



TUGAS AKHIR - SS 141501

**ANALISIS REGRESI LOGISTIK MULTINOMIAL
PADA FAKTOR-FAKTOR YANG MEMPENGARUHI
SUMBER AIR BERSIH RUMAH TANGGA
DI JAWA TIMUR**

**ZABRINNA ZENITHA ZAHROH
NRP 062114 4000 0120**

**Dosen Pembimbing :
Dr. Dra. Ismaini Zain, M.Si**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2018**



TUGAS AKHIR - SS 141501

**ANALISIS REGRESI LOGISTIK MULTINOMIAL
PADA FAKTOR-FAKTOR YANG MEMPENGARUHI
SUMBER AIR BERSIH RUMAH TANGGA
DI JAWA TIMUR**

**ZABRINNA ZENITHA ZAHROH
NRP 062114 4000 0120**

**Dosen Pembimbing :
Dr. Dra. Ismaini Zain, M.Si**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2018**

(Halaman ini sengaja dikosongkan)



FINAL PROJECT - SS 141501

**MULTINOMIAL LOGISTIC REGRESSION ANALYSIS
ON FACTORS AFFECTING HOUSEHOLD CLEAN
WATER SOURCE AT EAST JAVA**

**ZABRINNA ZENITHA ZAHROH
SN 062114 4000 0120**

**Supervisor
Dr. Dra. Ismaini Zain, M.Si**

**UNDERGRADUATE PROGRAMME
DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS, COMPUTING, AND DATA SCIENCE
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2018**

(Halaman ini sengaja dikosongkan)

LEMBAR PENGESAHAN

**ANALISIS REGRESI LOGISTIK MULTINOMIAL
PADA FAKTOR-FAKTOR YANG MEMPENGARUHI
SUMBER AIR BERSIH RUMAH TANGGA
DI JAWA TIMUR**

TUGAS AKHIR

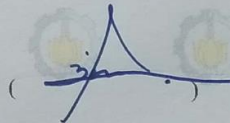
Diajukan untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Sains
pada
Program Studi Sarjana Departemen Statistika
Fakultas Matematika, Komputasi, dan Sains Data
Institut Teknologi Sepuluh Nopember

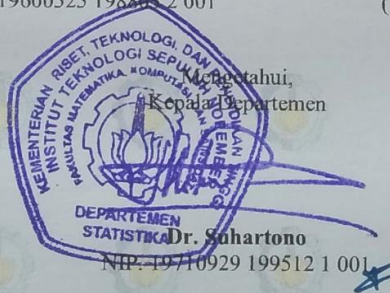
Oleh :

Zabrinna Zenitha Zahroh
NRP. 062114 4000 0120

Disetujui oleh Pembimbing:

Dr. Dra. Ismaini Zain, M.Si.
NIP. 19600525 198803 2 001




Ketua, Kepala Departemen
Dr. Suhartono
NIP. 19710929 199512 1 001

SURABAYA, AGUSTUS 2018

(Halaman ini sengaja dikosongkan)

ANALISIS REGRESI LOGISTIK MULTINOMIAL PADA FAKTOR-FAKTOR YANG MEMPENGARUHI SUMBER AIR BERSIH RUMAH TANGGA DI JAWA TIMUR

Nama Mahasiswa : Zabrinna Zenitha Zahroh
NRP : 062114 4000 0120
Departemen : Statistika
Dosen Pembimbing : Dr. Dra. Ismaini Zain, M.Si

Abstrak

Air merupakan sumber daya yang penting bagi kehidupan masyarakat guna memenuhi kebutuhannya seperti untuk minum, masak, mandi/cuci dan lainnya. Hal tersebut dijamin dalam program Sustainable Development Goals atau SDGs, dimana menjadi acuan bagi pembangunan negara, baik untuk negara maju ataupun negara berkembang guna menjawab keteringgalan pembangunan negara salah satunya dengan tercapainya ketersediaan air bersih bagi masyarakat. Pada penelitian ini, didapatkan hasil karakteristik yaitu variabel X_1 sebanyak 52% terletak didaerah pedesaan dan 48% terletak didaerah perkotaan. Variabel X_2 sebanyak 70% yaitu SD, 6% yaitu SMP, 18% yaitu SMA, dan 6% yaitu perguruan tinggi. Variabel X_3 sebanyak 91% bangunan rumah tangga berstatus milik sendiri dan 9% berstatus bukan milik sendiri. Variabel X_4 sebanyak 47% rumah tangga menggunakan perpipaan/hidran umum dan 53% tidak menggunakan penyaluran air. Variabel X_5 atau jumlah anggota keluarga di Jawa Timur rata-rata memiliki anggota keluarga sebanyak 3,4558 atau sebanyak 4 anggota keluarga dengan minimum 1 anggota keluarga dalam keluarga dan maximum 14 anggota keluarga dalam keluarga. Penelitian ini menghasilkan nilai akurasi dalam analisis model sebesar 67,3%.

Kata kunci : *Air Bersih, Pola Konsumsi, Regresi Logistik Multinomial.*

(Halaman ini sengaja dikosongkan)

MULTINOMIAL LOGISTIC REGRESSION ANALYSIS ON FACTORS AFFECTING HOUSEHOLD CLEAN WATER SOURCE AT EAST JAVA

Name : Zabrinna Zenitha Zahroh
SN : 062114 4000 0120
Department : Statistics
Supervisor : Dr. Dra. Ismaini Zain, M.Si

Abstract

Water is an important resource for people's lives to meet their needs such as for drinking, cooking, bathing / laundry and more. It is guaranteed in the Sustainable Development Goals or SDGs program, which becomes a reference for the development of the country, both for developed countries and developing countries in order to answer the country's lack of development by achieving the availability of clean water for the community. In this research, the characteristic of X_1 variable is 52% located in rural area and 48% is located in urban area. Variable X_2 as much as 70% that is SD, 6% that is junior, 18% is high school, and 6% that is college. The X_3 variable is 91% of self-owned households and 9% non-self-owned. Variable X_4 as much as 47% of households use general pipes / hydrants and 53% do not use water delivery. The variable X_5 or the number of family members in East Java has an average family member of 3.4558 or as many as 4 family members with a minimum of 1 family member in the family and a maximum of 14 family members in the family. This research yields accuracy value in model analysis equal to 67,3%.

Keywords: *Consumption Scheme, Multinomial Logistic Regression, Water Supply.*

(Halaman ini sengaja dikosongkan)

KATA PENGANTAR

Puji syukur penulis panjatkan atas rahmat dan hidayah yang diberikan Allah SWT sehingga penulis dapat menyelesaikan laporan Tugas Akhir yang berjudul **“Analisis Pola Konsumsi Rumah Tangga Di Jawa Timur Dalam Memenuhi Kebutuhan Air Bersih Menggunakan Regresi Logistik Multinomial”** dengan lancar.

Penulis menyadari bahwa Tugas Akhir ini dapat terselesaikan tidak terlepas dari bantuan dan dukungan berbagai pihak. Oleh karena itu, penulis menyampaikan terima kasih kepada :

1. Dr. Dra. Ismaini Zain, M.Si. selaku dosen pembimbing Tugas Akhir serta dosen wali yang dengan sangat sabar memeberi bimbingan, saran, dukungan serta motivasi selama penyusunan Tugas Akhir dan memberi arahan dalam proses belajar di Departemen Statistika.
2. Dr. Vita Ratnasari, S.Si., M.Si. dan Erma Oktania Permatasari, S.Si., M.Si. selaku dosen penguji yang telah banyak memberi masukan kepada penulis.
3. Dr. Suhartono Ketua Departemen Statistika, yang telah memberikan fasilitas, sarana, dan prasarana.
4. Dr. Sutikno, M.Si. selaku Ketua Program Studi Sarjana yang telah memberikan fasilitas, sarana, dan prasarana.
5. Kedua orang tua, atas segala do'a, nasehat, kasih sayang, dan dukungan yang diberikan kepada penulis demi kesuksesan dan kebahagiaan penulis.
6. Keluarga besar Statistika ITS yang selama ini telah membantu dan mendukung penulis selama masa perkuliahan berlangsung.
7. Keluarga RESPECT dan Sahabat-sahabat penulis yang tidak dapat disebutkan satu persatu nama kalian yang selama ini telah membantu, mendukung, dan mendengarkan keluh kesah penulis selama masa perkuliahan berlangsung.
8. Teman-teman seperjuangan TA dan teman-teman yang telah lulus terlebih dahulu yang pernah seperjuangan dalam proses belajar di Statistika ITS yang saling memberikan semangat.
9. Keluarga baru di luar Statistika ITS dan HIMASTA-ITS yang selama ini telah membantu, mendukung, dan mendengarkan keluh kesah penulis selama ini.

10. Semua pihak yang turut membantu dalam pelaksanaan Tugas Akhir yang tidak bisa penulis sebutkan satu persatu.

Besar harapan penulis untuk mendapatkan kritik dan saran yang membangun sehingga Tugas Akhir ini dapat memberikan manfaat bagi semua pihak yang terkait.

Surabaya, Juni 2018

Penulis

DAFTAR ISI

LEMBAR PENGESAHAN.....	Error! Bookmark not defined.
Abstrak.....	vii
Abstract	ix
DAFTAR TABEL	xv
DAFTAR GAMBAR.....	xvii
DAFTAR LAMPIRAN	xix
BAB I.....	1
PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan.....	4
1.4 Manfaat.....	4
BAB II.....	5
TINJAUAN PUSTAKA	5
2.1 Uji Independensi.....	5
2.2 Uji Korelasi.....	5
2.3 Regresi Logistik.....	6
2.4 Regresi Logistik Multinomial.....	7
2.4.1 Estimasi Parameter	8
2.4.2 Pengujian Signifikansi Parameter.....	9
2.5 Uji Kesesuaian Model.....	10
2.6 Interpretasi Koefisien Parameter	11
2.7 Pengukuran Akurasi Klasifikasi	13
2.8 Faktor-faktor yang Mempengaruhi Pemenuhan Kebutuhan Konsumsi Air Bersih.....	13
BAB III.....	15
METODOLOGI PENELITIAN	15
3.1 Sumber Data	15
3.2 Variabel Penelitian.....	15
3.3 Definisi Operasional	17
3.4 Langkah Penelitian	18
3.5 Diagram Alir.....	19
BAB IV.....	21
ANALISIS DAN PEMBAHASAN.....	21
4.1 Analisis Statistik Deskriptif.....	21
4.2 Pemodelan Regresi Logistik Multinomial	22

4.2.1 Uji Kesuaian Model.....	24
4.2.2 Pengujian Signifikansi Parameter.....	25
4.2.3 Interpretasi Analisis Regresi Logistik Multinomial	26
4.2.4 Ketepatan Klasifikasi.....	30
BAB V.....	31
KESIMPULAN DAN SARAN	31
5.1 Kesimpulan.....	31
5.2 Saran	32
LAMPIRAN	35
BIODATA PENULIS.....	39

DAFTAR TABEL

Tabel 2.1 Nilai Regresi Logistik untuk Variabel Prediktor	11
Tabel 3.1 Variabel Respon Penelitian.....	16
Tabel 3.2 Variabel Prediktor Penelitian.....	16
Tabel 3.3 Struktur Data Variabel Penelitian	17
Tabel 4.1 Statistik Deskriptif Variabel Numerik	22
Tabel 4.2 Chi-Square Test Variabel Kategorik.....	23
Tabel 4.3 Korelasi Variabel Numerik	23
Tabel 4.4 Uji Kesesuaian Model.....	24
Tabel 4.5 Nilai Pseudo R-Square.....	24
Tabel 4.6 Uji Serentak	25
Tabel 4.7 Uji Parsial	25
Tabel 4.8 Parameter Pembentuk Fungsi Logit.....	26
Tabel 4.9 Interpretasi Analisis Regresi Logistik Multinomial.....	28
Tabel 4.10 Keakurasian Klasifikasi	30

(Halaman ini sengaja dikosongkan)

DAFTAR GAMBAR

Gambar 2.1 Kerangka Konsep Penelitian	14
Gambar 3.1 Diagram Variabel Respon	15
Gambar 3.2 Diagram Alir Penelitian	19
Gambar 4.1 Statistik Deskriptif Variabel Kategorik	21

(Halaman ini sengaja dikosongkan)

DAFTAR LAMPIRAN

Lampiran 1 Surat Keterangan Pengambilan Data.....	35
Lampiran 2 Data Penelitian	36

(Halaman ini sengaja dikosongkan)

BAB I PENDAHULUAN

1.1 Latar Belakang

Air dikonsumsi untuk memenuhi kebutuhan sehari-hari seperti untuk minum, memasak, mandi/cuci dan lainnya. Air yang dapat dikonsumsi memiliki berbagai jenis dari air kemasan bermerek, air isi ulang, leding meteran, leding eceran, sumur bor/pompa, sumur terlindungi, sumur tak terlindungi, mata air terlindungi, mata air tak terlindungi, air permukaan yang terdiri dari sungai, danau, waduk, kolam dan irigasi, lalu ada pula air hujan. Pemenuhan akan konsumsi kebutuhan air bersih yang diperlukan oleh seluruh masyarakat dijamin dalam tujuan MDGs (*Millenium Development Goals*) yang dibuat oleh organisasi UNDP (*United Nations Development Programs*), yang menyebutkan bahwa kebutuhan air bersih dan sanitasi yang lebih baik masuk dalam tujuan ke-7, yaitu “*ensure environmental sustainability*” atau menjamin terciptanya kelestarian lingkungan secara global (ESCAP-UNDP, 2002). Indonesia yang dirasa masih kesulitan untuk mencapai pemenuhan akan kebutuhan air bersih. Maka dari itu dibuatlah kesepakatan baru sebagai pengganti MDGs yang berlaku tahun 2015-2030 yaitu *Sustainable Development Goals* atau SDGs, dimana menjadi acuan bagi pembangunan negara, baik untuk negara maju ataupun negara berkembang guna menjawab ketertinggalan pembangunan negara salah satunya dengan tercapainya ketersediaan air bersih bagi masyarakat, meskipun hal tersebut masih dianggap tidak mudah untuk dicapai (Giddings, 2002). Indonesia yang merupakan negara berkembang juga menggunakan program ini untuk memenuhi ketersediaan air bersih, namun di Indonesia masih banyak pulau-pulau yang kekurangan air bersih diantaranya pulau Jawa, Sulawesi, Bali, dan Nusa Tenggara. Dari keempat pulau tersebut Jawa memiliki krisis ketersediaan air yang paling tinggi, berdasarkan Badan Penanggulangan Bencana Daerah (BPBD) Jawa Timur memiliki krisis ketersediaan air paling tinggi dibandingkan Jawa Barat dan Jawa Tengah. Sebanyak 25 daerah di Jawa Timur yang mengalami krisis ketersediaan air bersih.

Survei Sosial Ekonomi Nasional (2016) menyebutkan jenis air bersih yang dapat digunakan dalam memenuhi konsumsi kebutuhan air antara lain air kemasan bermerek, air isi ulang, leding meteran, leding eceran, sumur bor/pompa, sumur terlindungi, sumur tak terlindungi, mata air terlindungi, mata air tak terlindungi, air permukaan yang terdiri dari sungai, danau, waduk, kolam dan irigasi, serta air hujan. Masyarakat dunia pasti mengkonsumsi air bersih dalam kebutuhan sehari-harinya dimana dari jenis air yang disebutkan oleh Survei Sosial Ekonomi Nasional (SUSENAS) dapat dikategorikan berdasarkan cara memperolehnya, yaitu baik dengan cara membeli atau tidak membeli. Jika suatu rumah tangga dalam memenuhi konsumsi air bersih dengan cara membeli maka akan ada pengeluaran untuk konsumsi air tersebut sedangkan jika tidak membeli mungkin rumah tangga tersebut memiliki sumber air tersendiri yang tidak perlu adanya pengeluaran. Dalam pemenuhan kebutuhan air bersih tersebut pasti memiliki faktor yang mempengaruhinya. Dari pernyataan tentang cara memperoleh air bersih terdapat suatu bahasan yang menarik untuk diteliti, yaitu tentang pola konsumsi rumah tangga dalam memenuhi kebutuhan air bersih serta ingin mengetahui karakteristik dan model dari pola konsumsi tersebut dengan faktor-faktor yang diduga mempengaruhi.

Menurut Mustikowati (2014) dalam penelitian sebelumnya menjelaskan bahwa faktor yang mempengaruhi konsumsi air antara lain pendapatan keluarga, jumlah anggota keluarga, kepemilikan tempat tinggal, kepemilikan sumber air lain diluar PDAM dan kualitas air PDAM. Lalu pada penelitian Winarna (2003) dalam penelitiannya faktor yang mempengaruhi konsumsi air bersih antara lain jumlah anggota keluarga, serta tingkat pendidikan kepala rumah tangga. Berdasarkan penelitian sebelumnya, faktor yang diduga mempengaruhi pola konsumsi rumah tangga dalam memenuhi kebutuhan air bersih antara lain klasifikasi perkotaan atau pedesaan, kepemilikan tempat tinggal, pendidikan terakhir kepala rumah tangga, pendapatan keluarga, jenis kegiatan pemakaian air, penyaluran air yang digunakan, jarak ke sanitasi, dan jumlah anggota keluarga.

Dimana pada penelitian ini, variabel sumber air bersih menjadi variabel respon dengan kategori berdasarkan sumber air bersihnya

sedangkan variabel prediktor sebagai faktor-faktor yang mempengaruhinya terdiri dari klasifikasi perkotaan atau pedesaan, kepemilikan tempat tinggal, pendidikan terakhir kepala rumah tangga, biaya pengeluaran keluarga, penyaluran air yang digunakan dan jumlah anggota keluarga, dimana semua data merupakan data kategorik kecuali variabel jumlah anggota keluarga dan biaya pengeluaran keluarga merupakan data rasio. Kemudian dari variabel-variabel diatas dapat didefinisikan bagaimana karakteristik pola konsumsi rumah tangga tersebut dalam memenuhi kebutuhan air bersih dengan menggunakan statistika deskriptif serta sebagai capaian dari penelitian ini adalah terbentuknya model antara sumber air bersih rumah tangga di Jawa Timur dalam memenuhi kebutuhan air bersih dengan faktor-faktor yang mempengaruhinya, karena pada variabel respon dikategorikan menjadi 4 berdasarkan sumber air bersih maka pada penelitian ini menggunakan metode regresi logistik multinomial. Menurut Afidah (2011) dalam penelitiannya memberikan hasil bahwa dengan regresi logistik multinomial diperoleh nilai akurasi sebesar 47,7% dalam mempresentasikan model yang dikarenakan prediktor yang kurang mewakili pola model, sedangkan menurut Parlinggoman (2016) pada penelitiannya memberikan hasil dengan regresi logistik multinomial diperoleh nilai akurasi sebesar 83,1%.

Berdasarkan uraian tersebut, penelitian ini untuk mengetahui karakteristik dan model analisis pada faktor-faktor yang mempengaruhi sumber air bersih rumah tangga di Jawa Timur dengan menggunakan analisis regresi logistik multinomial.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijabarkan, rumusan masalah yang akan diselesaikan dalam penelitian ini adalah sebagai berikut.

1. Bagaimana karakteristik analisis pada faktor-faktor yang mempengaruhi sumber air bersih rumah tangga di Jawa Timur?
2. Bagaimana pemodelan analisis pada faktor-faktor yang mempengaruhi sumber air bersih rumah tangga di Jawa Timur?

1.3 Tujuan

Berdasarkan rumusan masalah yang telah disusun, tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut.

1. Mendeskripsikan karakteristik analisis pada faktor-faktor yang mempengaruhi sumber air bersih rumah tangga di Jawa Timur.
2. Memodelkan analisis pada faktor-faktor yang mempengaruhi sumber air bersih rumah tangga di Jawa Timur.

1.4 Manfaat

Manfaat yang ingin dicapai dari penelitian ini adalah diperoleh model dari analisis sumber air bersih rumah tangga di Jawa Timur sehingga dapat diketahui hubungan dengan faktor-faktor yang mempengaruhi pola konsumsi air bersih tersebut. Model yang diperoleh diharapkan dapat menjadi bahan pertimbangan bagi pihak pemerintah dalam membuat suatu kebijakan di berbagai bidang salah satunya lingkungan dan kesehatan, serta mampu membantu mempercepat pencapaian target pemenuhan konsumsi air bersih bagi separuh penduduk sesuai target *Sustainable Development Goals* (SDGs).

BAB II TINJAUAN PUSTAKA

Bagian ini akan membahas beberapa kajian pustaka yang digunakan untuk menyelesaikan permasalahan mengenai analisis regresi logistik multinomial pada sumber air bersih rumah tangga di Jawa Timur.

2.1 Uji Independensi

Uji independensi merupakan suatu uji untuk mengetahui hubungan antar variabel. Hipotesis yang digunakan adalah sebagai berikut (Agresti, 1990).

H_0 : Tidak ada hubungan antar variabel

H_1 : Ada hubungan antar variabel

Statistik uji yang digunakan dalam pengujian ini adalah uji Chi-square sebagaimana terdapat pada persamaan 2.1.

$$X^2 = \sum_{i=1}^a \sum_{j=1}^b \frac{(n_{ij} - \hat{m}_{ij})^2}{\hat{m}_{ij}} \quad (2.1)$$

dimana :

n_{ij} = Jumlah pengamatan pada baris ke-i, kolom ke-j

\hat{m}_{ij} = Nilai harapan baris ke-i, kolom ke-j

i = banyaknya baris (1,2,...,a)

j = banyaknya kolom (1,2,...,b)

Penolakan hipotesis awal apabila $X^2 \geq X^2_{(db,1-\alpha)}$ dengan derajat bebas (i-1)(j-1).

2.2 Uji Korelasi

Analisis korelasi merupakan metode analisis data yang mengukur derajat hubungan antara dua variabel random X dan Y melalui sebuah bilangan yang disebut koefisien korelasi r. Kuadrat dari koefisien korelasi disebut koefisien determinasi yang merepresentasikan besarnya proporsi variasi dalam variabel y yang dijelaskan oleh variabel x dalam model. Hipotesis yang digunakan adalah sebagai berikut (Draper & Smith, 1992).

$H_0 : \rho=0$

$H_1 : \rho \neq 0$

Statistik Uji yang digunakan.

$$r = \frac{n \sum_{i=1}^n xy - \sum_{i=1}^n x \sum_{i=1}^n y}{\sqrt{\left[n \sum_{i=1}^n x^2 - \left(\sum_{i=1}^n x \right)^2 \right] \left[n \sum_{i=1}^n y^2 - \left(\sum_{i=1}^n y \right)^2 \right]}}$$

Besar r mengukur keeratan hubungan linier antara peubah-peubah acak Y dan X . Misal $r=1$ maka berkorelasi positif sempurna dan kemungkinan nilai Y dan X terletak pada suatu garis lurus dengan kemiringan yang positif pada bidang-bidang YX . Jika $r=0$ maka kedua peubah dikatakan tidak berkorelasi, artinya tidak berhubungan linier satu sama lain (Draper & Smith, 1992).

2.3 Regresi Logistik

Regresi adalah bagaimana satu variabel yaitu variabel dependen dipengaruhi oleh satu atau lebih variabel lain yaitu variabel prediktor dengan tujuan untuk memprediksi nilai rata-rata variabel respon didasarkan pada nilai variabel prediktor yang telah diketahui. Tujuan utama regresi adalah untuk memprediksi nilai variabel respon berdasarkan satu atau lebih variabel prediktor (Widarjono, 2010). Banyak kasus dalam analisis regresi dimana variabel responnya bersifat kualitatif. Variabel respon ini bisa mempunyai dua kelas atau kategori (biner) dan lebih dari 1 kelas (multinomial). Salah satu pendekatan yang digunakan untuk mengestimasi model regresi dengan variabel respon bersifat kualitatif adalah dengan model probabilitas logistik atau disingkat logit (Widarjono, 2010). Agresti (2007) menyatakan bahwa variabel dalam regresi logistik dapat berupa kategori atau kualitatif. Menurut Hosmer dan Lemeshow (2000), tujuan melakukan analisis data menggunakan regresi logistik adalah untuk mendapatkan model terbaik dan sederhana, namun model tersebut sejalan dengan tinjauan dari ilmu biologi untuk menjelaskan hubungan di antara hasil variabel respon dengan variabel prediktor.

2.4 Regresi Logistik Multinomial

Regresi logistik multinomial merupakan suatu metode analisis data yang digunakan untuk mencari hubungan antara variabel respon (y) yang bersifat polikotomus atau multinomial (Hosmer & Lemeshow, 2000). Variabel respon y terdiri lebih dari 2 kategori yang biasanya y dinotasikan dengan 0,1, atau 2. Hosmer dan Lemeshow (2000) menjelaskan bahwa model yang digunakan pada regresi logistik multinomial adalah pada persamaan 2.2.

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_b x_b)}{1 + \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_b x_b)} \quad (2.2)$$

Dilakukan transformasi logit terhadap $\pi(x)$, didapat persamaan fungsi logit yang dinyatakan didalam persamaan 2.3 dan persamaan 2.4.

$$P_1(x) = \ln\left(\frac{P(Y=1)1|x}{P(Y=1)0|x}\right) = \beta_{10} + \beta_{11}x_1 + \beta_{12}x_2 + \dots + \beta_{1b}x_b \quad (2.3)$$

$$= x'\beta_1$$

$$P_2(x) = \ln\left(\frac{P(Y=1)2|x}{P(Y=1)0|x}\right) = \beta_{20} + \beta_{21}x_1 + \beta_{22}x_2 + \dots + \beta_{2b}x_b \quad (2.4)$$

$$= x'\beta_2$$

Berdasarkan kedua fungsi logit tersebut maka didapatkan model regresi logistik yang terdapat pada persamaan 2.5, persamaan 2.6, dan persamaan 2.7.

$$\pi_0(x) = \frac{1}{1 + \exp P_1(x) + \exp P_2(x)} \quad (2.5)$$

$$\pi_1(x) = \frac{\exp P_1(x)}{1 + \exp P_1(x) + \exp P_2(x)} \quad (2.6)$$

$$\pi_2(x) = \frac{\exp P_2(x)}{1 + \exp P_1(x) + \exp P_2(x)} \quad (2.7)$$

2.4.1 Estimasi Parameter

Pada regresi linier, metode yang paling sering digunakan untuk mengestimasi parameter yang tidak diketahui adalah least square. Metode ini akan meminimumkan jumlah kuadrat residual. Metode ini akan menghasilkan estimator yang valid apabila digunakan pada data dengan asumsi IIDN yang terpenuhi. Sebaliknya, metode ini akan menghasilkan estimator yang bias jika diaplikasikan pada variabel respon yang bersifat polikotomus. Metode yang mengarah pada fungsi *least square* dalam model regresi linier disebut maximum likelihood dengan logaritma likelihood yang terdapat pada persamaan 2.8.

$$l(\beta) = \prod_{i=1}^n \pi_{1i}^{y_{1i}} \pi_{2i}^{y_{2i}} \pi_{3i}^{y_{3i}} \left(1 - \pi_{1i} - \pi_{2i} - \pi_{3i}\right)^{1 - y_{1i} - y_{2i} - y_{3i}} \quad (2.8)$$

dengan, $\sum_{i=1}^3 y_{ji} = 1$

Dilakukan transformasi logaritma, sehingga didapat persamaan logaritma likelihood dengan fungsi linier.

$$\begin{aligned}
 L(\beta) &= \log [l(\beta)] \\
 &= \sum_{i=1}^n (y_{1i}x_i\beta_1 + y_{2i}x_i\beta_2 + y_{3i}x_i\beta_3 \\
 &\quad + \ln(1 - \pi_{1i} - \pi_{2i} - \pi_{3i}))
 \end{aligned} \tag{2.9}$$

Estimasi parameter β dari persamaan 2.9 diperoleh dengan memaksimumkan $L(\beta)$. Yang perlu menjadi perhatian bahwa fungsi logaritma bersifat monoton naik sehingga jika fungsi log-likelihood mencapai maksimum, maka fungsi likelihood juga akan mencapai maksimum (Hosmer & Lemeshow, 2000). Nilai taksiran parameter β yang diturunkan dari fungsi log likelihood dapat diperoleh menggunakan iterasi Newton Raphson (Agresti, 2002), dengan melakukan penurunan dari $L(\beta)$ terhadap β_{jk} , maka diperoleh nilai β sebagaimana terdapat pada persamaan 2.10..

$$\frac{\partial L(\beta)}{\partial \beta_j} = \sum_{i=1}^n x_i (y_{ji} - \pi_{ji}) \tag{2.10}$$

dengan ; $j = 1, 2, \dots, b$

2.4.2 Pengujian Signifikansi Parameter

a. Uji Serentak

Pengujian ini dilakukan untuk memeriksa koefisien β secara serentak atau bersamaan terhadap variabel respon, dengan hipotesis sebagai berikut.

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_b = 0$$

$$H_1 : \text{paling sedikit ada satu } \beta_j \neq 0 ; j = 1, 2, \dots, b$$

Statistik Uji :

$$G = -2 \ln \frac{L_1}{L_0} \tag{2.11}$$

dimana,

L_1 = Likelihood tanpa variabel prediktor

L_0 = Likelihood dengan variabel prediktor

Statistik uji G yang didapat dari persamaan 2.11 merupakan likelihood ratio test dimana nilai G mengikuti distribusi X^2 sehingga H_0 ditolak apabila $G > X^2_{\alpha, v}$ atau P-value $< \alpha$, dengan v adalah banyaknya parameter dalam model tanpa β_0 (Hosmer & Lemeshow, 2000).

b. Uji Parsial

Untuk memeriksa kemaknaan koefisien β secara parsial dengan membandingkan dugaan β dengan penduga standar errornya, dengan hipotesis sebagai berikut (Hosmer & Lemeshow, 2000).

$H_0 : \beta_j = 0$ (Tidak ada pengaruh variabel prediktor secara parsial terhadap variabel respon)

$H_1 : \beta_j \neq 0 ; j = 1, 2, \dots, b$ (Ada pengaruh variabel prediktor secara parsial terhadap variabel respon)

Statistik Uji :

$$W^2 = \frac{\hat{\beta}_j^2}{SE(\hat{\beta}_j^2)} \quad (2.12)$$

Dengan $SE(\hat{\beta}_j^2)$ merupakan standar error koefisien dan $\hat{\beta}_j$ merupakan nilai koefisien dugaan variabel prediktor W^2 . Statistik uji W^2 disebut sebagai statistik uji Wald sebagaimana yang tertulis pada persamaan 2.12, mengikuti distribusi X^2 sehingga H_0 ditolak apabila $W^2 > X^2_{(\alpha, v)}$ atau P-value $< \alpha$, dimana v adalah banyaknya prediktor.

2.5 Uji Kesesuaian Model

Uji kesesuaian model dilakukan untuk menguji apakah model yang dihasilkan berdasarkan regresi logistik serentak sudah layak (tidak terdapat perbedaan antara hasil pengamatan dan kemungkinan hasil prediksi model) dimana uji tersebut sesuai dengan persamaan 2.13 (Hosmer & Lemeshow, 2000). Hipotesis yang digunakan adalah sebagai berikut.

H_0 : Model sesuai (tidak terdapat perbedaan antara hasil pengamatan dengan kemungkinan hasil prediksi model)

H_1 : Model tidak sesuai (terdapat perbedaan antara hasil pengamatan dengan kemungkinan hasil prediksi model)

Statistik Uji :

$$C = \frac{\sum_{k=1}^g \left(O_k - n'_k \bar{\pi}_k \right)^2}{\sum_{k=1}^g n'_k \bar{\pi}_k (1 - \bar{\pi}_k)} \quad (2.13)$$

dengan,

$O_k = \sum_{j=1}^{C_k} y_j =$ Jumlah nilai variabel respon pada grup ke-k

$\bar{\pi}_k = \sum_{j=1}^{C_k} \frac{m_j \pi(x_j)}{n'_k} =$ Rata-rata taksiran probabilitas

$g =$ Jumlah grup

$n'_k =$ Banyaknya subjek pada grup ke-k

$m_j =$ Banyaknya subjek yang memiliki nilai $\bar{\pi}_j$

$C_k =$ Kategori variabel respon

Tolak H_0 , apabila $\hat{C} > X^2_{\alpha, v}$ yang menyatakan bahwa model tidak sesuai (terdapat perbedaan antara hasil pengamatan dengan kemungkinan hasil prediksi model) (Hosmer & Lemeshow, 2000).

2.6 Interpretasi Koefisien Parameter

Intepretasi pada metode regresi logistik menggunakan nilai odds ratio. Untuk dapat menjelaskan odds ratio, digunakan dengan Tabel 2.1 dimana dicontohkan variabel respon dengan tiga kategori dan variabel prediktor dengan dua kategori.

Tabel 2.1 Nilai Regresi Logistik untuk Variabel Prediktor

Variabel Respon	Variabel X=0
(y) Y=0	$\pi_0(0) = \frac{1}{1 + \exp(\beta_{10} + \beta_{20})}$

Tabel 2.2 Nilai Regresi Logistik untuk Variabel Prediktor (Lanjutan)

Variabel Respon	Variabel X=0
(y) Y=1	$\pi_1(0) = \frac{\exp(\beta_{10})}{1 + \exp(\beta_{10}) + \exp(\beta_{20})}$
Y=2	$\pi_2(0) = \frac{\exp(\beta_{20})}{1 + \exp(\beta_{10}) + \exp(\beta_{20})}$

Tabel 2.1 Nilai Regresi Logistik untuk Variabel Prediktor (Lanjutan)

Variabel Respon	Bebas (x) X=1
(y) Y=0	$\pi_0(1) = \frac{1}{1 + \exp(\beta_{10} + \beta_{11}) + \exp(\beta_{20} + \beta_{21})}$
Y=1	$\pi_1(1) = \frac{\exp(\beta_{10} + \beta_{11})}{1 + \exp(\beta_{10} + \beta_{11}) + \exp(\beta_{20} + \beta_{21})}$
Y=2	$\pi_2(1) = \frac{\exp(\beta_{20} + \beta_{21})}{1 + \exp(\beta_{10} + \beta_{11}) + \exp(\beta_{20} + \beta_{21})}$

Didapatkan odds ratio sebagai berikut.

$$\psi_1 = \frac{\pi_1(1)/\pi_0(1)}{\pi_1(0)/\pi_0(0)} \quad (2.14)$$

$$\psi_2 = \frac{\pi_2(1)/\pi_0(1)}{\pi_2(0)/\pi_0(0)}$$

Nilai odds ratio ψ tersebut pada persamaan 2.14 digunakan untuk menunjukkan kecenderungan hubungan suatu variabel prediktor terhadap variabel respon. Dengan mengasumsikan $Y=0$ sebagai kategori kontrol, dengan kata lain sebagai acuan. Nilai odds ratio dari $Y=i$ merupakan nilai perbandingan dengan $Y=0$ untuk nilai kovariat $x=a$ dibandingkan $x=b$. (Hosmer & Lemeshow, 2000).

$$OR_j(a,b) = \frac{P(Y = j | x = a) / P(Y = 0 | x = a)}{P(Y = j | x = b) / P(Y = 0 | x = b)} \quad (2.15)$$

Untuk variabel non kategorik nilai odds ratio adalah sebagai berikut.

$$OR_j(c) = \frac{P(Y = j | x + c) / P(Y = 0 | x + c)}{P(Y = j | x) / P(Y = 0 | x)} \quad (2.16)$$

2.7 Pengukuran Akurasi Klasifikasi

Akurasi klasifikasi pada data imbalance dapat dihitung menggunakan geometrik mean (G-mean) dan Area Under ROC Curve (AUC) yang terdapat pada persamaan 2.17. G-mean untuk kasus mengelompokkan dengan jumlah kategori lebih dari dua. (Bekkar, Djeemaa, & Alitouch, 2013).

$$G_{mean} = \left(\prod_{i=1}^a R_i \right)^{1/a}$$

$$AUC = \frac{1}{c} \sum_{i=1}^a R_i \quad (2.17)$$

Dimana,

$$R_i = \frac{n_{ii}}{\sum_{l=1}^c n_{il}} ; i = 1, 2, \dots, a$$

Keterangan:

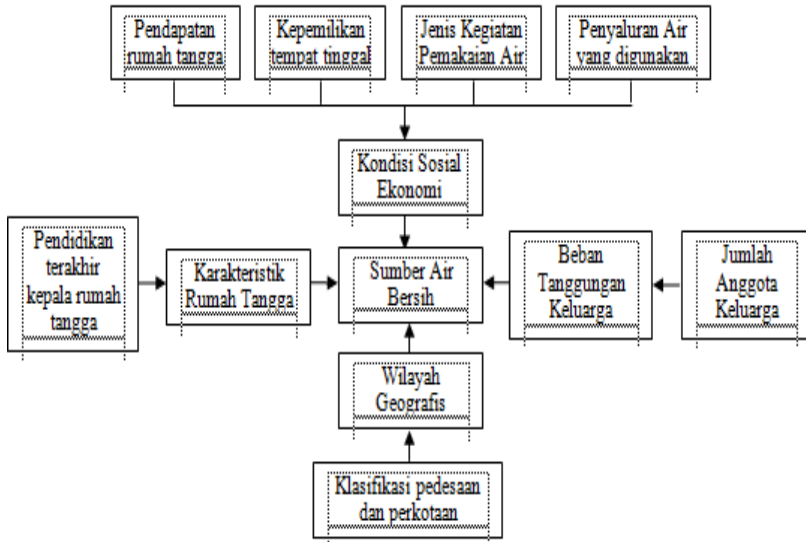
n_{ii} : jumlah prediksi $\hat{\pi}_i$ yang tepat diklasifikasikan ke π_i

n_{il} : jumlah prediksi $\hat{\pi}_i$ yang salah diklasifikasikan ke π_i

2.8 Faktor-faktor yang Mempengaruhi Pemenuhan Kebutuhan Konsumsi Air Bersih

Kebutuhan air adalah banyaknya jumlah air yang dibutuhkan untuk keperluan rumah tangga, industri, penggelontoran kota (PERPAMSI, 1994) serta menjadi prioritas kebutuhan air domestik, industri, pelayanan umum dan kebutuhan air untuk mengganti

kebocoran (Moegijantoro, 1996) dan untuk keperluan sehari-hari. Konsumsi air bersih dipengaruhi oleh beberapa faktor (Hall & James, 2001).



Gambar 2.1 Kerangka Konsep Penelitian

Penelitian yang terkait dengan faktor-faktor yang mempengaruhi konsumsi air bersih telah dilakukan oleh Winarna (2003) dalam penelitiannya faktor yang mempengaruhi konsumsi air bersih antara lain yaitu jumlah anggota keluarga, serta tingkat pendidikan kepala rumah tangga. Menurut Mutmainah (2011) faktor yang mempengaruhi pola konsumsi air bersih yaitu jumlah anggota keluarga serta kepemilikan sumur, Penyaluran air dan menurut Mustikowati (2014) dalam penelitiannya menjelaskan bahwa faktor yang mempengaruhi konsumsi air antara lain pendapatan keluarga, jumlah anggota keluarga, kepemilikan tempat tinggal, kepemilikan sumber air lain diluar PDAM dan kualitas air PDAM.

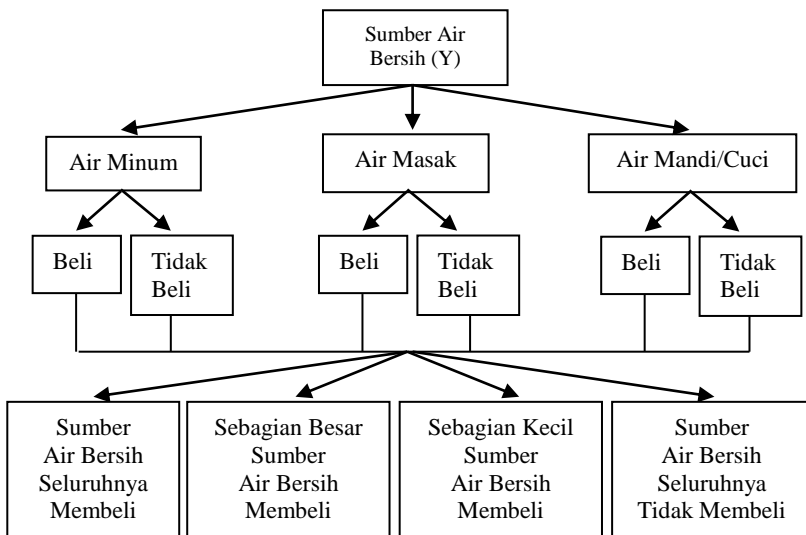
BAB III METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder, diambil dari Survey Ekonomi Sosial (SUSENAS) yang diperoleh di Badan Pusat Statistik (BPS) Propinsi Jawa Timur, dengan data SUSENAS tahun 2016. Unit pengamatan adalah rumah tangga.

3.2 Variabel Penelitian

Penelitian ini memiliki variabel respon berupa sumber air bersih rumah tangga. Berdasarkan SUSENAS, sumber air bersih tersebut terdiri dari air kemasan bermerek, air isi ulang, leding eceran, leding meteran, sumur bor/pompa, sumur terlindungi, sumur tak terlindungi, mata air terlindungi, mata air tak terlindungi, air permukaan, dan air hujan yang kemudian dipilah menjadi sumber air bersih yang membeli dan sumber air bersih tidak membeli. Sumber air bersih membeli terdiri dari air kemasan bermerek, air isi ulang, leding eceran, dan leding meteran sedangkan sisanya merupakan sumber air bersih tidak membeli, dimana dibagikan seperti pada Gambar 3.1.



Gambar 3.1 Diagram Variabel Respon

Berdasarkan Gambar 3.1 suatu rumah tangga mendapatkan sumber air bersih yaitu dengan membeli atau tidak membeli serta kegunaan sumber air bersih tersebut, maka variabel respon yang berupa sumber air bersih ini dikategorikan menjadi empat kategori. Kategori pertama yaitu sumber air bersih seluruhnya membeli yang artinya sumber air bersih untuk minum, masak, maupun mandi rumah tangga tersebut membeli, kategori kedua yaitu sebagian besar membeli yang artinya rumah tangga tersebut hanya membeli sumber air bersih untuk minum-masak, minum-mandi, atau masak-mandi, kategori tiga yaitu sebagian kecil membeli yang artinya rumah tangga tersebut hanya membeli sumber air bersih untuk kebutuhan minum atau masak atau mandi, dan kategori empat yaitu seluruhnya tidak membeli yang artinya rumah tangga tersebut tidak membeli sumber air bersih untuk kebutuhan minum, masak, maupun mandi. Berdasarkan keempat kategori tersebut, adapun rumah tangga yang terdapat pada tiap kategori disajikan pada Tabel 3.1.

Tabel 3.1 Variabel Respon Penelitian

Kategori	Keterangan	Jumlah (RT)
1	Sumber air bersih seluruhnya membeli	4.630
2	Sebagian besar sumber air bersih membeli	1.821
3	Sebagian kecil sumber air bersih membeli	4.358
4	Sumber air bersih seluruhnya tidak membeli	18.668

Variabel prediktor penelitian yang digunakan dalam analisis penelitian ditampilkan pada Tabel 3.2.

Tabel 3.2 Variabel Prediktor Penelitian

Variabel	Nama Variabel	Skala	Keterangan
X ₁	Klasifikasi desa atau kota	Kategorik	1 : Pedesaan 2 : Perkotaan
X ₂	Pendidikan Terakhir Kepala Rumah Tangga	Kategorik	1 : ≤ SD 2 : SMP 3 : SMA 4 : Perguruan Tinggi

Tabel 3.2 Variabel Prediktor Penelitian (Lanjutan)

Variabel	Nama Variabel	Skala	Keterangan
X ₃	Status Kepemilikan Bangunan	Kategorik	1 : Bukan Milik Sendiri 2 : Milik Sendiri
X ₄	Penyaluran Air Yang Digunakan	Kategorik	1: Perpipaian/Hidran umum 2: Tidak Semua
X ₅	Jumlah Anggota Keluarga	Rasio	

Variabel-variabel pada penelitian ini digunakan untuk analisis regresi logistik biner dengan struktur data yang ditampilkan pada Tabel 3.3.

Tabel 3.3 Struktur Data Variabel Penelitian

R	Variabel	Faktor-Faktor				
T	Dependen					
	Sumber Air Bersih (Y)	(X ₁)	(X ₂)	(X ₃)	(X ₄)	(X ₅)
1	Y ₁	X _{1,1}	X _{2,1}	X _{3,1}	X _{4,1}	X _{5,1}
2	Y ₂	X _{1,2}	X _{2,2}	X _{3,2}	X _{4,2}	X _{5,2}
3	Y ₃	X _{1,3}	X _{2,3}	X _{3,3}	X _{4,3}	X _{5,3}
4	Y ₄	X _{1,4}	X _{2,4}	X _{3,4}	X _{4,4}	X _{5,4}
5	Y ₅	X _{1,5}	X _{2,5}	X _{3,5}	X _{4,5}	X _{5,5}
.						
.						
.						
a	Y _a	X _{1,a}	X _{2,a}	X _{3,a}	X _{4,a}	X _{5,a}

3.3 Definisi Operasional

Berikut ini penjelasan mengenai masing-masing variabel yang digunakan dalam penelitian ini.

1. Variabel X₁ (Klasifikasi desa atau kota), merupakan variabel nominal klasifikasi tempat tinggal dengan 2 kategori yaitu tinggal di daerah perkotaan atau daerah pedesaan. Berdasarkan UU No.5 Tahun 1979, dikatakan wilayah pedesaan jika wilayah tersebut ditempati oleh sejumlah penduduk sebagai kesatuan masyarakat di dalamnya kesatuan masyarakat dan hukum yang mempunyai organisasi pemerintahan terendah langsung dibawah camat dan berhak menyelenggarakan rumah tangganya sendiri dalam ikatan Negara Kesatuan Republik Indonesia,

sedangkan berdasarkan UU No.22/1999 tentang Otonomi Daerah perkotaan adalah kawasan yang mempunyai kegiatan utama bukan pertanian dengan susunan fungsi kawasan sebagai tempat pelayanan jasa pemerintahan, pelayanan social dan kegiatan ekonomi.

2. Variabel X_2 (Pendidikan Terakhir Kepala Rumah Tangga), merupakan variabel nominal pendidikan terakhir yang ditempuh oleh kepala rumah tangga dengan 4 kategori yaitu \leq SD meliputi Tidak/belum pernah sekolah, Paket A, SDLB, SD, MI, SMP meliputi : Paket B, SMP, SMP LB, MTs, SMA meliputi : Paket C, SMLB, SMA, MA, SMK, MAK, dan Perguruan Tinggi meliputi : D1/D2, S1, S2, S3, Tamat&Lulus.
3. Variabel X_3 (Status Kepemilikan Bangunan), merupakan variabel nominal kepemilikan tempat tinggal dengan 2 kategori yaitu bukan milik sendiri yang terdiri dari kontrak/sewa, bebas sewa, dinas dan lainnya dan milik sendiri.
4. Variabel X_4 (Penyaluran Air Yang Digunakan), merupakan variabel nominal alat yang digunakan untuk menyalurkan air dengan 4 kategori yaitu pipa, hidran umum/terminal air, tidak tau, tidak semua.
5. Variabel X_5 (Jumlah Anggota Keluarga), merupakan banyaknya anggota keluarga yang ada dalam satu keluarga tersebut.

3.4 Langkah Penelitian

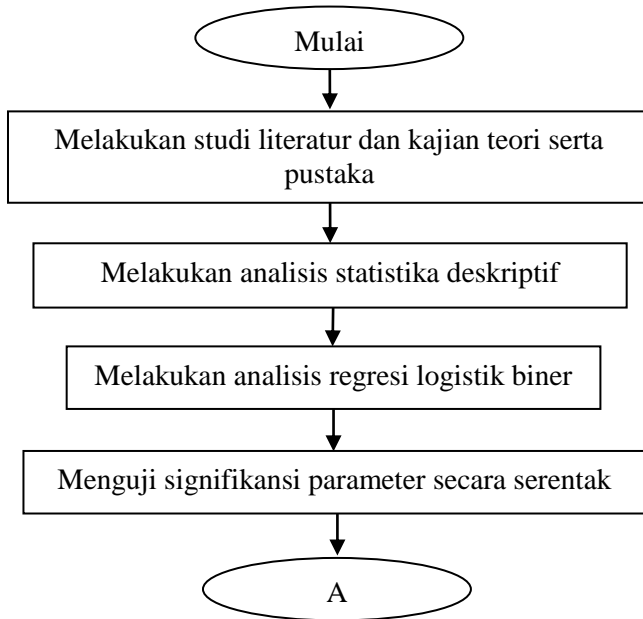
Langkah dalam penelitian ini untuk menganalisis pola konsumsi rumah tangga di Jawa Timur dalam memenuhi kebutuhan air bersih adalah sebagai berikut.

- a. Mendeskripsikan karakteristik analisis pada sumber air bersih rumah tangga di Jawa Timur dalam memenuhi kebutuhan air bersih maka digunakan diagram untuk variabel kategorik dan variabel numerik.
- b. Melakukan analisis model antara analisis pada sumber air bersih rumah tangga di Jawa Timur dengan faktor-faktor yang mempengaruhi dengan langkah sebagai berikut.
 - i. Menguji independensi antara variabel respon dengan masing-masing variabel prediktor dengan persamaan (2.1).

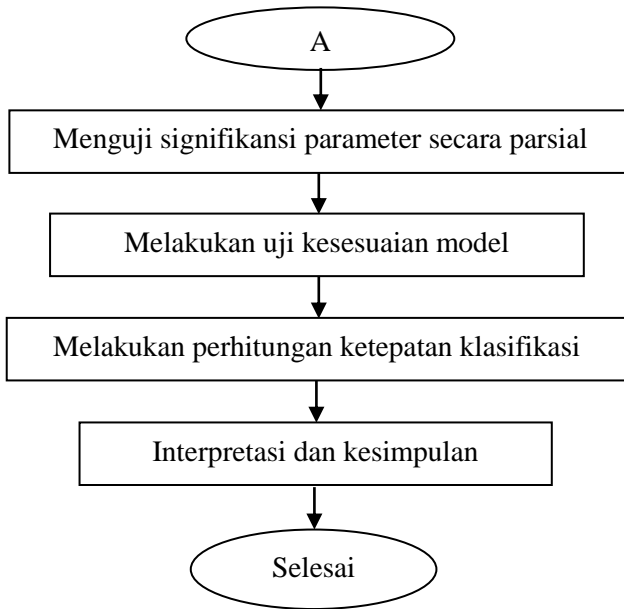
- ii. Melakukan analisis regresi logistik multinomial dengan menggunakan persamaan (2.2).
- iii. Melakukan estimasi parameter yang kemudian melakukan pengujian signifikansi parameter secara serentak dan parsial dengan menggunakan persamaan (2.11) untuk uji serentak dan persamaan (2.12) untuk uji parsial.
- iv. Melakukan uji kesesuaian model (goodness of fit) dengan menggunakan persamaan (2.13).
- v. Menghitung keakurasian klasifikasi menggunakan persamaan (2.15).
- vi. Menginterpretasikan analisis pada sumber air bersih rumah tangga di Jawa Timur dengan menggunakan odds ratio yang diperoleh dari persamaan (2.14)

3.5 Diagram Alir

Dalam penelitian ini langkah analisis digambarkan dalam diagram alir sebagai berikut.



Gambar 3.2 Diagram Alir Penelitian

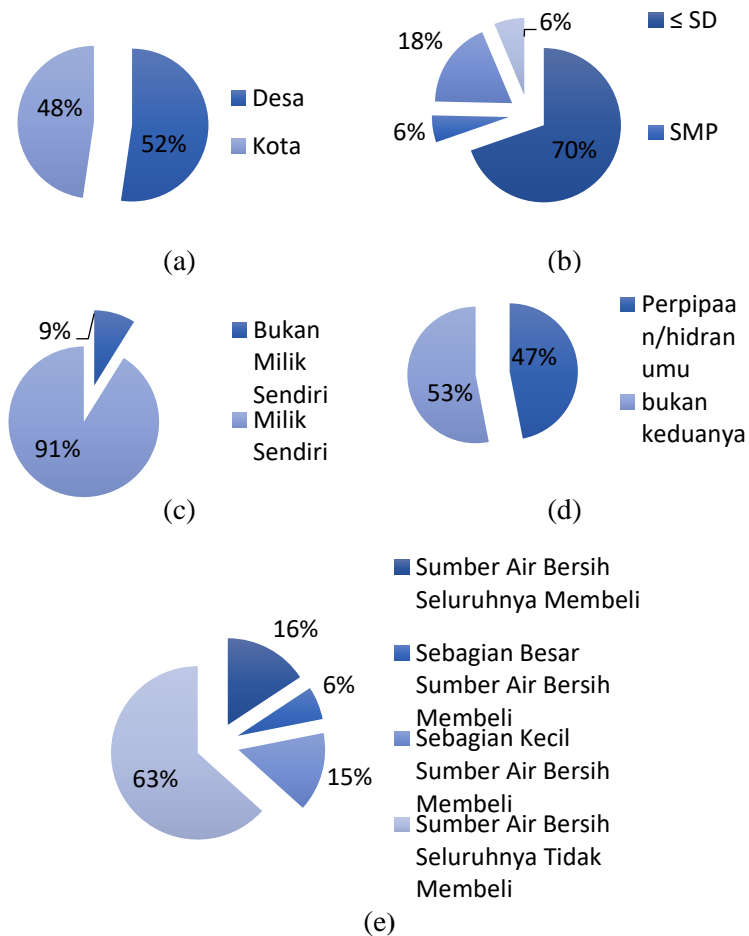


Gambar 3.2 Diagram Alir Penelitian (Lanjutan)

BAB IV ANALISIS DAN PEMBAHASAN

4.1 Analisis Statistik Deskriptif

Pada subbab ini, dilakukan analisis deskriptif terhadap variabel numerik dan variabel kategorik. Statistika deskriptif untuk variabel kategorik disajikan pada Gambar 4.1.



Gambar 4.1 Statistik Deskriptif Variabel Kategorik

Berdasarkan Gambar 4.1,(a) merupakan variabel klasifikasi desa/kota dengan sebanyak 52% rumah tangga berada pada kategori desa, sedangkan pada kategori kota sebanyak 48% rumah tangga berada pada kategori kota. Gambar (b) merupakan variabel pendidikan terakhir kepala rumah tangga dengan sebanyak 70% kepala rumah tangga pendidikan terakhirnya adalah sekolah dasar (SD), kepala rumah tangga dengan pendidikan terakhir SMP sebanyak 6%, kepala rumah tangga dengan pendidikan terakhir SMA sebanyak 18%, dan pendidikan terakhir kepala rumah tangga yaitu perguruan tinggi sebesar 6%. Pada gambar (c) yang merupakan variabel status kepemilikan bangunan dengan sebanyak 91% rumah tangga merupakan bangunan yang berstatus milik sendiri, sedangkan sisanya sebanyak 9% bangunan rumah tangga memiliki status bukan milik sendiri. Lalu gambar (d) merupakan variabel penyaluran air yang digunakan dengan sebanyak 47% rumah tangga menggunakan perpipaan/hidran umum dalam menyalurkan sumber air bersih menuju rumah masing-masing dan sebanyak 53% rumah tangga tidak menggunakan perpipaan/hidran umum atau tidak menggunakan penyaluran air dalam menyalurkan sumber air bersih menuju rumah masing-masing. Gambar (e) merupakan variabel respon dimana kategori sumber air bersih seluruhnya membeli sebesar 16%, kategori sebagian besar sumber air bersih membeli sebesar 6%, kategori sebagian kecil sumber air bersih membeli sebesar 15%, dan kategori sumber air bersih seluruhnya tidak membeli sebesar 63%.

Tabel 4.1 Statistik Deskriptif Variabel Numerik

Variabel	Mean	Minimum	Maximum
X_5	3,4558	1	14

Tabel 4.1, menunjukkan bahwa variabel X_5 atau jumlah anggota keluarga di Jawa Timur rata-rata memiliki anggota keluarga sebanyak 3,4558 atau sebanyak 4 anggota keluarga dengan minimum 1 anggota keluarga dalam keluarga dan maximum 14 anggota keluarga dalam keluarga.

4.2 Pemodelan Regresi Logistik Multinomial

Pada subbab ini dijelaskan mengenai analisis yang dilakukan untuk menentukan model regresi logistik multinomial pada analisis sumber air bersih rumah tangga di Jawa Timur. Sebelum melakukan

analisis pemodelan, terlebih dahulu dilakukan uji independensi variabel yang disajikan pada Tabel 4.2 dan Tabel 4.3.

Tabel 4.2 Chi-Square Test Variabel Kategorik

	Y				Total	Chi-Square Test		
	1	2	3	4				
X ₁	1	3501	1472	2868	7583	15424	0,000	
	2	1129	349	1490	11085			14053
X ₂	1	2474	937	2487	14645	20543	0,000	
	2	199	117	229	121			1666
	3	1237	555	1168	2438			5398
	4	720	212	474	464			1870
X ₃	1	710	387	587	935	2619	0,000	
	2	3920	1434	3771	17733			26858
X ₄	1	4438	940	1570	6876	13824	0,000	
	2	192	881	2788	11792			15653

Hasil analisis hubungan pada variabel kategorik terhadap variabel respon (Y).

Klasifikasi desa/kota (X₁) – Y,

H₀ : Tidak ada hubungan antar variabel

H₁ : Ada hubungan antar variabel

Berdasarkan Tabel 4.2 nilai *chi-square* test bernilai 0,000, dimana nilai tersebut lebih kecil dari alfa 5%. Maka, keputusannya adalah tolak H₀. Artinya terdapat hubungan antara variabel X₁ dengan Y.

Hal yang sama dengan variabel pendidikan terakhir kepala rumah tangga (X₂), Status kepemilikan bangunan (X₃), dan Penyaluran air (X₄) yang mana nilai *chi-square* test bernilai 0,000. Nilai tersebut lebih kecil dari alfa 5%. Maka, keputusan tolak H₀. Maka, variabel X₂, X₃, dan X₄ terdapat hubungan dengan Y.

Tabel 4.3 Korelasi Variabel Numerik

	X5	Y
X ₅	1	-0,019 (0,001)
Y	-0,019 (0,001)	1

Hasil analisis hubungan pada variabel kategorik terhadap variabel respon (Y).

Jumlah anggota keluarga (X_5) – Y,

H_0 : Tidak ada hubungan antar variabel

H_1 : Ada hubungan antar variabel

Berdasarkan Tabel 4.3 nilai signifikansi dari korelasi bernilai 0,001, dimana nilai tersebut lebih kecil dari alfa 5%. Maka, keputusannya adalah tolak H_0 . Artinya terdapat hubungan antara variabel X_5 dengan Y.

4.2.1 Uji Kesuaian Model

Sebelum melakukan analisis regresi logistik multinomial lebih lanjut, diuji dahulu apakah model yang terbentuk sudah sesuai dengan data (fit).

Tabel 4.4 Uji Kesesuaian Model

	Chi-Square	df	Sig.
Pearson	1150,301	765	0,000
Deviance	1132,358	765	0,000

Berdasarkan hasil pada Tabel 4.4 didapat nilai p-value (sig.) di baris hasil uji kelayakan metode Pearson yaitu sebesar 0,000. Nilai tersebut menyatakan bahwa uji kelayakan metode Pearson kurang dari alfa (5% = 0,05), maka tolak H_0 yang menjelaskan bahwa model tidak fit atau model tidak sesuai. Artinya model tidak dapat menjelaskan data.

Tabel 4.5 Nilai Pseudo R-Square

Cox and Snell	0,309
Nagelkerke	0,354
McFadden	0,179

Selain uji fit model, maka perlu memperhitungkan besarnya ukuran kebaikan model pada variabel bebas kualitatif. Berdasarkan Tabel 4.5 nilai tersebut diambil dari nilai Nagelkerke yaitu sebesar 0,354. Nilai tersebut mengindikasikan bahwa keragaman data variabel bebas mampu menjelaskan keragaman data variabel respon sebesar 35,4% sedangkan sisanya dijelaskan oleh variabel bebas lain yang ada di luar model. Nilai 0,354 tersebut sudah dianggap baik dalam menjelaskan nilai ukuran kebaikan model.

4.2.2 Pengujian Signifikansi Parameter

a. Uji Serentak

Pengujian ini dilakukan untuk memeriksa koefisien β secara serentak atau bersamaan terhadap variabel respon.

Tabel 4.6 Uji Serentak

Model	Model Fitting Criteria	Likelihood Ratio Test		
	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	13795,265			
Final	2891,955	10903,310	21	0,000

Berdasarkan Tabel 4.6, hasil uji serentak menunjukkan nilai pada baris final dengan nilai chi-square sebesar 10903,310 dengan nilai chi-square tabel sebesar 11,070 yang artinya minimal terdapat satu variabel bebas yang secara statistik signifikan, selain itu nilai p-value (sig.) sebesar 0,000. Dimana nilai tersebut lebih kecil dari alfa 5% atau 0,05, maka tolak H_0 yang artinya minimal terdapat satu variabel bebas yang secara statistik signifikan memengaruhi variabel respon. Karena, pada uji serentak menyatakan signifikan, maka dilanjutkan ke uji parsial.

b. Uji Parsial

Pengujian ini dilakukan untuk memeriksa kemaknaan koefisien β secara parsial. Dimana sebelumnya telah terlihat hasil uji serentak yang menjelaskan bahwa seluruh variabel prediktor berpengaruh terhadap variabel respon. Pada uji parsial akan dilihat pengaruh tiap variabel prediktor terhadap variabel respon.

Tabel 4.7 Uji Parsial

Effect	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.
X ₁	3113,854	3060,166	3	0,000
X ₂	2640,951	2545,250	9	0,000
X ₃	1013,137	960,878	3	0,000
X ₄	6416,345	6363,723	3	0,000
X ₅	298,701	26,846	3	0,000

Berdasarkan Tabel 4.7, uji parsial menunjukkan nilai chi-square pada kelima variabel bebas lebih besar dari nilai chi-square tabel yaitu 11,070 yang artinya kelima variabel bebas yang digunakan secara statistik memiliki pengaruh yang signifikan terhadap variabel respon, selain itu pada kolom signifikan untuk semua variabel berniali sebesar 0,000. Nilai-nilai tersebut lebih kecil dari alfa 5% atau lebih kecil dari 0,05, maka tolak H_0 yang artinya variabel bebas secara statistik signifikan memengaruhi variabel respon. Dengan demikian, kelima variabel bebas yang digunakan secara statistik memiliki pengaruh yang signifikan terhadap variabel respon.

4.2.3 Interpretasi Analisis Regresi Logistik Multinomial

Parameter pembentuk fungsi logit disajikan pada Tabel 4.8 yang didapat dari nilai Beta, sedangkan nilai odds ratio disajikan pada kolom $\exp(\text{Beta})$ pada Tabel 4.9.

Tabel 4.8 Parameter Pembentuk Fungsi Logit

Ya	Beta		
	1	2	3
Konstanta	-3,605	-2,633	-0,811
X_5	0,004	0,021	0,039
$[X_1=1]$	1,328	1,550	0,798
$[X_1=2]$			
$[X_2=1]$	-1,637	-1,423	-1,506
$[X_2=2]$	-1,731	-1,003	-1,335
$[X_2=3]$	-0,920	-0,564	-0,694
$[X_2=4]$			
$[X_3=1]$	0,724	1,054	0,665
$[X_3=2]$			
$[X_4=1]$	3,703	0,632	-0,026
$[X_4=2]$			

Berdasarkan Tabel 4.8 menghasilkan fungsi logit, sehingga didapatkan pula model regresi logistik.

$$\pi_1(x) = \frac{\exp P_1(x)}{1 + \exp P_1(x) + \exp P_2(x) + \exp P_3(x)}$$

$$\pi_2(x) = \frac{\exp P_2(x)}{1 + \exp P_1(x) + \exp P_2(x) + \exp P_3(x)}$$

$$\pi_3(x) = \frac{\exp P_3(x)}{1 + \exp P_1(x) + \exp P_2(x) + \exp P_3(x)}$$

Dengan nilai P_1 , P_2 , dan P_3 sebagai berikut.

$$\begin{aligned} P_1(x) &= \ln \left(\frac{P(Y=1)1 | x}{P(Y=1)0 | x} \right) \\ &= -3,605 + 1,328 X_{1.1} - 1,637 X_{2.1} - 1,731 X_{2.2} + \\ &\quad (-0,920 X_{2.3}) + 0,724 X_{3.1} + 3,703 X_{4.1} + \\ &\quad 0,004 X_5 \end{aligned}$$

$$\begin{aligned} P_2(x) &= \ln \left(\frac{P(Y=1)2 | x}{P(Y=1)0 | x} \right) \\ &= -2,633 + 1,550 X_{1.1} - 1,423 X_{2.1} - 1,003 X_{2.2} + \\ &\quad (-0,564 X_{2.3}) + 1,054 X_{3.1} + 0,632 X_{4.1} + \\ &\quad 0,021 X_5 \end{aligned}$$

$$\begin{aligned} P_3(x) &= \ln \left(\frac{P(Y=1)3 | x}{P(Y=1)0 | x} \right) \\ &= -0,811 + 0,798 X_{1.1} - 1,506 X_{2.1} - 1,335 X_{2.2} + \\ &\quad (-0,694 X_{2.3}) + 0,665 X_{3.1} - 0,026 X_{4.1} + \\ &\quad 0,039 X_5 \end{aligned}$$

Odds ratio merupakan nilai kecenderungan antara satu kategori dengan kategori lain pada variabel prediktor.

Tabel 4.9 Interpretasi Analisis Regresi Logistik Multinomial

Variabel	Y					
	Sumber Bersih Seluruhnya Membeli	Air - Sumber Bersih Seluruhnya Tidak Membeli	Sebagian Besar Sumber Bersih Seluruhnya Tidak Membeli	Air - Sebagian Besar Sumber Bersih Seluruhnya Tidak Membeli	Sebagian Kecil Sumber Bersih Seluruhnya Tidak Membeli	Air - Sebagian Kecil Sumber Bersih Seluruhnya Tidak Membeli
	Exp(B)		Exp(B)		Exp(B)	
X ₅	1,004		1,021		1,039	
X _{1,1}	3,772		4,701		2,221	
X _{1,2}						
X _{2,1}	0,195		0,241		0,222	
X _{2,2}	0,177		0,367		0,263	
X _{2,3}	0,398		0,569		0,499	
X _{2,4}						
X _{3,1}	2,063		2,869		1,944	
X _{3,2}						
X _{4,1}	40,555		1,882		0,975	
X _{4,2}						

Berdasarkan Tabel 4.9, pada variabel X₁ atau klasifikasi desa/kota dengan kategori kota sebagai kontrol, menunjukkan bahwa rumah tangga di daerah pedesaan memiliki kecenderungan resiko sumber air bersih seluruhnya membeli 3,772 lebih besar dibandingkan sumber air bersih seluruhnya tidak membeli, lalu untuk sebagian besar sumber air bersih membeli 4,701 lebih besar dibanding tidak membeli dan sebagian kecil sumber air bersih membeli 2,221 lebih besar dibandingkan sumber air bersih seluruhnya tidak membeli.

Pada variabel X₂ atau pendidikan terakhir kepala rumah tangga dengan kategori perguruan tinggi sebagai kontrol menunjukkan bahwa rumah tangga dengan pendidikan terakhir kepala rumah tangga yaitu \leq SD memiliki kecenderungan resiko sumber air bersih seluruhnya tidak membeli 1/0,195 lebih kecil dibanding sumber air bersih seluruhnya membeli, lalu

kecenderungan resiko sumber air bersih seluruhnya tidak membeli 1/0,241 lebih kecil dibandingkan sebagian besar sumber air bersih membeli dan kecenderungan resiko sumber air bersih seluruhnya tidak membeli 1/0,222 lebih kecil dibandingkan sebagian kecil sumber air bersih membeli. Rumah tangga dengan pendidikan terakhir kepala rumah tangga yaitu SMP memiliki kecenderungan resiko sumber air bersih seluruhnya tidak membeli 1/0,177 lebih kecil dibanding sumber air bersih seluruhnya membeli, lalu kecenderungan resiko sumber air bersih seluruhnya tidak membeli 1/0,367 lebih kecil dibandingkan sebagian besar sumber air bersih membeli dan kecenderungan resiko sumber air bersih seluruhnya tidak membeli 1/0,263 lebih kecil dibandingkan sebagian kecil sumber air bersih membeli. Rumah tangga dengan pendidikan terakhir kepala rumah tangga yaitu SMA memiliki kecenderungan resiko sumber air bersih seluruhnya tidak membeli 1/0,398 lebih kecil dibanding sumber air bersih seluruhnya membeli, lalu kecenderungan resiko sumber air bersih seluruhnya tidak membeli 1/0,569 lebih kecil dibandingkan sebagian besar sumber air bersih membeli dan kecenderungan resiko sumber air bersih seluruhnya tidak membeli 1/0,499 lebih kecil dibandingkan sebagian kecil sumber air bersih membeli.

Pada variabel X_3 atau status kepemilikan bangunan dengan kategori bangunan milik sendiri sebagai kontrol menunjukkan bahwa rumah tangga dengan status bangunan bukan milik sendiri memiliki kecenderungan resiko sumber air bersih seluruhnya membeli 2,063 lebih besar dibandingkan sumber air bersih seluruhnya tidak membeli, lalu kecenderungan resiko sebagian besar sumber air bersih membeli 2,869 lebih besar dibandingkan sumber air bersih seluruhnya tidak membeli dan kecenderungan resiko sebagian kecil sumber air bersih membeli 1,944 lebih besar dibandingkan sumber air bersih seluruhnya tidak membeli.

Pada variabel X_4 atau penyaluran air yang digunakan dengan kategori tidak menggunakan penyaluran air sebagai kontrol menunjukkan bahwa rumah tangga dengan penyaluran air menggunakan perpipaan/hidran memiliki kecenderungan resiko sumber air bersih seluruhnya membeli 40,555 lebih besar dibandingkan sumber air bersih seluruhnya tidak membeli dan

kecenderungan resiko sebagian besar sumber air bersih membeli 1,882 lebih besar dibandingkan sumber air bersih seluruhnya membeli.

Pada variabel X_5 atau jumlah anggota keluarga dengan jumlah sedikit atau banyak anggota keluarga memiliki kecenderungan resiko sumber air bersih seluruhnya membeli 1,039 lebih besar dibandingkan sebagian kecil sumber air bersih membeli.

4.2.4 Ketepatan Klasifikasi

Ketepatan klasifikasi untuk mengetahui atau mengevaluasi model yang telah terbentuk.

Tabel 4.10 Keakurasian Klasifikasi

Observed	Predicted				Percent Correct
	1	2	3	4	
1	1975	0	32	2623	42,7%
2	459	0	139	1223	0,0%
3	544	0	337	3477	7,7%
4	841	0	308	17519	93,8%
Overall Percentage	13,0%	0,0%	2,8%	84,3%	67,3%

Berdasarkan Tabel 4.10, ketepatan klasifikasi pada tiap kategori yaitu kategori 1 sebanyak 1975 observasi, kategori 2 sebanyak 0 observasi, katetgori 3 sebanyak 337 observasi, dan kategori 4 sebanyak 17519 observasi. Akurasi model yang didapatkan adalah sebesar 67,3%.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan analisis yang telah dilakukan, diperoleh kesimpulan sebagai berikut.

1. Proporsi rumah tangga dari variabel bebas yang dilihat dari Pie Chart, dimana pada variabel klasifikasi desa/kota sebanyak 52% terletak didaerah pedesaan dan 48% terletak didaerah perkotaan. Variabel pendidikan terakhir kepala rumah tangga sebanyak 70% yaitu SD, 6% yaitu SMP, 18% yaitu SMA, dan 6% yaitu perguruan tinggi. Variabel status kepemilikan bangunan, sebanyak 91% bangunan rumah tangga berstatus milik sendiri dan 9% berstatus bukan milik sendiri. Variabel penyaluran air yang digunakan, sebanyak 47% rumah tangga menggunakan perpipaan/hidran umum dan sebanyak 53% rumah tangga tidak menggunakan penyaluran air dengan perpipaan atau hidran umum. Variabel jumlah anggota keluarga di Jawa Timur rata-rata memiliki anggota keluarga sebanyak 3,4558 atau sebanyak 4 anggota keluarga dengan minimum 1 anggota keluarga dalam keluarga dan maximum 14 anggota keluarga dalam keluarga. Variabel respon yang terbagi menjadi empat kategori, sebanyak 16% rumah tangga sumber air bersih seluruhnya membeli, sebanyak 6% rumah tangga sebagian besar sumber air bersih membeli, sebanyak 15% rumah tangga sebagian kecil sumber air bersih membeli, dan sebanyak 63% rumah tangga sumber air bersih seluruhnya tidak membeli.
2. Model yang terbentuk untuk analisis regresi logistik multinomial pada sumber air bersih rumah tangga di Jawa Timur adalah sebagai berikut.

$$\begin{aligned}
 P_1(x) &= \ln \left(\frac{P(Y=1)1 | x}{P(Y=1)0 | x} \right) \\
 &= -3,605 + 1,328 \text{desa} - 1,637 \text{SD} - 1,731 \text{SMP} + \\
 &\quad (-0,920 \text{SMA}) + 0,724 \text{bukan milik sendiri} + \\
 &\quad 3,703 \text{ perpipaan / hidran umum} + \\
 &\quad 0,004 \text{ jumlah anggota keluarga}
 \end{aligned}$$

$$\begin{aligned}
 P_2(x) &= \ln \left(\frac{P(Y=1)2 | x}{P(Y=1)0 | x} \right) \\
 &= -2,633 + 1,550 \text{desa} - 1,423 \text{SD} - 1,003 \text{SMP} + \\
 &\quad (-0,564 \text{SMA}) + 1,054 \text{bukan milik sendiri} + \\
 &\quad 0,632 \text{ perpipaan / hidran umum} + \\
 &\quad 0,021 \text{ jumlah anggota keluarga}
 \end{aligned}$$

$$\begin{aligned}
 P_3(x) &= \ln \left(\frac{P(Y=1)3 | x}{P(Y=1)0 | x} \right) \\
 &= -0,811 + 0,798 \text{desa} - 1,506 \text{SD} - 1,335 \text{SMP} + \\
 &\quad (-0,694 \text{SMA}) + 0,665 \text{bukan milik sendiri} + \\
 &\quad (-0,026 \text{ perpipaan / hidran umum}) + \\
 &\quad 0,039 \text{ jumlah anggota keluarga}
 \end{aligned}$$

5.2 Saran

Berdasarkan analisis dan pembahasan serta kesimpulan yang diperoleh dapat dirumuskan saran, bahwa dari pembahasan diketahui bahwa rumah tangga lebih cenderung untuk membeli sumber air bersih, maka dari pemerintah dapat mengalokasikan dan mendistribusikan air bersih yang cukup untuk masyarakat diluar air bersih yang dapat diperoleh dengan tidak membeli.

Lalu pertimbangan untuk penelitian selanjutnya adalah pemodelan dapat menggunakan metode yang sesuai serta dapat mengatasi imbalance data sehingga pada saat pengolahan data menghasilkan hasil akhir yang lebih baik.

DAFTAR PUSTAKA

- Afidah, L. N. (2011). *Pola Tingkat Keparahan Korban Kecelakaan Lalu Lintas Dengan Menggunakan Regresi Logistik Multinomial (Studi Kasus Kecelakaan Lalu Lintas Di Surabaya)*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Agresti, A. (1990). *Categorical Data Analysis*. New York: John Wiley and Sons, Inc.
- Agresti, A. (2002). *Categorical Data Analysis*. New York: John Wiley and Sons, Inc.
- Agresti, A. (2007). *An Introduction to Categorical Data Analysis* (2 ed.). New York: John Wiley and Sons, Inc.
- Bappenas. (2004). *Laporan Perkembangan Pencapaian Tujuan Pembangunan Milenium (Millenium Development Goals)*. Jakarta: Bappenas.
- Batista, G. E., Bazzan, A. L., & Monard, M. C. (2003). Balancing Training Data for Automated Annotation of Keywords : a Case Study.
- Bekkar, M., Djeemaa, H., & Alitouch, T. (2013). Evaluation Measures for Models Assesment over Imbalanced Data Sets. *Journal of Information Engineering and Application*, 3, 27-38.
- BPS, B. P. (2016). *Survei Sosial Ekonomi Nasional*.
- Draper, N. R., & Smith, H. (1992). *Analisis Regresi Terapan*. Jakarta: PT Gramedia Pustaka Utama.
- ESCAP-UNDP. (2002). *ESCAP-UNDP Initiative for The Achievement of Millenium Development Goals in Asia and The Pacific*. Bangkok: UN ESCAP.
- Giddings, B. (2002). Environment, Economy, and Society: Fitting Them together Into Sustainable Development. *Sust. Dev. Sustainable Cities Research Institute*.
- Hall, & James, A. (2001). *Sistem Informasi Akutansi* (3 ed.). Jakarta: Salemba Empat.
- Hosmer, D. W., & Lemeshow, S. (2000). *Applied Logistic Regression*. New York: John Wiley and Sons, Inc.
- Moegijantoro. (1996). *Kebutuhan Air*. Surabaya: PT Empat Sekawan.

- Mustikowati, W. (2014). *Analisis Faktor-Faktor yang Mempengaruhi Konsumsi Air Bersih Golongan Pelanggan Rumah Tangga III Wilayah Pelayanan Cabang Timur PDAM Kota Semarang*. Semarang: Universitas Diponegoro.
- Mutmainah, H. (2011). *Faktor-Faktor Yang Mempengaruhi Permintaan Air Bersih dari Masyarakat Terhadap Perusahaan Daerah Air Minum Kabupaten Karanganyar*.
- Parlinggoman, C. (2016). *Pemodelan Faktor-Faktor Pemilihan Fasilitas Kesehatan Untuk Berobat Di Sulawesi Tenggara Tahun 2012 Menggunakan Regresi Logistik Multinomial*. Surabaya: Institut Teknologi Sepuluh Nopember.
- PERPAMSI. (1994). *Diklat Tenaga Teknik Penyediaan Air Minum*. Bandung: PERPAMSI dan ITB.
- Solberg, A., & Solberg, R. (1996). A Large-Scale Evaluation of Features for Automatic Detection of Oil Spills in ERS SAR Images. *In International Geoscience and Remote Sensing Symposium*, pp. 1484-1486.
- Walpole, R. E. (1995). *Pengantar Statistika Edisi ke-3*. Jakarta: Gramedia Pustaka Utama.
- Widarjono, A. (2010). *Analisis Statistika Multivariat Terapan (1 ed.)*. Yogyakarta: UPP STIM YKPN.
- Winarna, S. (2003). *Analisis Konsumsi Air Bersih Pelanggan Rumah Tangga Berdasarkan Faktor-Faktor Yang Mempengaruhinya (Studi Kasus pada PDAM Kabupaten Karanganyar)*. Semarang: Universitas Diponegoro.
- Yu, X., Zhou, M., Chen, X., Deng, L., & Wang, L. (2017). Using Class Imbalance Learning for Cross-Company Defect Prediction.

LAMPIRAN

Lampiran 1 Surat Keterangan Pengambilan Data



SURAT KETERANGAN

Yang bertanda tangan dibawah ini :

N a m a : Thomas Wunang Tjahjo, M.Sc, M.Eng.
 N I P : 19700329 1992 11 1 001
 Jabatan : Kepala Bidang Integrasi Pengolahan dan Diseminasi Statistik

Dengan ini menerangkan bahwa :

N a m a : Zabrinna Zenitha Zahroh
 Fakultas/Program Studi : Fakultas Matematika, Komputasi Dan Sains Data / Statistika
 N.R.P : 06211440000120
 Alamat Rumah : Perum Bumi Mas Blok E-1, Tunggalursari, Tulungagung
 Akademi / Universitas : Institut Teknologi Sepuluh Nopember (ITS)
 Telp (031) 594 3352, (031) 599 4251-55
 Fax (031) 592 2940

Di berikan kesempatan menggunakan data Badan Pusat Statistik (BPS) Provinsi Jawa Timur, dengan syarat menyebut judul publikasi dan sumbernya serta tidak untuk tujuan komersil. Data ini digunakan dalam rangka menyusun Tugas Akhir / Skripsi / Thesis / Disertasi dengan judul :

" Analisis Pola Konsumsi Rumah Tangga di Jawa Timur dalam Memenuhi Kebutuhan Air Bersih Menggunakan Regresi Logistik Multinomial "

Demikian surat keterangan ini dibuat dan agar dipergunakan sebagaimana mestinya

Surabaya, 24 Mei 2018

An. Kepala BPS Provinsi Jawa Timur
 Kepala Bidang IPDS

Thomas Wunang Tjahjo, M.Sc, M.Eng.

Lampiran 2 Data Penelitian

X1	X2	X3	X4	X5	X6	X7	Y
2	3	2	1	2	4	4	4
2	1	2	3	5	11	15	4
2	1	2	1	2	7	8	4
2	1	2	1	2	14	8	4
2	3	2	1	4	6	7	4
2	1	2	1	4	34	15	4
2	1	2	1	2	3	4	4
2	1	2	1	4	15	15	4
2	2	2	1	3	6	15	4
2	2	2	1	2	4	4	4
2	1	2	1	5	8	13	4
2	1	2	1	5	8	10	4
2	1	2	1	3	11	11	3
2	1	2	1	2	5	7	4
2	1	2	1	2	5	6	4
2	1	2	1	3	12	16	4
2	1	2	1	4	6	11	4
2	4	2	1	5	18	18	4
2	1	2	1	3	7	9	4
2	1	2	3	1	1	2	4
2	1	2	1	3	5	7	4
2	4	2	1	3	15	16	4
2	1	2	1	3	11	11	4
2	1	2	1	3	7	8	4
2	3	2	1	4	10	9	4
2	1	2	1	4	5	7	4
2	1	2	1	4	8	17	4
2	1	2	3	4	7	9	4

2	1	2	1	4	7	12	4
2	4	2	1	3	17	11	4
2	1	2	1	2	4	4	4
2	1	2	1	5	6	13	4
2	1	2	1	5	4	6	4
2	1	2	1	1	1	3	4
2	1	2	1	2	3	5	4
2	1	2	1	5	7	9	4
2	1	2	1	2	4	4	4
2	1	2	1	3	6	9	4
2	1	2	1	5	22	19	4
2	3	2	1	4	6	7	4
2	1	2	3	1	11	7	4
2	1	2	3	2	17	14	4
2	1	2	3	5	15	16	4
2	1	2	3	2	13	12	3
2	1	2	3	5	9	14	4
2	1	2	3	4	5	6	4
2	3	2	3	3	48	22	3
2	3	2	3	4	25	12	4
2	1	1	3	3	104	14	3

(Halaman ini sengaja dikosongkan)



BIODATA PENULIS

Penulis dengan nama lengkap Zabrinna Zenitha Zahroh dilahirkan di Tulungagung pada 15 Juni 1996 yang merupakan anak kedua dari pasangan (Alm) Bapak Gendrojo dan Ibu Nur Farida. Penulis menempuh pendidikan formal di Taman Kanak-kanak (TK) Batik Tulungagung (2000-2002), SDN Kampung Dalem 1 (2002-2008), SMPN 1 Tulungagung (2008-2011), dan SMAN 1 Boyolangu (2011-2014). Kemudian penulis diterima sebagai Mahasiswa Departemen Statistika ITS melalui jalur Mandiri-Kemitraan pada tahun 2014. Selama masa perkuliahan, penulis turut berpartisipasi dalam kepanitiaan seperti Pekan Raya Statistika (PRS) serikat kepanitiaan yang diselenggarakan oleh Badan Eksekutif Mahasiswa (BEM) ITS seperti Gerigi dan lainnya. Selain itu, penulis juga aktif dalam organisasi yang menaungi Departemen Statistika yaitu HIMASTA-ITS sebagai staff litbang HIMASTA-ITS 2015/2016 dan staff PSDM BEM-ITS 2015/2016. Selama menjalani masa perkuliahan. Sebagai salah satu syarat untuk memperoleh gelar Sarjana, penulis mengambil bidang Sosial Kependudukan pada peniltian Tugas Akhir dengan judul “Pola Konsumsi Rumah Tangga Di Jawa Timur Dalam Memenuhi Kebutuhan Air Bersih Menggunakan Regresi Logistik Multinomial”. Penulis menerima segala kritik, saran, masukan serta diskusi lebih lanjut mengenai Tugas Akhir ini, dapat menghubungi penulis melalui email zabrinna34@gmail.com atau nomor telepon 082257005375.