ²	XY
92	China	1.029	5.191	1.058841	26.94648	5.341539
27	Saudi Arabia	1.403	6.375	1.968409	40.64063	8.944125
150	Yemen	0.287	3.38	0.082369	11.4244	0.97006
65	Portugal	1.221	5.693	1.490841	32.41025	6.951153
43	Slovenia	1.258	6.118	1.582564	37.42992	7.696444
41	Lithuania	1.238	6.149	1.532644	37.8102	7.612462
130	Myanmar	0.71	4.36	0.5041	19.0096	3.0956
123	Tunisia	0.921	4.461	0.848241	19.90052	4.108581
112	Namibia	0.879	4.639	0.772641	21.52032	4.077681
9	Austria	1.376	7.246	1.893376	52.50452	9.970496
121	Mauritania	0.57	4.49	0.3249	20.1601	2.5593
144	Burundi	0.046	3.775	0.002116	14.25063	0.17365
38	Trinidad & Tobago	1.231	6.192	1.515361	38.34086	7.622352
137	Zambia	0.578	4.107	0.334084	16.86745	2.373846
98	Ivory Coast	0.569	4.944	0.323761	24.44314	2.813136
147	Botswana	1.041	3.488	1.083681	12.16614	3.631008
63	Northern Cyprus	1.263	5.718	1.595169	32.69552	7.221834
52	Latvia	1.187	5.94	1.408969	35.2836	7.05078
59	Kazakhstan	1.173	5.809	1.375929	33.74448	6.813957
128	Sierra Leone	0.268	4.374	0.071824	19.13188	1.172232
58	Honduras	0.642	5.86	0.412164	34.3396	3.76212
125	Iraq	1.043	4.437	1.087849	19.68697	4.627791
62	Paraguay	0.855	5.743	0.731025	32.98205	4.910265
54	Estonia	1.237	5.893	1.530169	34.72745	7.289641
80	Belarus	1.067	5.323	1.138489	28.33433	5.679641
13	Luxembourg	1.609	7.09	2.588881	50.2681	11.40781
122	Mozambique	0.204	4.466	0.041616	19.94516	0.911064
16	Germany	1.373	6.985	1.885129	48.79023	9.590405
73	Tajikistan	0.493	5.467	0.243049	29.88809	2.695231

36	Bahrain	1.362	6.199	1.855044	38.4276	8.443038
----	---------	-------	-------	----------	---------	----------

Tabel 3. Sampel data (30) happiness score yang diambil dari dataset

Untuk data populasinya tidak dapat dimuat dalah dokumen secara lengkap, namun akan kami lampirkan dalam lampiran.

	GDP per capita	Score
count	156.000000	156.000000
mean	0.905147	5.407096
std	0.398389	1.113120
min	0.000000	2.853000
25%	0.602750	4.544500
50%	0.960000	5.379500
75%	1.232500	6.184500
max	1.684000	7.769000

Tabel 4. Deskripsi ringkas tentang Populasi data

3.3 Koefisien Korelasi Data

Untuk mencari Koefisien korelasi data rumus yang digunakan adalah adalah :

$$r = \frac{n \sum XY - \sum X \sum Y}{\sqrt{[n(\sum X^2) - (\sum X)^2][n(\sum Y^2) - (\sum Y)^2]}}$$

Dalam perhitungan Koefisien korelasi data, kami menggunakan bahasa pemrograman python agar hasil perhitungan lebih presisi dan efisien, berikut adalah hasil perhitungan yang di dapatkan

$$\Sigma XY = 818.066182$$
 $\Sigma X = 141.203$
 $\Sigma Y = 843.507$
 $\Sigma X^2 = 152.41022900000002$
 $(\Sigma X)^2 = 19938.287209000002$
 $\Sigma Y^2 = 4752.974011$
 $(\Sigma Y)^2 = 711504.0590489999$
 $\Sigma Y = 0.7938828678781272$

$$r = \frac{156 \left(818.066182\right) - \left(141.203\right) \left(843.507\right)}{\sqrt{\left[156 \left(152.410229\right) - \left(19938.287209\right)\right] \left[156 \left(4752.974011\right) - \left(711504.059048\right)\right]}}$$

 $r \approx 0.7938828678781272$

3.4 Kesimpulan dari Nilai Koefisien Korelasi Data

Berdasarkan perhitungan diperoleh hasil koefisien korelasi data adalah $r \approx 0.7938828678781272$. Maka jika dilihat dari tabel 5 dapat diperoleh kesimpulan bahwa nilai koefisien relasi antara Nilai *Gross Domestic Product* atau GDP suatu negara dengan *Happiness Score* warga negara tersebut berada di antara interval 0.600-0.799 yang artinya keduanya memiliki tingkat hubungan yang Kuat. Sehingga salah satu variable ini bisa dijadikan sebagai *predictor*.

Interval Koefisien	Tingkat Hubungan
0,000-0,199	Sangat Rendah
0,200-0,399	Rendah
0,400-0,599	Sedang
0,600-0,799	Kuat
0,800-1,000	Sangat Kuat

Tabel 5. Interpretasi koefisien korelasi

3.5 Permodelan Regresi Linier

Dalam pemodelan regresi linier dapat dilakukan dengan menggunakkan cara di bawah ini. Regresi linier dapat diperoleh dengan persamaan:

$$\hat{Y} = a + bX$$

I. nilai a dapat diperoleh dari persamaan

$$a = \frac{\sum Y \sum X^2 - \sum X \sum XY}{n(\sum X^2) - (\sum X)^2}$$

Dalam perhitungan ini, kami menggunakan bahasa pemrograman python agar hasil perhitungan lebih presisi dan efisien, berikut adalah hasil perhitungan yang di dapatkan

$$a = \frac{(843.507)(152.410229) - (141.203)(818.066182)}{156(152.410229) - (711504.0590489999)} \approx 3.399345178292416$$

II. nilai *b* dapat diperoleh dari persamaan

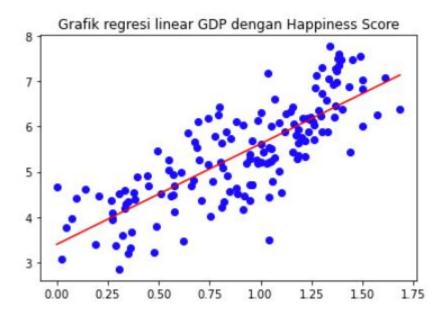
$$b = \frac{n\sum XY - \sum X\sum Y}{n(\sum X^2) - (\sum X)^2}$$

Dalam perhitungan ini, kami menggunakan bahasa pemrograman python agar hasil perhitungan lebih presisi dan efisien, berikut adalah hasil perhitungan yang di dapatkan

$$b = \frac{156(818.066182) - (141.203)(843.507)}{156(152.410229) - (711504.0590489999)} \approx 2.2181480010083603$$

Sehingga model regresi linier yang diperoleh adalah dari nilai $a \approx 3.399345178292416$ dan $b \approx 2.2181480010083603$ akan disubstitusi ke persamaan regresi linear yang dapat digunakan untuk memprediksi nilai Y, sehingga menjadi

$$\hat{Y} = 3.399345178292416 + 2.2181480010083603X$$



Jika diplot ke dalam Grafik, garis persamaan regresi linear dengan titik populasi tidak begitu tersebar jauh, sehingga persamaan regresi linear ini cukup reliabel

3.6 Prediksi Regresi Linier

Untuk menguji kehandalan model regresi ini, maka kita dapat menggunakan salah satu sample dengan studi kasus berikut :

"Pada tahun 2019, Gross Domestic Produk Indonesia adalah 0.931 per kapita. Berapakah index kesejahteraan penduduk Indonesia pada tahun 2019?"

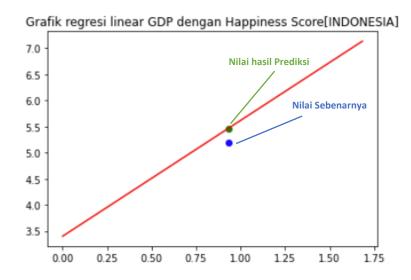
Dari contoh kasus ini, kita dapat memprediksi index kesejahteraan penduduk dengan menggunakan persamaan regresi linear,

$$\hat{Y} = 3.399345178292416 + (2.2181480010083603)(0.931)$$

$$\hat{Y} = 5.464441$$

Menurut hasil prediksi dari persamaan ini, Index kesejahteraan penduduk Indonesia adalah 5.464441.

Actual Happiness Score Indonesia = 5.192 Predicted Happiness Score Indonesia = 5.464441



Jika dilihat dari data sebenarnya, kesejahteraan penduduk Indonesia adalah 5.192. bisa dilihat pada grafik, hasil prediksi menunjukan nilai yang hampir mendekati nilai sebenarnya. Maka bisa disimpulkan, hasil prediksi dari persamaan regresi linear ini cukup reliabel.

3.7 Uji T

Kemudian Langkah selanjutnya adalah uji T, Hipotesis yang diajukan adalah

 $H_0: \beta = 0$; Nilai GDP tidak berpengaruh terhadap happiness score

 $H_1: \beta \neq 0$; Nilai GDP berpengaruh terhadap happiness score

Dengan parameter a = 0.05; r = 0.7938828678781272;

$$df = n - k = 156 - 2 = 154$$

Maka dapat diperoleh

$$t_{hitung} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.7938828678781272\sqrt{156-2}}{\sqrt{1-0.7938828678781272^2}} = 21.70000054478138$$

Dengan menggunakan tabel uji T untuk taraf signifikansi a=0.05 dan df=154 maka memperoleh nilai $t_{table}\approx 1.66$.

Membandingkan t_{hitung} dengan t_{table} ,

$$t_{hitung} > t_{table} \rightarrow 21.7 > 1.66$$

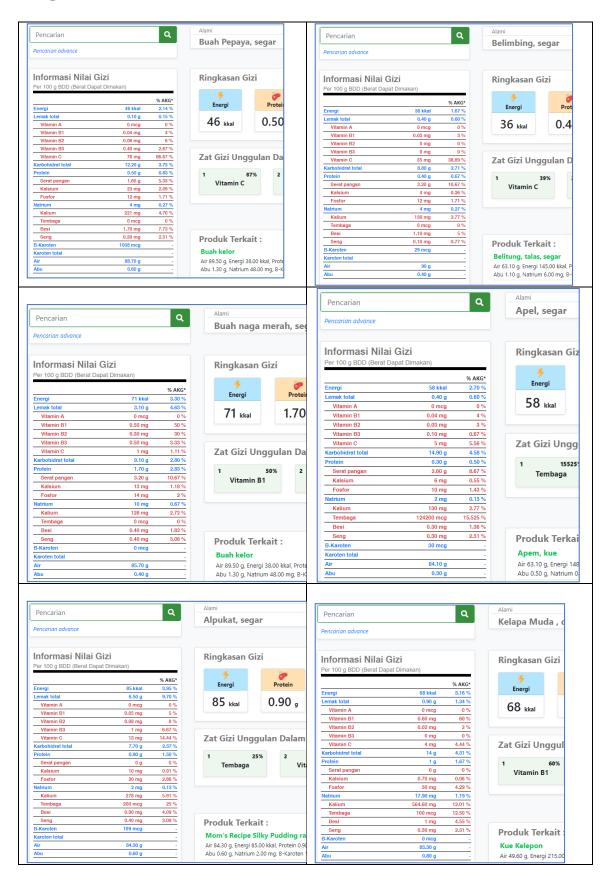
Kesimpulannya adalah nilai $t_{hitung} > t_{table}$ sehingga bisa dikatakan bahwa adanya pengaruh nyata dan signifikan variable predictor yaitu GDP terhadap variable response yaitu happiness score dengan taraf signifikansi 5%.

DAFTAR PUSTAKA

- [1] "Nilai Gizi," [Online]. Available: https://nilaigizi.com/ . [Diakses 30 April 2021].
- [2] "Heart Decease UCL," [Online]. Available: https://www.kaggle.com/ronitf/heart-disease-uci . [Diakses 30 April 2021].
- [3] "World Happiness Report," [Online]. Available: https://www.kaggle.com/unsdsn/world-happiness?select=2019.csv . [Diakses 30 April 2021].
- [4] "DataFrame," [Online]. Available: https://pandas.pydata.org/pandas-docs/stable/reference/frame.html [Diakses 30 April 2021].

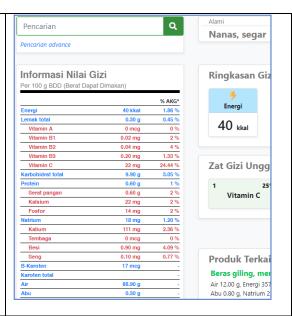
LAMPIRAN

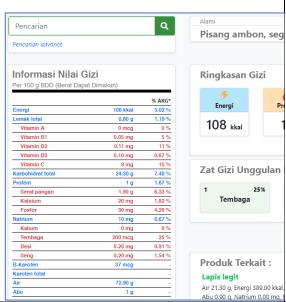
1. Lampiran BAB I – Penaksiran Data



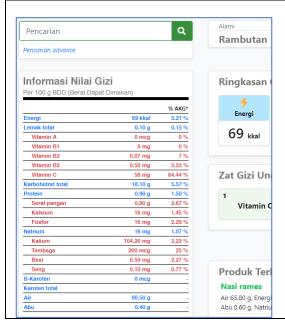


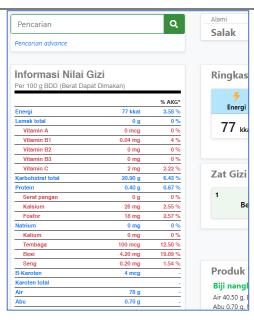


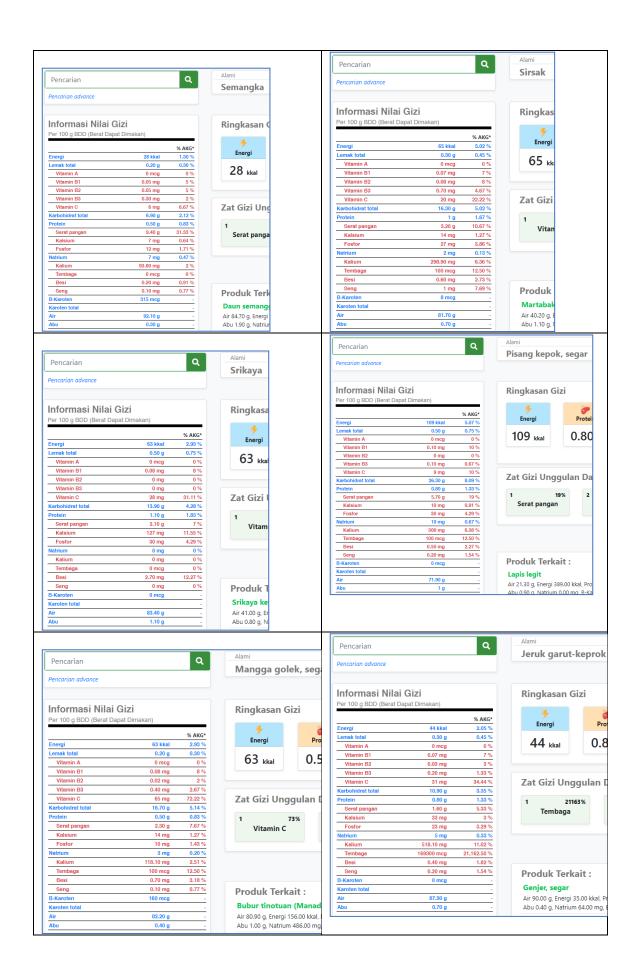


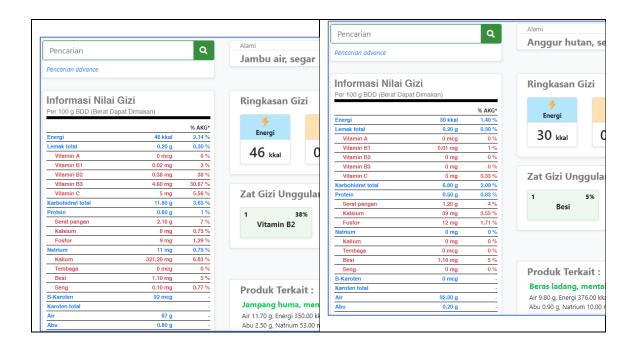






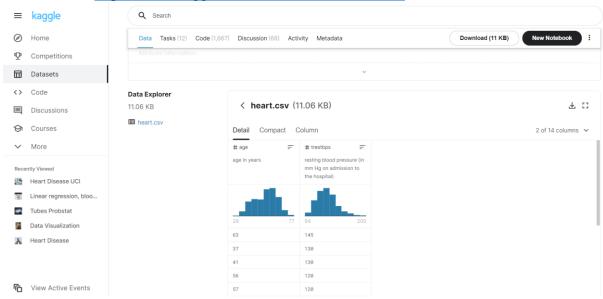






2. Lampiran BAB II – PENGUJIAN HIPOTESIS STATISTIK SATU ARAH DAN DUA ARAH

Link dataset: https://www.kaggle.com/ronitf/heart-disease-uci



1	Α	В	С	D	E	F	G	Н	I	J	K	L	М
1	Sampel	Umur	Tekanan Darah				Sampel	Tekanan Darah	(xi-xbar)	(xi- xbar)^2			
2	Α	50	120	AVG	124.862		Α	120	-4.86207	23.6397			
3	В	50	129	VARS	157.552		В	129	4.13793	17.1225			
4	С	50	140				С	140	15.1379	229.157			
5	D	50	144				D	144	19.1379	366.26			
6	E	49	130				E	130	5.13793	26.3983			
7	F	48	130				F	130	5.13793	26.3983			
8	G	47	110				G	110	-14.8621	220.881			
9	Н	47	112				Н	112	-12.8621	165.433			
10	ı	46	142				I	142	17.1379	293.709			
11	J	45	110				J	110	-14.8621	220.881			
12	K	44	140				K	140	15.1379	229.157			
13	L	44	108				L	108	-16.8621	284.329			
14	M	44	120				M	120	-4.86207	23.6397			
15	N	44	112				N	112	-12.8621	165.433			
16	0	44	120				0	120	-4.86207	23.6397			
17	Р	42	136				Р	136	11.1379	124.054			
18	Q	42	148				Q	148	23.1379	535.364			
19	R	41	112				R	112	-12.8621	165.433			
20	S	41	130				S	130	5.13793	26.3983			
21	Т	41	112				T	112	-12.8621	165.433			
22	U	41	110				U	110	-14.8621	220.881			
23	V	40	110				V	110	-14.8621	220.881			
24	W	38	138				W	138	13.1379	172.605			
25	X	38	138				X	138	13.1379	172.605			
26	Υ	35	138				Υ	138	13.1379	172.605			
27	Z	35	120				Z	120	-4.86207	23.6397			
28	AA	35	126				AA	126	1.13793	1.29489			
29	AB	34	118				AB	118	-6.86207	47.088			
30	AC	34	118				AC	118	-6.86207	47.088			
31	AVG	42.7241	124.862										
32	VARS	26.2783	157.552										
33													
34													

^{*}perhitungan sample dataset

Out[4]:

	age	trestbps
count	94.000000	94.000000
mean	43.553191	124.606383
std	4.094625	12.698504
min	34.000000	94.000000
25%	41.000000	115.750000
50%	44.000000	122.000000
75%	46.750000	133.500000
max	50.000000	152.000000

```
In [6]: variansi_bps = data_clean.var()['trestbps']
    average_bps = data_clean.mean()['trestbps']
    print("Variansi Populasi Tekanan darah = ", str(variansi_bps))
    print("Rata-rata Populasi Tekanan darah = ", str(average_bps))
```

Variansi Populasi Tekanan darah = 161.25200183024486 Rata-rata Populasi Tekanan darah = 124.6063829787234

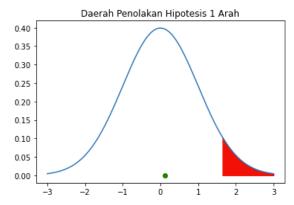
*perhitungan populasi dari dataset

```
In [7]: #perhitungan Z
zh = 124.862-124.606383
dzh = 12.55196 / (30**0.5)
print(zh/dzh)
```

0.11154209938670816

Perhitungan z

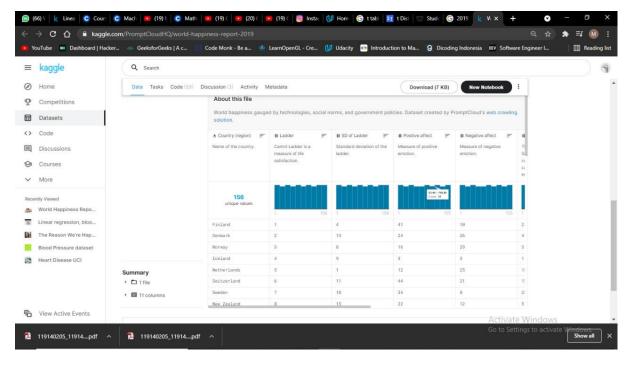
```
In [11]: x = np.arange(-3,3,0.0001)
z0 = 1.645
z_hypo = 0.1115420993
draw_z_score(x, x>z0, 0, 1, 'Daerah Penolakan Hipotesis 1 Arah', z_hypo)
```



Pengambaran grafik uji hipotesis 1 arah

3. Lampiran BAB III – PENGHITUNGAN KOEFISIEN KORELASI DAN PEMODELAN ANALISIS REGRESI LINIER SEDERHANA

Link dataset: https://www.kaggle.com/unsdsn/world-happiness?select=2019.csv



/_	Α	В	С	D	E	F	G	Н
1	Cou	intry or re	P per cap	Score	X^2	Y^2	XY	
2	92	China	1.029	5.191	1.05884	26.9465	5.34154	
3	27	Saudi Ara	1.403	6.375	1.96841	40.6406	8.94413	
4	150	Yemen	0.287	3.38	0.08237	11.4244	0.97006	
5	65	Portugal	1.221	5.693	1.49084	32.4102	6.95115	
6	43	Slovenia	1.258	6.118	1.58256	37.4299	7.69644	
7	41	Lithuania	1.238	6.149	1.53264	37.8102	7.61246	
8	130	Myanmar	0.71	4.36	0.5041	19.0096	3.0956	
9	123	Tunisia	0.921	4.461	0.84824	19.9005	4.10858	
10	112	Namibia	0.879	4.639	0.77264	21.5203	4.07768	
11	9	Austria	1.376	7.246	1.89338	52.5045	9.9705	
12	121	Mauritan	0.57	4.49	0.3249	20.1601	2.5593	
13	144	Burundi	0.046	3.775	0.00212	14.2506	0.17365	
14	38	Trinidad {	1.231	6.192	1.51536	38.3409	7.62235	
15	137	Zambia	0.578	4.107	0.33408	16.8674	2.37385	
16	98	Ivory Coa:	0.569	4.944	0.32376	24.4431	2.81314	
17	147	Botswana	1.041	3.488	1.08368	12.1661	3.63101	
18	63	Northern	1.263	5.718	1.59517	32.6955	7.22183	
19	52	Latvia	1.187	5.94	1.40897	35.2836	7.05078	
20	59	Kazakhsta	1.173	5.809	1.37593	33.7445	6.81396	
21	128	Sierra Leo	0.268	4.374	0.07182	19.1319	1.17223	
22	58	Honduras	0.642	5.86	0.41216	34.3396	3.76212	
23	125	Iraq	1.043	4.437	1.08785	19.687	4.62779	
24	62	Paraguay	0.855	5.743	0.73103	32.982	4.91027	
25	54	Estonia	1.237	5.893	1.53017	34.7274	7.28964	
26	80	Belarus	1.067	5.323	1.13849	28.3343	5.67964	
27	13	Luxembou	1.609	7.09	2.58888	50.2681	11.4078	
28	122	Mozambio	0.204	4.466	0.04162	19.9452	0.91106	
29	16	Germany	1.373	6.985	1.88513	48.7902	9.59041	
30	73	Tajikistar	0.493	5.467	0.24305	29.8881	2.69523	
31	36	Bahrain	1.362	6.199	1.85504	38.4276	8.44304	
32								
33								
24								

		GDP				
	Country or region	per capita (X)	Happiness Score (Y)	X^2	Y^2	XY
0	Finland	1.34	7.769	1.7956	60.35736	10.41046
1	Denmark	1.383	7.6	1.912689	57.76	10.5108
2	Norway	1.488	7.554	2.214144	57.06292	11.24035
3	Iceland	1.38	7.494	1.9044	56.16004	10.34172
4	Netherlands	1.396	7.488	1.948816	56.07014	10.45325
5	Switzerland	1.452	7.48	2.108304	55.9504	10.86096
6	Sweden	1.387	7.343	1.923769	53.91965	10.18474
7	New Zealand	1.303	7.307	1.697809	53.39225	9.521021
8	Canada	1.365	7.278	1.863225	52.96928	9.93447
9	Austria	1.376	7.246	1.893376	52.50452	9.970496
10	Australia	1.372	7.228	1.882384	52.24398	9.916816
11	Costa Rica	1.034	7.167	1.069156	51.36589	7.410678
12	Israel	1.276	7.139	1.628176	50.96532	9.109364
13	Luxembourg	1.609	7.09	2.588881	50.2681	11.40781
14	United Kingdom	1.333	7.054	1.776889	49.75892	9.402982
15	Ireland	1.499	7.021	2.247001	49.29444	10.52448
16	Germany	1.373	6.985	1.885129	48.79023	9.590405
17	Belgium	1.356	6.923	1.838736	47.92793	9.387588
18	United States	1.433	6.892	2.053489	47.49966	9.876236
19	Czech Republic	1.269	6.852	1.610361	46.9499	8.695188
20	United Arab Emirates	1.503	6.825	2.259009	46.58063	10.25798
21	Malta	1.3	6.726	1.69	45.23908	8.7438
22	Mexico	1.07	6.595	1.1449	43.49403	7.05665
23	France	1.324	6.592	1.752976	43.45446	8.727808
24	Taiwan	1.368	6.446	1.871424	41.55092	8.818128
25	Chile	1.159	6.444	1.343281	41.52514	7.468596
26	Guatemala	0.8	6.436	0.64	41.4221	5.1488
27	Saudi Arabia	1.403	6.375	1.968409	40.64063	8.944125
28	Qatar	1.684	6.374	2.835856	40.62788	10.73382
29	Spain	1.286	6.354	1.653796	40.37332	8.171244
30	Panama	1.149	6.321	1.320201	39.95504	7.262829
31	Brazil	1.004	6.3	1.008016	39.69	6.3252
32	Uruguay	1.124	6.293	1.263376	39.60185	7.073332
33	Singapore	1.572	6.262	2.471184	39.21264	9.843864
34	El Salvador	0.794	6.253	0.630436	39.10001	4.964882
35	Italy	1.294	6.223	1.674436	38.72573	8.052562
36	Bahrain	1.362	6.199	1.855044	38.4276	8.443038
37	Slovakia	1.246	6.198	1.552516	38.4152	7.722708
38	Trinidad & Tobago	1.231	6.192	1.515361	38.34086	7.622352
39	Poland	1.206	6.182	1.454436	38.21712	7.455492
40	Uzbekistan	0.745	6.174	0.555025	38.11828	4.59963
41	Lithuania	1.238	6.149	1.532644	37.8102	7.612462
42	Colombia	0.985	6.125	0.970225	37.51563	6.033125

43	Slovenia	1 250	£ 110	1.582564	27 //2002	7.696444
44	Nicaragua	1.258 0.694	6.118 6.105	0.481636	37.42992 37.27103	4.23687
45	Kosovo	0.882	6.103	0.481636	37.27103	5.3802
46	Argentina	1.092	6.086	1.192464	37.21	6.645912
47	Romania	1.162	6.07	1.350244	36.8449	7.05334
48	Cyprus	1.162	6.046	1.595169	36.55412	7.636098
49	Ecuador	0.912	6.028	0.831744	36.33678	5.497536
50	Kuwait	1.5	6.021	2.25	36.25244	9.0315
51	Thailand	1.05	6.008	1.1025	36.09606	6.3084
52	Latvia	1.187	5.94	1.408969	35.2836	7.05078
53	South Korea	1.301	5.895	1.692601	34.75103	7.669395
54	Estonia	1.237	5.893	1.530169	34.72745	7.289641
55	Jamaica	0.831	5.89	0.690561	34.6921	4.89459
56	Mauritius	1.12	5.888	1.2544	34.66854	6.59456
57	Japan	1.327	5.886	1.760929	34.645	7.810722
58	Honduras	0.642	5.86	0.412164	34.3396	3.76212
59	Kazakhstan	1.173	5.809	1.375929	33.74448	6.813957
60	Bolivia	0.776	5.779	0.602176	33.39684	4.484504
61	Hungary	1.201	5.758	1.442401	33.15456	6.915358
62	Paraguay	0.855	5.743	0.731025	32.98205	4.910265
63	Northern Cyprus	1.263	5.718	1.595169	32.69552	7.221834
64	Peru	0.96	5.697	0.9216	32.45581	5.46912
65	Portugal	1.221	5.693	1.490841	32.41025	6.951153
66	Pakistan	0.677	5.653	0.458329	31.95641	3.827081
67	Russia	1.183	5.648	1.399489	31.8999	6.681584
68	Philippines	0.807	5.631	0.651249	31.70816	4.544217
69	Serbia	1.004	5.603	1.008016	31.39361	5.625412
70	Moldova	0.685	5.529	0.469225	30.56984	3.787365
71	Libya	1.044	5.525	1.089936	30.52563	5.7681
72 73	Montenegro	1.051	5.523	1.104601 0.243049	30.50353 29.88809	5.804673
74	Tajikistan Croatia	0.493 1.155	5.467 5.432	1.334025	29.50662	2.695231 6.27396
75	Hong Kong	1.438	5.43	2.067844	29.4849	7.80834
76	Dominican Republic	1.015	5.425	1.030225	29.43063	5.506375
77	Bosnia and	1.015	3.123	1.050225	23. 13003	3.300373
	Herzegovina	0.945	5.386	0.893025	29.009	5.08977
78	Turkey	1.183	5.373	1.399489	28.86913	6.356259
79	Malaysia	1.221	5.339	1.490841	28.50492	6.518919
80	Belarus	1.067	5.323	1.138489	28.33433	5.679641
81	Greece	1.181	5.287	1.394761	27.95237	6.243947
82	Mongolia	0.948	5.285	0.898704	27.93123	5.01018
83	North Macedonia	0.983	5.274	0.966289	27.81508	5.184342
84	Nigeria	0.696	5.265	0.484416	27.72023	3.66444
85	Kyrgyzstan	0.551	5.261	0.303601	27.67812	2.898811
86	Turkmenistan	1.052	5.247	1.106704	27.53101	5.519844
87	Algeria	1.002	5.211	1.004004	27.15452	5.221422
88	Morocco	0.801	5.208	0.641601	27.12326	4.171608

۱ ۵۵	l					
89	Azerbaijan	1.043	5.208	1.087849	27.12326	5.431944
90	Lebanon	0.987	5.197	0.974169	27.00881	5.129439
91	Indonesia	0.931	5.192	0.866761	26.95686	4.833752
92	China	1.029	5.191	1.058841	26.94648	5.341539
93	Vietnam	0.741	5.175	0.549081	26.78063	3.834675
94	Bhutan	0.813	5.082	0.660969	25.82672	4.131666
95	Cameroon	0.549	5.044	0.301401	25.44194	2.769156
96	Bulgaria	1.092	5.011	1.192464	25.11012	5.472012
97	Ghana	0.611	4.996	0.373321	24.96002	3.052556
98	Ivory Coast	0.569	4.944	0.323761	24.44314	2.813136
99	Nepal	0.446	4.913	0.198916	24.13757	2.191198
100	Jordan	0.837	4.906	0.700569	24.06884	4.106322
101	Benin	0.393	4.883	0.154449	23.84369	1.919019
102	Congo (Brazzaville)	0.673	4.812	0.452929	23.15534	3.238476
103	Gabon	1.057	4.799	1.117249	23.0304	5.072543
104	Laos	0.764	4.796	0.583696	23.00162	3.664144
105	South Africa	0.96	4.722	0.9216	22.29728	4.53312
106	Albania	0.947	4.719	0.896809	22.26896	4.468893
107	Venezuela	0.96	4.707	0.9216	22.15585	4.51872
108	Cambodia	0.574	4.7	0.329476	22.09	2.6978
109	Palestinian Territories	0.657	4.696	0.431649	22.05242	3.085272
110 111	Senegal	0.45	4.681	0.2025	21.91176	2.10645
111	Somalia	0	4.668	0 773644	21.79022	0
113	Namibia	0.879	4.639	0.772641	21.52032	4.077681
114	Niger Burkina Faso	0.138	4.628	0.019044	21.41838	0.638664
115	Armenia	0.331	4.587	0.109561 0.7225	21.04057 20.78448	1.518297
116	1	0.85	4.559		20.78448	3.87515
117	Iran	1.1 0.38	4.548 4.534	1.21 0.1444	20.55716	5.0028 1.72292
118	Guinea			0.784996	20.33716	
119	Georgia Gambia	0.886 0.308		0.784996		
120	Kenya	0.508	4.516	0.094864	20.33426	1.390928 2.308608
121	Mauritania	0.512	4.309	0.202144	20.33108	2.5593
122	Mozambique	0.204	4.466	0.3249	19.94516	0.911064
123	Tunisia	0.921	4.461	0.848241	19.90052	
124	Bangladesh	0.562	4.456	0.315844	19.85594	2.504272
125	Iraq	1.043	4.437	1.087849	19.68697	
126	Congo (Kinshasa)	0.094	4.418	0.008836	19.51872	0.415292
127	Mali	0.385	4.39	0.148225	19.2721	1.69015
128	Sierra Leone	0.268	4.374	0.071824	19.13188	1.172232
129	Sri Lanka	0.949	4.366	0.900601	19.06196	4.143334
130	Myanmar	0.71	4.36	0.5041	19.0096	3.0956
131	Chad	0.35	4.35		18.9225	
132	Ukraine	0.82	4.332	0.6724	18.76622	3.55224
133	Ethiopia	0.336	4.286	0.112896	18.3698	1.440096
134	Swaziland	0.811	4.212	0.657721	17.74094	
I	J			- ·		

135	Uganda	0.332	4.189	0.110224	17.54772	1.390748
136	Egypt	0.913	4.166	0.833569	17.35556	3.803558
137	Zambia	0.578	4.107	0.334084	16.86745	2.373846
138	Togo	0.275	4.085	0.075625	16.68723	1.123375
139	India	0.755	4.015	0.570025	16.12023	3.031325
140	Liberia	0.073	3.975	0.005329	15.80063	0.290175
141	Comoros	0.274	3.973	0.075076	15.78473	1.088602
142	Madagascar	0.274	3.933	0.075076	15.46849	1.077642
143	Lesotho	0.489	3.802	0.239121	14.4552	1.859178
144	Burundi	0.046	3.775	0.002116	14.25063	0.17365
145	Zimbabwe	0.366	3.663	0.133956	13.41757	1.340658
146	Haiti	0.323	3.597	0.104329	12.93841	1.161831
147	Botswana	1.041	3.488	1.083681	12.16614	3.631008
148	Syria	0.619	3.462	0.383161	11.98544	2.142978
149	Malawi	0.191	3.41	0.036481	11.6281	0.65131
150	Yemen	0.287	3.38	0.082369	11.4244	0.97006
151	Rwanda	0.359	3.334	0.128881	11.11556	1.196906
152	Tanzania	0.476	3.231	0.226576	10.43936	1.537956
153	Afghanistan	0.35	3.203	0.1225	10.25921	1.12105
154	Central African					
	Republic	0.026	3.083	0.000676	9.504889	0.080158
155	South Sudan	0.306	2.853	0.093636	8.139609	0.873018
		T-1-1 1-4	1	•		

Tabel data populasi

```
#creating artificial table
subject = pd.Series(data3_clean['Country or region'])
subject = pd.Series(data3_clean['Country or region'])
x = pd.Series(data3_clean['GDP per capita'])
y = pd.Series(data3_clean['Score'])
y_square = pd.Series([i ** 2 for i in y])
x_square = pd.Series([i ** 2 for i in x])
x_y = pd.Series([x[i] * y[i] for i in range(len(x))])
data = pd.DataFrame({'Country or region' : subject, 'X' : x, 'Y' : y, 'X^2' : x_square, 'Y^2' : y_square, 'XY' : x_y})
rpint(data)
print(data)
 #data.to_excel('population3_data.xlsx')
n = 156
#y=mx+c
#y=mx+c

m = (n * data['XY'].sum() - data['X'].sum() * data['Y'].sum() ) / (n * data['X^2'].sum() - data['X'].sum() ** 2)

c = (data['Y'].sum() * data['X^2'].sum() - data['X'].sum() ** data['XY'].sum()) / (n * data['X^2'].sum() - data['X'].sum() ** 2)

print("b = ", m)

print("a = ", c)
                    Country or region
                                    Finland 1.340 7.769 1.795600 60.357361 10.410460
Denmark 1.383 7.600 1.912689 57.760000 10.510800
0
1
                                                   1.488
                                                                7.554
                                                                           2.214144
                                                                                            57.062916
                                                                                                              11.240352
                                      Norway
                                                               7.494 1.904400 56.160036 10.341720
7.488 1.948816 56.070144 10.453248
                                     Iceland
                                                   1.380
                              Netherlands 1.396
4
                                      Rwanda 0.359 3.334
                                                                           0.128881 11.115556
                                                                                                               1.196906
                                   Tanzania 0.476 3.231
152
                                                                           0.226576 10.439361
                                                                                                               1.537956
       Afghanistan 0.350 3.203 0.122500 10.259209 Central African Republic 0.026 3.083 0.000676 9.504889
153
                                                                                                                1.121050
                                                                                                                0.080158
                              South Sudan 0.306 2.853 0.093636
                                                                                             8.139609
                                                                                                                0.873018
[156 rows x 6 columns]
b = 2.2181480010083603
a = 3.399345178292416
```

Perhitungan variable a dan b pada regresi linear

```
In [43]: #print((m, c))
def get_prediction(x):
    return m"x + c

#inp = data
#print(inp)
x_plot = np.linspace(data['X'].min(), data['X'].max())
#indo_x = 0.931
#plt.plot(indo_x, get_prediction(indo_x), 'ro') # predictions (Red Dot)
plt.plot(inp['X'], inp['Y'], 'bo') # input data (Blue Dot)
plt.plot(x_plot, get_prediction(x_plot), 'r') # Line
plt.title('Grafik regresi linear GDP dengan Happiness Score')
plt.show()

Grafik regresi linear GDP dengan Happiness Score

8

7
6
6
6
7
```

Prediksi variable GDP dengan score

0.50 0.75 1.00 1.25 1.50

```
In [42]: t1 = 0.7938828678781272 * math.sqrt(154)
    t2 = math.sqrt( (1-0.7938828678781272))
    t = t1/t2
    print(t)
    21.70000054478138
In [ ]:
```

Perhitungan uji T