

**LAPORAN METODE STATISTIKA  
PREDIKSI GAYA HIDUP SEHAT MAHASISWA FTMM 2021**



**Disusun oleh :  
Kelompok 1**

<b>Raden Bagus Rifa'I Kacanegara</b>	<b>162112133047</b>
<b>Davano Al Raffi Abdul Jabbar</b>	<b>162112133083</b>
<b>Richard Hasannain M</b>	<b>162112133093</b>
<b>Shofia Ishma Najiyya</b>	<b>162112133087</b>
<b>Bunga Citra Surya Lestari</b>	<b>162112133107</b>

**MATA KULIAH METODE STATISTIKA  
PRODI TEKNOLOGI SAINS DATA  
FAKULTAS TEKNOLOGI MAJU DAN MULTIDISIPLIN  
UNIVERSITAS AIRLANGGA  
2022**

## DAFTAR ISI

<b>DAFTAR ISI</b>	<b>ii</b>
<b>DAFTAR TABEL</b>	<b>iv</b>
<b>DAFTAR GAMBAR</b>	<b>v</b>
<b>BAB I</b>	<b>1</b>
PENDAHULUAN	
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan Penelitian	2
1.4 Batasan Penelitian	2
<b>BAB II</b>	<b>3</b>
TINJAUAN PUSTAKA	3
TINJAUAN STATISTIKA	3
2.1 Regresi Linear Berganda	3
2.2 Uji Signifikansi Parameter	3
2.2.1 Uji Parsial	3
2.2.2 Uji Serentak	4
2.3 Ordinary Least Squares (OLS)	4
TINJAUAN NON STATISTIKA	5
2.4 Indeks Massa Tubuh	5
2.5 Definisi Sehat	5
2.6 Definisi Gaya Hidup	5
2.7 Gaya Hidup Sehat	6
2.8 Keseimbangan Gizi dan Minuman	6
2.9 Aktivitas Fisik	6
<b>BAB III</b>	<b>7</b>
METODOLOGI	7
3.1 Instrumen dan Metode Pengambilan Sampel	7
3.1.1 Instrumen Penelitian	7
3.1.2 Metode Pengambilan Sampel	7
3.2 Variabel Penelitian	7
<b>BAB IV</b>	<b>8</b>
HASIL DAN ANALISIS	8
4.1 Analisis Eksplorasi Data	8
4.1.1 Menyajikan secara visual hubungan data	8
4.1.2 Korelasi Antar Variabel	8
4.2 Analisis Regresi Linier Berganda	9

4.2.1 Regresi Variabel Dependen dan Independen	9
4.2.2 Uji Signifikansi Parameter	10
4.2.3 Uji Asumsi Error	11
4.2.4 Model Regresi Akhir	14
4.2.5 Interpretasi dan Kebaikan Model Regresi	15
<b>BAB V</b>	<b>16</b>
KESIMPULAN DAN SARAN	16
5.1 Kesimpulan	16
5.2 Saran	17
<b>DAFTAR PUSTAKA</b>	<b>19</b>
<b>LAMPIRAN</b>	<b>20</b>

## **DAFTAR TABEL**

<b>Tabel 4.1 Korelasi Antar Variabel</b>	<b>8</b>
<b>Tabel 4.2 Model Summary Regresi</b>	<b>9</b>
<b>Tabel 4.3 Regresi Variabel Dependen dan Independen</b>	<b>9</b>
<b>Tabel 4.4 Uji Signifikansi Parameter</b>	<b>10</b>
<b>Tabel 4.5 Variable Selection</b>	<b>14</b>
<b>Tabel 4.6 Interpretasi dan Kebaikan Model Regresi (Minitab)</b>	<b>15</b>

## **DAFTAR GAMBAR**

<b>Gambar 4.1 Histogram Weight</b>	<b>8</b>
<b>Gambar 4.2 Boxplot Variabel</b>	<b>8</b>
<b>Gambar 4.3 Uji Asumsi Error</b>	<b>11</b>
<b>Gambar 4.4 Uji Distribusi Normal</b>	<b>12</b>
<b>Gambar 4.5 Uji Variansi Konstan</b>	<b>14</b>
<b>Gambar 4.6 Uji Independensi Error</b>	<b>13</b>
<b>Gambar 4.7 Uji Multikolinieritas</b>	<b>13</b>
<b>Gambar 4.8 Intepretasi dan Kebaikan Model Regresi (Python)</b>	<b>15</b>

## **BAB 1**

### **PENDAHULUAN**

#### **1.1 Latar Belakang**

Gaya hidup sehat mempunyai peranan yang penting untuk meningkatkan kesehatan setiap individu. Gaya hidup sehat dapat dilakukan dengan cara mengonsumsi makanan yang seimbang, pola aktivitas atau olahraga secara teratur, tidur yang cukup, dan tidak merokok sehingga setiap individu akan bebas dari penyakit (Suryanto, 2011). Akan tetapi setiap individu memiliki gaya hidup yang berbeda-beda dalam mempertahankan kebugaran jasmani. Gaya hidup modern sekarang ini sebagian besar tidak sesuai dengan pola kesehatan yang dianjurkan sehingga setiap individu hendaknya mempunyai gaya hidup yang baik guna mendukung dan memperlancar kebugaran jasmani yang dimiliki (Hardiyono, 2011). Seperti aktivitas fisik yang tidak teratur, diet dan nutrisi tidak seimbang, mengonsumsi rokok serta alkohol (Shomoro et al., 2014).

Makanan, pola aktivitas/olahraga secara teratur, tidur yang cukup dan tidak merokok merupakan faktor yang penting bagi manusia untuk meningkatkan kebugaran jasmani. Kebugaran jasmani adalah derajat sehat dinamis seseorang untuk menjadi dasar kemampuan jasmani untuk dapat melaksanakan tugas yang dilaksanakan (Giriwijoyo & Sidik, 2013). Aktivitas fisik secara teratur dapat bermanfaat pada kesehatan (Wuest & Bucher dalam Kgodnapur et al., 2012). Manfaat kebugaran jasmani dapat meningkatkan kemampuan dalam melakukan tugas-tugas fisik tertentu, seperti kemampuan meningkatkan mobilisasi tubuh secara efisien, mengurangi resiko terhadap kelelahan fisik, psikologi dan stres yang terkait risiko kesehatan (Sonia dalam Shomoro & Mondal, 2014).

Aktivitas fisik/olahraga secara teratur dalam penelitian Sharkey (2003) dalam sehari cukup melakukan aktivitas fisik/olahraga secara teratur membutuhkan waktu 30 menit tanpa kelelahan sehingga masih cukup energi baik untuk waktu luang maupun keperluan mendadak. Kebiasaan merokok akan membahayakan kesehatan setiap individu karena mengandung bahan-bahan kimia didalamnya (Istiqomah, 2003). Akan tetapi setiap orang memiliki faktor kebiasaan sehingga akan berdampak terhadap kebugaran jasmani yang tidak baik. Indeks Massa Tubuh (IMT) atau Body Mass Index (BMI) adalah standar pengukuran merupakan angka yang dapat mengklasifikasikan seseorang ke dalam beberapa kategori tingkat obesitas. Indeks massa tubuh, merupakan perhitungan yang didapatkan dari membagi berat badan (dalam kilogram) dengan ukuran tinggi badan (dalam meter). Disini kami akan meneliti prediksi berat badan dengan beberapa variabel pengaruh yaitu ukuran tinggi badan (height), banyaknya makan berat dalam sehari, banyaknya makan buah/sayuran dalam sehari, banyaknya minum air dalam sehari menggunakan hitungan gelas, banyak jam tidur dalam sehari, dan banyak jam olahraga dalam seminggu.

## 1.2 Rumusan Masalah

Berdasarkan penjabaran latar belakang diatas maka penulis dapat merumuskan permasalahan yang akan menjadi fokus dalam penelitian yakni “Analisis faktor gaya hidup sehat mahasiswa FTMM 2021”

## 1.3 Tujuan Penelitian

Tujuan dari penelitian ini agar kami sebagai peneliti dapat mengetahui hasil survei Mahasiswa Unair Fakultas Teknologi Maju dan Multidisiplin terhadap gaya hidup sehat mereka. Jika rata - rata gaya hidup sehat mereka kurang ideal sesuai dengan standar BMI, maka kami sebagai peneliti ingin menghimbau kepada para mahasiswa tentang faktor faktor gaya hidup sehat sesuai dengan standar BMI yaitu pada kisaran antara 18.5 - 25.

## 1.4 Batasan Penelitian

Untuk menghindari pelebaran pokok pembahasan, penelitian hanya memfokuskan pada bmi, berapa kali kita makan dalam sehari, banyak mengonsumsi buah dan sayur dalam seminggu, banyak minum air dalam sehari dan berolahraga pada Mahasiswa Unair Fakultas Teknologi Maju dan Multidisiplin.

Kami mendapatkan jumlah sampel yang harus dicari sebanyak 83 orang dengan menggunakan rumus *Slovin*,.Rumus *Slovin* secara matematis ditulis:

$$n = \frac{N}{1+N \cdot e^2}$$

Keterangan:

n = jumlah sampel yang dicari

N = jumlah populasi (mahasiswa FTMM 2021 = 483)

e = margin of error (10% = 0.01)

## BAB 2

### TINJAUAN PUSTAKA

#### TINJAUAN STATISTIKA

##### 2.1 Regresi Linear Berganda

Regresi linear berganda merupakan model regresi yang melibatkan lebih dari satu variabel independen atau variabel prediktor. Analisis regresi linear berganda dilakukan untuk mengetahui arah dan seberapa besar pengaruh variabel independen ( $X_1, X_2, \dots, X_n$ ) terhadap variabel dependen (Y) (Ghozali, 2018). Model persamaan untuk menghitung regresi linier berganda yaitu :

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + e$$

keterangan :

Y = Variabel dependen / Variabel respon

$\beta_0$  = Konstanta

$\beta$  = Slope / Koefisien estimasi

X = Variabel independen / Variabel prediktor

e = error term

##### 2.2 Uji Signifikansi Parameter

Uji signifikan parameter ini dilakukan untuk mengetahui apakah taksiran parameter berpengaruh terhadap model atau tidak secara signifikan, serta mengetahui seberapa besar pengaruh masing-masing parameter tersebut (Dok, 2017). Cara pengujian signifikansi parameter terbagi menjadi dua, yaitu secara serentak dan secara parsial.

###### 2.2.1 Uji Parsial

Uji parsial atau individual adalah untuk menguji apakah suatu variabel bebas berpengaruh atau tidak terhadap variabel tidak bebas. Tujuan pengujian ini adalah untuk mengetahui apakah masing-masing variabel independen mempengaruhi variabel dependen secara signifikan. Terdapat langkah-langkah untuk uji signifikansi secara parsial :

###### Menentukan Tingkat Signifikansi

Contoh Tingkat signifikansi yang digunakan adalah  $\alpha = 0,05$

###### Uji Statistik



$$t_{hitung} = \frac{\hat{\beta}_l}{SE(\hat{\beta}_l)}$$

#### **Daerah Penolakan**

$H_0$  ditolak jika  $|t_{hitung}| \geq t_{(\alpha/2, n-k-1)}$  atau p-value  $< \alpha$

#### **Kesimpulan**

Jika  $H_0$  ditolak, maka dapat disimpulkan bahwa variabel ke-i berpengaruh signifikan terhadap model.

### **2.2.2 Uji Serentak**

Uji serentak dilakukan untuk mengetahui signifikansi parameter model regresi secara bersama-sama menggunakan analisis ANOVA (Analysis of Variance). Terdapat langkah - langkah uji signifikansi secara serentak

#### **Hipotesis**

$H_0 : \beta_1 = \beta_2 = \beta_3 = \dots = \beta_p = 0$

$H_1 : \text{paling tidak ada satu } \beta_j \neq 0, j=1,2,\dots,p$

#### **Menentukan tingkat signifikansi**

Tingkat signifikansi yang digunakan adalah  $\alpha=0,05$

#### **Daerah Penolakan $H_0$**

$H_0$  ditolak jika  $|t_{hitung}| \geq t_{(\alpha/2, n-k-1)}$  atau p-value  $< \alpha$

#### **Kesimpulan**

Jika  $H_0$  ditolak, maka dapat disimpulkan bahwa minimal terdapat satu variabel yang signifikan terhadap model.

### **2.3 OLS**

Ordinary Least Squares (OLS) adalah sebuah teknik regresi yang termasuk ke dalam *supervised learning*. Ini adalah *method* dari python yang digunakan untuk mengestimasi sebuah parameter yang tidak diketahui dengan membuat model yang akan meminimalkan *sum of the squared errors* antara variabel yang diamati atau variabel respons dengan variabel yang diprediksi.

## TINJAUAN NON STATISTIKA

### 2.4 Indeks Massa Tubuh

Indeks massa tubuh atau *body mass index* (BMI) adalah sebuah acuan yang saat ini digunakan untuk mendefinisikan karakteristik ukuran tinggi/berat dan untuk mengklasifikasikan (mengkategorikan) mereka ke dalam kelompok seperti *underweight*, *normal*, *overweight*, *obese*, *extremely obese*. Ini juga banyak digunakan sebagai prevalensi untuk beberapa masalah kesehatan yang ditimbulkan. BMI telah berguna dalam studi berbasis populasi berdasarkan penerimaan yang luas dalam mendefinisikan kategori tertentu dari massa tubuh sebagai masalah kesehatan (Frank, 2015). *Body Mass Index* di Indonesia dikelompokkan sebagai berikut :

- BMI dibawah angka 18,5 maka orang tersebut memiliki berat badan di bawah normal
- BMI 18,5 - 22,9 dikategorikan dalam kisaran normal
- BMI 23 - 24,9 dikategorikan mengalami kelebihan berat badan
- BMI diatas 25 termasuk dalam kategori obesitas

### 2.5 Definisi Sehat

Pengertian sehat menurut WHO adalah “*Health is a state of complete physical, mental and social well-being and not merely the absence of diseases or infirmity*”. Sehat adalah kondisi normal seseorang yang merupakan hak hidupnya. Sehat berhubungan dengan hukum alam yang mengatur tubuh, jiwa, dan lingkungan berupa udara segar, sinar matahari, bersantai, kebersihan serta pikiran, kebiasaan dan gaya hidup yang baik atau bisa dikatakan bahwa sehat adalah memperbaiki kondisi manusia, baik jasmani, rohani maupun akal, sosial dan bukan semata-mata memberantas penyakit (Al Fanjari, 1999). Kesehatan dalam pengertian paling luas merupakan suatu keadaan yang dinamis dimana individu menyesuaikan diri dengan perubahan-perubahan lingkungan internal (psikologis, intelektual, spiritual dan penyakit) dan eksternal (lingkungan fisik, sosial, dan ekonomi) dan merupakan sumber dari kesenangan, kenikmatan dan kebahagiaan yang tidak ternilai jika memiliki tubuh yang sehat, untuk mempertahankan hidup. (Indang Entjang, 1981:52)

### 2.6 Definisi Gaya Hidup

Menurut Sutisna dalam Heru Suprihhadi (2017) gaya hidup secara luas didefinisikan sebagai cara hidup yang diidentifikasi oleh bagaimana orang lain menghabiskan waktu mereka (aktivitas) dilihat dari pekerjaan, hobi, belanja, olahraga, dan kegiatan sosial serta interest (minat) terdiri dari makanan, keluarga, rekreasi dan juga opinion (pendapat) terdiri dari mengenai diri mereka sendiri, masalah-masalah sosial, bisnis, dan produk. Gaya hidup mencakup sesuatu yang lebih dari sekedar kelas sosial ataupun kepribadian seseorang. Gaya hidup seseorang dalam “kesehatan” dapat diubah dengan cara memberdayakan individu agar merubah gaya hidupnya, tetapi merubahnya bukan pada si individu saja, tetapi juga merubah lingkungan sosial dan kondisi kehidupan yang mempengaruhi pola perilakunya.

## **2.7 Gaya Hidup Sehat**

Gaya hidup sehat adalah suatu gaya hidup dengan memperhatikan faktor-faktor tertentu yang mempengaruhi kesehatan, antara lain makanan dan olahraga. Selain itu gaya hidup seseorang juga mempengaruhi tingkat kesehatannya, misalnya jika suka merokok dan minum minuman keras, tentu saja bukan pola hidup sehat (Anne, 2010: 1-2). Menurut Depkes RI (1997), gaya hidup sehat adalah segala upaya untuk menerapkan kebiasaan yang baik dalam menciptakan hidup yang sehat dan menghindari kebiasaan yang buruk yang dapat mengganggu kesehatan.

## **2.8 Keseimbangan Gizi dan Minuman**

Menurut Lasantha (2010:1) makanan yang baik digambarkan seperti 4 sehat 5 sempurna, pola seperti ini dimaksudkan untuk dapat melengkapi keseimbangan gizi, 4 sehat 5 sempurna terdiri dari nasi, sayur, lauk pauk, buah dan susu, dengan demikian diharapkan kandungan karbohidrat, lemak, protein, kalsium dan zat besi lainnya, dapat terpenuhi dalam susunan 4 sehat 5 sempurna. Karbohidrat, lemak, protein dan zat besi berperan sangat penting bagi tubuh untuk menopang aktivitas tubuh. Menurut pendapat Purnomo dan Abdul Kadir (1994:23) air yang sehat adalah air bersih, tidak berbau, tidak berwarna, tidak mengandung hama dan tidak mengandung zat-zat kimia yang berbahaya. Pemilihan makanan dan gizi yang tepat dan seimbang, cukup energi dan nutrisi, meliputi: karbohidrat, lemak, protein, vitamin, mineral dan air. Dengan proporsi makanan yaitu karbohidrat 60 %, protein 15 % dan lemak 25 %. Selain menjaga pola makan yang sehat juga harus meninggalkan kebiasaan buruk seperti, merokok, minuman beralkohol, obat-obatan terlarang, dan makan tidak teratur.

## **2.9 Aktivitas Fisik**

Aktivitas fisik merupakan suatu gerakan tubuh yang dihasilkan otot rangka dan membutuhkan energi, termasuk aktivitas yang dilakukan saat bekerja, bermain, melakukan pekerjaan rumah tangga, bepergian dan kegiatan rekreasi (WHO, 2017). Aktivitas fisik berbeda dengan olahraga karena olahraga merupakan suatu kegiatan fisik yang direncanakan, terstruktur, berulang dan bertujuan memperbaiki atau mempertahankan satu atau lebih komponen kebugaran fisik seseorang (WHO, 2010).

## **BAB 3**

### **METODOLOGI**

#### **3.1 Instrumen dan Metode Pengambilan Sampel**

##### **3.1.1 Instrumen Penelitian**

Suharsimi Arikunto (2002;126) mengemukakan bahwa instrumen penelitian adalah alat bantu yang dipilih dan digunakan oleh peneliti dalam kegiatan untuk mengumpulkan data atau informasi agar kegiatan pengumpulan data tersebut menjadi sistematis dan dipermudah olehnya. Instrumen yang digunakan dalam penelitian ini adalah instrumen yang telah tersedia dan teknik pengumpulan data yang digunakan pada penelitian analisis gaya hidup sehat ini adalah kuesioner. Responden diminta untuk mengisi jawaban atas pernyataan mengenai pola hidup yang dilakukan mahasiswa setiap hari.

##### **3.1.2 Metode Pengambilan Sampel**

Metode pengambilan sampel yang kami gunakan adalah Non Probability Sampling. Dalam sampel non-probabilitas, individu dipilih berdasarkan kriteria non-acak dan tidak setiap individu memiliki peluang untuk disertakan. Jenis Non Probability sampling yang kami gunakan adalah Convenience Sampling. sampel ini lebih mudah dan lebih murah untuk diakses, tetapi memiliki resiko bias pengambilan sampel yang lebih tinggi.

Dalam penelitian kami responden yang terpilih yaitu mahasiswa FTMM 2021 yang kami peroleh menggunakan data kuesioner yang kami sebar secara berkala kepada setiap Grup Prodi dalam FTMM 2021 sehingga kami memakai teknik Non Probability Sampling dengan jenis sampling yaitu Convenience Sampling.

#### **3.2 Variabel Penelitian**

Variabel penelitian yang kami gunakan ada 2 yaitu variabel Independen sebagai prediktor dan variabel dependen sebagai respon. Variabel dependen atau respon yang kami gunakan adalah ukuran berat badan (weight) dan variabel independen atau prediktor ada 6 yaitu :

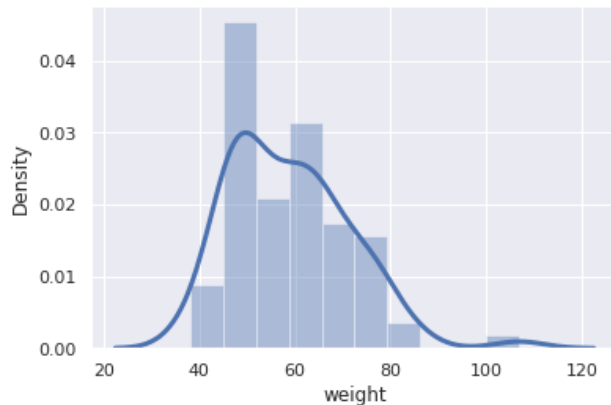
1. Ukuran tinggi badan (height)
2. Banyaknya makan berat dalam sehari
3. Banyaknya makan buah/sayuran dalam sehari
4. Banyaknya minum air dalam sehari menggunakan hitungan gelas
5. Banyak jam tidur dalam sehari
6. Banyak jam olahraga dalam seminggu.

## BAB 4

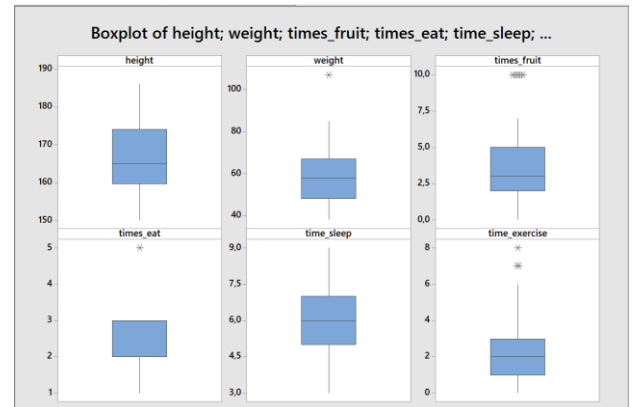
### HASIL DAN ANALISIS

#### 4.1 Analisis Eksplorasi Data

##### 4.1.1 Menyajikan secara visual hubungan data



Gambar 4.1 Histogram Weight



Gambar 4.2 Boxplot Variabel

Exploratory Data Analysis merupakan tahapan yang dilakukan untuk membuat beberapa visualisasi data yang bertujuan untuk mengetahui hubungan antara dua atau lebih variabel yang diteliti yaitu variabel respon dan variabel predictor. Pada analisis gaya hidup sehat ini variabel responnya adalah weight. Lalu terdapat boxplot yang menunjukkan distribusi dari semua variabel. Kita lihat bahwa weight, times\_fruit, times\_eat, dan time\_exercise memiliki outlier sehingga diduga memiliki distribusi yang kurang normal. Lalu untuk height dan time\_sleep memiliki distribusi yang normal dalam boxplot tersebut.

##### 4.1.2 Korelasi Antar Variabel

###### Correlations

	height	weight	times_eat	times_fruit	time_sleep
weight	0,525				
times_eat	-0,016	-0,161			
times_fruit	-0,120	0,020	0,311		
time_sleep	-0,073	-0,215	0,244	0,112	
time_exercise	-0,122	0,083	-0,096	0,319	-0,159

Tabel 4.1 Korelasi Antar Variabel

Diatas merupakan korelasi antar variabel independen dan dependen sehingga kita bisa mendapatkan hasil bahwa :

1. Hubungan yang kuat dan positif / searah (nilai korelasi  $>0.5$ ) adalah weight - height
2. Hubungan yang lemah dan negatif / berlawanan arah ( $-0.5 < \text{nilai korelasi} < 0$ ) times\_eat - height, times\_eat - weight, times\_fruit - height, time\_sleep - height, time\_sleep - weight, time\_exercise - height, time\_exercise - times\_eat
3. Hubungan yang lemah dan positif / searah ( $0 < \text{nilai korelasi} < 0.5$ ) adalah times\_fruit - weight, times\_fruit - times\_eat, times\_sleep - times\_eat, times\_sleep - times\_fruit, time\_exercise - weight, time\_exercise - times\_fruit

## 4.2 Analisis Regresi Linier Berganda

### 4.2.1 Regresi Variabel Dependen dan Independen

#### Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
10,1875	40,92%	34,53%	24,91%

**Tabel 4.2 Model Summary Regresi**

#### Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	-35,7	30,2	-1,18	0,240	
height	0,649	0,165	3,93	0,000	1,58
times_eat	-3,25	1,88	-1,73	0,088	1,24
times_fruit	0,680	0,499	1,36	0,177	1,32
time_sleep	-1,570	0,835	-1,88	0,064	1,13
time_exercise	-0,461	0,787	-0,59	0,560	1,63
gender					
Perempuan	-3,19	2,88	-1,11	0,271	1,63
is_smoking					
Ya	2,63	2,83	0,93	0,355	1,21
times_drink					
>=8	5,97	2,39	2,50	0,015	1,11

**Tabel 4.3 Regresi Variabel Dependen dan Independen**

Setelah kami menghimpun data yang kami peroleh dari kuesioner maka kami melakukan analisis linier berganda pada minitab sehingga didapatkan hasil semua persamaan seperti gambar diatas. Dari tabel tersebut bisa dilihat bahwa variabel tersebut memiliki nilai Variance Inflation Factor (VIF) di antara interval 1,11 sampai dengan 1,63 sehingga variabel-variabel tersebut tidak memiliki hubungan yang kuat (Multikolinieritas).

Estimasi Parameter :

- $\beta_0$  (Intercept) = -35,7
- $\beta_1$  (height) = 0,649

- $\beta_2$  (times\_eat) = -3,25
- $\beta_3$  (times\_fruit) = 0,68
- $\beta_4$  (times\_sleep) = -1,57
- $\beta_5$  (times\_exercise) = -0,461
- $\beta_6$  (gender) = -3,19
- $\beta_7$  (is\_smoking) = 2,63
- $\beta_8$  (times\_drink) = 5,97

Dengan koefisien tersebut, kita mampu membuat persamaan sebagai berikut:

$$Y = -35,7 + 0,649 X_1 - 3,25X_2 + 0,68 X_3 - 1,57X_4 - 0,461X_5 - 3,19 X_6 + 2,63X_7 + 5,97X_8$$

Model ini sudah diregresikan dengan variabel dependen (y) adalah weight. Dengan rangkuman bahwa model ini memiliki nilai R-Square sebesar 40,92%, adjusted R-square sebesar 34,53% dan predicted R-Square sebesar 24,91%. Nilai dari adjusted R-square menjelaskan seberapa baik model ini dalam menjelaskan variabel respon. Untuk model ini dapat kita simpulkan bahwa variabel respon dijelaskan oleh model ini sebesar 34,53%, sedangkan 65,47% sisanya dijelaskan oleh variabel lain yang tidak ada dalam model ini.

#### 4.2.2 Uji Signifikansi Parameter

##### Analysis of Variance

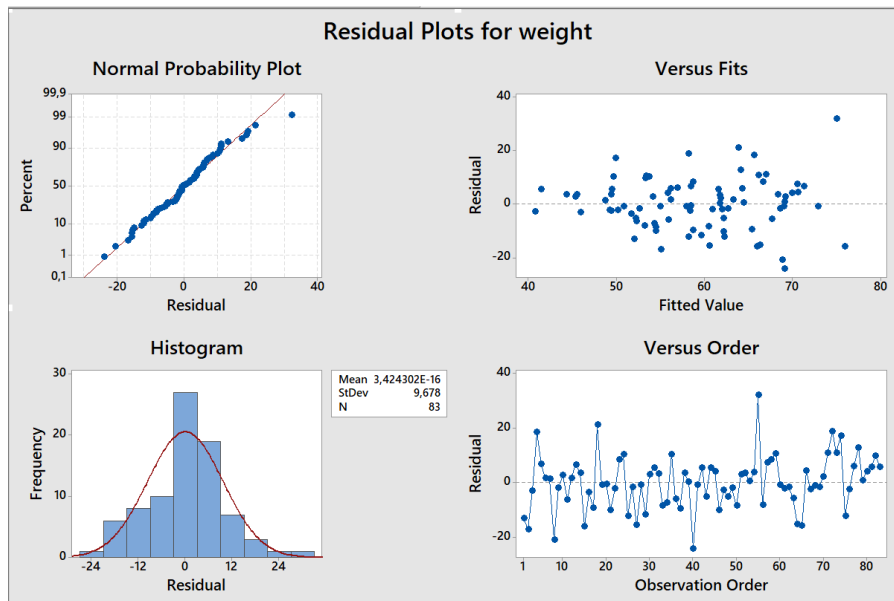
Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	8	5318,7	664,84	6,41	0,000
height	1	1605,6	1605,56	15,47	0,000
times_eat	1	309,6	309,58	2,98	0,088
times_fruit	1	192,6	192,56	1,86	0,177
time_sleep	1	366,4	366,40	3,53	0,064
time_exercise	1	35,7	35,65	0,34	0,560
gender	1	127,8	127,83	1,23	0,271
is_smoking	1	89,9	89,89	0,87	0,355
times_drink	1	647,4	647,45	6,24	0,015
Error	74	7680,2	103,79		
Total	82	12998,8			

**Tabel 4.4 Uji Signifikansi Parameter**

Karena dalam penelitian kami menggunakan alpha sebesar 10% atau 0,01 maka melalui uji serentak P-value keseluruhan yaitu 0,000 sudah memenuhi karena kurang dari alpha (10%). Melalui uji parsial variabel yang memiliki P-Value yang memenuhi atau P-Value tersebut kurang dari 10% atau 0,01 adalah height, times\_eat, times\_sleep, dan times\_drink. Sedangkan variabel yang memiliki P-Value yang tidak memenuhi atau

P-Value tersebut lebih dari 10% atau 0,01 adalah times\_fruit, time\_exercise, gender, dan is\_smoking.

### 4.2.3 Uji Asumsi Error



**Gambar 4.3 Uji Asumsi Error**

#### 1. Linieritas untuk Variabel Independen (X) dan Dependen (Y)

Ketika sifat umum hubungan antara  $y$  dan  $x$  membentuk kurva daripada garis lurus, prediksi yang diperoleh dapat menghasilkan hasil yang tidak masuk akal. Seringkali, transformasi data yang sesuai mengurangi relasi nonlinier ke relasi yang mendekati berbentuk linier. Dalam Normal Probability Normal terlihat scatter plot yang sesuai dengan garis regresi atau garis prediksi sehingga plot probabilitas normal dari residual.

#### 2. Uji Distribusi Normal

Untuk mengetahui apakah distribusi data pada setiap variabel normal atau tidak salah satu caranya dengan menggunakan Normal Q-Q plot, Histogram. Pada Histogram tersebut kita bisa melihat bahwa distribusi tersebut normal karena lengkungan garis regresinya di tengah dan tidak ada skewness. Selain itu Uji normalitas ini dapat menggunakan uji *Kolmogorov-Smirnov* (K-S). Dalam uji *Kolmogorov-Smirnov* (K-S) jika nilai probabilitas (p-value) lebih besar dari tingkat signifikansi ( $\alpha$ ) yaitu 0.05, maka data berdistribusi normal. Namun, jika sebaliknya, maka data tidak berdistribusi normal.



```
Statistic KS: 0.3693
P-value      : 0.0
Tolak  $H_0$ , error tidak berdistribusi normal
```

**Gambar 4.4 Uji Distribusi Normal**

Dari hasil uji *Kolmogorov-Smirnov* (K-S) didapatkan bahwa nilai probabilitas (p-value) yaitu 0.0, dimana nilai tersebut lebih kecil atau kurang dari tingkat signifikansi ( $\alpha$ ) yaitu 0.1. Maka dapat disimpulkan bahwa data tersebut tidak berdistribusi normal.

### 3. Uji Variansi Konstan

Bertujuan untuk menguji apakah dalam model regresi terjadi ketidaksamaan variansi dari residual satu pengamatan ke pengamatan yang lain. Jika variansi dari residual satu pengamatan ke pengamatan lain tetap, maka disebut homoskedastisitas dan jika berbeda disebut heteroskedastisitas. Model regresi yang baik adalah yang homoskedastisitas atau tidak terjadi heteroskedastisitas.

Menurut plot versus fits tersebut tidak terjadi heteroskedastisitas/ variansi erornya sudah konstan karena tidak ada pengumpulan data di tengah menyerupai bulatan, menyempit atau melebar menyerupai corong. Selain itu Uji variansi konstan ini dapat menggunakan uji *White Test*. Dalam uji *white test* jika nilai probabilitas (p-value) lebih besar dari tingkat signifikansi ( $\alpha$ ) yaitu 0.1, maka data terjadi heteroskedastisitas. Namun, jika sebaliknya, maka data bersifat homoskedastisitas.

```
{'Test Statistic': 18.047727412862155, 'Test Statistic p-value': 0.020871499445466155,
'F-Statistic': 2.570217667826968, 'F-Test p-value': 0.015646298498380614}
Gagal Tolak  $H_0$ , homoscedastisitas terjadi (residual tersebar merata)
```

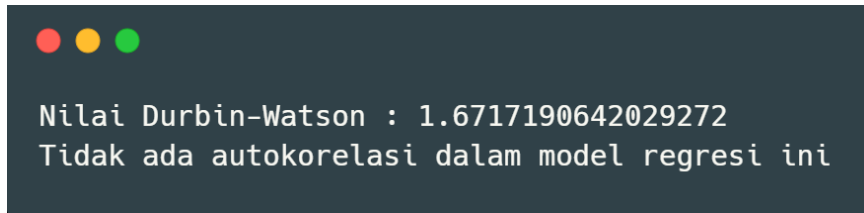
**Gambar 4.5 Uji Variansi Konstan**

Dari hasil uji *White Test* didapatkan bahwa nilai probabilitas (Tes Statistik p-value) yaitu 0.02087, dimana nilai tersebut lebih kecil atau kurang dari tingkat signifikansi ( $\alpha$ ) yaitu 0.1. Maka dapat disimpulkan bahwa data tersebut tidak terjadi heteroskedastisitas / bersifat homoskedastisitas yang berarti memiliki residual yang tersebar merata.

### 4. Uji Independensi Error

Bertujuan untuk menguji apakah dalam model regresi linier ada korelasi antara kesalahan pengganggu pada periode t dengan kesalahan pengganggu pada periode t-1.

Model regresi yang baik harus bebas dari asumsi autokorelasi. Melalui grafik Versus Order kita bisa melihat bahwa plot urutan waktunya tidak membentuk pola atau naik turun secara acak. Selain itu Uji independensi error ini dapat menggunakan uji *Durbin-Watson*. Dalam uji *Durbin-Watson* jika nilai tes statistiknya berada di interval 1,58 - 2,42 maka data tidak memiliki autokorelasi. Namun, jika sebaliknya, maka data memiliki autokorelasi.



```

Nilai Durbin-Watson : 1.6717190642029272
Tidak ada autokorelasi dalam model regresi ini
  
```

**Gambar 4.6 Uji Independensi Error**

Dari hasil uji didapatkan bahwa nilai statistik Durbin-Watson yaitu 1.672, dimana nilai tersebut berada pada interval uji Durbin-Watson yaitu 1,58 - 2,42. Maka dapat disimpulkan bahwa data tersebut tidak memiliki autokorelasi.

## 5. Uji Multikolinieritas

### Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	-35,7	30,2	-1,18	0,240	
height	0,649	0,165	3,93	0,000	1,58
times_eat	-3,25	1,88	-1,73	0,088	1,24
times_fruit	0,680	0,499	1,36	0,177	1,32
time_sleep	-1,570	0,835	-1,88	0,064	1,13
time_exercise	-0,461	0,787	-0,59	0,560	1,63
gender					
Perempuan	-3,19	2,88	-1,11	0,271	1,63
is_smoking					
Ya	2,63	2,83	0,93	0,355	1,21
times_drink					
>=8	5,97	2,39	2,50	0,015	1,11

**Gambar 4.7 Uji Multikolinieritas**

Analisis ini bertujuan untuk menguji apakah dalam model regresi ditemukan adanya korelasi antar variabel independen (X). Multikolinearitas dapat dilihat dari menganalisis matriks korelasi variabel-variabel independen. Kita bisa melihat bahwa nilai VIF dari masing2 variabel bernilai antara 1,11 - 1,63. karena nilai tersebut dibawah nilai 10 yang menjadi batasan multikolinieritas maka variabel-variabel tersebut tidak terjadi Multikolinieritas.

## 4.2.4 Model Regresi Akhir

### Stepwise Selection of Terms

Candidate terms: height; times\_eat; times\_fruit; time\_sleep; time\_exercise; gender; is\_smoking; times\_drink

	----Step 1----		----Step 2----		-----Step 3-----	
	Coef	P	Coef	P	Coef	P
Constant	-69,0		-70,4		-56,7	
height	0,771	0,000	0,760	0,000	0,738	0,000
times_drink			5,58	0,019	6,03	0,010
time_sleep					-1,746	0,031
S	10,7825		10,4809		10,2384	
R-sq	27,55%		32,39%		36,29%	
R-sq(adj)	26,66%		30,70%		33,87%	
R-sq(pred)	22,68%		26,18%		28,86%	
Mallows' Cp	11,74		7,67		4,79	

$\alpha$  to enter = 0,1;  $\alpha$  to remove = 0,1

**Tabel 4.5 Variable Selection**

Untuk mendapatkan model regresi akhir yang baik maka kita menggunakan selection untuk menghilangkan variabel yang memiliki P-Value lebih dari alpha (10%). Kami menggunakan metode stepwise untuk mendapatkan model regresi akhir yang baik. Dilihat dari stepwise selection ada beberapa langkah dalam pengambilan variabel yaitu sebagai berikut :

1. Memilih variabel yang memiliki P-Value yang paling rendah yaitu variabel height dengan P-Value sebesar 0,000
2. Memilih variabel yang memiliki P-Value terendah kedua yaitu variabel times\_drink dengan P-Value sebesar 0,015 lalu diregresikan kembali kedua variabel x dengan variabel y sehingga P-Value dari variabel times\_drink berubah menjadi 0,019
3. Memilih variabel yang memiliki P-Value terendah ketiga yaitu variabel time\_sleep dengan P-Value sebesar 0,064 lalu diregresikan kembali kedua variabel x dengan variabel y sehingga P-Value dari variabel times\_drink berubah menjadi 0,010 dan P-Value dari variabel time\_sleep berubah menjadi 0,031
4. Langkah berhenti karena jika memasukkan variabel lagi maka P-Value dari salah satu variabel akan melebihi alpha (10%)

## 4.2.5 Interpretasi dan Kebaikan Model Regresi

OLS Regression Results

```

=====
Dep. Variable:          weight    R-squared:          0.363
Model:                  OLS       Adj. R-squared:     0.339
Method:                 Least Squares   F-statistic:       15.00
Date:                   Tue, 14 Jun 2022   Prob (F-statistic): 8.11e-08
Time:                   00:22:08    Log-Likelihood:    -308.79
No. Observations:       83          AIC:               625.6
Df Residuals:           79          BIC:               635.3
Df Model:                3
Covariance Type:        nonrobust
=====

```

	coef	std err	t	P> t	[0.025	0.975]
const	-56.6608	22.807	-2.484	0.015	-102.057	-11.265
height	0.7382	0.132	5.573	0.000	0.475	1.002
time_sleep	-1.7465	0.794	-2.199	0.031	-3.327	-0.166
dum_drink	6.0279	2.286	2.637	0.010	1.477	10.578

```

=====
Omnibus:                 3.704    Durbin-Watson:          1.672
Prob(Omnibus):            0.157    Jarque-Bera (JB):        3.165
Skew:                     0.295    Prob(JB):                 0.205
Kurtosis:                 3.754    Cond. No.                 3.37e+03
=====

```

Gambar 4.8 Interpretasi dan Kebaikan Model Regresi (Python)

### Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
10,2384	36,29%	33,87%	28,86%

### Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	-56,7	22,8	-2,48	0,015	
height	0,738	0,132	5,57	0,000	1,01
time_sleep	-1,746	0,794	-2,20	0,031	1,01
times_drink >=8	6,03	2,29	2,64	0,010	1,01

Gambar 4.9 Interpretasi dan Kebaikan Model Regresi (Minitab)

Dengan menggunakan OLS Regression Results dalam estimasi parameter maka kita mendapatkan persamaan sebagai berikut :

$$Y = -56,6608 + 0,7382X_1 - 1,7465X_2 + 6,0279X_3 + e$$

Setelah itu kita dapat melihat model rangkuman yang berubah R-square jadi 36,29 %, R-square adjusted yaitu 33,87% dan R-square predicted yaitu 28,86 %

## BAB 5

### KESIMPULAN

#### 5.1 Kesimpulan

Berdasarkan analisis regresi linier berganda yang sudah kami lakukan untuk meneliti bagaimana pola hubungan antara variabel prediktor berupa height(X1), times\_eat(X2), times\_fruit(X3), times\_drink(X4), times\_sleep(X5), times\_exercise(X6) terhadap variabel respon berupa ukuran berat badan (Y), didapatkan hasil sebagai berikut:

1. Berdasarkan distribution plot, secara visual terlihat berdistribusi tidak normal. Secara visual, grafik distribusi tersebut terlihat memiliki skewness ke kanan atau Positive Skew.
2. Lalu berdasarkan boxplot, terlihat bahwa weight, times\_fruit, times\_eat, dan time\_exercise memiliki outlier sehingga diduga memiliki distribusi yang kurang normal. Lalu untuk height dan time\_sleep memiliki distribusi yang normal dalam boxplot.
3. Berdasarkan histogram tersebut menunjukkan bahwa variabel respon 'weight' diduga memiliki distribusi tidak normal karena kurva garis regresi tidak tepat di tengah dan terdapat skewness di kanan sehingga distribusi tersebut tidak normal yang menyebabkan hasil data tidak rata dan ada yang kurang sama. Hal tersebut bisa dijelaskan pada boxplot yang terdapat outlier di variabel respon 'weight'
4. Korelasi antara 'height' dengan 'weight' sebesar 0,525, 'times\_eat' dengan 'weight' sebesar -0,161, Lalu terdapat hubungan lemah namun positif untuk variabel 'times\_eat' dengan 'height' yaitu sebesar -0,016.
5. Berdasarkan uji serentak, didapatkan P-value keseluruhan dengan melihat regression yaitu 0,000. yang sudah memenuhi karena kurang dari alpha (10%). Sehingga gagal tolak  $H_0$  yang artinya minimal terdapat satu variabel yang memiliki hasil pengaruh yang signifikan terhadap model.
6. Berdasarkan uji parsial, semua variabel yang memiliki P-value memenuhi atau kurang dari 10% atau 0,01 adalah height, times\_eat, times\_sleep, dan times\_drink yang berarti Tolak  $H_0$  karena signifikan lalu variabel yang memiliki P-value lebih dari 10% atau 0,01 yaitu variabel times\_fruit, time\_exercise, \_gender, dan is\_smoking yang berarti gagal tolak  $H_0$  karena tidak signifikan.
7. Pada pengujian asumsi dengan lima metode, dimana uji pertama ada menguji linieritas untuk variabel independen dan dependen yang dimana diperoleh bahwa dapat disimpulkan yaitu ketika sifat umum antara y dan x membentuk kurva daripada garis lurus, prediksi yang diperoleh dapat menghasilkan hasil yang tidak masuk, Kemudian yang kedua adalah menguji autokorelasi dimana terdapat model regresi terjadi ketidaksamaan variansi dari residual satu pengamatan ke pengamatan yang lain. Jika variansi dari residual satu

pengamatan ke pengamatan lain tetap, maka disebut homoskedastisitas dan jika berbeda disebut heteroskedastisitas. Model regresi yang baik adalah yang homoskedastisitas atau tidak terjadi heteroskedastisitas. Metode yang ketiga adalah yaitu menguji heteroskedastisitas dimana dapat disimpulkan bahwa data tersebut tidak terjadi heteroskedastisitas / bersifat homoskedastisitas yang berarti memiliki residual yang tersebar merata. Metode yang keempat yakni menguji normalitas dan dari hasil uji *Kolmogorov-Smirnov* (K-S) didapatkan bahwa nilai probabilitas (p-value) yaitu 0,0, dimana nilai tersebut lebih kecil atau kurang dari tingkat signifikansi ( $\alpha$ ) yaitu 0.1. Maka dapat disimpulkan bahwa data tersebut tidak berdistribusi normal. Yang terakhir adalah menguji Multikolinieritas yang bertujuan untuk menguji apakah dalam model regresi ditemukan adanya korelasi antar variabel independen (X). Multikolinieritas bisa dilihat bahwa nilai VIF dari masing - masing variabel bernilai antara 1,11 - 1,63. Karena nilai tersebut dibawah nilai 10 yang menjadi sebuah batasan multikolinieritas maka variabel-variabel tersebut tidak terjadi multikolinieritas.

Karena kita menggunakan berat sebagai variabel respon sehingga kelompok kami bisa memberikan kategori gaya hidup sehat kepada anak dengan menggunakan rumus Indeks Massa Tubuh (IMT).

Contoh studi kasus :

Sofia : gender (perempuan), height (160), times\_eat (2), Times\_fruit (5), times\_drink(<8), is\_smoking (tidak), time\_sleep (6), time\_exercise (1). dengan melakukan persamaan yang sudah di-*selection* yaitu :

$$Y = -56,6608 + 0,7382X_1 - 1,7465 X_2 + 6,0279X_3 + e$$

Maka kita hanya mengambil  $X_1 = \text{height}$ ,  $X_2 = \text{time\_sleep}$ ,  $X_3 = \text{times\_drink}$ .

$Y = -56,6608 + 0,7382(160) - 1,7465 (6) + 6,0279(0)$  sehingga

$Y = 50,1082$ . Dari prediksi weight yang kita dapatkan maka kita akan mendapatkan BMI dengan rumus diatas sehingga kita memperoleh 19,53 yang berarti BMI seorang shofia dikategorikan dalam kisaran normal. Apabila membandingkan dengan BMI yang kita dapatkan dari data responden, sofia mendapatkan nilai BMI sebesar 15,23 yang berarti BMI sofia dikategorikan kurus. Perbedaan nilai dari BMI ini dikarenakan model kami kurang menjelaskan variabel respon dengan baik (nilai *adjusted R-Square*= 33,87%)

## 5.2 Saran

Adapun beberapa saran yang dapat bermanfaat sebagai acuan dan membangun pada penelitian kami atau pada peneliti yang ingin mengembangkan selanjutnya adalah sebagai berikut :

1. Mencari dan menelaah variabel prediktor yang sesuai untuk dijadikan pertanyaan agar R-Square dan Adjusted R-Square bisa meningkat lebih dari 50% agar regresi yang kita buat bisa memprediksi variabel respon dengan lebih tepat.

2. Melakukan transformasi data dengan berbagai metode seperti akar, kuadrat, logaritma dll apabila model yang didapatkan tidak menjelaskan variabel respon dengan baik (nilai  $R\text{-Square} < 25\%$ ).
3. Mencari data responden lebih banyak agar sesuai dengan rumus teknik sampling yang benar dan berjaga-jaga jika datanya ingin dihapus outliernya atau menghapus residual.

## DAFTAR PUSTAKA

Ahmad Syauqi Al Fanjari. (1999). Nilai Kesehatan Dalam Syari`at Islam / Ahmad Syauqi Al Fanjari (Ed 1, Cet 1.). JAKARTA: Bumi Aksara.

Nuttall, Frank. (2015). Body Mass Index. Nutrition Today. 50. 1.  
10.1097/NT.0000000000000092.

WHO. (1948). Health and Well-Being. Word Health Organization.  
<https://www.who.int/data/gho/data/major-themes/health-and-well-being#:~:text=The%20WHO%20constitution%20states%3A%20%22Health,of%20mental%20disorders%20or%20disabilities.>

Suprihhadi, H. (2017). PENGARUH MOTIVASI DAN LINGKUNGAN KERJA TERHADAP PRODUKTIVITAS KARYAWAN DIMODERASI BUDAYA KERJA. Jurnal Ilmu Dan Riset Manajemen (JIRM), 6(10).  
[https://scholar.google.co.id/citations?view\\_op=view\\_citation&hl=id&user=H7wZLIEAAAAJ&scstart=20&pagesize=80&citation\\_for\\_view=H7wZLIEAAAAJ:Qo2XoVZTnwC](https://scholar.google.co.id/citations?view_op=view_citation&hl=id&user=H7wZLIEAAAAJ&scstart=20&pagesize=80&citation_for_view=H7wZLIEAAAAJ:Qo2XoVZTnwC)

Ari, W. (2010). Gaya hidup dan gaya hidup sehat tantangan promosi kesehatan. Artikel perilaku hidup bersih dan sehat diakses pada tanggal 11 Maret 2016 dari <http://www.promosikesehatan.com>.

Indang Entjang. 1981. Ilmu Kesehatan Masyarakat. Bandung: Penerbit Alumni.

Anne Ahira. (2010). Pengertian Pola Hidup Sehat. Tersedia dalam <http://www.anneahira.com/pengertian-pola-hidup-sehat8691.htm>.  
[Diakses pada tanggal 13 Juni 2022].

Departemen Kesehatan RI. (1992). UU RI No.23 Tahun 1992 Tentang Kesehatan. Depkes RI.

----- (1997). Buku Panduan Manajemen Penyuluhan Kesehatan Masyarakat Tingkat Provinsi. Jakarta: Depkes RI.

----- (2002). Panduan Manajemen PHBS Menuju Kabupaten/Kota Sehat. Pusat Promosi Kesehatan. Jakarta: Depkes RI.

Soekidjo Notoatmodjo. (1993), Pengantar Pendidikan Kesehatan dan Ilmu Perilaku Kesehatan. Yogyakarta: Andi Offset

Lasantha. (2010). "Durasi tidur ideal anak sehari".  
<http://sehatzblog.blogspot.com> [diakses pada tanggal 13 Juni 2022]



**LAMPIRAN**  
**(KUESIONER, DATA)**

1. Tautan Spreadsheet:  
<https://docs.google.com/spreadsheets/d/1ijVmkFad9n1nGLn1ezU5bNTUcu2kY27L9Co6sZ8sNi8/edit?usp=sharing>
2. Tautan Google Forms:  
<https://bit.ly/Kelompok01MetStatSD-A1>
3. Tautan Google Colab:  
<https://colab.research.google.com/drive/1INcUsm2HewScwQTsmUtkWrBAD-zcCEUk?usp=sharing>