



TUGAS AKHIR - KS184822

**MODEL REGRESI *ROBUST* UNTUK ANALISIS
FAKTOR-FAKTOR YANG MEMPENGARUHI
PRESTASI BELAJAR MAHASISWA
JALUR UJIAN TULIS**

**KHOIROTUS NAINIYAH
NRP 062117 4500 0025**

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

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2019**



TUGAS AKHIR - KS184822

**MODEL REGRESI *ROBUST* UNTUK ANALISIS
FAKTOR-FAKTOR YANG MEMPENGARUHI
PRESTASI BELAJAR MAHASISWA
JALUR UJIAN TULIS**

**KHOIROTUS NAINIYAH
NRP 062117 4500 0025**

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

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2019**



FINAL PROJECT - KS184822

**ROBUST REGRESSION MODEL FOR ANALYSIS OF
THE FACTORS THAT AFFECTING STUDENT
LEARNING ACHIEVEMENTS
ON WRITING TEST**

**KHOIROTUS NAINIYAH
SN 062117 4500 0025**

**Supervisors
Dr. Suhartono
Dr. Dra. Ismaini Zain, M.Si.**

**UNDERGRADUATE PROGRAMME
DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS, COMPUTING, AND DATA SCIENCE
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2019**

LEMBAR PENGESAHAN

MODEL REGRESI *ROBUST* UNTUK ANALISIS FAKTOR-FAKTOR YANG MEMPENGARUHI PRESTASI BELAJAR MAHASISWA JALUR UJIAN TULIS

TUGAS AKHIR

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

Oleh :

Khoirotus Nainiyah
NRP. 062117 4500 0025

Disetujui oleh Pembimbing:

Dr. Suhartono

NIP. 19710929 199512 1 001

Dr. Dra. Ismaini Zain, M.Si.

NIP. 19600525 198803 2 001



Mengetahui,
Kepala Departemen Statistika

Dr. Suhartono
NIP. 19710929 199512 1 001

SURABAYA, JULI 2019

MODEL REGRESI *ROBUST* UNTUK ANALISIS FAKTOR-FAKTOR YANG MEMPENGARUHI PRESTASI BELAJAR MAHASISWA JALUR UJIAN TULIS

Nama Mahasiswa : Khoirotus Nainiyah
NRP : 062117 4500 0025
Departemen : Statistika-FMKSD-ITS
Dosen Pembimbing : Dr. Suhartono
Dr. Dra. Ismaini Zain, M.Si

Abstrak

Seleksi masuk PTN jalur ujian tulis memiliki 10 mata uji pada tahap seleksinya yang diharapkan menjadi prediksi tinggi untuk keberhasilan calon mahasiswa dalam mengikuti perkuliahan yang diukur dengan IPS dan IPK. ITS Surabaya merupakan salah satu PTN yang ikut melaksanakan jalur ujian tulis. Sehingga prodi yang ada di ITS akan digunakan sebagai acuan prodi nasional. Selain nilai mata uji, terdapat beberapa faktor yang dapat berpengaruh terhadap prestasi belajar mahasiswa diantaranya jenis kelamin, jenis sekolah, jurusan sekolah, bidikmisi dan asal daerah. Dari nilai mata uji dan karakteristik mahasiswa berdasarkan prodi diduga terdapat hubungan yang mempengaruhi indeks prestasi. Salah satu metode statistika untuk mengetahui hubungan tersebut adalah analisis regresi. Pada analisis regresi metode OLS memiliki asumsi yang harus terpenuhi untuk memperoleh model yang baik. Namun seringkali pada data analisis terdapat data outlier yang akan berpengaruh terhadap ketidakmampuan memenuhi asumsi, sehingga untuk mengatasi permasalahan outlier akan digunakan metode regresi robust. Terdapat beberapa metode robust yang telah banyak digunakan dan pada penelitian ini digunakan metode regresi robust LTS dan LWS. Penelitian ini menggunakan data sekunder hasil ujian tulis tahun 2014 yang pada kajian awal terdapat outlier dan akan di analisis hubungannya terhadap indeks prestasi. Hasil analisis didapatkan faktor yang berpengaruh terhadap prestasi belajar mahasiswa pada sebagian besar prodi yaitu nilai mata uji numerikal dan jenis kelamin, serta pemilihan model terbaik didapatkan hasil metode robust LTS.

Kata Kunci: Indeks Prestasi, OLS, Prodi ITS, Regresi Robust.

(Halaman ini sengaja dikosongkan)

ROBUST REGRESSION MODEL FOR ANALYSIS OF THE FACTORS THAT AFFECTING STUDENT LEARNING ACHIEVEMENTS ON WRITING TEST

Name : Khoirotus Nainiyah
Student Number : 062117 4500 0025
Department : Statistics-FMKSD-ITS
Supervisors : Dr. Suhartono
Dr. Dra. Ismaini Zain, M.Si

Abstract

The written test at public universities has 10 test subjects at the selection stage which are expected to be high predictions for the success of prospective students in attending lectures as measured by grade point. ITS Surabaya is one of the public universities that participated in the written test path. So that study programs at ITS will be used as a reference for national study programs. Besides on test subjects, there are several factors that can affect student learning achievement including gender, type of school, school majors, bidikmisi scholarship and regional origin. From the test scores and the characteristics of students based on study programs, it is assumed that there is a relationship that affects the achievement index. One of statistical method to find out the relationship is regression analysis. In OLS method regression analysis has assumptions that must be fulfilled to obtain a good model. However, often in the analysis data there are data outliers that will affect the inability to fulfill the assumptions, so the robust regression will be used to overcome the outlier problem. There are several robust methods that have been widely used and in this study, robust regression method of LTS and LWS are used. This study uses secondary data from the written exam results in 2014 which in the initial study there are outliers and will analyze the relationship to the achievement index. The results of the analysis found that the factors that affect student achievement in most study programs are the numerical value test and gender, and the best selection models obtained is the robust method of LTS.

Keywords: *Achievement Index, ITS Programme, OLS, Robust Regression.*

(This page is intentionally left blank)

KATA PENGANTAR

Alhamdulillahirobbil'Alamin, puji syukur yang tiada tara penulis panjatkan kehadiran Allah SWT yang telah melimpahkan rahmat serta hidayah-Nya sehingga penulis dapat menyelesaikan Tugas Akhir dengan judul "**Model Regresi *Robust* untuk Analisis Faktor-faktor yang Mempengaruhi Prestasi Belajar Mahasiswa Jalur Ujian Tulis**". Tak lupa sholawat serta salam semoga selalu terlimpahkan kepada junjungan kita, Nabi Muhammad SAW dan keluarganya.

Terselesaikannya laporan Tugas Akhir ini tentu tidak lepas dari bantuan serta doa banyak pihak. Oleh karena itu Penulis ucapkan rasa terimakasih tak terhingga kepada:

1. Keluargaku, Alm. Ayah dan Ibu tercinta yang telah memberi semangat untuk melanjutkan kuliah Lintas Jalur, Kakak dan Adik serta seluruh keluarga besar yang senantiasa memberikan doa serta dukungan baik moril maupun materiil.
2. Bapak Dr. Suhartono selaku dosen pembimbing Tugas Akhir sekaligus Kepala Departemen Statistika FMKSD ITS serta Ibu Dr. Dra. Ismaini Zain, M.Si. selaku co. pembimbing Tugas Akhir yang secara sabar telah memberikan bimbingan, arahan dan saran sejak penyusunan proposal hingga laporan Tugas Akhir.
3. Ibu Santi Puteri Rahayu, M.Si., Ph.D. dan Bapak Imam Safawi Ahmad, M.Si. selaku dosen penguji atas ilmu, saran dan masukan serta koreksian pada Tugas Akhir ini.
4. Ibu Dr. Santi Wulan Purnami, M.Si. selaku Kepala Program Studi Sarjana Departemen Statistika ITS dan Bapak Jerry D.T. Purnomo, M.Si., Ph.D. selaku sekretaris Program Studi Sarjana Departemen Statistika ITS.
5. Panitia pengkaji jalur tes tulis SBMPTN yang telah memperkenankan saya menggunakan data untuk Tugas Akhir.
6. Bapak Dr. Setiawan, MS. selaku dosen wali Penulis, serta seluruh dosen dan karyawan Departemen Statistika ITS.

7. Sahabat-sahabatku yang selalu memberi semangat dan doanya untuk kelancaran Tugas Akhir ini dari seminar hingga sidang.
8. Teman-teman seperjuangan PW 120, terutama Lintas Jalur angkatan 2017 yang selalu memberi semangat dan dorongan hingga terselesaikannya laporan ini.
9. Dan seluruh pihak yang tidak dapat disebutkan satu per satu.

Penulis menyadari bahwa laporan ini masih jauh dari sempurna. Oleh karena itu, kritik dan saran sangat diharapkan untuk perbaikan kedepan. Akhir kata semoga laporan Tugas Akhir ini dapat memberikan manfaat bagi berbagai pihak khususnya untuk Almamaterku ITS.

Surabaya, Juli 2019

Penulis

DAFTAR ISI

HALAMAN JUDUL	ii
TITLE PAGE	iii
LEMBAR PENGESAHAN	iv
ABSTRAK	v
ABSTRACT	vii
KATA PENGANTAR	ix
DAFTAR ISI	xi
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xv
DAFTAR LAMPIRAN	xvii
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan Penelitian.....	4
1.4 Manfaat Penelitian.....	4
1.5 Batasan Masalah.....	4
BAB II TINJAUAN PUSTAKA	7
2.1 Statistika Deskriptif.....	7
2.1.1 Koefisien Korelasi.....	7
2.2 Analisis Regresi.....	8
2.2.1 Multikolinearitas.....	8
2.2.2 <i>Outlier</i> dalam Regresi.....	9
2.2.3 <i>Ordinary Least Squares</i>	12
2.3 Regresi <i>Robust</i>	13
2.3.1 <i>Breakdown Point</i>	14
2.3.2 <i>Least Trimmed Squares</i>	14
2.3.3 <i>Least Winsorized Squares</i>	16
2.4 Pengujian Signifikansi Parameter.....	18
2.5 Asumsi IIDN.....	19
2.6 Pemilihan Model Terbaik.....	22
2.7 Prestasi Belajar.....	23
2.7.1 Faktor-faktor yang Mempengaruhi Prestasi Belajar.....	24

2.9 Penelitian Sebelumnya	26
BAB III METODOLOGI PENELITIAN	29
3.1 Sumber Data	29
3.2 Variabel Penelitian	30
3.3 Langkah Analisis	32
3.4 Langkah Analisis Studi Simulasi.....	35
BAB IV ANALISIS DAN PEMBAHASAN	37
4.1 Karakteristik Indeks Prestasi dan Nilai Mata Uji	37
4.1.1 Identifikasi Hubungan Indeks Prestasi dengan Nilai Mata Uji.....	43
4.2 Hasil Perbandingan Estimasi Model pada Studi Simulasi.....	47
4.3 Pemodelan Faktor-faktor yang Mempengaruhi Prestasi Belajar	56
4.3.1 Pemeriksaan Multikolinearitas	57
4.3.2 Deteksi <i>Outlier</i>	58
4.3.3 Analisis Asumsi IIDN	59
4.3.4 Pemodelan Analisis Regresi OLS	62
4.3.5 Pemodelan Analisis Regresi <i>Robust LTS</i>	68
4.3.6 Pemodelan Analisis Regresi <i>Robust LWS</i>	76
4.3.7 Pemilihan Model Terbaik.....	82
4.3.8 Pemodelan Data <i>Testing</i>	84
BAB V KESIMPULAN DAN SARAN	87
5.1 Kesimpulan	87
5.2 Saran	88
DAFTAR PUSTAKA	89
LAMPIRAN	93

DAFTAR TABEL

Tabel 3.1 Program Studi ITS	30
Tabel 3.2 Variabel Penelitian	30
Tabel 3.3 Struktur Data Penelitian.....	31
Tabel 4.1 Statistika Deskriptif Nilai IPS1, IPS2 dan IPK.....	38
Tabel 4.2 Statistika Deskriptif Rata-rata Nilai Mata Uji Seleksi Ujian Tulis	40
Tabel 4.3 Hasil Korelasi <i>Chi-square</i> Data Nominal.....	47
Tabel 4.4 Perbandingan Hasil Studi Simulasi	55
Tabel 4.5 Perhitungan Manual Studi Simulasi dengan Metode LTS.....	56
Tabel 4.6 Perhitungan Manual Studi Simulasi dengan Metode LWS	56
Tabel 4.7 Hasil Nilai VIF	57
Tabel 4.8 Hasil Multikolinearitas	58
Tabel 4.9 Hasil Deteksi <i>Outlier</i>	58
Tabel 4.10 Persentase <i>Outlier</i>	59
Tabel 4.11 Hasil Pengujian <i>Glejser</i>	60
Tabel 4.12 Prodi yang Tidak Heteroskedastisitas.....	60
Tabel 4.13 Hasil Pengujian <i>Kolmogorov-Smirnov</i>	61
Tabel 4.14 Hasil Pengujian Serentak Metode OLS	62
Tabel 4.15 Hasil Pengujian Parsial Metode OLS	63
Tabel 4.16 (a) Ringkasan Analisis OLS Model IPS1	65
Tabel 4.16 (b) Ringkasan Analisis OLS Model IPS2.....	66
Tabel 4.16 (c) Ringkasan Analisis OLS Model IPK	67
Tabel 4.17 Hasil Pengujian Serentak Metode LTS.....	69
Tabel 4.18 Hasil Pengujian Parsial Metode LTS.....	70
Tabel 4.19 (a) Ringkasan Analisis LTS Model IPS1	72
Tabel 4.19 (b) Ringkasan Analisis LTS Model IPS2	73
Tabel 4.19 (c) Ringkasan Analisis LTS Model IPK.....	74
Tabel 4.20 Hasil Pengujian Serentak Metode LWS	76
Tabel 4.21 Hasil Pengujian Parsial Metode LWS	77
Tabel 4.22 (a) Ringkasan Analisis LWS Model IPS1	79
Tabel 4.22 (b) Ringkasan Analisis LWS Model IPS2	80

Tabel 4.22 (c) Ringkasan Analisis LWS Model IPK	81
Tabel 4.23 Hasil Pemilihan Model Terbaik	82
Tabel 4.24 Pemodelan Data <i>Testing</i>	84

DAFTAR GAMBAR

Gambar 2.1	Grafik Ketika Terdapat <i>Outlier</i>	9
Gambar 2.2	Grafik Heteroskedastisitas	20
Gambar 2.3	Grafik Autokorelasi	21
Gambar 2.4	Grafik Distribusi Normal	22
Gambar 3.1	Diagram Alir Pemodelan Faktor-faktor yang Mempengaruhi Prestasi Belajar	33
Gambar 3.2	Diagram Alir Regresi <i>Robust</i>	34
Gambar 4.1	<i>Boxplot</i> Nilai IPS 1, IPS 2 dan IPK	38
Gambar 4.2	<i>Boxplot</i> Nilai Mata Uji Seleksi Ujian Tulis	39
Gambar 4.3	Mahasiswa Berdasarkan (a) Jenis Kelamin dan (b) Bidikmisi	41
Gambar 4.4	Mahasiswa Berdasarkan (a) Jenis Sekolah dan (b) Jurusan Sekolah.....	42
Gambar 4.5	Mahasiswa Berdasarkan (a) Asal Daerah dan (b) Pilihan Prodi	43
Gambar 4.6	<i>Scatterplot</i> IPS1 dan Nilai Mata Uji	44
Gambar 4.7	<i>Scatterplot</i> IPS2 dan Nilai Mata Uji	45
Gambar 4.8	<i>Scatterplot</i> IPK dan Nilai Mata Uji.....	46
Gambar 4.9	Korelasi <i>Pearson</i> Data Rasio	46
Gambar 4.10	Grafik OLS Tanpa <i>Outlier</i>	48
Gambar 4.11	Grafik <i>Robust</i> LTS dan LWS pada Data Tanpa <i>Outlier</i>	49
Gambar 4.12	Grafik OLS dengan <i>Outlier</i>	49
Gambar 4.13	Grafik OLS.....	50
Gambar 4.14	Grafik <i>Robust</i> LTS Dengan <i>Outlier</i>	50
Gambar 4.15	Grafik OLS Tanpa <i>Outlier</i> vs LTS Dengan <i>Outlier</i>	51
Gambar 4.16	Grafik <i>Robust</i> LWS.....	53
Gambar 4.17	Grafik OLS Tanpa <i>Outlier</i> vs LWS Dengan <i>Outlier</i>	54
Gambar 4.18	Grafik OLS, LTS dan LWS pada Data dengan <i>Outlier</i>	54

Gambar 4.19 Perbandingan akar MSE dengan Standar Deviasi.....	85
---	----

DAFTAR LAMPIRAN

Lampiran 1. Data Penelitian	93
Lampiran 2. <i>Syntax</i> R	96
Lampiran 3. Data Simulasi	98
Lampiran 4. <i>Output</i> Analisis Data Simulasi	99
Lampiran 5. <i>Output</i> Deteksi <i>Outlier</i>	101
Lampiran 6. <i>Output</i> Pengujian Asumsi IIDN	112
Lampiran 7. <i>Output</i> Analisis OLS	123
Lampiran 8. <i>Output</i> Analisis LTS	166
Lampiran 9. <i>Output</i> Analisis LWS	200
Lampiran 10. Surat Pernyataan Data	243

(Halaman ini sengaja dikosongkan)

BAB I

PENDAHULUAN

1.1 Latar Belakang

Seleksi masuk Perguruan Tinggi Negeri (PTN) pada program sarjana berdasarkan Peraturan Menteri Riset, Teknologi dan Pendidikan Tinggi Nomor 45 Tahun 2015 memiliki tiga jalur yaitu melalui seleksi nilai rapor, ujian tulis dan seleksi mandiri. Ujian tulis menjadi jalur yang banyak diminati oleh calon mahasiswa karena dapat diikuti oleh siswa lulusan tahun tersebut dan dua tahun sebelumnya (SBMPTN, 2019). Hal itu dapat diketahui dari peserta ujian tulis pada tahun 2018 yang mencapai 860.001 peserta, naik dari 797.738 peserta pada tahun 2017. Sedangkan untuk daya tampung tahun 2018 sebanyak 165.831 kursi, naik 10,7% dari tahun sebelumnya sebesar 148.066 kursi (Kompas, 2018).

Pada jalur ujian tulis terdapat 10 mata uji pada tahap seleksinya. Ke-sepuluh mata uji tersebut terbagi menjadi dua yaitu Tes Kemampuan dan Potensi Akademik (TKPA) dan Tes Kemampuan Dasar (TKD). Pada TKD terdapat kelompok ujian saintek dan kelompok ujian soshum. Nilai mata uji diharapkan menjadi prediksi yang cukup tinggi untuk keberhasilan calon mahasiswa dalam mengikuti perkuliahan nantinya atau dapat disebut prestasi belajar. Prestasi belajar diukur dengan indeks prestasi yang diperoleh dari hasil evaluasi pada tiap semester dan perhitungan kumulatif hingga mahasiswa menyelesaikan studinya. Indeks prestasi tiap semester di sebut Indeks Prestasi Semester (IPS) sedangkan secara kumulatif di disebut Indeks Prestasi Kumulatif (IPK). IPS 1 dan IPS 2 dikenal sebagai indeks prestasi tahap persiapan yang diduga cukup mencerminkan prestasi belajar mahasiswa dalam perkuliahan selanjutnya. Sehingga sebagai pembanding juga akan digunakan IPK mahasiswa untuk mengetahui faktor-faktor apasaja yang dapat mempengaruhi indeks prestasi awal dan akhir masa perkuliahan mahasiswa.

Program studi (prodi) yang ditawarkan pada jalur ujian tulis sangat banyak dan dikelompokkan sesuai kelompok ujian.

Sehingga ingin pula di ketahui bagaimana faktor-faktor yang mempengaruhi prestasi belajar mahasiswa berdasarkan prodi pada PTN di Indonesia. Setiap PTN tentunya memiliki banyak prodi, ada yang sangat diminati, cukup diminati dan kurang diminati. Institut Teknologi Sepuluh Nopember (ITS) Surabaya sebagai salah satu PTN yang ikut melaksanakan jalur ujian tulis juga memiliki banyak prodi. Sebelum pengembangan prodi pada akhir tahun 2017, jumlah prodi yang ada di ITS sebanyak 27 prodi (ITS, 2019). Dalam penelitian ini, sebagai acuan prodi yang digunakan adalah prodi yang ada di ITS. Selain prodi dan nilai mata uji pada jalur ujian tulis, terdapat beberapa faktor yang dapat berpengaruh terhadap prestasi belajar mahasiswa. Menurut Dalyono (2007) terdapat dua faktor utama yang mempengaruhi prestasi belajar yaitu faktor internal mahasiswa meliputi kesehatan, intelegensi, minat, bakat, motivasi dan cara belajar serta faktor eksternal mahasiswa meliputi keluarga, sekolah, masyarakat dan lingkungan sekitar. Pada jalur ujian tulis faktor-faktor tersebut adalah karakteristik mahasiswa meliputi jenis kelamin, jenis sekolah, jurusan sekolah, bidikmisi, asal daerah dan pilihan prodi.

Dari karakteristik mahasiswa, nilai mata uji jalur tulis berdasarkan prodi diduga terdapat hubungan yang mempengaruhi indeks prestasi mahasiswa. Salah satu metode statistika untuk mengetahui hubungan tersebut adalah analisis regresi. Pada analisis regresi sederhana, metode *Ordinary Least Squares* (OLS) sering digunakan untuk menduga koefisien parameternya. Metode ini memiliki asumsi yang harus terpenuhi untuk memperoleh model yang baik (Draper & Smith, 1998). Namun seringkali pada data analisis terdapat data *outlier* yang akan berpengaruh terhadap ketidakmampuan memenuhi asumsi tersebut, sehingga untuk mengatasi permasalahan *outlier* akan digunakan metode regresi *robust*. Metode *robust* cenderung mengabaikan residual yang berhubungan dengan *outlier* yang besar sehingga estimator yang dihasilkan tetap bekerja baik meskipun terdapat data *outlier* (Atkinson & Riani, 2000). Terdapat beberapa metode *robust* diantaranya *M-estimation*, *S-estimation*, *MM-estimation*, *Least*

Mean Squares (LMS), *Least Trimmed Squares* (LTS) dan *Least Winsorized Squares* (LWS) (Rousseeuw & Leroy, 1987). Pada penelitian ini akan digunakan metode *robust* LTS dan LWS karena metode ini memiliki *breakdown point* yang tinggi. *Breakdown point* merupakan suatu ukuran ke-*robust*-an dari suatu estimator terhadap *outlier* (Gad & Qura, 2016).

Penelitian tentang regresi *robust* pernah dilakukan oleh Matdoan (2017) dan Herawati dkk. (2011). Matdoan (2017) meneliti tentang perbandingan regresi kuantil dengan regresi *robust* LTS pada kasus faktor-faktor yang mempengaruhi penyebaran penyakit malaria di Indonesia. Herawati dkk. (2011) meneliti tentang perbandingan analisis ketegaran regresi *robust* terhadap letak pencilan. Sedangkan penelitian tentang indeks prestasi pernah dilakukan oleh Shodiqin dkk. (2018) dan Antari (2016). Shodiqin dkk. (2018) meneliti tentang perbandingan metode regresi *robust* LTS dan estimasi-MM pada data ujian tulis masuk terhadap IPK mahasiswa Universitas PGRI Semarang dan Antari (2016) meneliti tentang pemodelan Indeks Prestasi Persiapan (IPP) mahasiswa ITS dari jalur SNMPTN menggunakan regresi kuantil. Sebagai pembeda dengan penelitian sebelumnya, pada penelitian ini akan digunakan regresi OLS dan regresi *robust* untuk mendapatkan faktor-faktor yang mempengaruhi prestasi belajar. IPS1, IPS2 dan IPK akan dianalisis secara univariat sebagai variabel dependen terhadap karakteristik mahasiswa dan nilai mata uji jalur ujian tulis. Selain itu penelitian ini menggunakan prodi yang ada di ITS sebagai acuan penggunaan kategori prodi secara nasional. Sehingga akan di dapatkan model dari faktor-faktor yang mempengaruhi indeks prestasi belajar mahasiswa.

1.2 Rumusan Masalah

Berdasarkan latar belakang mengenai pengaruh karakteristik mahasiswa serta nilai mata uji jalur ujian tulis terhadap indeks prestasi berdasarkan prodi secara nasional. Dengan kajian awal terdapat data *outlier* sehingga penggunaan regresi OLS menjadi

tidak efisien. Maka permasalahan yang ingin dianalisis pada penelitian ini adalah sebagai berikut.

1. Bagaimana karakteristik indeks prestasi dan nilai mata uji mahasiswa jalur ujian tulis?
2. Bagaimana perbandingan estimasi model dengan metode OLS dan regresi *robust* menggunakan studi simulasi?
3. Bagaimana pemodelan faktor-faktor yang mempengaruhi prestasi belajar mahasiswa jalur ujian tulis dengan metode OLS dan regresi *robust*?

1.3 Tujuan Penelitian

Mengacu pada rumusan masalah, maka tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut.

1. Memperoleh karakteristik indeks prestasi dan nilai mata uji mahasiswa jalur ujian tulis
2. Mendapatkan perbandingan estimasi model dengan metode OLS dan regresi *robust* menggunakan studi simulasi
3. Memperoleh faktor-faktor yang mempengaruhi prestasi belajar mahasiswa jalur ujian tulis dengan metode OLS dan regresi *robust*.

1.4 Manfaat Penelitian

Manfaat yang diharapkan dari hasil penelitian ini yaitu dapat digunakan sebagai bahan masukan/pertimbangan kepada instansi dan lembaga terkait untuk pertimbangan kebijakan sistem penerimaan PTN jalur ujian tulis serta dapat dijadikan sebagai referensi untuk penelitian-penelitian selanjutnya yang berkaitan dengan prestasi belajar mahasiswa.

1.5 Batasan Masalah

Batasan dalam penelitian ini yaitu data yang digunakan adalah data mahasiswa yang diterima pada jalur ujian tulis tahun 2014. Indeks prestasi yang digunakan adalah IPS 1, IPS 2 dan IPK yang akan di analisis secara univariat sebagai variabel dependen terhadap nilai mata uji serta karakteristik mahasiswa sebagai

variabel independen. Analisis berdasarkan prodi dibatasi dengan menggunakan prodi yang ada di ITS sebagai acuan analisis prodi secara nasional dengan kelompok ujian saintek. Selain itu metode regresi *robust* yang digunakan hanya metode *robust* LTS dan *robust* LWS.

(Halaman ini sengaja dikosongkan)

BAB II TINJAUAN PUSTAKA

Pada bab ini akan dibahas mengenai kajian pustaka yang digunakan untuk menyelesaikan permasalahan mengenai faktor-faktor yang mempengaruhi prestasi belajar mahasiswa jalur ujian tulis yaitu statistika deskriptif, korelasi, multikolinearitas, deteksi *outlier*, analisis regresi meliputi metode regresi OLS, *robust* LTS dan *robust* LWS, pengujian signifikansi parameter, asumsi IIDN serta definisi prestasi belajar dan faktor-faktor yang mempengaruhinya.

2.1 Statistika Deskriptif

Statistika deskriptif merupakan penyajian data sedemikian rupa hingga dapat memberikan informasi. Statistika deskriptif memiliki dua pengukuran yaitu ukuran pemusatan dan ukuran penyebaran. Cara penyajian dalam statistika deskriptif dapat menggunakan tabel, diagram, grafik, *scatterplot* maupun *boxplot* (Walpole *et al.*, 2012).

2.1.1 Koefisien Korelasi

Korelasi merupakan salah satu analisis yang digunakan untuk mengetahui hubungan linier antara dua variabel. Terdapat beberapa macam analisis korelasi bergantung pada skala data pada variabel yang digunakan. Jika dua variabel sama-sama berskala rasio maka analisis korelasi yang digunakan yaitu korelasi *pearson* yang akan di tunjukkan pada Persamaan 2.1 dengan kriteria nilai korelasi anatar 0 hingga 1 dimana nilai korelasi yang mendekati 1 menandakan semakin kuatnya hubungan antara kedua variabel sedang nilai korelasi yang semakin mendekati 0 menandakan sebaliknya yaitu semakin lemahnya hubungan kedua variabel (Draper & Smith, 1998).

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (2.1)$$

Sedangkan jika skala data variabel yang digunakan merupakan data kategorik akan digunakan uji *chi-square* dari tabel kontingensinya dengan rumus yang akan ditunjukkan pada Persamaan 2.2 dan hipotesis sebagai berikut (Agresti, 2002).

H_0 : Tidak ada hubungan antara variabel 1 dan 2 ($P_{ij} = P_i \cdot P_j$)

H_1 : Ada hubungan antara variabel 1 dan 2 ($P_{ij} \neq P_i \cdot P_j$)

Statistik uji:

$$\chi_{hitung}^2 = \sum_{i=1}^I \sum_{j=1}^J \frac{(O_{ij} - E_{ij})^2}{E_{ij}}, \quad (2.2)$$

dengan O_{ij} adalah frekuensi pengamatan pada sel (i,j) dan E_{ij} nilai ekspektasi pada sel (i,j) . E_{ij} didapatkan dari perhitungan $E_{ij} = \frac{O_{i.} \cdot O_{.j}}{O_{..}}$.

Apabila nilai $\chi_{hitung}^2 > \chi_{(1-\alpha);(I-1)(J-1)}^2$ maka diberi keputusan tolak H_0 .

2.2 Analisis Regresi

Analisis regresi merupakan salah satu metode statistika untuk menentukan hubungan sebab akibat antara satu variabel dengan variabel lainnya. Dimana variabel terikat disebut variabel dependen yaitu variabel yang dipengaruhi oleh variabel lainnya dan dinotasikan dengan y . Sedangkan variabel bebas disebut variabel independen yaitu variabel yang mempengaruhi dan dinotasikan dengan x (Draper & Smith, 1998). Terdapat beberapa metode dalam analisis regresi diantaranya metode *Ordinary Least Squares* (OLS) dan regresi *robust*. Metode OLS memiliki beberapa asumsi yang harus terpenuhi yaitu kenormalan, identik dan independen. Apabila terdapat *outlier* pada data analisis maka metode ini tidak dapat bekerja dengan baik dan mengganggu pemenuhan asumsi, sehingga penggunaan regresi *robust* dilakukan untuk menangani hal tersebut (Atkinson & Riani, 2000).

2.2.1 Multikolinearitas

Analisis multikolinieritas bertujuan untuk mengetahui ada tidaknya hubungan antara variabel independen atau bebas. Analisis ini dilakukan sebagai syarat digunakannya regresi berganda

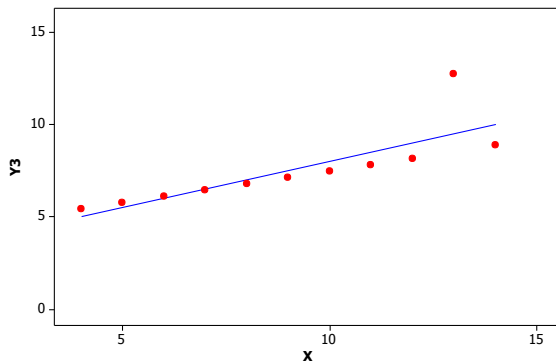
dimana regresi yang baik adalah regresi yang terbebas dari masalah multikolinearitas. Penentuan ada atau tidaknya multikolinearitas salah satunya dengan melihat nilai *Variance Inflation Factor* (VIF). Gejala multikolinearitas terjadi ketika nilai $VIF \geq 10$. Perhitungan nilai VIF menggunakan Persamaan (2.3) (Gujarati & Porter, 2009).

$$VIF_k = \frac{1}{1 - R^2}, \quad (2.3)$$

dengan $k = 1, 2, \dots, p-1$, p adalah banyaknya parameter dan R^2 adalah koefisien regresi yang dihasilkan dari variabel independen x_i dengan variabel independen lain x_j ($x_i \neq x_j$).

2.2.2 *Outlier* dalam Regresi

Outlier atau pencilan adalah suatu pengamatan yang tidak mengikuti pola dan terletak jauh dari pusat data (Atkinson & Riani, 2000). Pada tahun 1973, Anscombe menunjukkan pentingnya melihat sekumpulan data secara grafis sebelum memulai menganalisis menurut jenis hubungan tertentu dan ketidakcukupan sifat statistik dasar untuk menggambarkan kumpulan data yang realistis serta efek *outlier* pada sifat statistik.



Gambar 2.1 Grafik Ketika Terdapat *Outlier*.
(Sumber: Anscombe, 1973)

Gambar 2.1 memperlihatkan bahwa ketika data memiliki *outlier* meskipun distribusinya linier, *outlier* tersebut akan memberi pengaruh untuk menurunkan koefisien korelasi yang dimiliki. Sehingga diperlukan regresi yang *robust* sebagai analisisnya. *Outlier* dapat muncul karena disebabkan oleh banyak hal, diantaranya kesalahan dalam memasukkan data, kesalahan pengukuran, analisis dan lain sebagainya. Data *outlier* akan mengganggu proses analisis regresi karena dapat menyebabkan residual yang besar dari model yang terbentuk dan varians pada data juga menjadi lebih besar sehingga taksiran intervalnya memiliki rentang yang lebar (Soemartini, 2007). Selain itu adanya *outlier* dapat berpengaruh memberikan nilai penduga parameter yang bersifat bias sehingga mengakibatkan interpretasi hasil yang tidak valid.

Model regresi yang diperoleh dengan OLS mensyaratkan asumsi bahwa residual dari model yang dihasilkan harus berdistribusi normal. Tetapi dengan adanya *outlier* menyebabkan asumsi kenormalan tidak terpenuhi. Dalam analisis regresi, terdapat satu variabel dependen yang digambarkan pada diagram pencar sebagai arah y dan satu atau beberapa variabel independen digambarkan sebagai arah x . Keberadaan data *outlier* mungkin terdapat pada arah y atau pada arah x atau pada arah keduanya.

a. *Leverage Value*

Nilai *leverage* hanya menggambarkan *outlier* yang terjadi pada variabel independen. *Leverage* menginformasikan seberapa jauh *outlier* tersebut dari nilai *mean* data variabel independen. Pada analisis regresi dengan beberapa variabel independen, nilai *leverage* di peroleh dari matriks $\mathbf{H} = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$, dimana suatu pengamatan ke- i dapat dikatakan sebagai *high leverage point* jika $h_{ii} > 2p/n$ dengan h_{ii} di tunjukkan pada Persamaan (2.4).

$$h_{ii} = \mathbf{x}_i^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{x}_i, \quad (2.4)$$

untuk \mathbf{x}_i^T menyatakan setiap baris ke- i dari \mathbf{X} , p menyatakan banyaknya parameter dan n menyatakan banyaknya pengamatan (Myers, 1990).

b. Studentized Residual

Untuk mendeteksi *outlier* terhadap variabel dependen yaitu menggunakan *externally studentized residual* atau biasa disebut *R-student* yang didefinisikan pada Persamaan (2.5) (Montgomery *et al.*, 2012).

$$t_i = \frac{\varepsilon_i}{s_{-i}\sqrt{1-h_{ii}}}, \quad (2.5)$$

dimana s_{-i} adalah simpangan baku yang dihitung tanpa mengikutsertakan pengamatan ke- i .

$$s_{-i} = \sqrt{\frac{(n-p)s^2 - \varepsilon_i^2 / (1-h_{ii})}{n-p-1}}, \quad (2.6)$$

dengan $s^2 = \sum_{i=1}^n \frac{(y_i - \hat{y}_i)^2}{n-p}$.

Suatu pengamatan dikatakan sebagai *outlier* apabila pengamatan tersebut memiliki nilai $|t_i| > t_{\alpha/2; n-(p+1)}$.

c. Difference fitted value FITS

Difference fitted value FITS (*DfFITS*) merupakan suatu ukuran untuk mendeteksi pengaruh yang ditimbulkan oleh pengamatan ke- i terhadap nilai taksiran y_i . *DfFITS* diperkenalkan oleh Belsley *et al.* (1980) untuk mendeteksi *outlier* yang dideteksi dari nilai \hat{y}_i serta gabungan nilai *leverage* (h_{ii}) dengan residual yang sudah terstandarkan (Rousseeuw & Leroy, 1987). Nilai *DfFITS* dihitung dengan Persamaan (2.7).

$$DfFITS_i = t_i \sqrt{\frac{h_{ii}}{1-h_{ii}}}, \quad (2.7)$$

Suatu pengamatan ke- i data diidentifikasi sebagai *outlier* apabila:

$$|DfFITS| > 1 \text{ untuk } n \leq 30,$$

$$|DfFITS| > 2\sqrt{(p/n)} \text{ untuk } n > 30.$$

2.2.3 Ordinary Least Squares

Ordinary Least Squares (OLS) sering digunakan dalam estimasi parameter model regresi linear. Metode OLS pada prinsipnya adalah meminimumkan jumlah kuadrat dari kesalahan yang disebut dengan jumlah kuadrat residual terhadap garis regresi. Estimator yang dihasilkan OLS akan bersifat tak bias dan efisien (*Best Linear Unbiased Estimator/BLUE*) jika komponen sisaan atau residual memenuhi beberapa asumsi klasik, yaitu: kenormalan, kehomogenan varians dan tidak terjadi autokorelasi (Weisberg, 2014). Jika terdapat pelanggaran terhadap asumsi tersebut, estimator yang diperoleh bersifat bias dan tidak efisien sehingga model regresi yang diperoleh tidak cocok (*fit*) terhadap data yang dimodelkan. Model linier dengan beberapa variabel independen dapat dituliskan dengan Persamaan (2.8).

$$y_i = \beta_0 + \sum_{k=1}^p \beta_k x_{ik} + \varepsilon_i, \quad (2.8)$$

dengan,

y_i : nilai observasi variabel dependen ke- i

β_0 : nilai intersep model regresi

x_{ik} : nilai observasi variabel prediktor ke- k pada pengamatan ke- i

β_k : koefisien regresi variabel prediktor ke- k

ε_i : residual pada pengamatan ke- i dengan asumsi identik, independen dan berdistribusi normal.

Persamaan (2.8) dapat dinyatakan dalam bentuk matriks sebagai berikut:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad (2.9)$$

dengan,

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{p1} \\ 1 & x_{12} & x_{22} & \cdots & x_{p2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & x_{2n} & \cdots & x_{pn} \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}, \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}.$$

Selanjutnya jumlah kuadrat residual atau *Sum of Square Error* (SSE) dinyatakan pada Persamaan (2.10).

$$\begin{aligned} SSE &= \sum_{i=1}^n \varepsilon_i^2 = \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} \\ &= (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \\ &= \mathbf{y}^T \mathbf{y} - 2\mathbf{y}^T \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\beta}^T \mathbf{X}^T \mathbf{X}\boldsymbol{\beta}. \end{aligned} \quad (2.10)$$

Dalam hal ini, estimasi kuadrat terkecil harus memenuhi:

$$\frac{\partial SSE}{\partial \hat{\boldsymbol{\beta}}} = -2\mathbf{X}^T \mathbf{y} + 2\mathbf{X}^T \mathbf{X}\hat{\boldsymbol{\beta}} = 0,$$

$$\mathbf{X}^T \mathbf{y} = \mathbf{X}^T \mathbf{X}\hat{\boldsymbol{\beta}},$$

sehingga nilai taksiran $\boldsymbol{\beta}$ adalah:

$$\hat{\boldsymbol{\beta}} = \mathbf{b} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}, \quad (2.11)$$

dengan,

$\boldsymbol{\varepsilon}$: vektor residual berukuran ($n \times 1$)

\mathbf{y} : vektor variabel dependen berukuran ($n \times 1$)

\mathbf{X} : matrik variabel independen berukuran ($n \times p$)

$\boldsymbol{\beta}$: vektor parameter regresi berukuran ($p \times 1$).

2.3 Regresi *Robust*

Regresi *robust* merupakan metode regresi yang digunakan ketika distribusi dari sisaan tidak normal atau adanya beberapa *outlier* yang berpengaruh pada model. Metode ini digunakan untuk menganalisis data yang dipengaruhi *outlier* karena metode ini cenderung mengabaikan residual yang berhubungan dengan *outlier* yang besar, sehingga estimator yang dihasilkan tetap bekerja baik meskipun terdapat data *outlier* (Atkinson & Riani, 2000). Metode

regresi *robust* yang digunakan pada penelitian ini yaitu metode *robust* LTS dan metode *robust* LWS karena memiliki *breakdown point* yang tinggi.

2.3.1 Breakdown Point

Breakdown point adalah jumlah observasi minimal yang dapat menggantikan sejumlah observasi awal yang berakibat pada nilai estimator yang dihasilkan sangat berbeda dari estimator sebenarnya. Dengan kata lain, *breakdown point* sebagai suatu ukuran ke-*robust*-an dari suatu estimator terhadap *outlier*. Semakin besar nilai persentase dari *breakdown point* pada suatu estimator, maka estimator tersebut semakin *robust*. Kemungkinan *breakdown point* terkecil adalah $1/n$ yang cenderung 0% ketika ukuran sampel n menjadi besar. Hal tersebut terjadi pada estimasi *least squares*. Jika estimator regresi *robust* memiliki 50% *breakdown point* maka 50% dari data dapat berisi *outlier* dan koefisien akan tetap dapat digunakan (Gad & Qura, 2016). Regresi *robust* yang mempunyai *high breakdown point* adalah regresi *robust* dengan metode LTS dan LWS.

2.3.2 Least Trimmed Squares

Estimasi *Least Trimmed Squares* (LTS) adalah metode dengan *high breakdown point* yang dikenalkan oleh Rousseeuw (1984). Tingkat ke-*robust*-an terbaik akan tercapai saat $h = n/2$ yang berarti *breakdown point* yang dimiliki mencapai 50% (Rousseeuw & Leroy, 1987). Prosedur estimasi LTS mirip dengan estimasi OLS, tetapi dalam prosedur LTS, kecocokan tidak begitu banyak dipengaruhi oleh *outlier*, karena titik data yang sesuai dengan persentase tertentu dari residual tertinggi berdasarkan estimasi OLS awal dihapus (Mutan, 2009). Metode LTS akan meminimumkan jumlah kuadrat sebanyak h dengan $h \leq n$ residual (fungsi objektif).

$$\min \sum_{i=1}^h \varepsilon_i^2, \quad (2.12)$$

dimana $h = \frac{(n+p+1)}{2}$, dengan,

ε_i^2 : kuadrat residual yang diurutkan dari terkecil ke terbesar

$$\varepsilon_1^2 < \varepsilon_2^2 < \dots < \varepsilon_i^2 < \dots < \varepsilon_h^2 < \dots < \varepsilon_n^2$$

n : banyak sampel

p : banyaknya parameter.

Jumlah h menunjukkan sejumlah subset data dengan kuadrat fungsi objektif terkecil. Nilai residual pada LTS di dapatkan dengan menggunakan algoritma LTS. Menurut Rousseeuw dan Van Driessen (1998) algoritma LTS adalah gabungan FAST-LTS dan *C-step*. Pembentukan estimasi parameter LTS dilakukan hingga *Final Weighted Least Square* (FWLS) dengan fungsi pembobot pada Persamaan (2.13).

$$w_i = \begin{cases} 0 & , \frac{|\varepsilon_i|}{S_{LTS}} > r \\ 1 & , \text{lainnya} \end{cases}, \quad (2.13)$$

dengan $r = 3$ dan

$$S_{LTS} = d_{h,n} \sqrt{\frac{1}{h} \sum_{i=1}^n \varepsilon_i^2}, \quad (2.14)$$

$$d_{h,n} = \frac{1}{\sqrt{1 - \frac{2n}{hc_{h,n}} \phi\left(\frac{1}{c_{h,n}}\right)}}, \quad (2.15)$$

$$c_{h,n} = \frac{1}{\Phi^{-1}\left(\frac{(h+n)}{2n}\right)}, \quad (2.16)$$

dengan,

n : banyak pengamatan

Φ : fungsi kumulatif normal standar

ϕ : fungsi densitas normal standar.

Algoritma dalam pembentukan model dengan metode estimasi LTS dilakukan dengan tahapan sebagai berikut.

- i. Menghitung estimasi parameter $b_{(0)}$ melalui OLS
- ii. Menghitung n residual kuadrat $\varepsilon_i^2 = (\hat{y}_i - x_i b_{(0)})^2$ melalui OLS ($b_{(0)}$)
- iii. Menghitung $h = \frac{(n+p+1)}{2}$ pengamatan dengan nilai ε_i^2 terkecil
- iv. Menghitung $\sum_{i=1}^h \varepsilon_i^2$
- v. Melakukan estimasi parameter baru (b_{new}) dari h pengamatan melalui OLS
- vi. Menghitung n residual kuadrat baru $\varepsilon_{i_new}^2 = (\hat{y}_i - x_i b_{new})^2$ yang bersesuaian dengan OLS (b_{new})
- vii. Menghitung h_{new} pengamatan dengan nilai $\varepsilon_{i_new}^2$ terkecil
- viii. Menghitung $\sum_{i=1}^{h_{new}} \varepsilon_{i_new}^2$
- ix. Melakukan *C-steps* (tahap v sampai viii) untuk mendapatkan fungsi objektif yang kecil dan konvergen
- x. Menghitung bobot untuk masing-masing pengamatan dengan fungsi FWLS
- xi. Mengestimasi parameter dengan *Weighted Least Square*.

2.3.3 Least Winsorized Squares

Metode *Least Winsorized Squares* (LWS) diperkenalkan oleh Winsor (1946) sebagai salah satu alternatif untuk menyelesaikan masalah pada perhitungan nilai statistik jika terdapat *outlier* pada data/observasi atau dapat dikatakan untuk mengurangi efek *outlier*. Metode LWS diterapkan dengan mengganti residual yang paling ekstrim dengan residual terdekat berikutnya dalam sampel dengan cara yang berulang (Mutan,

2009). Estimator pada metode LWS ditunjukkan pada Persamaan (2.17).

$$\min \sum_{i=1}^h (\varepsilon_i^2) + (n-h)(\varepsilon_h^2). \quad (2.17)$$

Untuk mengurangi kompleksitas notasi dan ekspresi yang panjang digunakan bentuk *canonical* dari model regresi linier (Liu & Gao, 2011) dan untuk mendapatkan estimasi LWS dapat digunakan algoritma *stepwise* (Yale & Forsythe, 1976). Bentuk *canonical* regresi linier ditunjukkan pada Persamaan (2.18).

$$\mathbf{Y} = \mathbf{Z}\mathbf{a} + \boldsymbol{\varepsilon}, \quad (2.18)$$

dimana $\mathbf{Z} = \mathbf{X}\mathbf{Q}$, $\mathbf{a} = \mathbf{Q}'\boldsymbol{\beta}$ dan $\mathbf{Q} = (q_1, q_2, \dots, q_p)$ adalah matriks orthogonal dari eigen vektor q_1, q_2, \dots, q_p sesuai dengan *eigenvalue* $\lambda_1, \lambda_2, \dots, \lambda_p \geq 0$ dari matriks $\mathbf{X}^T\mathbf{X}$ dan penggunaan bentuk *canonical* tidak akan berpengaruh terhadap *Mean Square Error* (MSE) dari estimator. Algoritma dalam pembentukan model dengan metode estimasi LWS dilakukan dengan tahapan sebagai berikut (Kashid & Jadhav, 2014).

- i. Menghitung estimasi OLS dan nilai prediksi dari y_i , $i=1, 2, \dots, n$
- ii. Menghitung jumlah poin (h) untuk *winsorized* di setiap nilai ekstrim
- iii. Mendapatkan nilai residual $\varepsilon_i = y_i - \hat{y}_i$ dan diurutkan ($\varepsilon_1 \leq \varepsilon_2 \leq \dots \leq \varepsilon_n$)
- iv. Menghitung estimasi kuadrat terkecil menggunakan n observasi \mathbf{Y}^T dan \mathbf{X} , dengan:

$$Y_i^T = \hat{Y}_i + \varepsilon_i^T \text{ dan } \varepsilon_i^T = \begin{cases} \varepsilon_{h+1}, & i=1, 2, \dots, h \\ \varepsilon_i, & i=g+1, \dots, n-h \\ \varepsilon_{n-h}, & i=n-g+1, \dots, n \end{cases} \quad (2.19)$$
- v. Mengulangi langkah iv untuk menentukan nilai iterasi (b). Pada masing-masing iterasi data awal dari variabel respon

(\mathbf{Y}) yang telah dimodifikasi sebagai $\mathbf{Y} = \mathbf{Y}'$ (saat iterasi pertama), $\mathbf{Y} = \mathbf{Y}''$ (saat iterasi kedua), $\mathbf{Y} = \mathbf{Y}'''$ (saat iterasi ketiga) dan seterusnya dengan residual baru ($\varepsilon', \varepsilon'', \varepsilon'''$ dan seterusnya)

- vi. Modifikasi data pada akhir iterasi b^{th} yang dinotasikan dengan $(\mathbf{Y}^*, \mathbf{X})$. Standarisasi data yang telah dimodifikasi sedemikian rupa sehingga \mathbf{Y}^*, \mathbf{X} menunjukkan korelasi antara variabel respon yang dimodifikasi dan keseluruhan variabel regresi
- vii. Mengubah data modifikasi yang terstandarisasi menjadi bentuk *canonical* dengan menggunakan matriks eigen vektor (\mathbf{Q}) dari matriks.

2.4 Pengujian Signifikansi Parameter

Pengujian signifikansi parameter digunakan untuk mengetahui ada tidaknya hubungan antar parameter didalam model regresi yang dilakukan secara serentak (simultan) dan secara parsial (Draper & Smith, 1998). Pengujian signifikansi parameter pada metode OLS serta regresi *robust* pada dasarnya sama, yang membedakan adalah pada regresi *robust* terdapat pembobotan (*weighted*) pada pengujiannya. Berikut adalah pengujian signifikansi parameter secara serentak dan parsial.

a. Uji Serentak

Uji ini digunakan untuk melihat signifikansi parameter secara bersama-sama dalam model dengan menggunakan nilai F_{hitung} sebagai statistik uji. Hipotesis yang digunakan adalah sebagai berikut.

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

$$H_1: \text{minimal ada satu } \beta_k \neq 0; k = 1, 2, \dots, p$$

Statistik uji:

$$F_{hitung} = \frac{\left[\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2 \right] / p}{\left[\sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \right] / [n - (p + 1)]}, \quad (2.20)$$

Keputusan yang akan diambil yaitu H_0 ditolak apabila $F_{hitung} > F_{\alpha, k; n - (p + 1)}$ atau $P_{value} < \alpha$.

b. Uji Parsial

Uji parsial atau yang sering disebut uji individu dilakukan untuk mengetahui signifikansi parameter β terhadap respon menggunakan statistik uji t . Hipotesis yang digunakan adalah sebagai berikut.

$$H_0 : \beta_k = \beta_{k0} = 0$$

$$H_1 : \beta_k \neq 0 \quad ; k = 0, 1, \dots, p$$

Statistik uji:

$$t = \frac{\hat{\beta}_k}{SE(\hat{\beta}_k)}, \quad (2.21)$$

dengan $SE(\hat{\beta}_k) = s\sqrt{g_{kk}}$ yang merupakan standar *error* dari koefisien $\hat{\beta}_k$. g_{kk} adalah elemen diagonal ke- k matriks $(\mathbf{X}^T \mathbf{X})^{-1}$ dan $s = \sqrt{MSE}$. Keputusan yang akan diambil yaitu H_0 ditolak apabila $|t| > t_{\alpha/2; n - (p + 1)}$ atau $P_{value} < \alpha$.

2.5 Asumsi IIDN

Terdapat beberapa asumsi yang harus dipenuhi dalam analisis regresi, asumsi tersebut adalah sebagai berikut.

a. Residual Identik (tidak ada heteroskedastisitas)

Uji heteroskedastisitas bertujuan 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. Situasi heteroskedastisitas akan menyebabkan penaksiran koefisien-koefisien regresi menjadi tidak efisien dan hasil taksiran dapat menjadi kurang atau melebihi dari yang semestinya dengan ilustrasi yang dapat dilihat pada Gambar 2.2. Hipotesis yang digunakan adalah sebagai berikut.

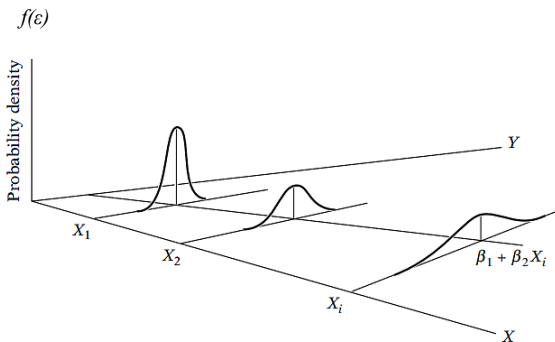
H_0 : Tidak terdapat heteroskedastisitas

H_1 : Terdapat heteroskedastisitas

Uji heteroskedastisitas dapat dilakukan dengan menggunakan uji *Glejser* yang dirumuskan pada Persamaan (2.22) (Gujarati & Porter, 2009).

$$|\boldsymbol{\varepsilon}| = \mathbf{X}\boldsymbol{\beta} + \mathbf{v}, \quad (2.22)$$

dengan \mathbf{v} : matriks residual baru dari hasil regresi absolut residual dengan variabel prediktor. Keputusan tolak H_0 dilakukan ketika $F_{hitung} > F_{\alpha, k; n-(p+1)}$ atau $P_{value} < \alpha$ yang artinya model dinyatakan terdapat masalah heteroskedastisitas.



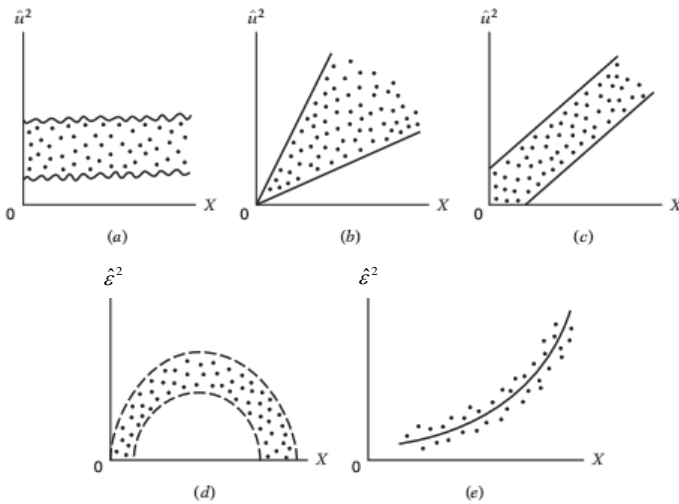
Gambar 2.2 Grafik Heteroskedastisitas.

Sumber : (Gujarati & Porter, 2009)

b. Residual Independen (tidak ada autokorelasi)

Autokorelasi dalam konsep regresi linier berarti komponen residual berkorelasi berdasarkan urutan waktu (pada data *time series*) dan urutan ruang (pada data *cross section*) atau terjadi

ketergantungan antara residual ke- i dengan ke- j (Gujarati & Porter, 2009). Autokorelasi merupakan pelanggaran terhadap asumsi independen. Residual dikatakan independen ketika plot menyebar secara acak atau random. Ketika plot data menyebar membentuk pola maka dapat dikatakan tidak independen atau terjadi autokorelasi. Pola plot data yang dapat mengindikasikan terjadi autokorelasi diantaranya ditunjukkan pada Gambar 2.3.



Gambar 2.3 Grafik Autokorelasi.
Sumber : (Gujarati & Porter, 2009)

Namun pada penelitian ini data yang digunakan merupakan data acak dimana faktor-faktor yang mempengaruhi indeks prestasi antara mahasiswa satu dengan lainnya tidak saling bergantung, selain itu data pada penelitian ini bukan merupakan deret waktu sehingga tidak ada pengujian yang valid untuk menyimpulkan bahwa residual independen. Maka pada penelitian ini residual akan diasumsikan telah independen.

c. Residual Berdistribusi Normal

Uji distribusi normal digunakan untuk menguji apakah model regresi variabel bebas dan variabel terikat keduanya mempunyai

distribusi normal atau tidak. Pada regresi linear diasumsikan bahwa tiap ε_i didistribusikan secara random dengan $\varepsilon_i \sim N(0, \sigma^2)$ (Draper & Smith, 1998). Umumnya asumsi normalitas tidak akan terpenuhi jika terdapat data yang mengandung *outlier*. *Outlier* akan membuat bentuk sebaran data menjadi tidak simetrik dan lebih menjulur ke arah pencilan (Herawati dkk., 2011). Ketika asumsi ini tidak terpenuhi juga akan menyebabkan nilai residual besar. Ilustrasi distribusi normal dapat dilihat pada Gambar 2.4. Salah satu uji yang dapat digunakan adalah uji *Kolmogorov-Smirnov* dengan rumus pada Persamaan (2.23).

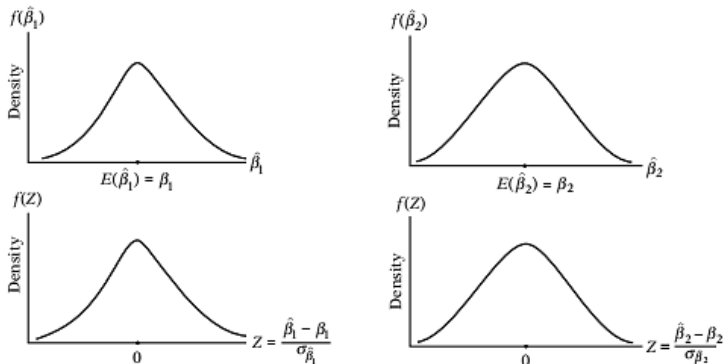
$$D = \max |F_0(\varepsilon_i) - F_n(\varepsilon_i)|, \quad (2.23)$$

dengan $i=1,2,\dots,n$. $F_0(\varepsilon_i)$ adalah fungsi distribusi frekuensi kumulatif relatif dari distribusi normal residual dan $F_n(\varepsilon_i)$ adalah distribusi frekuensi kumulatif residual pengamatan sebanyak n sampel dengan menggunakan kurva normalitas. Hipotesis dalam uji distribusi normal adalah sebagai berikut.

H_0 : Data berdistribusi normal

H_1 : Data tidak berdistribusi normal

Apabila nilai $D > D_{tabel}$ atau $P_{value} < \alpha$ maka asumsi normalitas residual tidak terpenuhi.



Gambar 2.4 Grafik Distribusi Normal.

Sumber : (Gujarati & Porter, 2009)

2.6 Pemilihan Model Terbaik

Pemilihan model terbaik dapat dilakukan dengan menggunakan nilai MSE dan koefisien determinasi (R^2). MSE merupakan rata-rata kuadrat nilai residual pengamatan dengan prediksi sedangkan R^2 merupakan metode untuk mengukur seberapa jauh kemampuan model dalam menerangkan variasi dari variabel independen. Nilai R^2 adalah antara 0 dan 1. Metode terbaik dipilih dengan kriteria semakin besar nilai R^2 dan semakin kecil nilai MSE yang ditunjukkan pada Persamaan (2.24) dan (2.25) (Gujarati & Porter, 2009).

$$MSE = \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n - (p + 1)}, \quad (2.24)$$

$$R^2 = \frac{\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2}. \quad (2.25)$$

2.7 Prestasi Belajar

Menurut Winkel (2009) prestasi belajar adalah hasil suatu penilaian yang dinyatakan dalam bentuk nilai. Prestasi belajar juga merupakan suatu bukti keberhasilan belajar atau kemampuan seseorang mahasiswa dalam melakukan kegiatan belajarnya sesuai dengan bobot yang dicapainya. Umumnya prestasi belajar dinyatakan dengan angka, huruf atau kalimat dan terdapat dalam periode tertentu (Muslimin, 2012). Berdasarkan Peraturan Rektor Institut Teknologi Sepuluh Nopember Nomor 15 Tahun 2018 tentang Peraturan Akademik ITS tahun 2018, ukuran keberhasilan proses dan hasil belajar mahasiswa dinyatakan dengan Indeks Prestasi (IP) yang dihitung dengan Persamaan (2.26).

$$IP = \frac{\sum_{j=1}^m K_j \times M_j}{\sum_{j=1}^m K_j}, \quad (2.26)$$

dengan,

M : nilai numerik hasil evaluasi masing-masing mata kuliah

K : besar sks masing-masing mata kuliah

m : jumlah mata kuliah yang telah diambil.

Ukuran keberhasilan kegiatan proses dan hasil belajar dalam satu semester dinyatakan dengan Indeks Prestasi Semester (IPS) yaitu IP yang dihitung dari semua mata kuliah yang diambil dalam semester yang bersangkutan dan pada awal masa perkuliahan dikenal adanya Indeks Prestasi Tahap Persiapan yaitu nilai Indeks pada semester 1 dan 2, sedangkan Indeks Prestasi Kumulatif (IPK) dihitung pada akhir masa studi.

2.7.1 Faktor-faktor yang Mempengaruhi Prestasi Belajar

Terdapat beberapa faktor yang mempengaruhi prestasi belajar, menurut Dalyono (2007) faktor tersebut adalah faktor internal yang merupakan faktor dari diri mahasiswa serta faktor eksternal yang merupakan faktor dari luar diri mahasiswa. Faktor internal meliputi kesehatan, intelegensi dan bakat, minat dan motivasi serta cara belajar. Sedangkan faktor eksternal mencakup keluarga, sekolah, masyarakat dan lingkungan sekitar. Faktor internal pada penelitian ini akan diwakili oleh variabel jenis kelamin, nilai mata uji jalur ujian tulis serta kategori bidikmisi sedangkan faktor eksternal akan diwakili variabel asal daerah, jenis sekolah dan jurusan sekolah.

a. Jalur Ujian Tulis

Jalur ujian tulis merupakan salah satu pola penerimaan mahasiswa baru yang dilakukan oleh PTN di lingkungan Kementerian Riset, Teknologi, dan Pendidikan Tinggi (Kemristekdikti) dan Kementerian Agama secara bersama dibawah koordinasi oleh Lembaga Tes Masuk Perguruan Tinggi (LTMP). Sebanyak 85 PTN ikut dalam pelaksanaan jalur ujian

tulis dengan total 3.167 pilihan prodi. Jalur ujian tulis memiliki banyak manfaat yaitu lebih efisien, murah, dan fleksibel dengan adanya mekanisme lintas wilayah. Dari tahun ke tahun skema penerimaan jalur ujian tulis selalu diperbaiki dan dikembangkan sesuai dengan perkembangan zaman dan kemajuan teknologi. Seleksi jalur ujian tulis mengujikan 10 mata uji pada tahap seleksinya untuk masing-masing kelompok ujian saintek dan soshum (SBMPTN, 2019).

b. Program studi

Program Studi adalah kesatuan kegiatan pendidikan dan pembelajaran yang memiliki kurikulum dan metode pembelajaran tertentu dalam satu jenis pendidikan akademik, pendidikan profesi, dan/atau pendidikan vokasi. ITS memiliki 27 program studi yang dinaungi oleh 5 fakultas dan pada akhir tahun 2017 dikembangkan menjadi 35 program studi dengan 10 fakultas (ITS, 2019).

c. Jenis Kelamin

Terdapat pengaruh langsung peran gender terhadap prestasi belajar. Siswa perempuan lebih banyak yang berprestasi daripada siswa laki-laki. Siswa perempuan cenderung mempunyai kepribadian rapi dalam belajar, motivasi untuk belajar juga lebih tinggi, sedangkan siswa laki-laki cenderung agak malas belajar dan bersikap acuh terhadap motivasi belajar. Namun, dalam hal tantangan kesuksesan akademik siswa laki-laki lebih tinggi, sedangkan perempuan lebih bersifat aktif belajar karena pengaruh kondisi setempat yang membentuk agar rajin belajar (Zahroh, 2008).

d. Jenis Sekolah

Terdapat tiga kategori jenis sekolah yang ada di Indonesia, yaitu umum (SMA), madrasah (MA) dan sekolah kejuruan (SMK) dengan kurikulum yang berbeda pada masing-masing jenis sekolah. SMA dan MA biasanya didesain untuk lebih menekankan pembelajaran teori dan lebih diarahkan untuk melanjutkan tingkat pendidikan yang lebih tinggi. Sedangkan

SMK lebih banyak praktik secara langsung untuk meningkatkan keahlian siswanya sehingga akan lebih siap untuk langsung terjun ke dunia kerja. Hal tersebut dapat berdampak terhadap prestasi belajar yang dimiliki (Julina, 2015).

e. Bidikmisi

Bidikmisi merupakan program bantuan biaya pendidikan yang diberikan pemerintah bagi calon mahasiswa yang tidak mampu secara ekonomi dan memiliki potensi akademik baik untuk menempuh pendidikan di perguruan tinggi pada program studi unggulan sampai lulus tepat waktu (Ristekdikti, 2019). Mahasiswa penerima bidikmisi cenderung memiliki prestasi yang lebih tinggi dibanding mahasiswa non bidikmisi. Terdapat beberapa penyebab perbedaan tersebut diantaranya karena mahasiswa bidikmisi memiliki motivasi tinggi untuk meningkatkan prestasinya karena mereka dituntut agar lulus tepat waktu dan memiliki tekad yang kuat untuk memutus rantai kemiskinan di keluarganya (Sucahyo, 2014).

f. Asal Daerah

Mahasiswa ada yang berasal dari daerah tempat PTN yang di tuju maupun luar daerah bahkan hingga luar pulau. Hal tersebut akan mempengaruhi tingkat adaptasi terhadap lingkungan baru. Mahasiswa yang berasal dari desa biasanya memiliki prestasi yang rendah dibandingkan yang berasal dari kota. Hal tersebut dapat dikarenakan keterbatasan biaya. Selain itu sikap individu serta lingkungan pada penduduk pedesaan terhadap tingkat pendidikan masih kurang memprioritaskan dibanding penduduk perkotaan. Sehingga perbedaan wilayah dapat berpengaruh terhadap prestasi belajar (Julina, 2015).

2.8 Penelitian Sebelumnya

Terdapat beberapa penelitian sebelumnya yang membahas mengenai analisis indeks prestasi serta analisis pada data yang mengandung *outlier*, diantaranya oleh Antari (2016), Matdoan (2017), Shodiqin dkk. (2018) dan Herawati dkk. (2011). Antari (2016) meneliti tentang pemodelan Indeks Prestasi Persiapan (IPP)

mahasiswa ITS dari jalur SNMPTN menggunakan regresi kuantil, dari analisis tersebut diperoleh hasil nilai IPP mahasiswa ITS banyak memiliki data *outlier* serta mahasiswa bidikmisi cenderung memiliki IPP lebih tinggi dibanding mahasiswa non bidikmisi. Matdoan (2017) meneliti tentang perbandingan regresi kuantil dengan regresi *robust* LTS pada kasus faktor-faktor yang mempengaruhi penyebaran penyakit malaria di Indonesia dan diperoleh hasil regresi *robust* LTS dan regresi kuantil lebih baik dibandingkan metode OLS. Shodiqin dkk. (2018) meneliti tentang perbandingan metode regresi *robust* LTS dan estimasi-MM pada data ujian tulis masuk terhadap IPK mahasiswa UPGRIS, dari analisis diperoleh hasil regresi *robust* LTS lebih baik daripada estimasi-MM dan penelitian oleh Herawati dkk. (2011) tentang perbandingan analisis ketegaran regresi *robust* terhadap letak pencilan dengan hasil metode estimasi-MM lebih baik dibandingkan metode OLS.

(Halaman ini sengaja dikosongkan)

BAB III

METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Panitia Pengkaji Jalur Tes Tulis. Data penelitian merupakan data mahasiswa yang lolos pada seleksi masuk Perguruan Tinggi Negeri tahun 2014 secara nasional. Sehingga unit sampling dalam penelitian ini adalah mahasiswa. Data IPS1, IPS2 dan IPK mahasiswa digunakan sebagai variabel dependen yang akan di analisis secara univariat terhadap variabel independen. Variabel independen yang digunakan terdiri dari 10 mata uji pada jalur tes tulis yaitu 6 mata uji TKPA (verbal, numerikal, figural, matematika dasar, B. Indonesia dan B. Inggris) dan 4 mata uji TKD saintek (Matematika IPA, Fisika, Kimia, Biologi) karena pada prodi ITS sebagai acuan prodi nasional merupakan prodi saintek. Selain itu pada variabel independen juga akan digunakan data jenis kelamin, jenis sekolah, jurusan sekolah, asal daerah, bidikmisi dan pilihan prodi.

Sebelum melakukan analisis pada data mahasiswa jalur tes tulis akan dilakukan tahap *preprocessing* data terlebih dahulu untuk menghilangkan data *missing*. Data *missing* merupakan data yang tidak terisi (*blank*) ataupun data yang bernilai nol. Hal tersebut dapat disebabkan mahasiswa tidak melakukan daftar ulang atau mengundurkan diri. Selain itu dapat pula disebabkan karena institusi yang bersangkutan belum melaporkan. Pada data awal terdapat sebanyak 79.947 data. Selanjutnya di lakukan pemilihan prodi secara nasional yang disesuaikan dengan prodi ITS sebagai acuan dan diperoleh sejumlah 10.835 data, data ini tanpa menyertakan prodi Desain Produk karena prodi tersebut hanya terdapat di ITS dan pada tahun tersebut hanya menerima calon mahasiswa melalui jalur ujian yang dilaksanakan oleh ITS yaitu UMDESAIN. Tahap selanjutnya menghapus data *missing* hingga diperoleh data untuk analisis sebanyak 9.177 data. Dari data tersebut akan dibagi berdasarkan prodi ITS sebagai acuan prodi

nasional. Pada analisis regresi nantinya data pada masing-masing prodi akan dibagi menjadi data *testing* dan data *training* dengan perbandingan 1 : 3.

Berikut adalah daftar prodi ITS yang akan digunakan sebagai acuan analisis secara nasional dalam penelitian ini beserta jumlah data yang akan digunakan sebagai analisis.

Tabel 3.1 Program Studi ITS

No.	Program Studi	Jumlah Data	No.	Program Studi	Jumlah Data
1.	Fisika	654	14.	Teknik Sipil	943
2.	Matematika	740	15.	Arsitektur	437
3.	Statistika	183	16.	Teknik Lingkungan	221
4.	Kimia	775	17.	Teknik Geomatika	28
5.	Biologi	762	18.	PWK	148
6.	Teknik Mesin	714	19.	T. Geofisika	90
7.	Teknik Elektro	839	20.	Desain Interior	17
8.	Teknik Kimia	417	21.	Teknik Perkapalan	117
9.	Teknik Fisika	30	22.	Teknik Sistem Perkapalan	65
10.	Teknik Industri	449	23.	Teknik Kelautan	79
11.	Teknik Material dan Metalurgi	86	24.	Transportasi Laut	13
12.	Manajemen Bisnis	18	25.	Teknik Informatika	871
13.	Teknik Multimedia dan Jaringan	18	26.	Sistem Informasi	463

3.2 Variabel Penelitian

Variabel yang digunakan pada penelitian ini akan di uraikan pada Tabel 3.2.

Tabel 3.2 Variabel Penelitian

Variabel Dependen	Keterangan	Skala Data
Y_1 : IPS1	Indeks Prestasi Mahasiswa saat Semester 1	Rasio
Y_2 : IPS2	Indeks Prestasi Mahasiswa saat Semester 2	Rasio
Y_3 : IPK	Indeks Prestasi Mahasiswa Secara Kumulatif (diperoleh dari IPS 1 hingga IPS 8)	Rasio

Tabel 3.2 Variabel Penelitian (lanjutan)

Variabel Independen	Keterangan	Skala Data
X ₁ : Verbal	Tes Mata Uji Verbal	Rasio
X ₂ : Numerikal	Tes Mata Uji Numerikal	Rasio
X ₃ : Figural	Tes Mata Uji Figural	Rasio
X ₄ : Matematika Dasar	Tes Mata Uji Matematika Dasar	Rasio
X ₅ : B. Indonesia	Tes Mata Uji B. Indonesia	Rasio
X ₆ : B. Inggris	Tes Mata Uji B. Inggris	Rasio
X ₇ : Matematika IPA	Tes Mata Uji Matematika IPA	Rasio
X ₈ : Fisika	Tes Mata Uji Fisika	Rasio
X ₉ : Kimia	Tes Mata Uji Kimia	Rasio
X ₁₀ : Biologi	Tes Mata Uji Biologi	Rasio
X ₁₁ : Jenis Kelamin	0 : Laki-laki	Nominal
	1 : Perempuan	
X ₁₂ : Jenis Sekolah	0 : MA	Nominal
	1 : SMA	
	2 : SMK	
	3 : SMTA Lain-lain	
X ₁₃ : Jurusan Sekolah	0 : IPA	Nominal
	1 : IPS 2 : Lainnya (Agama dan Bahasa)	
X ₁₄ : Asal Daerah	0 : Jawa	Nominal
	1 : Luar Jawa	
X ₁₅ : Bidikmisi	0 : Tidak Bidikmisi	Nominal
	1 : Bidikmisi	
X ₁₆ : Pilihan Prodi	1 : Pilihan 1	Nominal
	2 : Pilihan 2	
	3 : Pilihan 3	

Struktur data analisis univariat dalam penelitian ini di tunjukkan pada Tabel 3.3.

Tabel 3.3 Struktur Data Penelitian

Subjek	Y ₁	Y ₂	Y ₃	X ₁	X ₂	X ₃	...	X ₁₆
1	Y _{1,1}	Y _{2,1}	Y _{3,1}	X _{1,1}	X _{2,1}	X _{3,1}	...	X _{16,1}
2	Y _{1,2}	Y _{2,2}	Y _{3,2}	X _{1,2}	X _{2,2}	X _{3,2}	...	X _{16,2}
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
183*)	Y _{1,n}	Y _{2,n}	Y _{3,n}	X _{1,n}	X _{2,n}	X _{3,n}	...	X _{16,n}

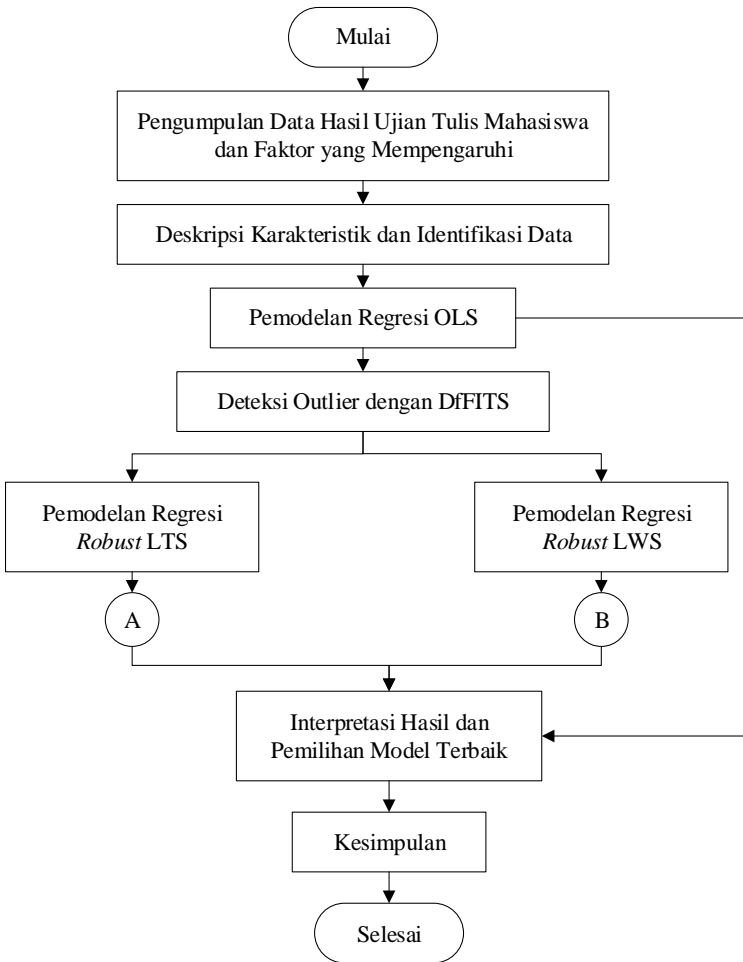
*) sebagai contoh adalah Prodi Statistika sedangkan pada prodi lainnya akan menyesuaikan jumlah pengamatan sesuai Tabel 3.1.

3.3 Langkah analisis

Secara umum langkah analisis faktor-faktor yang mempengaruhi prestasi belajar mahasiswa berdasarkan tujuan yang ingin di capai yaitu sebagai berikut.

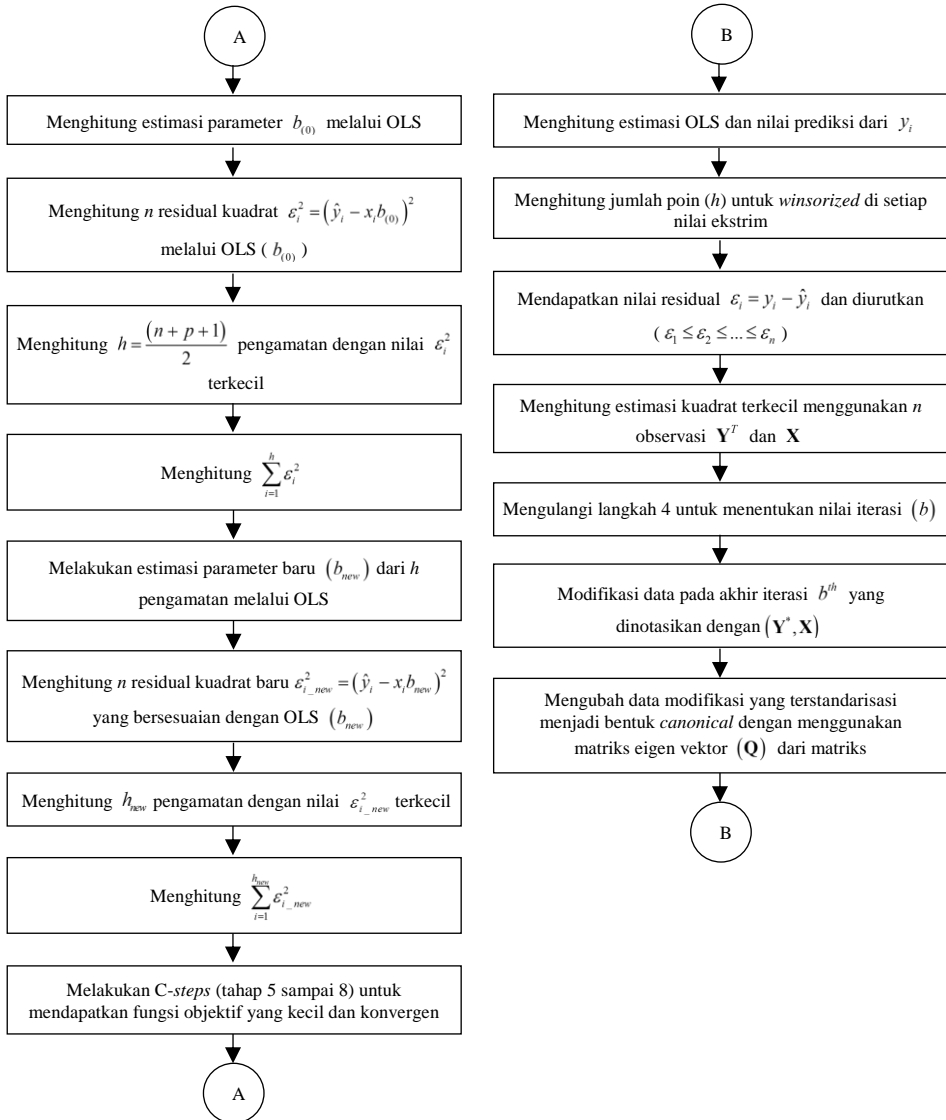
1. Melakukan identifikasi data dengan analisis statistika deskriptif untuk memperoleh karakteristik indeks prestasi dan nilai mata uji mahasiswa jalur ujian tulis, selanjutnya membuat diagram pencar serta analisis korelasi dengan rumus pada Persamaan 2.1 dan 2.2 antar masing-masing variabel prediktor dengan variabel respon untuk mengetahui pola hubungannya
2. Melakukan analisis dan membandingkan estimasi model dengan metode OLS dan metode *robust* menggunakan studi simulasi yang akan dijelaskan langkah-langkahnya pada Subbab 3.4.
3. Memodelkan faktor-faktor yang mempengaruhi prestasi belajar mahasiswa jalur ujian tulis:
 - a. Melakukan analisis regresi OLS faktor-faktor yang mempengaruhi prestasi belajar mahasiswa yaitu estimasi parameter dan pengujian asumsi residual sesuai tinjauan pustaka pada Subbab 2.4 dan 2.5.
 - b. Mendeteksi *outlier* dari residual analisis regresi OLS dengan menggunakan nilai *DFFITs* sesuai Persamaan 2.7.
 - c. Melakukan pemodelan regresi *robust* LTS dan LWS pada data faktor-faktor yang mempengaruhi prestasi belajar dengan langkah-langkah pada Subbab 2.3.2 dan 2.3.3.
 - d. Mendapatkan faktor-faktor yang mempengaruhi prestasi belajar mahasiswa
 - e. Mendapatkan nilai kebaikan model yaitu MSE (Persamaan 2.24) dan R^2 (Persamaan 2.25)
 - f. Melakukan pemilihan model terbaik dengan membandingkan nilai MSE dan R^2 dari ketiga metode
 - g. Menginterpretasikan hasil analisis dan memperoleh kesimpulan.

Langkah-langkah analisis secara umum dalam penelitian ini dapat digambarkan dengan diagram alir pada Gambar 3.1.



Gambar 3.1 Diagram Alir Pemodelan Faktor-faktor yang Mempengaruhi Prestasi Belajar.

Secara lebih rinci diagram alir pada masing-masing metode *robust* diuraikan sebagai berikut.



Gambar 3.2 Diagram Alir Regresi *Robust*.

3.4 Langkah Analisis Studi Simulasi

Studi simulasi digunakan untuk memberikan contoh perbandingan efek masing-masing metode terhadap adanya *outlier* pada data. Studi simulasi dilakukan dengan langkah-langkah sebagai berikut.

1. Menetapkan parameter regresi $\beta_0 = 0$ dan $\beta_1 = 1$ (Maharani dkk., 2014), menetapkan ukuran pengamatan ($n = 100$) serta tingkat persentase *outlier* (dalam simulasi ini sebesar 5%)
2. Membangkitkan variabel independen (x) berdistribusi uniform dengan nilai minimum 0 dan maksimum 20 $\sim U(0,20)$
3. Membangkitkan residual (ε) berdistribusi normal standar $\sim N(0,1)$ tanpa kontaminasi *outlier* sebanyak n pengamatan dan membangkitkan data *outlier* (ε_{out}) sebanyak 5% dari n pengamatan (5 data) dengan distribusi normal $\sim N(50,1)$
4. Mengganti sejumlah 5 data residual dari residual tanpa adanya kontaminasi *outlier* (ε) dengan data *outlier* (ε_{out}) pada urutan nilai variabel bebas (x) tertinggi. Data residual hasil penggantian merupakan residual yang terkontaminasi (ε')
5. Menghitung nilai y_i dari persamaan regresi $y_i = x_i + \varepsilon_i$ dengan $i = 1, 2, 3, \dots, 100$ lalu mencari nilai estimasi β_0 dan β_1 dengan menggunakan metode OLS sehingga diperoleh model regresi tanpa adanya kontaminasi *outlier* sebagai berikut:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad , i = 1, 2, 3, \dots, 100$$

6. Menghitung nilai y_i dari persamaan regresi $y_i = x_i + \varepsilon'_i$ dengan $i = 1, 2, 3, \dots, 100$ kemudian mencari nilai estimasi β_0 dan β_1 dengan menggunakan metode OLS sehingga diperoleh model regresi dengan kontaminasi *outlier* sebagai berikut:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon'_i \quad , i = 1, 2, 3, \dots, 100$$

7. Mencari nilai estimasi β_0 dan β_1 dengan metode LTS pada data residual yang terkontaminasi *outlier* (ε'). Pada metode ini data *outlier* akan dihapus sehingga n pengamatan berkurang sebanyak \leq persen *outlier* dan dilakukan iterasi hingga

diperoleh model dengan estimasi parameter yang konvergen sebagai berikut:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon'_i \quad , i = 1, 2, 3, \dots, 92$$

8. Mencari nilai estimasi β_0 dan β_1 dengan metode LWS pada data residual yang terkontaminasi *outlier* (ε'). Pada metode ini data *outlier* akan diganti dengan residual terdekat yang tidak *outlier* dengan menggunakan persentase *trimm* sebesar 10% dan dilakukan iterasi hingga diperoleh model dengan estimasi parameter yang konvergen sebagai berikut:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon'_i \quad , i = 1, 2, 3, \dots, 100$$

9. Membandingkan hasil analisis pada ketiga metode dan menarik kesimpulan.

BAB IV

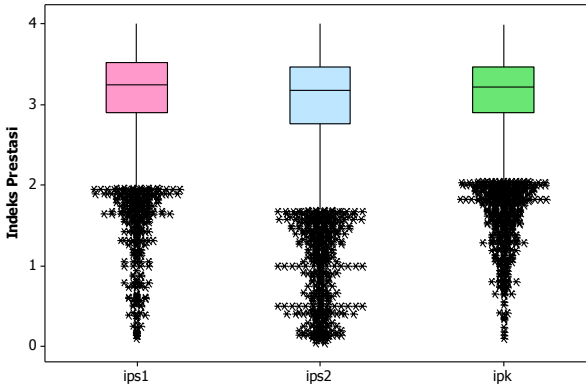
ANALISIS DAN PEMBAHASAN

Pada bab ini akan ditunjukkan hasil analisis berdasarkan tujuan yang ingin dicapai menggunakan metode dan data yang telah dijelaskan pada bab sebelumnya. Pembahasan hasil analisis meliputi karakteristik indeks prestasi dan nilai mata uji, analisis korelasi, hasil analisis pada studi simulasi, pemeriksaan multikolinearitas, deteksi *outlier* pada residual data, analisis regresi serta asumsinya kemudian pemilihan metode terbaik dari ketiga metode regresi yaitu OLS, *robust LTS* dan *robust LWS*.

4.1 Karakteristik Indeks Prestasi dan Nilai Mata Uji

Indeks prestasi merupakan hasil belajar yang dapat menggambarkan kemampuan mahasiswa dalam proses belajar. Pada awal proses belajar, kemampuan tersebut dapat tercermin dari IPS 1 dan IPS 2 sedangkan pada akhir proses belajar di tunjukkan dengan IPK. Indeks prestasi tersebut dapat ditunjukkan dengan statistika deskriptif sehingga dapat dibandingkan bagaimana prestasi belajar mahasiswa berdasarkan karakteristiknya. Selain itu hasil nilai ujian tulis juga dapat menggambarkan kemampuan mahasiswa dalam menjawab soal saat ujian. Nilai ujian tulis dapat digambarkan pula dengan statistika deskriptif berdasarkan karakteristiknya untuk mengetahui perbedaan nilai hasil ujian pada tiap karakteristik.

Karakteristik indeks prestasi berdasarkan Gambar 4.1 terlihat perbedaan nilai IPS1, IPS2 dan IPK cukup kecil. IPS1 memiliki nilai sedikit lebih tinggi dari IPS2 maupun IPK yang dapat dilihat dari nilai median maupun nilai kuantil 1 dan kuantil 3 yang dimiliki. Sedangkan dari *Inter Quartil Range* (IQR) dari ketiga nilai, IPS2 memiliki IQR yang paling lebar sedangkan IPK paling sempit. Hal ini memperlihatkan bahwa persebaran data pada IPS2 lebih besar dibanding indeks prestasi lainnya. Pada ketiga nilai indeks prestasi memusat pada angka 2 hingga 4 namun di dominasi pada sekitaran angka 3. Dapat dilihat pula pada ketiga indeks terdapat banyak *outlier* dan berada dibawah kuantil 1.



Gambar 4.1 *Boxplot* Nilai IPS1, IPS2 dan IPK.

Selanjutnya statistika deskriptif nilai indeks prestasi meliputi rata-rata, standar deviasi serta median berdasarkan masing-masing karakteristik mahasiswa.

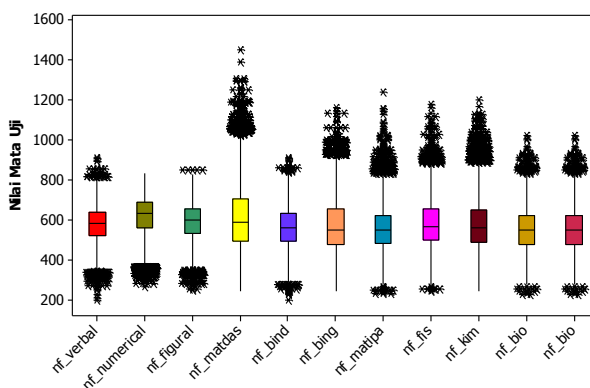
Tabel 4.1 Statistika Deskriptif Nilai IPS1, IPS2 dan IPK

Variabel	Ket.	Mean			StDev			Median		
		IPS1	IPS2	IPK	IPS1	IPS2	IPK	IPS1	IPS2	IPK
Jenis Kelamin	0	3,08	2,91	3,02	0,57	0,71	0,55	3,17	3,07	3,15
	1	3,25	3,19	3,23	0,45	0,56	0,43	3,33	3,30	3,31
Jenis Sekolah	0	3,16	3,04	3,07	0,50	0,63	0,52	3,23	3,17	3,17
	1	3,18	3,06	3,14	0,51	0,65	0,49	3,26	3,20	3,25
	2	2,93	2,77	2,88	0,63	0,78	0,63	3,02	2,96	3,03
Jurusan Sekolah	3	3,01	2,80	2,88	0,87	1,30	0,91	3,08	3,11	3,21
	0	3,17	3,04	3,13	0,51	0,65	0,50	3,25	3,19	3,23
	1	2,59	2,40	2,55	0,76	0,89	0,69	2,79	2,61	2,71
Asal Daerah	2	2,95	2,79	2,73	0,41	0,54	0,58	2,99	2,81	2,90
	0	3,25	3,13	3,22	0,45	0,57	0,43	3,32	3,25	3,30
Bidikmisi	1	3,07	2,94	3,02	0,58	0,73	0,56	3,16	3,10	3,14
	0	3,14	3,00	3,09	0,55	0,69	0,53	3,23	3,16	3,21
Pilihan Prodi	1	3,25	3,14	3,20	0,41	0,54	0,42	3,30	3,25	3,27
	1	3,12	2,98	3,09	0,56	0,70	0,54	3,20	3,14	3,22
	2	3,17	3,04	3,12	0,51	0,66	0,50	3,26	3,18	3,23
	3	3,18	3,07	3,12	0,50	0,63	0,50	3,26	3,21	3,23

*Penjelasan kategori karakteristik dijelaskan pada Bab III

Berdasarkan Tabel 4.1 dapat diketahui bahwa rata-rata indeks prestasi tertinggi yaitu pada mahasiswa perempuan namun keragaman nilai pada mahasiswa laki-laki lebih beragam. Dilihat

dari median mahasiswa laki-laki maupun perempuan memiliki median diatas 3. Selanjutnya mahasiswa yang berasal dari SMA memiliki rata-rata lebih tinggi, namun tingkat keragaman nilai mahasiswa dari SMTA lainnya memiliki keragaman lebih besar. Berdasarkan jurusan sekolah, rata-rata tertinggi berada pada kategori IPA, namun tingkat keragaman lebih besar berada pada kategori IPS. Dilihat dari median, hanya mahasiswa yang berasal dari IPA yang memiliki median diatas 3. Pada variabel asal daerah, mahasiswa dari Pulau Jawa memiliki rata-rata nilai lebih tinggi dibandingkan yang berasal dari luar Jawa. Namun tingkat keragaman nilai pada mahasiswa luar Jawa lebih besar. Dari median kedua kategori memiliki nilai diatas 3. Pada variabel bidikmisi, mahasiswa yang menerima beasiswa bidikmisi memiliki rata-rata nilai lebih tinggi dibandingkan yang tidak memperoleh bidikmisi. Namun tingkat keragaman nilai mahasiswa non-bidikmisi lebih beragam dan dari median kedua kategori memiliki nilai diatas 3. Selanjutnya pada variabel pilihan prodi, mahasiswa yang diterima pada pilihan 3 memiliki rata-rata nilai lebih tinggi dibandingkan yang diterima pada pilihan lainnya, namun tingkat keragaman nilai pada pilihan 1 lebih besar dibandingkan pilihan lainnya. Dilihat dari median pada ketiga pilihan prodi juga memiliki nilai lebih dari 3.



Gambar 4.2 *Boxplot* Nilai Mata Uji Seleksi Ujian Tulis.

Gambar 4.2 memperlihatkan *boxplot* dari kesepuluh mata uji bahwa pada masing-masing mata uji memiliki IQR yang berbeda-beda. Hal ini menunjukkan persebaran nilai pada masing-masing mata uji yang cukup beragam. Persebaran yang paling lebar terdapat pada mata uji matematika dasar dengan nilai maksimum yang paling tinggi dibanding kesembilan mata uji lainnya sedangkan pada mata uji verbal memiliki persebaran paling kecil. Pada masing-masing mata uji diketahui memiliki *outlier*.

Statistika deskriptif terhadap mata uji berdasarkan karakteristiknya ditunjukkan pada Tabel 4.2. Nilai tersebut merupakan rata-rata kesepuluh nilai mata uji jalur ujian tulis.

Tabel 4.2 Statistika Deskriptif Rata-rata Nilai Mata Uji Seleksi Ujian Tulis

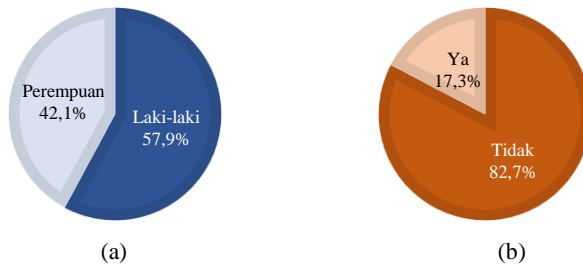
Variabel	Ket.	Mean	StDev	Min	Med	Max
Jenis	0	584,78	66,35	390,11	584,88	848,37
Kelamin	1	580,04	58,13	378,95	579,87	826,87
Jenis	0	579,52	54,84	411,80	581,47	772,27
Sekolah	1	589,76	61,57	378,95	589,27	848,37
	2	528,39	53,74	395,03	526,34	714,51
	3	568,9	89,1	450,7	573,2	751,1
Jurusan	0	584,78	62,34	378,95	584,74	848,37
Sekolah	1	510,83	44,52	409,47	509,04	640,19
	2	516,23	44,04	450,01	509,84	597,08
Asal	0	617,97	55,01	401,56	617,37	848,37
Daerah	1	552,38	52,95	378,95	551,18	831,39
Bidikmisi	0	583,68	64,08	378,95	582,60	848,37
	1	578,51	57,73	389,20	583,13	757,78
Pilihan	1	585,80	74,79	389,20	583,46	848,37
Prodi	2	586,39	58,43	390,11	589,81	774,94
	3	575,00	49,17	378,95	575,74	708,14

*Penjelasan kategori karakteristik dijelaskan pada Bab III

Berdasarkan Tabel 4.2 dapat diketahui bahwa mahasiswa laki-laki cenderung memiliki nilai ujian yang lebih tinggi dibanding mahasiswa perempuan dengan keragaman yang juga lebih besar. Hal tersebut juga dapat dilihat dari nilai median yang lebih besar dibanding nilai rata-ratanya. Mahasiswa yang berasal dari SMA memiliki rata-rata nilai ujian yang lebih tinggi di banding kategori lainnya. Kemudian mahasiswa yang berasal dari jurusan IPA memiliki rata-rata nilai ujian yang paling tinggi dengan keragaman yang juga paling tinggi. Selanjutnya mahasiswa yang berasal dari

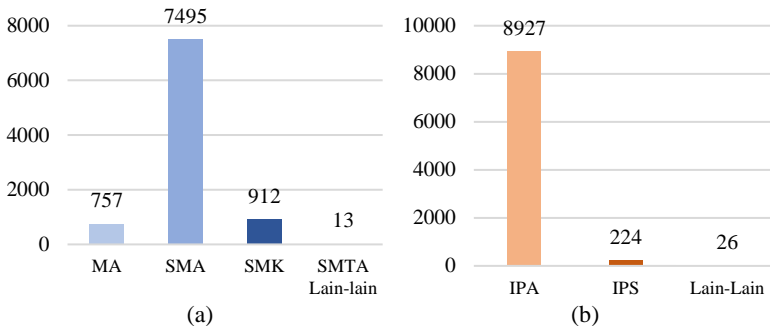
Pulau Jawa memiliki rata-rata nilai ujian yang lebih tinggi dibandingkan mahasiswa dari luar Jawa. Begitu pula dengan mahasiswa yang mendapat beasiswa bidikmisi memiliki rata-rata nilai mata uji yang lebih tinggi. Sedangkan pada pilihan prodi, mahasiswa pada pilihan ke-2 memiliki rata-rata nilai ujian lebih tinggi dibandingkan pilihan 1 maupun 3. Namun keragaman nilai pada pilihan 1 lebih tinggi dibandingkan kedua pilihan lainnya.

Statistika deskriptif perbandingan jumlah mahasiswa pada masing-masing karakteristik akan dijelaskan selanjutnya.



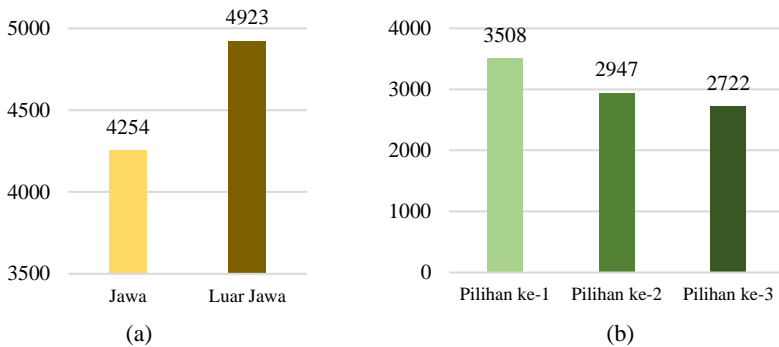
Gambar 4.3 Mahasiswa Berdasarkan (a) Jenis Kelamin dan (b) Bidikmisi.

Mahasiswa yang diterima pada jalur ujian tulis berdasarkan jenis kelamin lebih banyak mahasiswa laki-laki di banding mahasiswa perempuan (Gambar 4.3a). Persentase tersebut 57,9% atau sebanyak 5.310 mahasiswa dibanding 42,1% atau sebanyak 3.867 mahasiswa. Hal ini dapat disebabkan karena sebagian besar prodi ITS yang dalam penelitian ini digunakan sebagai acuan, merupakan prodi teknik sehingga persentase mahasiswa laki-laki lebih besar dibanding mahasiswa perempuan karena pada umumnya peminat prodi teknik adalah laki-laki. Sedangkan pada kategori bidikmisi hanya 17,3% atau sebanyak 1.590 mahasiswa mendapat beasiswa bidikmisi dan 83,7% sisanya tidak mendapat beasiswa bidikmisi (Gambar 4.3b). Hal ini menunjukkan bahwa hanya sebagian kecil mahasiswa yang berasal dari keluarga kurang mampu, selain itu kuota penerima bidikmisi yang di sediakan dari pemerintah memang tidak terlalu banyak jika dibanding semua mahasiswa yang masuk pada perguruan tinggi sehingga persentase penerima bidikmisi juga tidak begitu besar.



Gambar 4.4 Mahasiswa Berdasarkan (a) Jenis Sekolah dan (b) Jurusan Sekolah.

Berdasarkan jenis sekolah mahasiswa yang diterima jalur ujian tulis sebagian besar berasal dari SMA dengan persentase lebih dari 80% mahasiswa (Gambar 4.4a). Sedangkan mahasiswa yang berasal dari SMK maupun MA memiliki jumlah yang lebih sedikit yakni sebanyak 912 dan 757 mahasiswa dan mahasiswa yang berasal dari SMTA lain-lain hanya 13 mahasiswa. Kategori SMTA lain-lain dapat disebabkan karena tidak tercatatnya nama sekolah maupun jurusan sekolah secara jelas dan spesifik sehingga di kategorikan tersendiri. Banyaknya mahasiswa yang berasal dari SMA disebabkan sebagian besar sekolah menengah di Indonesia merupakan SMA, kemudian diikuti SMK sebagai sekolah kejuruan sedangkan sekolah MA hanya ada sebagian kecil karena sekolah ini lebih di tekankan pada pendidikan agama Islam dalam pembelajarannya. Berdasarkan jurusan sekolah, jumlah mahasiswa yang diterima pada jalur ujian tulis hampir keseluruhan berasal dari jurusan IPA. Hal tersebut dapat dilihat pada grafik Gambar 4.4b, jumlah mahasiswa yang berasal dari jurusan IPA sebanyak 8.927 mahasiswa. Sedangkan yang berasal dari jurusan IPS jauh lebih sedikit yakni 224 mahasiswa dan pada kategori lain-lain yaitu jurusan agama dan bahasa hanya 26 mahasiswa. Banyaknya mahasiswa yang berasal dari IPA dikarenakan prodi teknik memiliki dasar ilmu yaitu sains yang dipelajari mahasiswa pada jurusan IPA sewaktu di SMA, sehingga sebagian besar peminat prodi teknik memang berasal dari mahasiswa IPA.

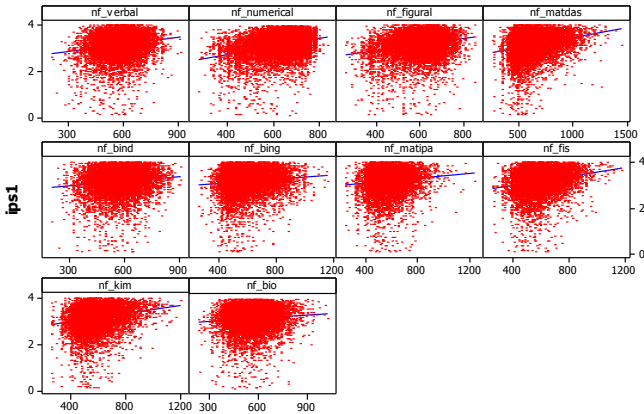


Gambar 4.5 Mahasiswa Berdasarkan (a) Asal Daerah dan (b) Pilihan Prodi.

Sebagian besar mahasiswa berasal dari luar Jawa dengan perbedaan yang tidak terlalu besar (Gambar 4.5a), hal tersebut menunjukkan bahwa prodi teknik saintek sangat diminati baik di Pulau Jawa maupun luar Jawa. Sedangkan pada kategori pilihan (Gambar 4.5b), mahasiswa yang diterima pada pilihan pertama lebih besar dibandingkan yang diterima pada pilihan kedua ataupun ketiga. Sebanyak 3.508 mahasiswa diterima pada pilihan pertama. Sedangkan pada pilihan kedua sebanyak 2.947 mahasiswa dan pada pilihan ketiga sebanyak 2.722 mahasiswa. Perbedaan jumlah pada pilihan kedua dan ketiga tidak terlalu besar. Persentase diterimanya mahasiswa pada pilihan pertama memiliki persentase paling besar, kemudian pilihan kedua dan terakhir yang ketiga. Sehingga dapat dikatakan calon mahasiswa memiliki peluang paling besar untuk diterima pada pilihan pertama.

4.1.1 Identifikasi Hubungan Indeks Prestasi dengan Nilai Mata Uji

Pola hubungan antara indeks prestasi dengan masing-masing nilai mata uji dapat diidentifikasi dengan menggunakan *scatterplot* serta nilai korelasi. *Scatterplot* akan menunjukkan hubungan kedua variabel secara visual sedangkan pada analisis korelasi akan menunjukkan hubungan dari segi pengujian. *Scatterplot* pada IPS1, IPS2 dan IPK dengan masing-masing nilai mata uji disajikan pada Gambar 4.6.

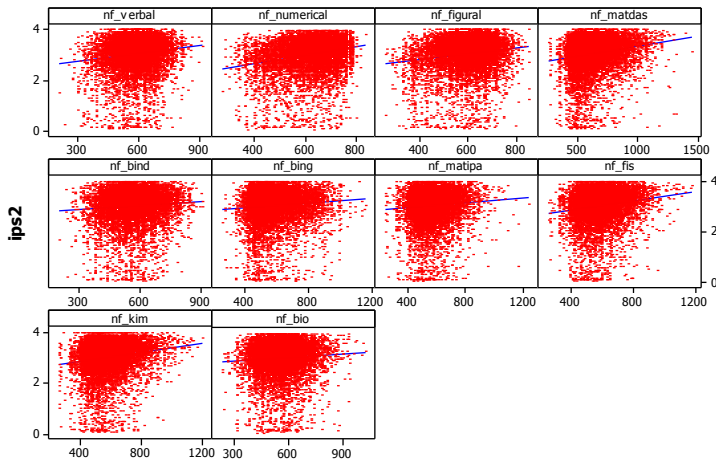


Gambar 4.6 Scatterplot IPS1 dan Nilai Mata Uji.

Gambar 4.6 menunjukkan *scatterplot* IPS1 dengan masing-masing nilai mata uji dapat dilihat secara visual bahwa hubungan IPS1 dengan nilai mata uji yang cenderung membentuk garis linier horizontal menandakan lemahnya hubungan antara IPS1 dengan nilai mata uji, sedangkan garis linier yang cenderung ke kanan atas pada beberapa hubungan IPS1 dengan nilai mata uji menandakan hubungan yang cukup kuat. Dari kesepuluh nilai mata uji, sebagian besar membentuk garis linier yang cenderung ke arah kanan atas terhadap nilai IPS1 yaitu pada nilai verbal, numerikal, figural, matematika dasar, fisika dan kimia. Selain itu, dari plot data juga dapat dilihat sebagian besar plot berkumpul pada bagian atas dan terdapat banyak *outlier* yang dapat diketahui dari jauhnya jarak plot dengan garis regresi.

Selanjutnya Gambar 4.7 menunjukkan *scatterplot* IPS2 dengan masing-masing nilai mata uji ujian tulis, dimana sebagian besar hubungan IPS2 dengan nilai mata uji juga cenderung membentuk garis linier yang cenderung ke arah kanan atas, diantaranya terhadap nilai verbal, numerikal, figural, matematika dasar, fisika dan juga kimia serta dari plot data juga terlihat sebagian besar plot berkumpul pada bagian atas namun dari bagian tengah terdapat banyak pula plot data yang menyebar hingga ke

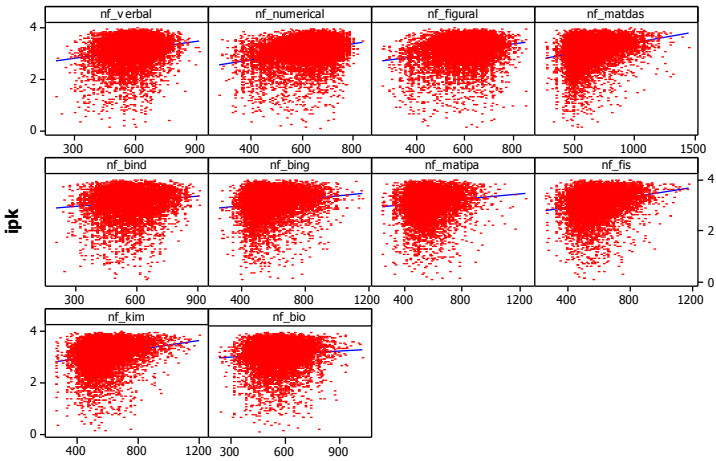
bagian bawah sehingga dapat diketahui *outlier* pada data ini lebih banyak dibanding data pada IPS1.



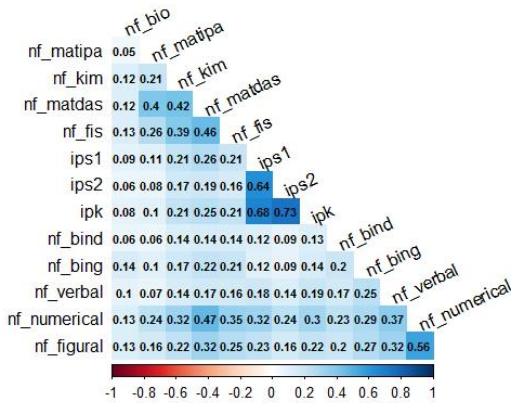
Gambar 4.7 Scatterplot IPS2 dan Nilai Mata Uji.

Gambar 4.8 menunjukkan *scatterplot* hubungan IPK dengan masing-masing nilai mata uji ujian tulis yang memperlihatkan sebagian besar hubungan IPK dengan nilai mata uji hampir sama dengan hubungan nilai mata uji terhadap IPS1 dan IPS2 sebelumnya. Yakni garis linier yang cenderung horizontal menandakan hubungan yang lemah sedangkan garis linier yang cenderung ke kanan atas menandakan hubungan yang cukup kuat. Dari plot data juga terlihat sebagian besar plot berkumpul pada bagian atas dan juga terdapat banyak *outlier* yang berada jauh dari garis linier.

Kemudian untuk mengetahui hubungan antar variabel yaitu IPS1, IPS2 dan IPK dengan masing-masing nilai mata uji menggunakan analisis korelasi akan di tunjukkan hasil korelasi *pearson* pada Gambar 4.9.



Gambar 4.8 Scatterplot IPK dan Nilai Mata Uji.



Gambar 4.9 Korelasi *Pearson* Data Rasio.

Gambar 4.9 memperlihatkan hubungan/korelasi antar variabel mata uji ujian tulis pada masing-masing variabel dimana nilai korelasi yang mendekati 1 menunjukkan hubungan yang sangat kuat yang digambarkan dengan warna biru tua sedangkan nilai korelasi yang mendekati 0 menunjukkan hubungan yang sangat lemah yang digambarkan dengan warna biru muda, selain itu dapat dilihat dari

indikator warna bahwa keseluruhan variabel tidak ada yang memiliki warna coklat atau mendekati warna coklat sehingga dapat disimpulkan bahwa hubungan antar variabel adalah positif.

Selanjutnya hubungan korelasi antar variabel *dummy* menggunakan uji independensi *chi-square* akan ditunjukkan pada Tabel 4.3.

Tabel 4.3 Hasil Korelasi *Chi-square* Data Nominal

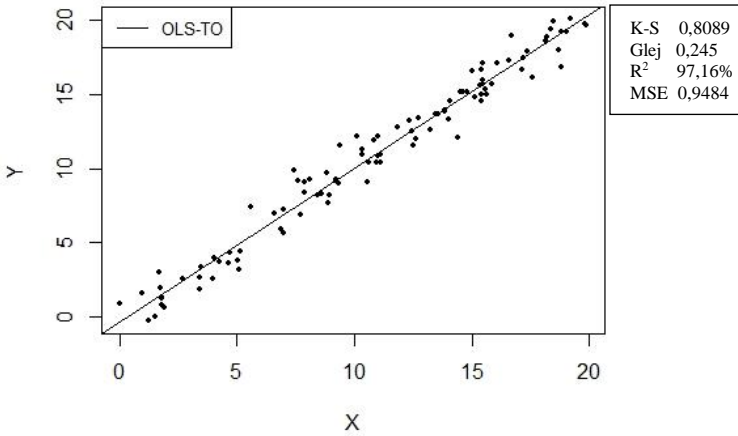
Variabel	Asal Daerah	Bidikmisi	Jenis Kelamin	Jenis Sekolah	Jurusan Sekolah	Pilihan Prodi
Asal Daerah	1	0,3748	0,0438	$< 2,2 \times 10^{-16}$	$< 2,2 \times 10^{-16}$	$1,04 \times 10^{-6}$
Bidikmisi	0,3748	1	$< 2,2 \times 10^{-16}$	$2,5 \times 10^{-16}$	0,0059	$2,37 \times 10^{-10}$
Jenis Kelamin	0,0438	$< 2,2 \times 10^{-16}$	1	$< 2,2 \times 10^{-16}$	$1,16 \times 10^{-13}$	$1,96 \times 10^{-13}$
Jenis Sekolah	$< 2,2 \times 10^{-16}$	$2,5 \times 10^{-16}$	$< 2,2 \times 10^{-16}$	1	0,0005	0,00049
Jurusan Sekolah	$< 2,2 \times 10^{-16}$	0,0059	$1,16 \times 10^{-13}$	0,0005	1	0,00049
Pilihan Prodi	$1,04 \times 10^{-6}$	$2,37 \times 10^{-10}$	$1,96 \times 10^{-13}$	0,00049	0,00049	1

Tabel 4.3 menunjukkan P_{value} analisis korelasi *chi-square* antar dua variabel *dummy*. Dapat dilihat dari nilai korelasi pada masing-masing variabel *dummy*, dengan menggunakan taraf signifikan ($\alpha=10\%$). Pada variabel asal daerah dengan bidikmisi diperoleh nilai sebesar 0,3748, dimana nilai ini lebih besar dari taraf signifikan yang berarti gagal tolak H_0 sehingga dapat disimpulkan hubungan antara variabel asal daerah dan bidikmisi tidak signifikan. Sedangkan pada variabel *dummy* lainnya diperoleh P_{value} yang kurang dari taraf signifikan yang berarti antar variabel tersebut memiliki hubungan yang signifikan.

4.2 Hasil Perbandingan Estimasi Model pada Studi Simulasi

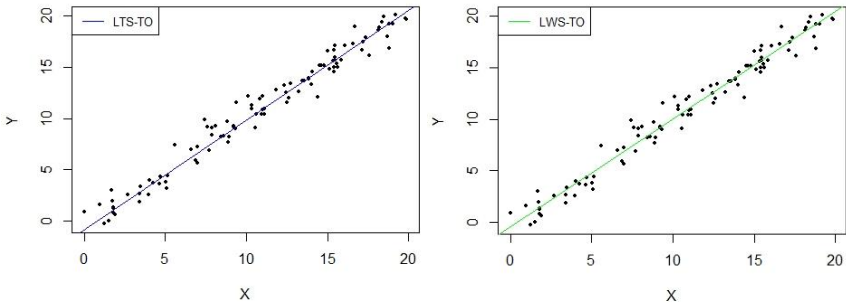
Pemodelan regresi akan menggunakan studi simulasi sebagai contoh sederhana untuk mempermudah memahami efek *outlier* pada data penelitian serta perbedaan hasil dari metode yang berbeda terhadap adanya *outlier* yang selanjutnya akan dibandingkan hasilnya. Hasil analisis pada studi simulasi akan ditunjukkan berikut ini.

Analisis pada simulasi data akan ditunjukkan pada beberapa tahapan dengan langkah-langkah yang telah di jelaskan pada Subbab 3.4. Grafik dari masing-masing metode dan hasil analisisnya ditunjukkan berikut ini.



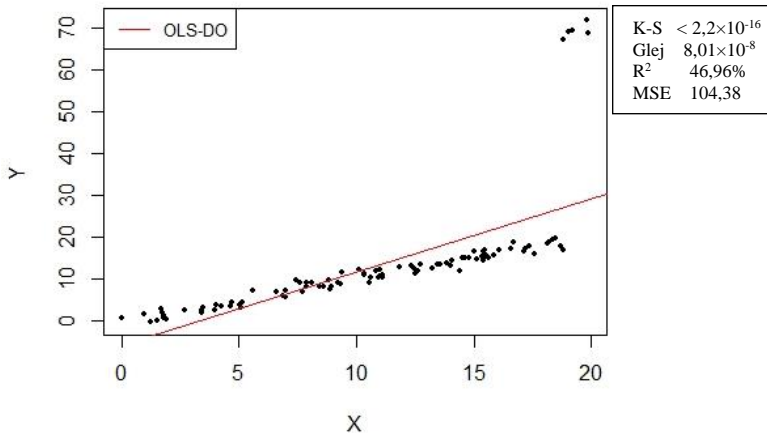
Gambar 4.10 Grafik OLS Tanpa *Outlier*.

Gambar 4.10 menunjukkan grafik regresi data simulasi dengan metode OLS ketika tidak terdapat *outlier* di dalamnya. Hasil simulasi memperlihatkan dimana plot data menyebar disekitar garis regresi dan membentuk garis linier. Pada analisis distribusi normal didapatkan hasil bahwa data berdistribusi normal dengan P_{value} sebesar 0,8089 dimana nilai ini lebih besar dari taraf signifikan 10% kemudian pengujian residual identik didapatkan hasil tidak terjadi heteroskedastisitas. Selanjutnya hasil analisis regresi diperoleh model: $\hat{y}_i = -0,35582 + 1,03591X$ dengan R^2 sebesar 97,16% yang berarti 97,16% variabilitas data dapat dijelaskan oleh variabel didalam model sedangkan sisanya hanya 2,84% yang dijelaskan oleh variabel lain diluar model. Kemudian MSE yang didapatkan sebesar 0,9484. Selanjutnya akan di gunakan metode *robust* pada data tanpa *outlier* yang di tampilkan pada Gambar 4.11.



Gambar 4.11 Grafik *Robust* LTS dan LWS pada Data Tanpa *Outlier*.

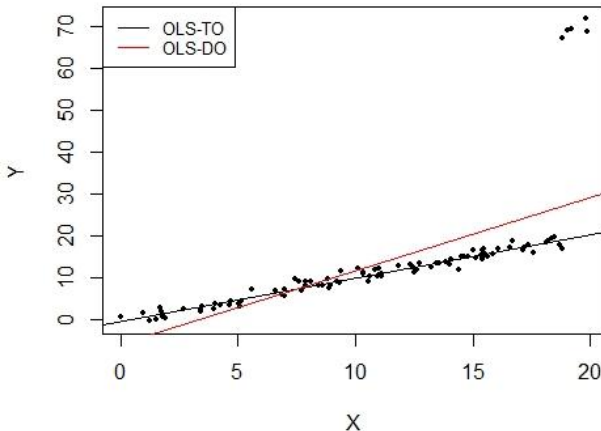
Gambar 4.11 menunjukkan hasil secara visual metode *robust* jika diterapkan pada data yang tidak mengandung *outlier*, terlihat pada kedua metode memiliki garis regresi yang linier. Selanjutnya dengan menggunakan data yang mengandung *outlier* akan ditunjukkan grafik hasil analisis dengan metode OLS pada Gambar 4.12.



Gambar 4.12 Grafik OLS dengan *Outlier*.

Gambar 4.12 menunjukkan garis regresi dengan metode OLS pada data yang telah terkontaminasi *outlier*. Terlihat garis regresi pada kedua ujung sisi menjauhi plot data dan cenderung tertarik oleh nilai *outlier*. Pengujian distribusi normal memberikan hasil bahwa data tidak berdistribusi normal dengan P_{value} yang diperoleh

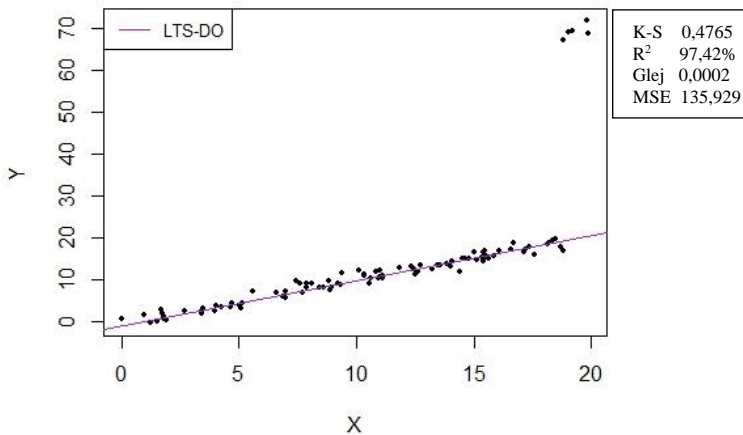
sebesar $< 2,2 \times 10^{-16}$ dimana kurang dari taraf signifikan 10% kemudian pengujian residual identik didapatkan hasil terjadi heteroskedastisitas. Hal tersebut menunjukkan bahwa adanya *outlier* dapat mempengaruhi pemenuhan asumsi pada metode OLS. Selanjutnya hasil analisis regresi didapatkan model: $\hat{y}_i = -5,5976 + 1,7488X$ dengan R^2 sebesar 46,96% yang berarti hanya 46,96% variabilitas data dapat dijelaskan oleh model sedangkan 53,04% dijelaskan oleh variabel lain diluar model. Selanjutnya grafik perbandingan metode OLS dengan data tanpa kontaminasi *outlier* dan dengan kontaminasi *outlier* ditunjukkan pada Gambar 4.13.



Gambar 4.13 Grafik OLS.

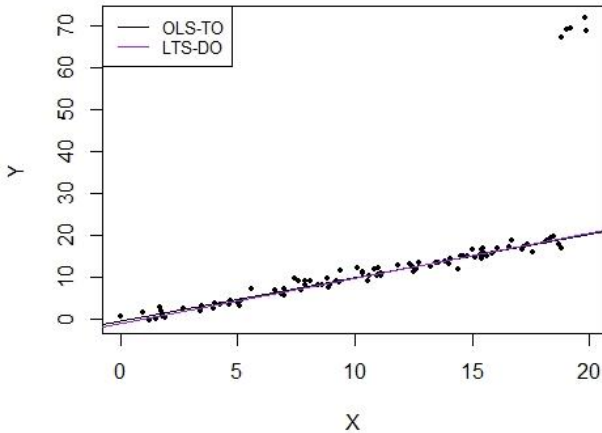
Terlihat perbandingan garis regresi OLS-Dengan *Outlier* semakin menjauhi garis linier terhadap grafik OLS-Tanpa *Outlier* dan cenderung tertarik oleh nilai *outlier* (Gambar 4.13). Hal tersebut mengindikasikan bahwa adanya *outlier* sangat berpengaruh terhadap analisis regresi yang dihasilkan dan dapat membuat model yang terbentuk tidak *fit/cocok* sehingga menimbulkan interpretasi dan kesimpulan yang salah. Untuk mengatasi hal tersebut maka diperlukan metode yang *robust* terhadap adanya *outlier*, beberapa diantaranya yaitu dengan proses

trimmed (metode *robust* LTS) dan *winsorized* (metode *robust* LWS). Selanjutnya hasil dari analisis regresi *robust* akan di tunjukkan berikut ini.



Gambar 4.14 Grafik *Robust* LTS Dengan *Outlier*.

Gambar 4.14 menunjukkan grafik metode *robust* LTS terhadap data yang telah terkontaminasi *outlier*. Pengujian distribusi normal memberikan hasil bahwa dengan metode *robust* LTS data telah berdistribusi normal yang ditunjukkan dari P_{value} yang diperoleh yaitu sebesar 0,4765 dimana nilai tersebut lebih dari taraf signifikan 10% kemudian pengujian residual identik didapatkan hasil terjadi heteroskedastisitas. Secara visual terlihat garis regresi dengan metode LTS tidak terpengaruh terhadap adanya nilai *outlier*. Hasil analisis regresi diperoleh model: $\hat{y}_i = -0,80721 + 1,0718X$ dengan R^2 sebesar 97,42% yang artinya sebesar 97,42% variabilitas data dapat dijelaskan oleh variabel di dalam model sedangkan sisanya sebesar 2,58% dijelaskan oleh variabel lain diluar model dan MSE yang diperoleh sebesar 135,929. Selanjutnya perbandingan grafik OLS-Tanpa *Outlier* dengan grafik *robust* LTS-Dengan *Outlier* akan ditunjukkan pada Gambar 4.15.

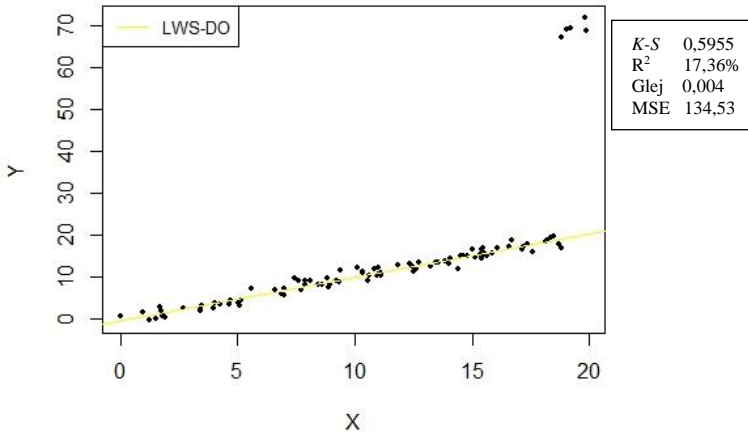


Gambar 4.15 Grafik OLS Tanpa *Outlier* vs LTS Dengan *Outlier*.

Terlihat dari perbandingan kedua grafik OLS-Tanpa *Outlier* dengan metode *robust* LTS dengan kontaminasi *outlier* menunjukkan bahwa garis regresi *robust* LTS hampir sama dengan grafik OLS (Gambar 4.15). Hal tersebut menunjukkan bahwa metode LTS hampir sama baiknya dengan metode OLS saat tidak terdapat data yang terkontaminasi *outlier* sehingga hal tersebut memperlihatkan bahwa metode LTS cukup baik di terapkan ketika terdapat data *outlier*. Selanjutnya untuk mengetahui efek dari metode *robust* lainnya yaitu metode LWS akan ditunjukkan hasil analisisnya sebagai berikut.

Gambar 4.16 menunjukkan garis regresi dengan menggunakan metode *robust* LWS terlihat berada di sekitar plot data dan membentuk garis linier. Pengujian distribusi normal memberikan hasil data telah berdistribusi normal yang dapat dilihat dari P_{value} yang didapatkan yaitu sebesar 0,5955 dimana nilai tersebut lebih besar dari taraf signifikan 10%. Selanjutnya hasil analisis regresi dengan menggunakan *trimm* sebesar 0,1 didapatkan model: $\hat{y}_i = -0,449 + 1,0463X$ dengan R^2 sebesar 17,36% yang berarti 17,36% variabilitas data dapat dijelaskan oleh variabel

dalam model sedangkan sisanya dijelaskan oleh variabel lain diluar model dan nilai MSE dari hasil analisis ini sebesar 134,53.

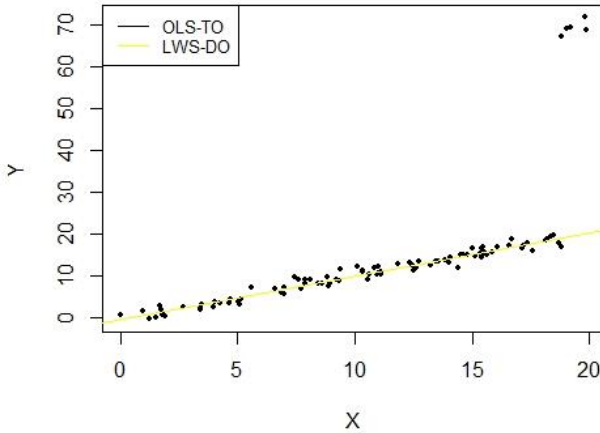


Gambar 4.16 Grafik *Robust* LWS.

Gambar 4.16 menunjukkan garis regresi dengan menggunakan metode *robust* LWS terlihat berada di sekitar plot data dan membentuk garis linier. Pengujian distribusi normal memberikan hasil data telah berdistribusi normal yang dapat dilihat dari P_{value} yang didapatkan yaitu sebesar 0,5955 dimana nilai tersebut lebih besar dari taraf signifikan 10% kemudian pengujian residual identik didapatkan hasil terjadi heteroskedastisitas. Selanjutnya hasil analisis regresi dengan menggunakan *trimm* sebesar 0,1 didapatkan model: $\hat{y}_i = -0,449 + 1,0463X$ dengan R^2 sebesar 17,36% yang berarti 17,36% variabilitas data dapat dijelaskan oleh variabel dalam model sedangkan sisanya dijelaskan oleh variabel lain diluar model dan nilai MSE dari hasil analisis ini sebesar 134,53. Selanjutnya perbandingan grafik OLS-Tanpa *Outlier* dengan *robust* LWS dengan kontaminasi *outlier* akan ditunjukkan pada Gambar 4.17.

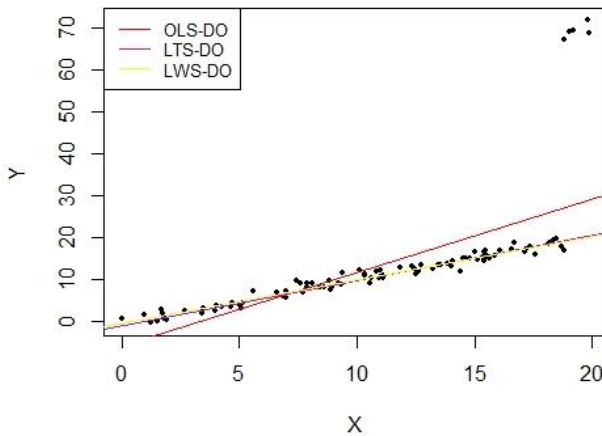
Terlihat grafik dengan metode *robust* LWS dengan kontaminasi *outlier* dibandingkan grafik OLS-Tanpa *Outlier* terlihat sama (Gambar 4.17) sehingga garis OLS tidak terlihat

secara visual. Hal tersebut membuktikan bahwa metode *robust* akan lebih efektif digunakan jika terdapat *outlier* pada data penelitian.



Gambar 4.17 Grafik OLS Tanpa *Outlier* vs LWS Dengan *Outlier*.

Selanjutnya akan ditunjukkan hasil perbandingan secara visual dari ketiga metode pada data yang terkontaminasi *outlier* pada Gambar 4.18.



Gambar 4.18 Grafik OLS, LTS dan LWS pada Data dengan *Outlier*.

Gambar 4.18 memperlihatkan grafik dari ketiga metode pada data yang terkontaminasi *outlier*, terlihat grafik LTS dan LWS cenderung membentuk garis linier disekitar plot data. Selanjutnya hasil perbandingan analisis dari ketiga metode akan di tunjukkan pada Tabel 4.4.

Tabel 4.4 Perbandingan Hasil Studi Simulasi

Metode	Tanpa <i>Outlier</i>			Dengan <i>Outlier</i>		
	OLS	LTS	LWS	OLS	LTS	LWS
<i>K-S</i>	0,8089	0,6002	0,7695	$< 2,2 \times 10^{-16}$	0,4765	0,5955
<i>Glejser</i>	0,2459	0,3654	0,289	$8,01 \times 10^{-8}$	0,00028	0,0004
β_0	-0,3558	-0,80974	-0,449	-5,5976	-0,80721	-0,449
β_1	1,0359	1,06854	1,0463	1,7488	1,0718	1,0463
R^2	0,9716	0,9812	0,9584	0,4696	0,9742	0,1736
MSE	0,9484	1,079013	1,22644	104,382	135,929	134,5348

Tabel 4.4 memperlihatkan perbandingan nilai dari hasil analisis pada masing-masing metode pada studi simulasi. Pengujian distribusi normal dengan menggunakan ketiga metode pada data yang terkontaminasi maupun yang tidak terkontaminasi *outlier* diperoleh hasil data tidak berdistribusi normal pada metode OLS ketika data terkontaminasi *outlier*. Metode *robust* yang diterapkan pada data tanpa *outlier* menghasilkan kesimpulan yang sama baik seperti metode OLS tanpa kontaminasi. Selanjutnya metode OLS yang digunakan pada data yang terkontaminasi *outlier* membuat asumsi normal dan identik tidak terpenuhi. Hal tersebut dapat dilihat dari P_{value} yang didapatkan jauh lebih kecil dari taraf signifikan 10%. Nilai intersep yang didapatkan pada metode OLS dengan kontaminasi *outlier* juga memiliki nilai yang jauh berbeda dengan metode lainnya dimana mencapai -5,5976 sedangkan pada metode lainnya berkisar pada 0 hingga -1. Nilai estimasi parameter β_1 juga terlihat paling jauh berbeda. Sedangkan dari R^2 metode LTS memiliki nilai paling tinggi dibandingkan yang lainnya yaitu sebesar 97,42% sedangkan metode LWS memiliki R^2 paling kecil diantara ketiga metode. Dari nilai MSE yang didapatkan kedua metode *robust* memiliki MSE cukup besar dan lebih tinggi dibandingkan metode OLS baik pada data kontaminasi *outlier* maupun tidak. Dari hasil analisis tersebut dapat disimpulkan bahwa adanya *outlier* pada data mengakibatkan

asumsi normal pada metode OLS tidak dapat terpenuhi sehingga model yang didapatkan tidak cocok digunakan dan menggunakan metode *robust* membuat asumsi normal terpenuhi meskipun terdapat data *outlier* didalamnya karena pada metode *robust* cenderung mengabaikan data *outlier* sehingga model yang dihasilkan lebih baik.

Selanjutnya akan ditunjukkan hasil perhitungan manual dari metode regresi *robust* dengan beberapa tahap iterasi sebagai berikut.

a. Metode *Robust* LTS

Tabel 4.5 Perhitungan Manual Studi Simulasi dengan Metode LTS

Tahap	n	h	β_0	β_1	$\sum \varepsilon_{i\text{LTS}}^2$
0	100		-5,59761	1,748823	63158,2
1	100	52	-2,46417	1,188298	1403,3
2	52	28	-3,11814	1,231026	1017,01
3	28	16	-3,62999	1,29644	528,6
⋮	⋮	⋮	⋮	⋮	⋮
20	-0,80721	10,718	135,9

b. Metode *Robust* LWS

Tabel 4.6 Perhitungan Manual Studi Simulasi dengan Metode LWS

Tahap	n	h	β_0	β_1	$\sum \varepsilon_{i\text{LWS}}^2$
0	100		-5,59761	1,748823	63158,28
1	90	10%	6,97205	1,056315	57,48709
2	81	10%	-0,82343	1,048566	0,088441
3	73	10%	-10,0137	1,042426	91,28205
⋮	⋮	⋮	⋮	⋮	⋮
20	-0,449	1,0463	134,5348

4.3 Pemodelan Faktor-faktor yang Mempengaruhi Prestasi Belajar

Pemodelan faktor-faktor yang mempengaruhi indeks prestasi mahasiswa jalur ujian tulis dilakukan berdasarkan data masing-masing prodi dengan menggunakan data *training* sehingga akan didapatkan model pada setiap prodi. Pemodelan akan dilakukan dengan menggunakan metode OLS serta metode *robust* LTS dan LWS sebagai pembandingnya. Setelah melakukan analisis OLS diperoleh hasil untuk analisis asumsi residual (IIDN) dan analisis

multikolinieritas yang akan ditunjukkan terlebih dahulu untuk mengetahui efek dari adanya *outlier* yang telah dideteksi sebelumnya.

4.3.1 Pemeriksaan Multikolinieritas

Terjadinya multikolinieritas dapat mengakibatkan analisis regresi kurang baik. Analisis multikolinieritas menggunakan kriteria nilai VIF (dijelaskan pada persamaan 2.3) dimana $VIF \geq 10$ mengindikasikan adanya masalah multikolinieritas. Hasil analisis multikolinieritas pada data Prodi Statistika akan ditunjukkan pada Tabel 4.7 sedangkan hasil analisis multikolinieritas pada prodi lainnya di tujukkan pada Lampiran 6.

Tabel 4.7 Hasil Nilai VIF

Prodi	Variabel	VIF
Statistika	X1 : Verbal	1,2250
	X2 : Numerikal	1,4314
	X3 : Figural	1,3415
	X4 : Matematika Dasar	2,0086
	X5 : B. Indonesia	1,2528
	X6 : B. Inggris	1,3600
	X7 : Matematika IPA	1,4712
	X8 : Fisika	1,4482
	X9 : Kimia	1,5256
	X10 : Biologi	1,3962
	X11 : Jenis Kelamin	1,1937
	X12 : Jenis Sekolah	1,2576
	X13 : Jurusan Sekolah	-
	X14 : Asal Daerah	1,5606
	X15 : Bidikmisi	1,1031
	X16 : Pilihan Prodi	1,4836
PTN	1,2757	

Tabel 4.7 menunjukkan bahwa VIF pada masing-masing variabel independen Prodi Statistika memiliki nilai kurang dari 10 sehingga dapat diberi kesimpulan tidak terjadi pelanggaran multikolinieritas pada data analisis. Selanjutnya dengan pengujian yang sama hasil analisis multikolinieritas pada prodi lainnya berdasarkan Lampiran 6 menunjukkan bahwa sebagian besar prodi memiliki VIF kurang dari 10 pada variabel independennya namun pada beberapa prodi terdapat VIF yang lebih besar dari 10 yang

menandakan terjadinya pelanggaran asumsi tersebut. Prodi yang terindikasi multikolinearitas dapat dilihat pada Tabel 4.8.

Tabel 4.8 Hasil Multikolinearitas

Program Studi
Manjemen Bisnis
Teknik Multimedia dan Jaringan
Desain Interior
Transportasi Laut

4.3.2 Deteksi *Outlier*

Outlier dapat berpengaruh terhadap hasil analisis sehingga memberikan kesimpulan yang salah. Oleh sebab itu, permasalahan *outlier* penting untuk ditindaklanjuti. Deteksi *outlier* pada analisis faktor-faktor yang mempengaruhi indeks prestasi mahasiswa jalur ujian tulis menggunakan DFFITS dengan kriteria yang telah dijelaskan pada Subbab 2.2.2 poin (c). Selanjutnya akan ditunjukkan hasil deteksi *outlier* data Prodi Statistika pada Tabel 4.9. Sedangkan hasil analisis DFFITS pada masing-masing prodi lainnya dapat dilihat pada Lampiran 5.

Tabel 4.9 Hasil Deteksi *Outlier*

Prodi	$2\sqrt{(p/n)}$	IPS1		IPS2		IPK	
		Obs. ke-	DFFITS	Obs. ke-	DFFITS	Obs. ke-	DFFITS
Statistika	0,78	48	-1,146	38	-1,409	18	-1,151
		50	-0,881	40	-0,855	19	-1,143
		59	0,932	71	-2,989	38	-1,178
		68	-1,009	80	-1,167	44	1,009
		89	-1,064	89	-1,619	47	-0,921
		90	-0,893	100	-0,842	71	-1,778
		108	0,827	103	0,842	89	-1,805
		116	-0,807			102	-1,01

Tabel 4.9 menunjukkan letak observasi hasil analisis regresi secara univariat nilai IPS1, IPS2 dan IPK terhadap variabel independen (nilai mata uji dan karakteristik mahasiswa) pada data Prodi Statistika. Nilai *absolute* DFFITS pada observasi tersebut melebihi kriteria batas $2\sqrt{(p/n)}$ sehingga mengindikasikan residual data pada observasi mengandung *outlier*. *Outlier* tersebut dapat berpengaruh terhadap hasil analisis regresi yang akan mempengaruhi pemenuhan asumsi klasik (IIDN) pada metode OLS sehingga dibutuhkan metode yang *robust* terhadap *outlier*.

Hasil deteksi *outlier* pada masing-masing prodi berdasarkan Lampiran 5 memperlihatkan bahwa pada masing-masing prodi terdeteksi adanya *outlier* dengan jumlah yang bervariasi. Persentase jumlah *outlier* pada masing-masing prodi akan ditunjukkan pada Tabel 4.10.

Tabel 4.10 Persentase *Outlier*

Program Studi	Persentase <i>Outlier</i>			Program Studi	Persentase <i>Outlier</i>		
	IPS1	IPS2	IPK		IPS1	IPS2	IPK
Fisika	5,05%	5,3%	5,05%	Teknik Sipil	4,5%	4,9%	4,9%
Matematika	4,06%	4,06%	4,7%	Arsitektur	6,9%	6,9%	7,9%
Statistika	6,6%	5,7%	6,6%	T. Lingkungan	4,8%	3,4%	8,2%
Kimia	4,5%	3,9%	4,65%	T. Geomatika	28,6%	21,4%	21,4%
Biologi	4,9%	3,7%	4,3%	PWK	10,2%	5,1%	8,2%
Teknik Mesin	5,9%	5,9%	6,3%	T. Geofisika	11,7%	11,7%	8,3%
Teknik Elektro	4,5%	5,01%	4,3%	Desain Interior	58,8%	58,8%	58,8%
Teknik Kimia	3,2%	5,4%	5,04%	T. Perkapalan	7,69%	5,13%	8,97%
Teknik Fisika	20%	23,3%	13,3%	T. Sist. Perkapalan	13,9%	18,6%	20,9%
Teknik Industri	4,35%	4,01%	6,35%	Teknik Kelautan	9,6%	9,6%	11,5%
T. Material dan Metalurgi	15,8%	12,3%	12,3%	Transportasi Laut	53,85%	53,8%	61,5%
Manajemen Bisnis	61,1%	77,8%	66,7%	Teknik Informatika	5%	6,55%	6,9%
T. Multimedia dan Jaringan	72,2%	66,7%	66,7%	Sistem Informasi	3,95%	5,5%	5,8%

Tabel 4.10 menunjukkan persentase *outlier* pada beberapa prodi memiliki persentase lebih dari 50% yaitu pada prodi Manajemen Bisnis, Teknik Multimedia dan Jaringan, Desain Interior dan Transportasi Laut. Persentase *outlier* yang cukup tinggi akan mempengaruhi hasil analisis regresi sederhana (OLS) menjadi tidak baik, *outlier* akan berdampak terhadap rata-rata dari data analisis sehingga membuat rata-rata dapat menjadi sangat tinggi atau sangat rendah sehingga asumsi distribusi normal tidak dapat terpenuhi.

4.3.3 Analisis Asumsi IIDN

Asumsi klasik merupakan syarat pada metode regresi OLS agar model yang terbentuk dapat dikatakan sesuai. Pelanggaran terhadap asumsi ini akan mengakibatkan model kurang *fit/cocok* terhadap data analisis sehingga perlu metode lain untuk menganalisisnya. Berikut akan di tunjukkan hasil analisis asumsi klasik pada metode OLS.

a. Pemeriksaan Asumsi Residual Identik

Pengujian asumsi residual identik menggunakan uji *Glejser* dengan kriteria pengujian yang telah dijelaskan pada Subbab 2.5 poin (a). Analisis pada data Prodi Statistika diperoleh hasil yang akan ditunjukkan pada Tabel 4.11.

H_0 : Tidak terdapat heteroskedastisitas

H_1 : Terdapat heteroskedastisitas

Statistik uji:

Tabel 4.11 Hasil Pengujian *Glejser*

Prodi	Variabel	F_{hitung}	P_{value}
Statistika	IPS1	1,49	0,1088
	IPS2	0,95	0,5151
	IPK	1,48	0,1098

Hasil pengujian heteroskedastisitas pada data Prodi Statistika yang ditunjukkan Tabel 4.11 dengan menggunakan taraf signifikan ($\alpha=10\%$) didapatkan hasil gagal tolak H_0 pada ketiga analisis variabel dependen karena F_{hitung} bernilai kurang dari $F_{(0,1;18;103)}$ (1,513) begitu pula P_{value} yang diperoleh lebih besar dari 0,1 yang artinya tidak terdapat pelanggaran asumsi identik. Selanjutnya dengan pengujian yang sama pada data prodi lainnya yang ditunjukkan pada Lampiran 6, diperoleh hasil pengujian heteroskedastisitas bahwa pada beberapa prodi tidak terjadi heteroskedastisitas atau varians data telah homogen yaitu pada prodi yang akan di tunjukkan Tabel 4.12.

Tabel 4.12 Prodi yang Tidak Heteroskedastisitas

IPS1	IPS2	IPK
Fisika	Statistika	Statistika
Statistika	Teknik Fisika	Teknik Fisika
Teknik Fisika	Manajemen Bisnis	Teknik Geomatika
T. Material & Metalurgi	T. Multimedia & Jaringan	T. Multimedia & Jaringan
Manajemen Bisnis	Teknik Geomatika	Teknik Geofisika
T. Multimedia & Jaringan	Teknik Geofisika	Desain Interior
Teknik Lingkungan	Desain Interior	Teknik Kelautan
Teknik Geomatika	Teknik Perkapalan	Transportasi Laut
Teknik Geofisika	Teknik Sistem Perkapalan	
Desain Interior	Teknik Kelautan	
Teknik Sistem Perkapalan	Transportasi Laut	
Teknik Kelautan		
Transportasi Laut		

b. Pemeriksaan Asumsi Residual Independen

Pengujian asumsi residual pada penelitian ini akan di asumsikan telah independen karena data penelitian yang digunakan merupakan data acak sehingga tidak saling bergantung. Selain itu data penelitian bukan merupakan data deret waktu sehingga tidak ada pengujian valid untuk menyimpulkan bahwa asumsi independen terpenuhi.

c. Pemeriksaan Asumsi Residual Berdistribusi Normal

Pengujian asumsi distribusi normal menggunakan uji *Kolmogorov-Smirnov* dengan hipotesis serta kriteria pengujian yang telah dijelaskan pada Subbab 2.5 poin (c). Kemudian hasil pengujian distribusi normal pada Prodi Statistika akan disajikan pada Tabel 4.13.

H_0 : Data berdistribusi normal

H_1 : Data tidak berdistribusi normal

Statistik uji:

Tabel 4.13 Hasil Pengujian *Kolmogorov-Smirnov*

Prodi	Variabel	D	P_{value}
Statistika	IPS1	0,31046	$1,222 \times 10^{-10}$
	IPS2	0,24186	$1,266 \times 10^{-6}$
	IPK	0,28211	$7,373 \times 10^{-9}$

Hasil pengujian distribusi normal pada data Prodi Statistika yang ditunjukkan Tabel 4.13 diperoleh keputusan dengan menggunakan taraf signifikan ($\alpha=10\%$) yaitu tolak H_0 karena D_{hitung} lebih besar dari D_{tabel} (0,0902) begitu pula P_{value} yang didapatkan kurang dari 0,1 sehingga dapat disimpulkan bahwa terdapat pelanggaran asumsi normal atau dapat dikatakan data tidak berdistribusi normal. Selanjutnya dengan pengujian yang sama pada masing-masing prodi lainnya diperoleh hasil hanya satu model yang berdistribusi normal yaitu pada prodi Geomatika (dapat dilihat pada Lampiran 6) dengan analisis regresi pada variabel dependen IPS2, sedangkan hasil pengujian lainnya menunjukkan data tidak berdistribusi normal.

4.3.4 Pemodelan Analisis Regresi OLS

Pemodelan dengan metode OLS dilakukan sebelum melakukan analisis metode *robust* agar dapat dibandingkan nilai R^2 dan MSE sebagai pembanding pemilihan model terbaik.

a. Pengujian Serentak

Pemodelan dengan metode OLS pada data prodi Statistika menggunakan data *training* sebanyak 122 data. Analisis secara serentak dengan perumusan hipotesis yang telah dijelaskan pada Subbab 2.4 poin (a) dan menggunakan taraf signifikan ($\alpha=10\%$) diperoleh hasil pada Tabel 4.14.

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_{18} = 0$$

$$H_1 : \text{minimal ada satu } \beta_k \neq 0; k = 1, 2, \dots, 18$$

Statistik uji:

Tabel 4.14 Hasil Pengujian Serentak Metode OLS

Prodi	Model	F_{hitung}	P_{value}
Statistika	IPS1	4,201	$1,497 \times 10^{-6}$
	IPS2	3,041	0,0002
	IPK	4,276	$1,095 \times 10^{-6}$

Berdasarkan Tabel 4.14 diperoleh nilai F_{hitung} sebesar 4,2 dan P_{value} $1,497 \times 10^{-6}$ pada model IPS1, sedangkan pada model IPS2 diperoleh F_{hitung} sebesar 3,04 dengan P_{value} 0,0002 serta pada model IPK diperoleh F_{hitung} sebesar 4,276 dengan P_{value} $1,095 \times 10^{-6}$. Dari ketiga model dapat diberi keputusan tolak H_0 karena $F_{hitung} > F_{(0,1;18;103)}(1,514)$ dan $P_{value} < \alpha$ sehingga dapat disimpulkan minimal terdapat satu variabel prediktor yang berpengaruh terhadap masing-masing model indeks prestasi mahasiswa prodi Statistika.

b. Pengujian Parsial

Selanjutnya pemodelan secara parsial untuk mengetahui variabel mana saja yang berpengaruh signifikan terhadap indeks prestasi dengan kriteria pengujian sesuai perumusan pada Subbab 2.4 poin (b) dan menggunakan taraf signifikan ($\alpha=10\%$) diperoleh hasil pada Tabel 4.15.

$$H_0 : \beta_k = \beta_{k0} = 0$$

$$H_1 : \beta_k \neq 0 ; k = 0, 1, \dots, 18$$

Statistik uji:

Tabel 4.15 Hasil Pengujian Parsial Metode OLS

variabel	IPS1				IPS2				IPK			
	Estimasi	SE	t	<i>P</i> value	Estimasi	SE	t	<i>P</i> value	Estimasi	SE	t	<i>P</i> value
β_0	1,535*	0,5434	2,824	0,006	0,35799	0,8982	0,399	0,691	1,21528*	0,681	1,785	0,077
β_1	0,00006	0,0004	0,152	0,88	0,00093	0,0006	1,473	0,144	0,00017	0,0005	0,362	0,718
β_2	0,00124*	0,0005	2,506	0,014	0,00144*	0,0008	1,752	0,083	0,0009	0,0006	1,446	0,151
β_3	-0,00009	0,0004	-0,211	0,833	-0,00055	0,0007	-0,769	0,443	-0,00043	0,0005	-0,802	0,425
β_4	0,00007	0,0002	0,312	0,756	0,00026	0,0004	0,693	0,49	0,00025	0,0003	0,861	0,391
β_5	-0,00034	0,0003	-1,149	0,253	-0,00019	0,0005	-0,376	0,708	-0,00024	0,0004	-0,648	0,519
β_6	0,00018	0,0003	0,629	0,531	-0,00039	0,0005	-0,821	0,413	0,00018	0,0004	0,517	0,607
β_7	0,00017	0,0003	0,597	0,552	0,00018	0,0005	0,372	0,711	0,00019	0,0004	0,523	0,602
β_8	0,00078*	0,0003	2,576	0,011	0,00108*	0,0005	2,164	0,033	0,00111*	0,0004	2,915	0,004
β_9	0,00114*	0,0003	4,331	0,000	0,00123*	0,0004	2,846	0,005	0,00082*	0,0003	2,499	0,014
β_{10}	-0,00039	0,0004	-1,098	0,275	-0,00035	0,0006	-0,602	0,549	0,00017	0,0004	0,375	0,709
D ₁₁ (1)	0,1076*	0,063	1,707	0,091	0,28655*	0,1042	2,751	0,007	0,29137*	0,079	3,689	0,000
D ₁₂ (1)	-0,1539	0,0959	-1,606	0,111	0,00893	0,1584	0,056	0,955	-0,16738	0,1201	-1,393	0,167
D ₁₂ (2)	0,2232	0,2479	0,901	0,37	-0,06251	0,4098	-0,153	0,879	-0,14319	0,3107	-0,461	0,646
D ₁₄ (1)	0,08079	0,0699	1,155	0,251	0,16672	0,1156	1,442	0,152	0,11326	0,0877	1,292	0,199
D ₁₅ (1)	0,042	0,0758	0,554	0,581	-0,02254	0,1253	-0,18	0,858	0,06752	0,095	0,711	0,479
D ₁₆ (2)	-0,05207	0,0676	-0,77	0,443	-0,0427	0,1118	-0,382	0,703	-0,07787	0,0848	-0,919	0,36
D ₁₆ (3)	-0,1172	0,0857	-1,368	0,174	-0,00961	0,1416	-0,068	0,946	-0,2447*	0,1074	-2,278	0,025
PTNnon	0,1382*	0,0762	1,813	0,073	0,38612*	0,126	3,064	0,003	0,02195	0,0955	0,23	0,819
R ²	42,33%				34,7%				42,77%			

*) Signifikan pada $\alpha = 10\%$

Berdasarkan Tabel 4.15 dapat diketahui variabel yang berpengaruh signifikan pada masing-masing analisis regresi univariat indeks prestasi pada Prodi Statistika yaitu nilai mata uji numerikal, fisika, kimia, jenis kelamin dan kategori PTN pada model IPS1 dan IPS2 kemudian nilai mata uji fisika, kimia, jenis kelamin dan pilihan prodi pada model IPK. Selanjutnya koefisien determinasi (R^2) yang didapatkan pada masing-masing model yaitu sebesar 42,33% pada model IPS1, 34,7% pada model IPS2 dan 42,77% pada model IPK. Nilai tersebut mengartikan bahwa keragaman data yang dapat di jelaskan oleh variabel didalam model sebesar 42,33% pada model IPS1, 34,7% pada model IPS2 dan 42,77% pada model IPK sedangkan sisanya dijelaskan oleh variabel-variabel yang tidak masuk kedalam model. Selanjutnya model pada masing-masing indeks prestasi dituliskan sebagai berikut.

$$\widehat{IPS1} = 1,535 + 0,00006X_1 + 0,00124X_2 - 0,00009X_3 + 0,00007X_4 - 0,00034X_5 \\ + 0,00018X_6 + 0,00017X_7 + 0,00078X_8 + 0,00114X_9 - 0,00039X_{10} \\ + 0,1076D_{11}(1) - 0,1539D_{12}(1) + 0,2232D_{12}(2) + 0,0808D_{14}(1) \\ + 0,042D_{15}(1) - 0,05207D_{16}(2) - 0,1172D_{16}(3) + 0,1382PTN(\text{non})$$

$$\widehat{IPS2} = 0,358 + 0,00093X_1 + 0,00144X_2 - 0,00055X_3 + 0,00026X_4 - 0,00019X_5 \\ - 0,00039X_6 + 0,00018X_7 + 0,00108X_8 + 0,00123X_9 - 0,00035X_{10} \\ + 0,2865D_{11}(1) + 0,00893D_{12}(1) - 0,0625D_{12}(2) + 0,16672D_{14}(1) \\ - 0,02254D_{15}(1) - 0,0427D_{16}(2) - 0,00961D_{16}(3) + 0,38612PTN(\text{non})$$

$$\widehat{IPK} = 1,2153 - 0,00017X_1 + 0,0009X_2 - 0,00043X_3 + 0,00025X_4 - 0,00024X_5 \\ + 0,00018X_6 + 0,00019X_7 + 0,00111X_8 + 0,00082X_9 + 0,00017X_{10} \\ + 0,2914D_{11}(1) - 0,16738D_{12}(1) - 0,14319D_{12}(2) + 0,11326D_{14}(1) \\ + 0,06752D_{15}(1) - 0,07787D_{16}(2) - 0,24469D_{16}(3) + 0,02195PTN(\text{non})$$

Selanjutnya dengan cara yang sama didapatkan hasil pada masing-masing prodi lainnya yang akan disajikan secara ringkas variabel mana saja yang berpengaruh signifikan terhadap indeks prestasi mahasiswa pada Tabel 4.16.

Tabel 4.16 (a) Ringkasan Analisis OLS Model IPS1

Program Studi	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	X ₁₂	X ₁₃	X ₁₄	X ₁₅	X ₁₆	R ²
Fisika	-	+	+	+	-	+	-	+	+	-	+	++		-	+	--	11,99%
Matematika	+	+	+	+	-	+	+	+	+	+	+	++	-	+	+	-+	26,04%
Statistika	+	+	-	+	-	+	+	+	+	-	+	-+		+	+	--	42,33%
Kimia	+	+	+	+	+	-	+	+	+	+	+	+++	-	+	+	++	24,15%
Biologi	+	+	-	+	+	+	-	+	+	+	+	--	-	+	+	++	24,36%
Teknik Mesin	+	+	+	+	+	-	-	+	+	-	+	--	--	-	+	++	17,22%
Teknik Elektro	+	+	+	+	+	-	-	+	+	+	+	+++	--	+	+	-+	22,08%
Teknik Kimia	+	-	-	+	-	-	+	+	+	-	+	+-	-	+	-	-+	11,11%
Teknik Industri	-	+	-	+	+	+	+	+	+	+	+	--	-	+	+	++	33,81%
T. Material & Metalurgi	+	+	+	+	-	-	-	+	+	-	+	+		-	-	++	49,94%
Teknik Sipil	-	+	+	+	+	-	+	+	+	+	+	+++	++	-	+	++	23,81%
Arsitektur	+	+	+	+	+	-	+	-	+	+	+	++	-	+	+	++	24%
Teknik Lingkungan	-	+	-	-	+	-	-	+	+	+	+	++		+	-	-+	19,97%
PWK	+	+	+	+	-	+	-	+	+	+	+	+++	-	+	+	++	58,67%
Teknik Geofisika	+	+	+	+	-	-	+	+	+	-	+	--	-	+	+	-+	44,81%
Teknik Perkapalan	+	+	-	+	+	-	+	+	-	-	+	+-	-	+	+	-+	43,56%
Teknik Sistem Perkapalan	+	+	-	+	-	+	+	-	+	-	+	+	-	-	+	--	59,04%
Teknik Kelautan	+	-	+	+	-	+	-	+	+	+	+	++		-	+	--	38,25%
Teknik Informatika	+	+	+	+	+	-	+	+	+	-	+	+-	+-	-	+	++	28,51%
Sistem Informasi	-	+	+	-	+	-	-	-	+	+	+	+++	+-	-	+	-+	18,28%
Teknik Fisika	+	+	-	+	+	+	+	+	-	+	+	+		-	+	++	71,56%
Manajemen Bisnis	-	-	+	-	+	-	+	+	+	+	-		-		+	--	98,39%
T. Multimedia & Jaringan	+	+	-	+	-	+	-	-	+	-	+	--				+-	89,63%
Teknik Geomatika	+	+	+	-	+	+	-	+	+	+	+	+			+	--	77,6%
Desain Interior	+	+	-	-	+	-	+	-	-	+	-				+	++	93,6%
Transportasi Laut	+	-	-	-	-	-	+	+	-	+							92,23%

Keterangan :

■	Signifikan pada $\alpha = 10\%$
■	Tidak Signifikan
■	Variabel tidak digunakan

Simbol :

+/- : Tanda Koefisien Parameter

*Jumlah tanda (+/-) menandakan banyaknya variabel *dummy*

Tabel 4.16 (b) Ringkasan Analisis OLS Model IPS2

Program Studi	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	X ₁₂	X ₁₃	X ₁₄	X ₁₅	X ₁₆	R ²
Fisika	+	+	+	+	-	-	-	+	+	-	+	++		+	+	--	9,43%
Matematika	+	+	+	+	-	+	+	+	+	+	+	++	-	+	+	++	16,94%
Statistika	+	+	-	+	-	-	+	+	+	-	+	+-		+	-	--	34,7%
Kimia	+	+	-	+	+	-	+	+	+	-	+	+++	-	+	+	++	15,46%
Biologi	+	+	-	+	+	+	-	-	-	+	+	++	-	+	+	++	13,38%
Teknik Mesin	+	+	+	-	+	-	-	+	+	-	+	++	++	-	+	++	14,75%
Teknik Elektro	+	+	+	+	+	-	-	+	-	+	+	+-	--	-	+	+-	15,31%
Teknik Kimia	+	-	-	-	-	-	+	+	+	-	+	--	-	+	-	+-	11,08%
Teknik Industri	-	+	-	+	+	-	+	+	+	-	+	--	-	-	+	++	22,15%
T. Material & Metalurgi	+	-	+	+	-	-	+	-	+	-	+	+		-	-	+-	35,68%
Teknik Sipil	-	+	+	+	+	+	+	+	+	+	+	+++	+	+	+	++	21,02%
Arsitektur	+	+	+	-	+	+	+	+	+	+	+	++	-	+	-	++	16,71%
Teknik Lingkungan	+	-	-	+	-	+	-	+	-	-	+	+-		+	-	+-	17,07%
PWK	+	+	+	-	-	+	-	-	+	+	+	+++	+	-	+	+-	38%
Teknik Geofisika	+	-	+	+	-	-	+	+	-	-	+	--	-	+	+	++	30,97%
Teknik Perkapalan	+	+	-	-	-	-	+	+	+	-	+	++	-	-	+	--	38,6%
Teknik Sistem Perkapalan	-	+	-	+	-	+	+	-	+	-	-	+	-	+	+	--	45,1%
Teknik Kelautan	+	-	-	+	+	+	-	+	+	-	+	--		+	+	--	44,58%
Teknik Informatika	+	+	+	+	+	-	+	+	+	-	+	+++	+-	-	-	++	16,6%
Sistem Informasi	+	+	+	-	+	-	-	-	+	-	+	++	++	-	+	--	18,74%
Teknik Fisika	+	+	-	-	-	-	-	+	-	-	+	-		-	-	++	77,51%
Manajemen Bisnis	-	-	+	-	+	-	+	+	+	+	-		-		+	--	95,38%
T. Multimedia & Jaringan	+	+	-	+	-	-	-	-	+	+	+	--				+-	93,09%
Teknik Geomatika	+	+	-	-	-	-	-	-	+	+	+	+			+	+-	53,9%
Desain Interior	+	-	-	-	+	+	+	-	+	+	-				+	++	95,06%
Transportasi Laut	+	-	-	-	-	-	+	+	-	+							35,77%

Keterangan :

- Signifikan pada $\alpha = 10\%$
- Tidak Signifikan
- Variabel tidak digunakan

Simbol :

- +/- : Tanda Koefisien Parameter
- *Jumlah tanda (+/-) menandakan banyaknya variabel *dummy*

Tabel 4.16 (c) Ringkasan Analisis OLS Model IPK

Program Studi	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	X ₁₂	X ₁₃	X ₁₄	X ₁₅	X ₁₆	R ²
Fisika	+	+	+	+	-	-	-	+	+	-	+	++		-	+	--	13,48%
Matematika	+	+	+	+	+	-	-	+	+	+	+	++	-	-	+	--	21,78%
Statistika	+	+	-	+	-	+	+	+	+	+	+	--		+	+	--	42,77%
Kimia	+	+	+	+	-	+	+	+	+	+	+	+++	-	+	+	++	23,73%
Biologi	+	+	-	+	+	+	-	-	+	+	+	+-	-	+	+	++	14,63%
Teknik Mesin	+	+	+	-	+	-	-	+	+	+	+	++	--	-	+	--	19,62%
Teknik Elektro	+	+	+	+	-	+	-	+	+	+	+	+-	+-	-	+	+-	26,48%
Teknik Kimia	-	-	-	+	+	+	+	+	+	-	+	--	-	+	-	+-	17,15%
Teknik Industri	+	+	-	+	+	+	+	+	+	+	+	--	-	-	+	++	38,3%
T. Material & Metalurgi	+	+	+	+	-	-	-	+	+	-	+	+		-	-	++	47,4%
Teknik Sipil	-	+	-	+	+	+	+	+	+	-	+	+++	+-	+	+	+-	27,48%
Arsitektur	+	+	+	+	+	+	+	-	+	-	+	++	-	+	-	++	23,83%
Teknik Lingkungan	+	+	-	+	-	+	-	+	+	-	+	--		-	-	+-	26,08%
PWK	+	-	+	-	+	-	-	-	+	+	+	+++	-	-	-	+-	43,21%
Teknik Geofisika	+	+	+	-	+	-	+	+	-	-	+	--	-	-	+	++	46,9%
Teknik Perkapalan	+	+	-	-	-	+	+	+	-	-	+	+-	-	-	+	--	50,23%
Teknik Sistem Perkapalan	+	-	-	+	-	+	+	+	+	-	+	+	-	-	+	--	61,17%
Teknik Kelautan	+	-	-	+	-	+	-	+	+	-	+	+-		-	+	--	47,24%
Teknik Informatika	+	+	+	+	+	-	+	+	+	-	+	+-	--	-	+	++	21,22%
Sistem Informasi	+	+	+	-	+	-	-	-	+	-	+	+++	--	-	-	+-	26,28%
Teknik Fisika	+	-	-	+	+	+	-	+	+	+	+	-		-	-	++	74,45%
Manajemen Bisnis	+	-	-	+	-	-	+	+	+	-	-		-		-	--	95,09%
T. Multimedia & Jaringan	+	+	-	+	-	-	-	-	+	+	+	--				+-	91,93%
Teknik Geomatika	+	+	-	-	+	-	-	-	-	-	+	+			+	+-	53,66%
Desain Interior	+	+	-	-	+	-	+	-	-	+	-				+	++	95,45%
Transportasi Laut	-	+	-	-	+	-	+	+	-	+							74,39%

Keterangan :

	Signifikan pada $\alpha = 10\%$
	Tidak Signifikan
	Variabel tidak digunakan

Simbol :

- +/- : Tanda Koefisien Parameter
- *Jumlah tanda (+/-) menandakan banyaknya variabel *dummy*

Tabel 4.16 menunjukkan variabel independen yang signifikan beserta nilai koefisien determinasi yang didapatkan pada masing-masing analisis regresi univariat terhadap IPS1, IPS2 dan IPK dengan metode OLS. Pada beberapa prodi terlihat tidak menggunakan variabel lengkap untuk analisis dikarenakan nilai variabel pada prodi tersebut terdiri atas satu kategori sehingga tidak dapat difaktorkan. Sebagian besar data analisis yang digunakan menggunakan data *training* namun karena keterbatasan data pada beberapa prodi maka pada prodi tertentu tidak digunakan proses *training-testing*. Prodi tersebut yaitu Teknik Fisika, Manajemen Bisnis, Teknik Multimedia dan Jaringan, Teknik Geomatika, Desain Interior dan Transportasi Laut, dimana sebagian besar merupakan prodi yang hanya ada di ITS sehingga data yang di dapatkan tidak mencukupi jika dilakukan proses *training-testing*.

Secara visual, Tabel 4.16(a) menunjukkan variabel yang signifikan terhadap IPS1 di banyak prodi yaitu nilai mata uji numerikal kemudian nilai kimia, sedangkan pada karakteristik yaitu jenis kelamin dan jurusan sekolah. Pada Tabel 4.16(b) terlihat nilai mata uji numerikal dan matematika dasar merupakan mata uji yang paling banyak signifikan terhadap IPS2 dan pada karakteristik hanya jenis kelamin yang paling banyak signifikan di banyak prodi. Selanjutnya pada Tabel 4.16(c) terlihat mata uji numerikal dan verbal yang paling banyak signifikan terhadap IPK sedangkan pada karakteristik, jenis kelamin paling mendominasi. Secara keseluruhan jenis kelamin menjadi variabel signifikan yang paling banyak disusul dengan variabel mata uji numerikal dan nilai mata uji matematika IPA menjadi variabel yang paling sedikit signifikan di banyak prodi. Tanda koefisien parameter mengartikan hubungan variabel independen terhadap variabel dependen dimana tanda (+) menunjukkan bahwa ketika nilai variabel independen bertambah satu-satuan akan meningkatkan nilai variabel dependen sebesar koefisien tersebut, sedangkan tanda (-) berarti sebaliknya.

4.3.5 Pemodelan Analisis Regresi *Robust LTS*

Pemodelan regresi *robust Least Trimmed Squares (LTS)* digunakan sebagai salah alternatif untuk menanggulangi

permasalahan *outlier* pada data yang menyebabkan asumsi klasik regresi OLS tidak terpenuhi.

a. Pengujian Serentak

Pemodelan analisis regresi *robust* LTS pada Prodi Statistika juga menggunakan data *training* seperti pada metode OLS. Pengujian signifikansi parameter pada metode ini dilakukan dengan menghapus data yang terkontaminasi *outlier* lalu dilakukan iterasi hingga diperoleh estimasi parameter yang konvergen. Dengan menggunakan taraf signifikan ($\alpha=10\%$), hasil pengujian secara serentak ditunjukkan pada Tabel 4.17.

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_{18} = 0$$

$$H_1 : \text{minimal ada satu } \beta_k \neq 0; k = 1, 2, \dots, 18$$

Statistik uji:

Tabel 4.17 Hasil Pengujian Serentak Metode LTS

Prodi	Model	F _{hitung}	P _{value}
Statistika	IPS1	5,247	$3,25 \times 10^{-8}$
	IPS2	4,194	$2,01 \times 10^{-6}$
	IPK	6,504	$4,21 \times 10^{-10}$

Berdasarkan Tabel 4.17 dapat dilihat F_{hitung} yang diperoleh pada masing-masing model indeks prestasi yaitu pada analisis regresi dengan variabel dependen IPS1 diperoleh nilai 5,247, pada IPS2 diperoleh nilai 4,194 dan pada IPK diperoleh nilai 6,504 dimana nilai tersebut lebih besar dari F_{tabel}(0,58). Selain itu P_{value} yang didapatkan kurang dari taraf signifikan 0,1 pada masing-masing model sehingga memberikan keputusan untuk tolak H₀ yang berarti minimal terdapat satu variabel prediktor yang berpengaruh terhadap indeks prestasi.

b. Pengujian Parsial

Selanjutnya hasil analisis secara parsial pada data Prodi Statistika menggunakan metode *robust* LTS dengan taraf signifikan ($\alpha=10\%$) akan ditunjukkan pada Tabel 4.18.

$$H_0 : \beta_k = \beta_{k0} = 0$$

$$H_1 : \beta_k \neq 0 ; k = 0, 1, \dots, 18$$

Statistik uji:

Tabel 4.18 Hasil Pengujian Parsial Metode LTS

Variabel	IPS1				IPS2				IPK			
	Estimasi	SE	t	<i>P</i> value	Estimasi	SE	t	<i>P</i> value	Estimasi	SE	t	<i>P</i> value
β_0	1,826*	0,474	3,852	0,000	1,399*	0,6584	2,125	0,036	1,77*	0,4710	3,759	0,000
β_1	0,00012	0,0003	0,363	0,718	0,00017	0,0004	0,404	0,687	0,00003	0,0003	0,082	0,935
β_2	0,00094*	0,0004	2,13	0,036	0,00079	0,0006	1,317	0,191	0,00102*	0,0004	2,357	0,021
β_3	0,00037	0,0004	1,021	0,310	-0,00027	0,0005	-0,546	0,587	-0,00047	0,0004	-1,212	0,229
β_4	0,00016	0,0002	0,831	0,408	0,00033	0,0003	1,283	0,202	0,00019	0,0002	0,967	0,336
β_5	-0,00024	0,0003	-0,939	0,350	0,00018	0,0003	0,521	0,604	-0,00048*	0,0003	-1,836	0,070
β_6	-0,00023	0,0002	-0,964	0,337	-0,00007	0,0003	-0,199	0,843	0,00003	0,0003	0,118	0,906
β_7	0,00005	0,0003	0,205	0,838	0,0001	0,0003	0,293	0,770	0,0003	0,0003	1,186	0,238
β_8	0,00119*	0,0003	4,518	0,000	0,00064*	0,0004	1,828	0,071	0,00083*	0,0003	3,17	0,002
β_9	0,00063*	0,0002	2,731	0,008	0,00082*	0,0003	2,778	0,007	0,00086*	0,0002	3,814	0,000
β_{10}	-0,00062*	0,0003	-2,051	0,043	-0,00051	0,0004	-1,273	0,206	-0,00012	0,0003	-0,398	0,691
D ₁₁ (1)	0,05734	0,054	1,063	0,291	0,2796*	0,0734	3,811	0,000	0,2619*	0,0541	4,844	0,000
D ₁₂ (1)	-0,06026	0,080	-0,753	0,453	0,02781	0,1108	0,251	0,802	-0,09228	0,0793	-1,163	0,248
D ₁₂ (2)	0,1458	0,2075	0,703	0,484	-0,1623	0,2750	-0,59	0,556	-0,1479	0,2031	-0,728	0,468
D ₁₄ (1)	-0,0438	0,06	-0,73	0,467	0,07122	0,0778	0,915	0,362	0,1376*	0,0619	2,224	0,029
D ₁₅ (1)	0,02114	0,0676	0,313	0,755	-0,05392	0,0863	-0,625	0,533	-0,01041	0,0630	-0,165	0,869
D ₁₆ (2)	-0,09668	0,0582	-1,661	0,100	-0,06066	0,0772	-0,786	0,434	-0,09206	0,0576	-1,598	0,113
D ₁₆ (3)	-0,09265	0,0752	-1,232	0,221	0,09619	0,0976	0,985	0,327	-0,1975*	0,0772	-2,557	0,012
PTNnon	0,175*	0,0641	2,731	0,008	0,3774*	0,0867	4,353	0,000	0,04013	0,0625	0,642	0,523
R ²	49,59%				44,02%				55,73%			

*) Signifikan pada $\alpha = 10\%$

Pada model IPS1 dapat dilihat dari Tabel 4.18 bahwa variabel yang berpengaruh signifikan terhadap IPS1 yaitu mata uji numerikal, fisika, kimia, biologi dan kategori PTN dengan keragaman data yang dapat dijelaskan oleh variabel didalam model sebesar 49,59% dan sisanya dijelaskan variabel lain diluar model. Pada model IPS2 diperoleh variabel yang berpengaruh yaitu mata uji fisika, kimia, jenis kelamin dan kategori PTN dengan keragaman data yang dapat dijelaskan oleh variabel didalam model sebesar 44,02%. Sedangkan pada model IPK, variabel yang berpengaruh yakni nilai mata uji numerikal, B. Indonesia, fisika, kimia, jenis kelamin, asal daerah dan pilihan prodi dengan keragaman data yang dapat dijelaskan oleh variabel didalam model sebesar 55,73%. Pada metode LTS data *outlier* akan di hapus sebanyak minimum persentase *outlier* yang telah ditunjukkan pada Tabel 4.10 lalu dilakukan iterasi hingga didapatkan estimasi yang konvergen. Model pada masing-masing indeks prestasi akan dituliskan berikut ini.

$$\widehat{\text{IPS1}} = 1,826 + 0,00012X_1 + 0,00094X_2 + 0,00037X_3 + 0,00016X_4 - 0,00024X_5 - 0,00023X_6 + 0,00005X_7 + 0,00119X_8 + 0,00063X_9 - 0,00062_{10} + 0,05734D_{11}(1) - 0,06026D_{12}(1) + 0,1458D_{12}(2) - 0,0438D_{14}(1) + 0,02114D_{15}(1) - 0,09668D_{16}(2) - 0,09265D_{16}(3) + 0,175\text{PTN}(\text{non})$$

$$\widehat{\text{IPS2}} = 1,399 + 0,00017X_1 + 0,00079X_2 - 0,00027X_3 + 0,00033X_4 + 0,00018X_5 - 0,00007X_6 + 0,0001X_7 + 0,00064X_8 + 0,00082X_9 - 0,00051X_{10} + 0,2796D_{11}(1) + 0,02781D_{12}(1) - 0,1623D_{12}(2) + 0,07122D_{14}(1) - 0,05392D_{15}(1) - 0,06066D_{16}(2) + 0,09619D_{16}(3) + 0,37740\text{PTN}(\text{non})$$





$$\widehat{\text{IPK}} = 1,77 + 0,00003X_1 + 0,00102X_2 - 0,00047X_3 + 0,00019X_4 - 0,00048X_5 + 0,00003X_6 + 0,0003X_7 + 0,00083X_8 + 0,00086X_9 - 0,00012X_{10} + 0,2619D_{11}(1) - 0,09228D_{12}(1) - 0,1479D_{12}(2) + 0,1376D_{14}(1) - 0,01041D_{15}(1) - 0,09206D_{16}(2) - 0,1975D_{16}(3) + 0,04013\text{PTN}(\text{non})$$

Kemudian dengan cara yang sama didapatkan hasil analisis regresi menggunakan metode *robust* LTS pada masing-masing prodi lainnya yang selanjutnya akan ditunjukkan pada Tabel 4.19.

Tabel 4.19 (a) Ringkasan Analisis LTS Model IPS1

Program Studi	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	X ₁₂	X ₁₃	X ₁₄	X ₁₅	X ₁₆	R ²
Fisika	+	+	+	+	+	+	+	-	+	-	+	++		-	+	--	12,44%
Matematika	+	+	+	+	-	-	-	+	+	+	+	++	-	-	+	++	27,7%
Statistika	+	+	+	+	-	-	+	+	+	-	+	+		-	+	--	49,59%
Kimia	+	+	+	+	+	-	+	+	+	+	+	+++	-	+	-	++	24,98%
Biologi	+	+	-	+	-	+	+	+	+	+	+	+	-	+	+	++	33,05%
Teknik Mesin	+	+	-	+	+	-	+	+	-	-	+	++	--	-	+	++	22,04%
Teknik Elektro	+	+	+	+	+	+	+	+	+	+	+	+++	--	+	+	+	26,89%
Teknik Kimia	+	-	+	+	-	-	+	+	+	+	+	+	-	+	-	++	19,88%
Teknik Industri	-	+	-	+	+	+	+	+	+	+	+	--	-	+	+	++	31,62%
T. Material & Metalurgi	+	+	+	+	-	-	+	+	+	-	+	-		-	-	+	84,5%
Teknik Sipil	-	+	+	+	+	-	+	+	+	+	+	+++	++	-	+	++	30,41%
Arsitektur	+	+	+	-	-	-	+	+	+	-	+	+	+	+	+	++	18,51%
Teknik Lingkungan	-	+	-	+	+	-	-	+	+	-	+	+		+	-	--	18,49%
PWK	+	+	+	+	-	+	+	+	+	+	+	+-	-	+	+	+	68,51%
Teknik Geofisika	-	+	+	-	-	+	+	+	-	-	+	--	-	+	+	++	69,23%
Teknik Perkapalan	+	-	-	+	+	-	-	+	-	-	+	++	+	+	+	++	49,8%
Teknik Sistem Perkapalan	-	+	-	+	-	+	+	+	+	-	+	+	-	+	+	--	75,96%
Teknik Kelautan	+	-	-	+	+	+	-	+	+	+	+	++		+	+	+	40,05%
Teknik Informatika	+	+	+	+	+	-	+	+	+	-	+	+-	+	+	+	++	35,63%
Sistem Informasi	+	+	-	-	+	-	-	-	+	-	+	++	+	-	+	+	20,72%
Teknik Fisika																	-
Manajemen Bisnis																	-
T. Multimedia & Jaringan																	-
Teknik Geomatika																	-
Desain Interior																	-
Transportasi Laut																	-

Keterangan :

	Signifikan pada $\alpha = 10\%$
	Tidak Signifikan
	Variabel tidak digunakan
	Data analisis kurang dari banyak variabel





Simbol :

+/- : Tanda Koefisien Parameter
 *Jumlah tanda (+/-) menandakan banyaknya variabel *dummy*

Tabel 4.19 (b) Ringkasan Analisis LTS Model IPS2

Program Studi	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	X ₁₂	X ₁₃	X ₁₄	X ₁₅	X ₁₆	R ²	
Fisika	+	+	-	+	-	+	-	-	+	+	+	++		-	-	--	18,17%	
Matematika	+	+	-	+	-	-	+	+	+	+	+	++	-	+	+	++	20,8%	
Statistika	+	+	-	+	+	-	+	+	+	-	+	+/-		+	-	-+	44,02%	
Kimia	+	+	+	+	+	-	-	+	+	+	+	+++	-	+	+	++	14%	
Biologi	+	+	-	+	+	+	-	+	+	+	+	--	-	+	+	++	19,25%	
Teknik Mesin	+	+	-	-	+	-	-	+	+	+	+	++	--	-	+	++	15,2%	
Teknik Elektro	+	+	+	+	+	-	-	+	+	+	+	+++	--	-	+	-+	26,16%	
Teknik Kimia	+	-	-	+	-	-	-	+	+	-	+	--	-	+	-	++	15,76%	
Teknik Industri	+	+	+	-	+	+	+	+	+	+	+	+	+/-	-	+	++	43,02%	
T. Material & Metalurgi	+	-	-	+	-	-	+	+	+	-	+	-		+	+	--	59,82%	
Teknik Sipil	+	+	-	+	+	-	+	+	+	+	+	+++	++	+	+	++	30,38%	
Arsitektur	+	+	+	-	+	+	+	+	+	-	+	-+	-	+	-	++	27,6%	
Teknik Lingkungan	+	-	-	+	-	+	-	+	+	-	+	-+		+	-	-+	29,66%	
PWK	+	+	-	+	-	+	-	-	+	+	+	+++	-	+	+	--	69,83%	
Teknik Geofisika	-	+	+	+	+	+	+	+	+	-	+	++	-	+	-	++	59,83%	
Teknik Perkapalan	+	+	-	-	-	-	+	-	+	-	+	++	-	+	+	--	56,26%	
Teknik Sistem Perkapalan	-	-	-	+	-	+	+	-	+	-	-	-	-	+	+	--	81,86%	
Teknik Kelautan	+	-	+	+	+	+	-	+	+	-	+	--		+	+	--	44,53%	
Teknik Informatika	+	+	+	+	+	-	+	+	+	+	+	+++	-+	+	+	++	23,38%	
Sistem Informasi	+	+	+	-	+	-	+	-	+	-	+	++	--	-	+	++	22,11%	
Teknik Fisika																	-	
Manajemen Bisnis																		-
T. Multimedia & Jaringan																		-
Teknik Geomatika																		-
Desain Interior																		-
Transportasi Laut																		-

Keterangan :

	Signifikan pada $\alpha = 10\%$
	Tidak Signifikan
	Variabel tidak digunakan
	Data analisis kurang dari banyak variabel





Simbol :

+/- : Tanda Koefisien Parameter
 *Jumlah tanda (+/-) menandakan banyaknya variabel *dummy*

Tabel 4.19 (c) Ringkasan Analisis LTS Model IPK

Program Studi	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	X ₁₂	X ₁₃	X ₁₄	X ₁₅	X ₁₆	R ²	
Fisika	+	+	+	+	-	+	-	+	+	+	++	++		-	+	--	17%	
Matematika	+	+	+	+	-	-	-	-	+	+	+	++	-	+	+	--	22,32%	
Statistika	+	+	-	+	-	+	+	+	+	-	+	--		+	-	--	55,73%	
Kimia	+	+	+	+	-	+	+	+	+	+	+	+++	-	+	+	++	37,16%	
Biologi	+	+	-	+	-	+	+	+	+	+	+	++	-	+	-	++	20,36%	
Teknik Mesin	+	+	-	-	-	-	+	+	+	+	+	--	--	-	+	--	20,35%	
Teknik Elektro	+	+	+	+	-	-	-	+	+	+	+	+++	--	-	+	--	34,31%	
Teknik Kimia	+	-	-	+	-	+	+	+	+	-	+	--	-	-	-	++	26,93%	
Teknik Industri	-	+	+	+	+	+	-	+	+	-	+	+-	-	+	-	++	58,59%	
T. Material & Metalurgi	-	+	+	+	+	-	-	+	+	+	+	++		-	-	--	75,3%	
Teknik Sipil	-	+	+	+	+	-	+	+	+	-	+	+++	++	+	+	++	40,15%	
Arsitektur	+	+	-	-	+	+	+	+	+	-	+	++	-	+	-	++	18,73%	
Teknik Lingkungan	-	-	+	-	+	+	-	+	+	-	+	++		+	-	--	30,73%	
PWK	+	-	+	-	-	+	-	+	+	-	+	+++	-	+	-	++	75,75%	
Teknik Geofisika	+	+	+	+	+	-	-	-	+	-	+	--	-	+	+	++	54,46%	
Teknik Perkapalan	+	-	-	-	-	+	+	+	+	-	+	++	-	-	+	++	67,64%	
Teknik Sistem Perkapalan	-	-	-	+	-	+	+	-	+	-	-	+	-	-	+	--	62,68%	
Teknik Kelautan	+	-	-	+	+	+	-	+	+	-	+	++		-	-	--	64,45%	
Teknik Informatika	+	+	+	+	+	+	+	+	+	+	+	+-	--	+	-	+-	32,32%	
Sistem Informasi	+	+	+	+	+	-	-	-	+	-	+	++	--	+	+	++	16,96%	
Teknik Fisika																	-	
Manajemen Bisnis																		-
T. Multimedia & Jaringan																		-
Teknik Geomatika																		-
Desain Interior																		-
Transportasi Laut																		-

Keterangan :

	Signifikan pada $\alpha = 10\%$
	Tidak Signifikan
	Variabel tidak digunakan
	Data analisis kurang dari banyak variabel

Simbol :

+/- : Tanda Koefisien Parameter

*Jumlah tanda (+/-) menandakan banyaknya variabel *dummy*

Tabel 4.19 menunjukkan variabel-variabel independen yang berpengaruh signifikan terhadap indeks prestasi pada analisis regresi univariat menggunakan metode *robust* LTS pada masing-masing prodi. Selain itu didapatkan nilai R^2 pada masing-masing prodi, dimana nilai R^2 pada metode *robust* LTS lebih besar jika dibandingkan nilai R^2 pada metode OLS sehingga tingkat keragaman data yang dapat dijelaskan variabel di dalam model jauh lebih baik. Pada hasil analisis secara visual terlihat warna abu-abu pada beberapa prodi yang mengartikan bahwa pada prodi tersebut tidak dapat dilakukan analisis menggunakan metode *robust* LTS dikarenakan jumlah observasi yang dibutuhkan untuk analisis metode ini kurang dari banyaknya variabel yang ada sehingga hasil analisis tidak didapatkan. Penyebabnya dikarenakan analisis perhitungan pada metode ini yaitu dengan menghilangkan observasi yang terdeteksi *outlier* sehingga ketika terdapat banyak *outlier* maka akan membuat jumlah observasi berkurang sebanyak data *outlier* tersebut. Dapat dilihat pula bahwa variabel yang signifikan lebih banyak ketika menggunakan metode *robust* LTS yang terlihat dari banyaknya warna biru tua yang mendominasi.

Tabel 4.19(a) menunjukkan nilai mata uji numerikal paling banyak signifikan terhadap IPS1 begitupula mata uji fisika. Sedangkan pada karakteristik hanya jenis kelamin yang paling berpengaruh signifikan pada banyak prodi. Hasil yang diperoleh terhadap IPS2 juga hampir sama (Tabel 4.19(b)) dimana nilai numerikal dan fisika paling banyak signifikan sedangkan pada karakteristik selain jenis kelamin, pilihan prodi, jenis sekolah serta asal daerah juga banyak yang signifikan. Kemudian pada Tabel 4.19(c) terlihat tidak hanya nilai numerikal dan fisika yang paling banyak berpengaruh signifikan namun nilai kimia juga cukup banyak yang signifikan. Sedangkan karakteristik yang berpengaruh signifikan didominasi oleh kategori jenis kelamin, lalu pilihan prodi, jenis sekolah dan jurusan sekolah. Secara keseluruhan variabel jenis kelamin dan nilai numerikal menjadi variabel yang berpengaruh signifikan pada banyak prodi sedangkan nilai mata uji figural yang paling sedikit.

4.3.6 Pemodelan Analisis Regresi *Robust* LWS

Metode alternatif lainnya yang dapat digunakan sebagai solusi untuk permasalahan *outlier* pada data analisis yaitu dengan menggunakan metode *Least Winsorized Squares* (LWS). Analisis dengan metode LWS dalam penelitian ini menggunakan nilai *trimm* sebesar 10%.

a. Pengujian Serentak

Pemodelan regresi *robust* LWS secara serentak pada data Prodi Statistika dengan menggunakan taraf signifikan ($\alpha=10\%$) didapatkan hasil yang akan ditunjukkan pada Tabel 4.20.

$$H_0: \beta_1 = \beta_2 = \dots = \beta_{18} = 0$$

$$H_1: \text{minimal ada satu } \beta_k \neq 0; k = 1, 2, \dots, 18$$

Statistik uji:

Tabel 4.20 Hasil Pengujian Serentak Metode LWS

Prodi	Model	F _{hitung}	P _{value}
	IPS1	3,7	$1,2 \times 10^{-5}$
Statistika	IPS2	3,093	0,00016
	IPK	3,864	$6,1 \times 10^{-6}$

Berdasarkan Tabel 4.20 dapat dilihat F_{hitung} yang diperoleh pada masing-masing model indeks prestasi yaitu IPS1, IPS2 dan IPK dengan metode LWS didapatkan nilai secara berturut-turut yaitu 3,7, 3,093 dan 4,758 dimana nilai tersebut lebih besar dari F_{tabel} (0,58). Selain itu juga P_{value} yang didapatkan kurang dari taraf signifikan 0,1 sehingga memberikan keputusan untuk tolak H₀ pada semua model yang berarti minimal terdapat satu variabel prediktor yang berpengaruh terhadap indeks prestasi.

b. Pengujian Parsial

Selanjutnya hasil analisis secara parsial pada data Prodi Statistika menggunakan metode *robust* LWS dengan taraf signifikan ($\alpha=10\%$) akan ditunjukkan pada Tabel 4.21.

$$H_0: \beta_k = \beta_{k0} = 0$$

$$H_1: \beta_k \neq 0 ; k = 0, 1, \dots, 18$$

Statistik uji:

Tabel 4.21 Hasil Pengujian Parsial Metode LWS

Variabel	IPS1				IPS2				IPK			
	Estimasi	SE	t	P _{value}	Estimasi	SE	t	P _{value}	Estimasi	SE	t	P _{value}
β_0	1,116	0,7684	1,453	0,149	-1,97872	1,5819	-1,251	0,214	0,54446	0,9478	0,574	0,567
β_1	0,00014	0,0005	0,308	0,759	0,00199*	0,0008	2,391	0,019	0,00035	0,0006	0,606	0,546
β_2	0,00106	0,0007	1,572	0,119	0,00052	0,0012	0,443	0,659	0,0007	0,0008	0,821	0,414
β_3	-0,00009	0,0005	-0,166	0,869	0,00011	0,0009	0,115	0,908	-0,00047	0,0007	-0,71	0,480
β_4	0,00004	0,0003	0,131	0,896	0,00013	0,0005	0,264	0,793	0,00019	0,0003	0,555	0,580
β_5	-0,00028	0,0004	-0,767	0,445	0,00055	0,0007	0,838	0,404	-0,00017	0,0004	-0,386	0,701
β_6	0,00028	0,0004	0,8	0,426	-0,00019	0,0006	-0,319	0,751	0,00036	0,0004	0,811	0,419
β_7	0,0005	0,0004	1,333	0,186	0,00088	0,0007	1,331	0,186	0,00053	0,0005	1,172	0,244
β_8	0,00074*	0,0004	1,736	0,086	0,00215*	0,0007	2,893	0,005	0,00154*	0,0005	2,86	0,005
β_9	0,00165*	0,0004	4,571	0,000	0,00188*	0,0007	2,717	0,008	0,00097*	0,0004	2,225	0,028
β_{10}	-0,0004	0,0004	-0,929	0,355	-0,00057	0,0008	-0,747	0,457	0,00046	0,0005	0,86	0,392
D ₁₁ (1)	0,1112	0,0673	1,652	0,102	0,35542*	0,1223	2,905	0,004	0,29216*	0,0827	3,532	0,001
D ₁₂ (1)	-0,2314*	0,1093	-2,118	0,037	-0,04124	0,1649	-0,25	0,803	-0,25556*	0,1527	-1,673	0,097
D ₁₂ (2)	0,09806	0,2603	0,377	0,707	-0,2256	0,4221	-0,534	0,594	-0,28084	0,3273	-0,858	0,393
D ₁₄ (1)	0,06373	0,0752	0,847	0,399	0,31964*	0,1512	2,114	0,037	0,11722	0,0938	1,25	0,214
D ₁₅ (1)	0,01527	0,0791	0,193	0,847	0,02413	0,1417	0,17	0,865	0,05395	0,1039	0,519	0,605
D ₁₆ (2)	-0,06282	0,0721	-0,872	0,385	0,01019*	0,1314	0,078	0,938	-0,09713	0,0903	-1,076	0,285
D ₁₆ (3)	-0,1053	0,0889	-1,184	0,239	0,0994	0,1597	0,622	0,535	-0,23822*	0,1107	-2,151	0,034
PTNnon	0,1605*	0,0803	1,999	0,048	0,36303*	0,1393	2,606	0,011	0,01001	0,1019	0,098	0,922
R ²	39,27%				35,09%				40,31%			

*) Signifikan pada $\alpha = 10\%$

Tabel 4.21 menunjukkan hasil analisis regresi dengan metode LWS diperoleh variabel yang berpengaruh terhadap IPS1 yaitu nilai mata uji fisika, kimia, jenis sekolah dan kategori PTN dengan keragaman data yang dapat dijelaskan variabel dalam model sebesar 39,27% dan sisanya dijelaskan variabel lain diluar model. Pada model IPS2 diperoleh variabel signifikan nilai mata uji numerikal, fisika, kimia, jenis kelamin, asal daerah dan pilihan prodi dengan tingkat keragaman yang dapat di jelaskan oleh variabel dalam model sebesar 35,09% dan sisanya dijelaskan variabel lain di luar model. Serta pada variabel IPK didapatkan nilai mata uji fisika, kimia, jenis kelamin, jenis sekolah dan pilihan prodi dengan tingkat keragaman data yang dapat dijelaskan oleh variabel dalam model sebesar 40,31% sedangkan sisanya dijelaskan variabel lain yang tidak masuk dalam model. Pada metode ini data *outlier* akan diganti dengan nilai terdekat yang tidak *outlier* sebanyak persentase *trimm* yang digunakan dan dilakukan iterasi hingga didapatkan estimasi parameter yang konvergen. Berikut model pada masing-masing indeks prestasi.

$$\widehat{IPS1} = 1,116+0,00014X_1 + 0,00106X_2 - 0,00009X_3 + 0,00004X_4 - 0,00028X_5 \\ + 0,00028X_6 + 0,0005X_7 + 0,00074X_8 + 0,00165X_9 - 0,0004X_{10} \\ + 0,1112D_{11}(1) - 0,2314D_{12}(1) + 0,09806D_{12}(2) + 0,06373D_{14}(1) \\ + 0,01527D_{15}(1) - 0,06282D_{16}(2) - 0,1053D_{16}(3) + 0,1605PTN(\text{non})$$

$$\widehat{IPS2} = -1,9787+0,00199X_1 + 0,00052X_2 + 0,00011X_3 + 0,00013X_4 + 0,00055X_5 \\ - 0,00019X_6 + 0,00088X_7 + 0,00215X_8 + 0,00188X_9 - 0,00057X_{10} \\ + 0,35542D_{11}(1) - 0,04124D_{12}(1) - 0,2256D_{12}(2) + 0,31964_{14}(1) \\ + 0,02413D_{15}(1) + 0,01019D_{16}(2) + 0,0994D_{16}(3) + 0,363PTN(\text{non})$$

$$\widehat{IPK} = 0,54446+0,00035X_1 + 0,0007X_2 - 0,00047X_3 + 0,00019X_4 - 0,00017X_5 \\ + 0,00036X_6 + 0,00053X_7 + 0,00154X_8 + 0,00097X_9 + 0,00046X_{10} \\ + 0,29216D_{11}(1) - 0,25556D_{12}(1) - 0,28084D_{12}(2) + 0,11722D_{14}(1) \\ + 0,05395D_{15}(1) - 0,09713D_{16}(2) - 0,23822D_{16}(3) + 0,1001PTN(\text{non})$$

Selanjutnya dengan cara yang sama didapatkan variabel yang signifikan pada masing-masing prodi yang akan ditunjukkan ringkasannya pada Tabel 4.22.

Tabel 4.22 (a) Ringkasan Analisis LWS Model IPS1

Program Studi	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	X ₁₂	X ₁₃	X ₁₄	X ₁₅	X ₁₆	R ²
Fisika	+	+	+	+	-	-	-	-	+	-	+	++		+	+	--	12,85%
Matematika	+	+	+	+	-	+	+	+	+	+	+	++	-	+	+	-+	23,33%
Statistika	+	+	-	+	-	+	+	+	+	-	+	++		+	+	--	39,27%
Kimia	+	+	+	+	+	-	+	+	+	+	+	+++	-	+	-	++	21,12%
Biologi	+	+	-	+	+	+	-	+	+	+	+	++	-	+	+	++	21,46%
Teknik Mesin	+	+	+	+	+	-	-	+	+	-	+	--	--	-	+	++	16,39%
Teknik Elektro	+	+	+	+	+	+	-	+	+	+	+	+++	--	+	+	-+	21,99%
Teknik Kimia	+	-	-	+	-	-	+	+	+	+	+	++	-	+	-	++	10,2%
Teknik Industri	-	+	+	+	+	+	+	+	+	+	+	--	-	+	+	++	19,45%
T. Material & Metalurgi	+	+	+	+	-	-	+	+	+	-	+	-		+	+	++	47,25%
Teknik Sipil	-	+	+	+	+	-	+	+	+	+	+	+++	++	-	+	++	23,79%
Arsitektur	+	+	+	+	+	-	+	+	+	+	+	++	-	+	+	++	12,11%
Teknik Lingkungan	-	+	-	+	+	-	-	+	+	+	+	++		+	-	++	18,74%
PWK	+	+	+	-	+	-	+	+	+	+	+	+++	-	+	+	++	58,67%
Teknik Geofisika	-	+	+	-	-	+	+	+	+	-	+	+-	-	+	+	-+	52,3%
Teknik Perkapalan	+	+	-	+	+	-	+	+	-	-	+	++	-	+	-	-+	42,66%
Teknik Sistem Perkapalan	+	+	-	+	-	+	+	+	+	-	+	+	-	+	+	--	42,89%
Teknik Kelautan	+	-	+	+	-	+	-	+	+	+	+	++		+	+	--	26,64%
Teknik Informatika	+	+	+	+	+	-	+	+	+	-	+	++	++	-	+	++	26,09%
Sistem Informasi	+	+	+	-	+	+	-	-	+	+	+	++	++	-	+	-+	18,79%
Teknik Fisika	+	+	-	+	+	+	+	+	-	+	+	-		-	+	++	80,42%
Manajemen Bisnis	+	-	-	+	-	+	+	-	+	-	+		+		-	++	42,6%
T. Multimedia & Jaringan	+	+	-	+	+	+	+	+	-	-	-	--				--	94,07%
Teknik Geomatika	+	+	+	-	+	+	+	+	+	+	+	-			+	-+	77,16%
Desain Interior	+	+	-	+	+	-	+	-	-	+	-				+	++	47,28%
Transportasi Laut	-	-	-	-	+	-	+	+	-	+							95,71%

Keterangan :

	Signifikan pada $\alpha = 10\%$
	Tidak Signifikan
	Variabel tidak digunakan

Simbol :

+/- : Tanda Koefisien Parameter

*Jumlah tanda (+/-) menandakan banyaknya variabel *dummy*

Tabel 4.22 (b) Ringkasan Analisis LWS Model IPS2

Program Studi	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	X ₁₂	X ₁₃	X ₁₄	X ₁₅	X ₁₆	R ²
Fisika	+	+	+	+	-	-	-	+	+	-	+	++		+	+	--	10,01%
Matematika	+	+	-	+	-	+	+	+	+	+	+	++	-	+	+	++	16,45%
Statistika	+	+	+	+	+	-	+	+	+	-	+	--		+	+	++	35,09%
Kimia	+	+	-	+	+	-	+	+	+	-	+	+++	-	+	+	++	11,95%
Biologi	+	+	-	+	+	+	-	+	-	+	+	-+	-	+	+	++	12,97%
Teknik Mesin	+	+	+	+	+	-	-	+	-	+	+	-+	--	-	+	-+	14,63%
Teknik Elektro	+	+	+	+	+	+	-	+	+	+	+	+++	--	+	+	-+	15,29%
Teknik Kimia	+	-	-	-	-	-	+	+	+	-	+	--	-	+	-	-+	10,91%
Teknik Industri	-	+	-	+	+	-	+	+	+	-	+	--	-	-	+	++	12,53%
T. Material & Metalurgi	-	+	+	+	-	-	+	+	+	+	+	+		-	+	--	36,74%
Teknik Sipil	-	+	-	+	+	+	+	+	+	+	+	+++	++	+	+	-+	21,02%
Arsitektur	+	+	+	-	-	+	+	+	+	-	+	++	-	-	-	++	13,54%
Teknik Lingkungan	+	-	-	+	-	+	-	+	-	-	+	-+		+	-	--	21,76%
PWK	+	+	+	-	-	+	-	-	+	+	+	+++	+	-	+	+-	38%
Teknik Geofisika	-	+	+	-	-	-	-	+	-	-	+	--	-	+	+	++	36,67%
Teknik Perkapalan	+	+	-	-	-	-	+	+	-	-	+	++	-	-	+	--	37,78%
Teknik Sistem Perkapalan	-	+	-	+	-	+	+	-	+	-	-	+	-	+	+	--	67,22%
Teknik Kelautan	+	-	+	+	-	+	-	+	+	-	+	--		+	+	--	50,78%
Teknik Informatika	+	+	+	+	+	-	+	+	-	-	+	+++	+-	-	+	++	15,58%
Sistem Informasi	+	+	+	-	+	-	-	-	+	-	+	++	+-	-	+	--	17,89%
Teknik Fisika	+	+	-	+	+	+	-	+	+	-	+	-		-	-	+-	79,24%
Manajemen Bisnis	-	-	+	-	-	+	+	+	+	+	+		-		-	--	68,71%
T. Multimedia & Jaringan	+	+	-	+	+	-	+	+	-	-	-	--				--	87,84%
Teknik Geomatika	+	+	-	-	+	+	-	+	+	-	+	-			+	+-	89,94%
Desain Interior	+	-	-	-	+	+	+	-	-	+	-				+	+-	63,39%
Transportasi Laut	+	-	+	-	-	-	+	+	-	+							45,35%

Keterangan :

	Signifikan pada $\alpha = 10\%$
	Tidak Signifikan
	Variabel tidak digunakan

Simbol :

+/- : Tanda Koefisien Parameter

*Jumlah tanda (+/-) menandakan banyaknya variabel *dummy*

Tabel 4.22 (c) Ringkasan Analisis LWS Model IPK

Program Studi	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀	X ₁₁	X ₁₂	X ₁₃	X ₁₄	X ₁₅	X ₁₆	R ²
Fisika	+	+	+	+	+	-	-	+	+	-	+	++		+	+	--	14,12%
Matematika	+	+	+	+	-	-	-	+	+	+	+	+-	-	+	+	--	19,22%
Statistika	+	+	-	+	-	+	+	+	+	+	+	++		+	+	--	40,31%
Kimia	+	+	+	+	-	+	+	+	+	+	+	+++	-	+	-	++	19,43%
Biologi	+	+	-	+	+	+	-	+	+	+	+	+-	-	+	+	++	14,22%
Teknik Mesin	+	+	+	+	+	-	-	+	+	+	+	+-	--	-	+	--	19,89%
Teknik Elektro	+	+	+	+	-	+	-	+	+	+	+	---	+-	-	+	+-	26,28%
Teknik Kimia	-	-	-	+	+	+	+	+	+	-	+	--	-	+	-	+-	17,32%
Teknik Industri	+	+	+	+	+	-	+	+	+	+	+	--	-	-	+	++	25,8%
T. Material & Metalurgi	-	+	+	+	+	-	-	+	+	+	+	-		-	+	++	32,91%
Teknik Sipil	-	+	-	+	+	+	+	+	+	-	+	+++	+-	+	+	+-	27,48%
Arsitektur	+	+	+	+	+	+	+	+	+	-	+	++	-	+	-	++	16,76%
Teknik Lingkungan	+	+	-	+	-	+	-	+	+	-	+	--		-	-	+-	26,27%
PWK	+	-	+	-	+	-	-	-	+	+	+	+++	-	-	-	++	44,13%
Teknik Geofisika	+	+	+	-	+	-	+	+	+	-	+	--	-	-	+	++	37,58%
Teknik Perkapalan	+	+	-	-	-	+	+	+	-	-	+	+-	-	-	+	--	49,76%
Teknik Sistem Perkapalan	+	-	-	+	-	+	+	+	+	-	+	+	-	-	+	--	39,86%
Teknik Kelautan	+	-	+	+	-	+	-	+	+	-	+	++		-	+	++	42,42%
Teknik Informatika	+	+	+	+	+	+	+	+	+	-	+	+-	--	-	+	++	19,62%
Sistem Informasi	+	+	+	-	+	-	-	-	+	-	+	++	--	-	-	+-	21,08%
Teknik Fisika	+	-	-	+	+	+	-	+	+	+	+	-		-	-	+-	77,42%
Manajemen Bisnis	+	+	-	+	-	+	+	-	-	-	+		+		-	++	61,87%
T. Multimedia & Jaringan	+	+	-	+	+	-	+	+	-	-	-	--				--	85,52%
Teknik Geomatika	+	+	-	-	+	-	+	+	-	-	+	-			+	+-	65,87%
Desain Interior	+	-	-	+	+	-	+	-	-	+	-				-	++	13,76%
Transportasi Laut	+	-	-	-	+	+	+	+	+	+							72,1%

Keterangan :

	Signifikan pada $\alpha = 10\%$
	Tidak Signifikan
	Variabel tidak digunakan

Simbol :

+/- : Tanda Koefisien Parameter

*Jumlah tanda (+/-) menandakan banyaknya variabel *dummy*

Tabel 4.22 menunjukkan variabel-variabel yang berpengaruh signifikan pada setiap prodi hasil analisis metode *robust* LWS dan juga nilai R^2 pada masing-masing model. Jika dibandingkan dengan metode OLS dan *robust* LTS, hasil analisis pada metode ini tidak terlalu baik. Hal tersebut dapat dilihat dari banyaknya variabel yang berpengaruh signifikan lebih sedikit serta nilai R^2 yang cenderung tidak jauh berbeda dengan metode OLS. Secara visual terlihat pada Tabel 4.22(a) bahwa variabel yang signifikan pada banyak prodi terhadap IPS1 yaitu jenis kelamin dan nilai mata uji fisika. Pada Tabel 4.22(b) terlihat hanya jenis kelamin yang signifikan pada banyak prodi terhadap IPS2. Kemudian pada Tabel 4.22(c) terlihat selain jenis kelamin, jenis sekolah dan nilai mata uji numerikal paling banyak signifikan pada banyak prodi terhadap IPK. Secara keseluruhan nilai mata uji numerikal dan fisika menjadi variabel yang signifikan pada banyak prodi sedangkan pada karakteristiknya jenis kelamin dan jenis sekolah cukup mendominasi. Selain itu, variabel yang paling sedikit signifikan yaitu nilai mata uji figural dan matematika IPA.

4.3.7 Pemilihan Model Terbaik

Selanjutnya untuk mengetahui model terbaik dari ketiga metode yakni regresi OLS, *robust* LTS dan *robust* LWS akan digunakan nilai R^2 serta MSE dari ketiga metode pada masing-masing model sesuai prodi. Pemilihan kriteria model terbaik yaitu dengan melihat nilai R^2 yang lebih besar dan MSE yang lebih kecil, selain itu juga dengan melihat jumlah variabel signifikan yang paling banyak diantara ketiga metode. Hasil analisis pemilihan model terbaik akan disajikan pada Tabel 4.23.

Tabel 4.23 Hasil Pemilihan Model Terbaik

IPSI						IPSI2						IPK					
OLS		LTS		LWS		OLS		LTS		LWS		OLS		LTS		LWS	
R ²	MSE	R ²	MSE	R ²	MSE	R ²	MSE	R ²	MSE	R ²	MSE	R ²	MSE	R ²	MSE	R ²	MSE
11,99%	0,200	12,44%	0,205	12,85%	0,198	9,43%	0,295	18,17%	0,320	10,01%	0,293	13,48%	0,185	17%	0,191	14,12%	0,184
26,04%	0,185	27,7%	0,196	23,33%	0,191	16,94%	0,251	20,8%	0,257	16,45%	0,250	21,78%	0,166	22,32%	0,173	19,22%	0,171
42,33%	0,09	49,59%	0,103	39,27%	0,096	34,7%	0,247	44,02%	0,27	35,09%	0,25	42,77%	0,142	55,73%	0,155	40,31%	0,144
24,15%	0,202	24,98%	0,214	21,12%	0,207	15,46%	0,288	14%	0,317	11,95%	0,296	23,73%	0,188	37,16%	0,204	19,43%	0,194
24,36%	0,190	33,05%	0,230	21,46%	0,193	13,38%	0,253	19,25%	0,270	12,97%	0,254	14,63%	0,199	20,36%	0,215	14,22%	0,200
17,22%	0,289	22,04%	0,313	16,39%	0,289	14,75%	0,417	15,2%	0,449	14,63%	0,417	19,62%	0,273	20,35%	0,317	19,89%	0,272
22,08%	0,187	26,89%	0,194	21,99%	0,188	15,31%	0,396	26,16%	0,420	15,29%	0,396	26,48%	0,188	34,31%	0,202	26,28%	0,189
11,11%	0,175	19,88%	0,182	10,2%	0,177	11,08%	0,449	15,76%	0,497	10,91%	0,449	17,15%	0,206	26,93%	0,231	17,32%	0,207
33,81%	0,115	31,62%	0,116	19,45%	0,131	22,15%	0,236	43,02%	0,272	12,53%	0,253	38,3%	0,131	58,59%	0,146	25,8%	0,146
49,94%	0,189	84,5%	0,271	47,25%	0,194	35,68%	0,543	59,82%	0,784	36,74%	0,523	47,4%	0,163	75,3%	0,249	32,91%	0,175
23,81%	0,219	30,41%	0,224	23,79%	0,212	21,02%	0,367	30,38%	0,393	21,02%	0,367	27,48%	0,212	40,15%	0,226	27,48%	0,220
24%	0,144	18,51%	0,174	12,11%	0,155	16,71%	0,368	27,6%	0,412	13,54%	0,370	23,83%	0,182	18,73%	0,214	16,76%	0,189
19,97%	0,127	18,49%	0,133	18,74%	0,129	17,07%	0,185	29,66%	0,194	21,76%	0,183	26,08%	0,142	30,73%	0,200	26,27%	0,142
58,67%	0,099	68,51%	0,140	58,67%	0,099	38%	0,286	69,83%	0,350	38%	0,286	43,21%	0,172	75,75%	0,254	44,13%	0,171
44,81%	0,225	69,23%	0,282	52,3%	0,188	30,97%	0,283	59,83%	0,406	36,67%	0,274	46,9%	0,146	54,46%	0,164	37,58%	0,155
43,56%	0,185	49,8%	0,306	42,66%	0,188	38,6%	0,320	56,26%	0,414	37,78%	0,324	50,23%	0,126	67,64%	0,187	49,76%	0,128
59,04%	0,102	75,96%	0,209	42,89%	0,011	45,1%	0,173	81,86%	0,364	67,22%	0,010	61,17%	0,153	62,68%	0,285	39,86%	0,019
38,25%	0,160	40,05%	0,194	26,64%	0,169	44,58%	0,230	44,53%	0,300	50,78%	0,227	47,24%	0,201	64,45%	0,324	42,42%	0,193
28,51%	0,285	35,63%	0,310	26,09%	0,292	16,6%	0,522	23,38%	0,596	15,58%	0,530	21,22%	0,305	32,32%	0,342	19,62%	0,310
18,28%	0,467	20,72%	0,528	18,79%	0,474	18,74%	0,517	22,11%	0,623	17,89%	0,526	26,28%	0,190	16,96%	0,217	21,08%	0,201
71,56%	0,051	-	-	80,42%	0,031	77,51%	0,155	-	-	79,24%	0,111	74,45%	0,061	-	-	77,42%	0,042
98,39%	0,091	-	-	42,64%	1,130	95,38%	0,064	-	-	67,45%	0,310	95,09%	0,098	-	-	61,04%	0,563
89,63%	0,804	-	-	94,0%	0,384	93,09%	0,465	-	-	87,84%	0,596	91,93%	0,537	-	-	85,52%	0,719
77,6%	0,134	-	-	77,16%	0,123	53,9%	1,120	-	-	89,94%	0,248	53,66%	0,500	-	-	65,87%	0,301
93,6%	0,123	-	-	47,28%	0,682	95,06%	0,151	-	-	63,39%	0,835	95,45%	0,129	-	-	13,76%	2,081
92,23%	0,053	-	-	95,71%	0,026	35,77%	0,342	-	-	45,35%	0,291	74,39%	0,056	-	-	72,1%	0,058

Selanjutnya Tabel 4.23 memperlihatkan bahwa dari perbandingan nilai R^2 yang lebih tinggi didominasi pada metode *robust* LTS, sedangkan MSE terkecil cenderung pada metode *robust* LWS. Sehingga metode *robust* dapat dikatakan memberi hasil yang lebih baik dibandingkan metode OLS. Namun dilihat dari banyaknya variabel yang signifikan diperoleh jumlah jauh lebih banyak pada metode *robust* LTS sehingga model terbaik yang dipilih pada penelitian ini yaitu metode *robust* LTS.

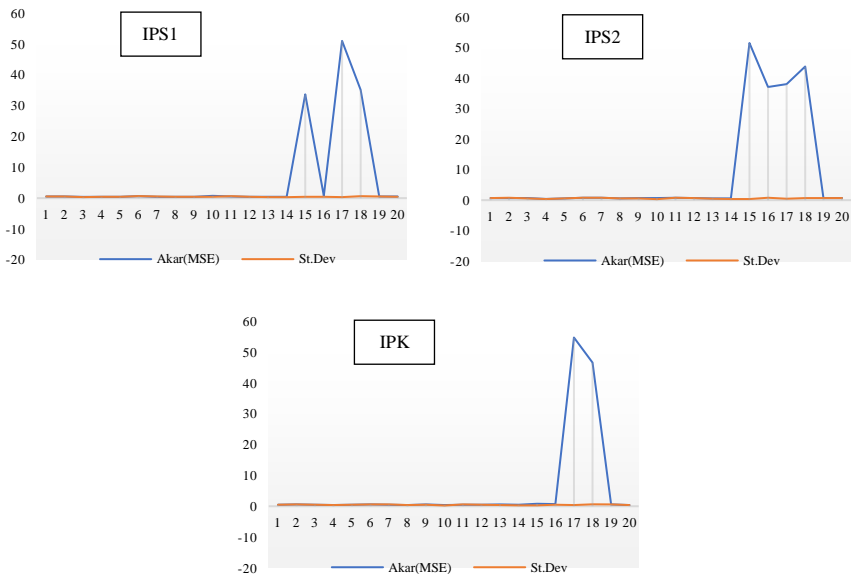
4.3.8 Pemodelan Data *Testing*

Berdasarkan hasil pemilihan model terbaik diperoleh metode *robust* LTS yang memiliki nilai R^2 jauh lebih tinggi dan variabel signifikan yang lebih banyak. Sehingga pada pemodelan data *testing* akan digunakan model dari metode *robust* LTS sebagai analisisnya. Pemodelan dengan menggunakan data *testing* dimaksudkan untuk mengevaluasi kebaikan model yang diperoleh dari hasil pemodelan data *training*. *Mean Square Error* akan digunakan sebagai penilaian kebaikan pada analisis masing-masing prodi yang akan disajikan pada Tabel 4.24.

Tabel 4.24 Pemodelan Data *Testing*

Program Studi	IPS1		IPS2		IPK	
	\sqrt{MSE}	St.Dev	\sqrt{MSE}	St.Dev	\sqrt{MSE}	St.Dev
Fisika	0,5023	0,5251	0,6888	0,6798	0,5281	0,5426
Matematika	0,5101	0,5470	0,7344	0,7384	0,5751	0,5958
Statistika	0,4139	0,3823	0,7151	0,6279	0,5437	0,4806
Kimia	0,4173	0,3990	0,4047	0,4121	0,3966	0,3958
Biologi	0,4759	0,4660	0,5756	0,5602	0,4621	0,4670
Teknik Mesin	0,6638	0,6654	0,7952	0,7940	0,6354	0,6028
Teknik Elektro	0,4865	0,5047	0,7773	0,7368	0,5396	0,5567
Teknik Kimia	0,4417	0,4226	0,5794	0,5271	0,3967	0,3690
Teknik Industri	0,4417	0,4179	0,7046	0,5715	0,6062	0,5205
T. Material & Metalurgi	0,7507	0,4116	0,6586	0,3232	0,4040	0,2442
Teknik Sipil	0,5851	0,6173	0,7846	0,7768	0,5381	0,5808
Arsitektur	0,4493	0,4276	0,7191	0,6499	0,4854	0,4582
Teknik Lingkungan	0,4522	0,3780	0,5903	0,5000	0,5620	0,4064
PWK	0,4595	0,3179	0,5427	0,3733	0,4776	0,3319
Teknik Geofisika	33,651	0,4916	51,490	0,3454	0,7591	0,2898
Teknik Perkapalan	0,7926	0,4318	37,066	0,7430	0,6675	0,4581
T. Sistem Perkapalan	50,971	0,3282	38,009	0,4787	54,693	0,3636
Teknik Kelautan	35,077	0,6281	43,824	0,6700	46,582	0,6362
Teknik Informatika	0,5397	0,5611	0,7263	0,7288	0,5931	0,6079
Sistem Informasi	0,4970	0,4912	0,7246	0,6896	0,4052	0,4071

Berdasarkan hasil pemodelan data *testing* pada 20 prodi dengan menggunakan metode *robust* LTS, dapat dilihat pada Tabel 4.24 bahwa nilai \sqrt{MSE} yang didapatkan dari masing-masing model prodi hasil data prediksi relatif lebih kecil dibanding standar deviasi data aktual dan hanya pada beberapa prodi terlihat nilai \sqrt{MSE} jauh lebih tinggi. Hal tersebut menandakan kesalahan hasil analisis pada data observasi dengan prediksi cukup kecil sehingga model *robust* LTS dapat dikatakan sudah baik. Untuk memudahkan analisis akan ditunjukkan grafik perbandingan \sqrt{MSE} dengan standar deviasi secara visual pada Gambar 4.18.



Gambar 4.19 Perbandingan akar MSE dengan Standar Deviasi.

Gambar 4.19 menunjukkan bahwa pada beberapa prodi terlihat nilai \sqrt{MSE} jauh lebih tinggi dibandingkan nilai standar deviasinya, namun jika dibandingkan dari keseluruhan prodi, sebagian besar terlihat memiliki standar deviasi yang lebih tinggi dibanding \sqrt{MSE} .

(Halaman ini sengaja dikosongkan)

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil analisis terhadap permasalahan yang telah dijelaskan pada rumusan masalah maka kesimpulan yang didapatkan sebagai berikut.

1. Karakteristik indeks prestasi mahasiswa jalur ujian tulis yaitu nilai IPS1 lebih tinggi dibanding IPS2 dan IPK, nilai mata uji yang memiliki persebaran data paling lebar adalah matematika dasar. Sebagian besar mahasiswa jalur ujian tulis didominasi oleh mahasiswa laki-laki, dari sekolah SMA, jurusan IPA, berasal dari luar Jawa, non bidikmisi dan diterima pada pilihan pertama
2. Perbandingan estimasi model dengan studi simulasi diperoleh hasil metode *robust* LTS lebih baik dibandingkan OLS pada data yang terkontaminasi *outlier* dengan R^2 sebesar 97,42%
3. Pemodelan regresi didapatkan hasil pada seluruh prodi terdapat *outlier*, hanya pada 4 prodi terjadi multikolinearitas, sebagian besar terjadi heteroskedastisitas dan hanya terdapat satu prodi yang berdistribusi normal. Pemilihan model terbaik diperoleh hasil metode LTS merupakan metode terbaik dengan R^2 yang jauh lebih tinggi dan jumlah variabel signifikan yang lebih banyak dibanding metode lainnya. Selanjutnya didapatkan variabel signifikan pada masing-masing metode yaitu:
 - a. Hasil metode OLS menunjukkan bahwa variabel yang signifikan adalah jenis kelamin, nilai mata uji numerikal dan jurusan sekolah
 - b. Hasil metode *robust* LTS menunjukkan bahwa variabel yang signifikan adalah jenis kelamin, nilai mata uji numerikal, nilai mata uji fisika dan pilihan prodi
 - c. Hasil metode *robust* LWS menunjukkan bahwa variabel yang signifikan adalah jenis kelamin, nilai mata uji numerikal, nilai mata uji fisika dan jenis sekolah.

5.2 Saran

Berdasarkan hasil analisis yang telah dilakukan serta meninjau batasan masalah yang digunakan pada penelitian ini, saran yang dapat diberikan adalah sebagai berikut.

1. Kepada pihak sumber data dan perguruan tinggi yang ikut melaksanakan seleksi jalur ujian tulis agar selalu melaporkan perkembangan nilai mahasiswanya untuk meminimalkan terjadinya ketidaklengkapan data.
2. Kepada pihak-pihak terkait ataupun pemerintah diharapkan dari penelitian ini dapat dijadikan masukan dalam penetapan kebijakan mengenai penerimaan mahasiswa baru jalur tes tulis di Perguruan Tinggi Negeri tahun-tahun mendatang.
3. Pada penelitian selanjutnya diharapkan dari penelitian ini dapat dijadikan referensi. Selain itu, karena keterbatasan peneliti pada penelitian ini diharapkan penelitian selanjutnya dapat dikembangkan baik dari segi analisis, metode, maupun data sehingga memberikan informasi yang lebih baik.

DAFTAR PUSTAKA

- Agresti, A. (2002). *Categorical Data Analysis, 2nd Edition*. New York: John Wiley & Sons, Inc.
- Anscombe, F.J. (1973). Graphs in Statistical Analysis. *The American Statistician*, 27(1), 17-21.
- Antari, I.S.W. (2016). *Pemodelan Indeks Prestasi Kumulatif Tahap Persiapan Mahasiswa ITS Jalur Seleksi Nasional Masuk Perguruan Tinggi Negeri Menggunakan Pendekatan Regresi Kuantil*. Tugas Akhir FMIPA. Surabaya: ITS.
- Atkinson, A. & Riani, M. (2000). *Robust Diagnostic Regression Analysis*. New York: Springer Science+Business Media.
- Belsley, D.A., Kuh, E. & Welsch, R.E. (1980). *Regression diagnostics: identifying influential data and sources of collinearity*. *Wiley Series in Probability and Mathematical Statistics*. New York: John Wiley and Sons Inc.
- Dalyono, M. (2007). *Psikologi Pendidikan*. Jakarta: Rineka Cipta.
- Draper, N.R. & Smith, H. (1998). *Applied Regression Analysis, 3rd Edition*. New York: John Wiley & Sons, Inc.
- Gad, A.M. & Qura, M.E. (2016). Regression Estimation in the Presence of Outliers: A Comparative Study. *International Journal of Probability and Statistics*, 5(3), 65-72.
- Gujarati, D.N. & Porter, D.C. (2009). *Basic Econometrics, 5th Edition*. New York: McGraw-Hill Companies, Inc.
- Herawati, N., Nisa, K. dan Setiawan, E. (2011). Analisis Ketegaran Regresi Robust Terhadap Letak Pencilan: Studi Perbandingan. *Bulletin of Mathematics*, 3(1), 49-60.
- ITS. (2018). *Fakta dan Sejarah Institut Teknologi Sepuluh Nopember Surabaya*. Diakses pada 19 Februari 2019 dari <https://www.its.ac.id>.
- Julina. (2015). *Analisis Pengaruh Daerah Asal dan Jenis Sekolah Terhadap Daya Saing Belajar Mahasiswa Fakultas Ekonomi dan Ilmu Sosial UIN Suska Riau*. Penelitian Dosen Fakultas Ekonomi dan Ilmu Sosial. Riau: UIN Suska.

- Kashid, D.N. & Jadhav, N.H. (2014). Robust Winsorized Shrinkage Estimators for Linear Regression Model. *Journal of Modern Applied Statistical Methods*, 13(2), Article 7, 131-150.
- Kompas. (2019). *Jadwal Pegumuman SBMPTN dipercepat 165.831 peserta dinyatakan lulus*. Diakses pada 03 Desember 2018 dari <https://edukasi.kompas.com>.
- Liu, X. Q. & Gao, F. (2011). Linearized Ridge Regression Estimator in Linear Regression. *Communications in Statistics—Theory and Methods*, 40(12), 2182-2192.
- Maharani, I.F., Satyahadewi, N. Dan Kusnandar D. (2014). Metode Ordinary Least Squares dan Least Trimmed Squares dalam Mengestimasi Parameter Regresi ketika Terdapat Outlier. *Buletin Ilmiah Mat. Stat. Dan Terapannya (Bimaster)*, 3(3), 163-168.
- Matdoan, M.Y. (2017). *Perbandingan Estimasi Regresi Quantil dengan Regresi Robust Least Trimmed Square (LTS) (Studi Kasus : Faktor-faktor yang mempengaruhi penyebaran penyakit malaria di Indonesia)*. Tesis FMIPA. Surabaya: Pascasarjana Statistika ITS.
- Muslimin, Z.I. (2012). Prestasi Belajar Mahasiswa Ditinjau dari Jalur Penerimaan Mahasiswa Baru, Asal Sekolah dan Skor Tes Potensi Akademik. *Jurnal Penelitian Psikologi*, 3(1), 381-393.
- Montgomery, D.C., Peck, E.A. & Vining, G.G. (2012). *Introduction to Linear Regression Analysis, 5th Edition*. New York: John Wiley & Sons.
- Mutan, O.C. (2009). A Monte Carlo Comparison of Regression Estimators When The Error Distribution is Long Tailed Symmetric. *Journal of Modern Applied Statistical Methods*, 8(1), 161-172.
- Myers, R.H. (1990). *Classical and Modern Regression with Applications*. Boston: PWS-KENT Publishing Company.
- Ristekdikti. (2019). *Laman Pendaftaran Bidikmisi*. Diakses pada 20 Februari dari <https://bidikmisi.belmawa.ristekdikti.go.id>.

- Rousseeuw, P.J. (1984). Least Median of Squares Regression. *Journal of the American Statistical Association*, 79(388), 871-880.
- Rousseeuw, P.J. & Leroy, A.M. (1987). *Robust Regression and Outlier Detection*. Wiley Series in Probability and Mathematical Statistics. New York: John Wiley and Sons Inc.
- Rousseeuw, P.J. & Van Driessen, K. (1998). *Computing LTS Regression for Large Data Sets*. Technical reports. Belgium: University of Antwerpen.
- SBMPTN. (2019). *Seleksi Bersama Masuk Perguruan Tinggi Negeri (SBMPTN) Tahun 2019*. Diakses pada 24 Januari 2019 dari <https://sbmptn.ltmt.ac.id>.
- Shodiqin, A., Aini, A.N. dan Rubowo, M.R. (2018). Perbandingan Dua Metode Regresi Robust yakni Metode Least Trimmed Squares (LTS) dengan Metode Estimator-MM (Studi Kasus Data Ujian Tulis Masuk Terhadap Hasil IPK Mahasiswa UPGRIS). *Jurnal Ilmiah Teknosains*, 4(1), 35-42.
- Soemartini. (2007). *Pencilan (Outlier)*. Bandung: Universitas Padjajaran.
- Sucahyo, A.Y. (2014). Perbandingan Prestasi Belajar Akademik antara Mahasiswa Bidikmisi dan Mahasiswa Non Bidikmisi. *Jurnal Pendidikan Olahraga dan Kesehatan*, 2(1), 266-268.
- Walpole, R.E., Myers, R.H., Myers, S.L. & Ye, K. (2012). *Probability & Statistics for Engineers & Scientists, 9th Edition*. USA: Pearson Education, Inc.
- Weisberg, S. (2014). *Applied Linear Regression, 4th Edition*. New Jersey: John Wiley & Sons, Inc.
- Winkel, W.S. (2009). *Psikologi Pengajaran*. Yogyakarta: Media Abadi.
- Yale, C., & Forsythe, A.B. (1976). Winsorized Regression. *Technometrics*, 18(3), 291-300.
- Zahroh, F. (2008). *Pengaruh Gender Terhadap Motivasi Memilih Sekolah dan Prestasi Belajar Siswa di SMK PGRI*

Turen Malang. Skripsi Program Studi Pendidikan Matematika. Malang: Universitas Negeri Malang.

LAMPIRAN

Lampiran 1. Data Penelitian

No.	Verbal	numerikal	figural	matdas	bind	bing
1	537,9877	663,3856	500,1139	671,5196	674,4444	450,3592
2	621,8908	647,5904	613,8095	479,8328	615,0098	508,5945
3	575,2780	442,2527	566,4363	495,8067	615,0098	625,0652
4	510,0200	458,0479	500,1139	463,8589	579,3490	406,6827
5	510,0200	497,5359	500,1139	399,9632	508,0274	639,6240
6	705,7938	544,9215	452,7408	527,7545	615,0098	683,3005
7	696,4713	687,0784	774,8781	783,3369	543,6882	639,6240
8	426,1170	521,2287	661,1826	543,7284	531,8013	537,7122
9	547,3103	758,1569	519,0632	591,6501	615,0098	464,9180
10	593,9231	584,4096	490,6393	575,6762	555,5751	654,1829
11	593,9231	537,0239	632,7587	447,8850	424,8189	581,3887
12	491,3749	402,7647	481,1647	479,8328	508,0274	523,1533
13	631,2133	639,6928	452,7408	495,8067	615,0098	494,0357
14	621,8908	702,8736	500,1139	543,7284	519,9143	508,5945
15	528,6651	631,7952	623,2841	479,8328	603,1228	508,5945
16	668,5036	560,7168	509,5886	511,7806	555,5751	610,5064
17	631,2133	726,5664	613,8095	463,8589	436,7058	697,8594
18	640,5359	560,7168	547,4871	447,8850	626,8967	537,7122
19	537,9877	521,2287	443,2662	447,8850	448,5927	435,8003
20	435,4395	694,9760	708,5557	543,7284	555,5751	319,3296
21	491,3749	560,7168	500,1139	527,7545	555,5751	566,8298
22	621,8908	639,6928	642,2333	559,7023	508,0274	741,5359
23	565,9554	702,8736	651,7080	559,7023	555,5751	450,3592
24	575,2780	497,5359	424,3169	495,8067	615,0098	494,0357
25	575,2780	544,9215	661,1826	479,8328	496,1405	552,2710
26	547,3103	631,7952	566,4363	543,7284	496,1405	683,3005
27	426,1170	726,5664	462,2154	655,5457	650,6705	435,8003
28	519,3426	568,6144	490,6393	431,9111	448,5927	654,1829
29	621,8908	679,1808	433,7916	591,6501	531,8013	494,0357
30	528,6651	758,1569	613,8095	639,5718	448,5927	581,3887
31	556,6328	537,0239	566,4363	543,7284	615,0098	566,8298
32	584,6005	371,1742	462,2154	383,9893	377,2712	552,2710
33	528,6651	434,3551	661,1826	463,8589	615,0098	464,9180
34	631,2133	568,6144	594,8602	511,7806	603,1228	537,7122
⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮
9177	695,1681	560,1790	577,3882	453,9323	669,1078	483,9361

Lampiran 1. Data Penelitian (lanjutan)

No.	Matipa	fisika	kimia	biologi	JK	Jenis Sekolah	Jurusan Sekolah
1	473,7897	583,7609	674,8944	466,3871	p	SMA	IPA
2	697,4317	531,8684	532,8272	652,5677	p	SMA	IPA
3	569,6362	324,2983	461,7936	523,6734	l	SMA	IPA
4	649,5084	441,0565	518,6205	495,0303	l	MA	IPA
5	505,7385	531,8684	547,0339	452,0655	p	SMA	IPA
6	409,8919	518,8952	404,9667	537,9950	p	SMA	IPA
7	713,4061	661,5997	589,6541	609,6029	p	SMA	IPA
8	585,6107	454,0296	731,7213	523,6734	p	SMA	IPA
9	681,4573	492,9490	532,8272	452,0655	l	SMA	IPA
10	409,8919	454,0296	575,4474	666,8892	p	SMA	IPA
11	585,6107	467,0027	447,5869	695,5324	p	SMA	IPA
12	537,6874	402,1371	404,9667	595,2813	p	SMA	IPA
13	441,8408	609,7071	618,0675	452,0655	p	SMA	IPA
14	457,8152	583,7609	433,3801	437,7439	l	SMA	IPA
15	441,8408	544,8415	717,5146	409,1008	p	SMA	IPA
16	457,8152	583,7609	532,8272	452,0655	l	MA	IPA
17	825,2271	583,7609	461,7936	452,0655	p	SMA	IPA
18	585,6107	544,8415	476,0003	580,9598	l	SMA	IPA
19	441,8408	609,7071	476,0003	580,9598	l	SMK	IPA
20	793,2783	596,7340	561,2406	466,3871	l	SMA	IPA
21	425,8664	428,0833	547,0339	652,5677	l	MA	IPA
22	457,8152	518,8952	547,0339	623,9245	l	SMA	IPA
23	665,4828	518,8952	461,7936	566,6382	p	SMA	IPA
24	537,6874	596,7340	632,2742	552,3166	l	SMA	IPA
25	409,8919	609,7071	518,6205	638,2461	p	SMA	IPA
26	585,6107	389,1639	561,2406	681,2108	l	SMA	IPA
27	425,8664	778,3578	774,3415	638,2461	l	SMA	IPA
28	489,7641	531,8684	660,6877	681,2108	p	SMA	IPA
29	649,5084	402,1371	603,8608	366,1360	p	SMA	IPA
30	585,6107	544,8415	419,1734	308,8497	p	SMA	IPA
31	585,6107	700,5191	575,4474	466,3871	p	MA	IPA
32	745,3550	415,1102	433,3801	466,3871	l	SMK	IPA
33	441,8408	428,0833	362,3465	595,2813	l	SMA	IPA
34	505,7385	583,7609	404,9667	695,5324	l	SMA	IPA
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
9177	516,4843	483,4746	487,9256	515,3685	l	SMA	IPA

Lampiran 1. Data Penelitian (lanjutan)

No.	Asal Daerah	Bidikmisi	Pilihan Prodi	IPS1	IPS2	IPK
1	luar jawa	0	3	2,97	3,29	3,34
2	luar jawa	0	1	2,64	3,21	2,95
3	luar jawa	0	1	2,58	2,5	3,06
4	luar jawa	0	3	2,79	3,03	3,22
5	luar jawa	0	1	3,58	3,45	3,45
6	luar jawa	0	3	3,23	3,05	2,87
7	luar jawa	0	1	3,79	3,64	3,77
8	luar jawa	1	1	2,79	3,22	2,91
9	luar jawa	0	2	3	3,03	2,89
10	luar jawa	0	2	3,16	3,32	3,29
11	luar jawa	0	2	2,98	2,79	2,96
12	luar jawa	0	2	3,53	3,48	3,51
13	luar jawa	0	2	3,28	3,13	3,28
14	luar jawa	1	3	3,32	3,08	3,24
15	luar jawa	0	2	3,68	3,81	3,6
16	luar jawa	0	3	2,71	2,76	2,95
17	luar jawa	1	1	3,58	3,62	3,49
18	luar jawa	0	3	2,73	2,87	2,88
19	luar jawa	0	2	2,82	2,68	2,32
20	luar jawa	0	1	3,18	3,05	3,29
21	luar jawa	0	2	3,63	3,18	3,4
22	luar jawa	0	1	3,25	3,32	2,98
23	luar jawa	0	1	3,58	3,45	3,55
24	luar jawa	0	2	2,79	3,58	3,18
25	luar jawa	0	3	3,34	3,58	3,45
26	luar jawa	0	1	2,76	3,03	3,1
27	luar jawa	0	2	3,53	3,5	3,06
28	luar jawa	0	2	3,61	3,37	3,58
29	luar jawa	0	2	3,32	3,37	3,4
30	luar jawa	1	2	3,32	3,08	3,05
31	luar jawa	1	1	3,68	3,81	3,84
32	luar jawa	0	1	2,26	2,85	2,7
33	luar jawa	1	3	3,1	3,15	2,92
34	luar jawa	1	2	2,55	3,1	2,88
:	:	:	:	:	:	:
:	:	:	:	:	:	:
9177	luar jawa	0	1	3,1	3,04	3,7

Lampiran 2. Syntax R

```

#KORELASI DATA RASIO
korelasi=read.csv('E:/korelasi1.csv', sep=";", header=TRUE)
head(korelasi)
library(corrplot)
library(RcolorBrewer)
kor<-cor(korelasi)
corrplot(kor, method = "color",
          type = "lower", order = "hclust", number.cex = 0.7,
          addCoef.col = "black", tl.col = "black", tl.srt = 25,
          diag = FALSE)
#KORELASI DATA NOMINAL
korelasi=read.csv('E:/korelasi2.csv', sep=";", header=TRUE)
head(korelasi)
.Table <- xtabs(~variabel(i)+variabel(j), data = korelasi)
.Table
chisq.test(.Table, correct = FALSE)
fisher.test(.Table) #jika terdapat data dengan E<5
fisher.test(.Table, simulate.p.value=TRUE)

#DATA TRAINING DAN TESTING
datastat=read.csv('E:/DATA ANALISIS TA/statistika.csv', sep=";",
header=TRUE)
n=nrow(datastat)
n.train=floor(n*(2/3))
n.test=n-n.train
train.index=sample(seq(n),size=n.train)
data.train=datastat[train.index,]
data.test=datastat[-train.index,]
write.csv(data.train, "E:/DATA ANALISIS TA/DATA TEST-
TRAIN/stattraining.csv")
write.csv(data.test, "E:/DATA ANALISIS TA/DATA TEST-
TRAIN/stattesting.csv")

library(MASS)
library(robust)
library(robustbase)
library(fda)
#Data menyesuaikan masing-masing prodi
stattrain=read.csv('E:/DATA ANALISIS TA/DATA TEST-
TRAIN/stattraining.csv', header=TRUE)
head(stattrain)
str(stattrain)
stattrain$Pilihan_Pro <- as.factor(stattrain$Pilihan_Pro)
str(stattrain)

#ANALISIS REGRESI OLS
ols<-lm(ips ~ nf_verbal+nf_numerical+nf_figural+nf_matdas+nf_bind
+nf_bing+nf_matipa+nf_fis+nf_kim+nf_bio+Asal_Daerah+Bdidikmisi+Jeni
s_Kel+Jenis_Sek+Jurusan_Sek+Pilihan_Pro, data = stattrain)
summary(ols)

#DETEKSI OUTLIER
n <- nrow(stattrain)
p <- length(ols$coefficients)
cv <- 2*sqrt(p/n)
cv
data.dffits=dffits(ols)
dffits=abs(data.dffits)
data.dffits[which(dffits>cv)]

```

Lampiran 2. Syntax R (lanjutan)

```
#PENGUJIAN ASUMSI RESIDUAL
library(car)
ks.test(residuals(ols), "pnorm")
glejser<-lm(abs(residuals(ols)) ~
nf_verbal+nf_numerical+nf_figural+nf_matdas+nf_bind+nf_bing+nf_mat
ipa+nf_fis+nf_kim+nf_bio+Asal_Daerah+Bidikmisi+Jenis_Kel+Jenis_Sek
+Jurusan_Sek+Pilihan_Pro, data = stattrain)
summary(glejser)
vif(ols)

#ANALISIS REGRESI ROBUST LTS
set.seed(123)
lts<-ltsReg(ips ~ nf_verbal+nf_numerical+nf_figural+nf_matdas
+nf_bind+nf_bing+nf_matipa+nf_fis+nf_kim+nf_bio+Asal_Daerah
+Bidikmisi+Jenis_Kel+Jenis_Sek+Jurusan_Sek+Pilihan_Pro, data =
stattrain)
summary(lts)

#ANALISIS REGRESI ROBUST LWS
lws<-lmwinsor(ips ~ nf_verbal+nf_numerical+nf_figural+nf_matdas
+nf_bind+nf_bing+nf_matipa+nf_fis+nf_kim+nf_bio+Asal_Daerah
+Bidikmisi+Jenis_Kel+Jenis_Sek+Jurusan_Sek+Pilihan_Pro, data =
stattrain)
summary(lws)
```

Lampiran 3. Data Simulasi

No.	ε	x	y	$\varepsilon_{outlier}$	εl	$y l$
1	-0,17817	19,86635	19,68817	48,78421	48,78421	68,65056
2	-0,09677	19,81425	19,71748	51,89142	51,89142	71,70568
3	0,863309	19,19187	20,05518	50,08801	50,08801	69,27988
4	0,197889	19,01253	19,21042	49,95751	49,95751	68,97004
5	0,40255	18,81126	19,21381	48,46522	48,46522	67,27649
6	-1,95516	18,79389	16,83873	-	-1,95516	16,83873
7	-0,68236	18,68255	18,00019	-	-0,68236	18,00019
8	1,400935	18,49308	19,89402	-	1,400935	19,89402
9	0,981604	18,37914	19,36074	-	0,981604	19,36074
10	0,682418	18,19897	18,88139	-	0,682418	18,88139
11	0,433924	18,13573	18,56965	-	0,433924	18,56965
12	-1,46805	17,56459	16,09653	-	-1,46805	16,09653
13	0,501481	17,3515	17,85298	-	0,501481	17,85298
14	0,204001	17,21139	17,41539	-	0,204001	17,41539
15	-0,51745	17,13742	16,61998	-	-0,51745	16,61998
16	2,211878	16,70872	18,9206	-	2,211878	18,9206
17	0,679028	16,58241	17,26144	-	0,679028	17,26144
18	1,049414	16,0619	17,11132	-	1,049414	17,11132
19	-0,13184	15,83578	15,70394	-	-0,13184	15,70394
20	-0,59997	15,61172	15,01175	-	-0,59997	15,01175
21	-0,2074	15,57283	15,36543	-	-0,2074	15,36543
22	0,474488	15,49191	15,9664	-	0,474488	15,9664
23	1,611452	15,48617	17,09763	-	1,611452	17,09763
24	0,223718	15,45481	15,67853	-	0,223718	15,67853
25	1,202898	15,42458	16,62748	-	1,202898	16,62748
26	-0,89479	15,41849	14,5237	-	-0,89479	14,5237
27	-0,35985	15,39012	15,03027	-	-0,35985	15,03027
28	0,220899	15,34597	15,56687	-	0,220899	15,56687
29	-0,35501	15,14105	14,78604	-	-0,35501	14,78604
30	1,557109	15,01278	16,56989	-	1,557109	16,56989
31	0,317542	14,82101	15,13855	-	0,317542	15,13855
32	0,543999	14,61674	15,16074	-	0,543999	15,16074
33	0,674797	14,5034	15,1782	-	0,674797	15,1782
34	-2,31786	14,41173	12,09388	-	-2,31786	12,09388
⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮
100	0,835386	0,045565	0,88095	-	0,835386	0,88095

Lampiran 4. Output Analisis Data Simulasi

```

#ANALISIS REGRESI
Call:
lm(formula = y ~ x)
Residuals:
    Min       1Q   Median       3Q      Max
-2.47959 -0.68487 -0.03485  0.56361  2.55385
Coefficients: Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.35582    0.21691   -1.64    0.104
x            1.03591    0.01789   57.89 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.9739 on 98 degrees of freedom
Multiple R-squared:  0.9716,    Adjusted R-squared:  0.9713
F-statistic: 3351 on 1 and 98 DF,  p-value: < 2.2e-16

Call:
lm(formula = y1 ~ x)
Residuals:
    Min       1Q   Median       3Q      Max
-10.431  -4.817  -1.761   1.278   42.652
Coefficients: Estimate Std. Error t value Pr(>|t|)
(Intercept)  -5.5976    2.2756  -2.460  0.0157 *
x             1.7488    0.1877   9.316  3.7e-15 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 10.22 on 98 degrees of freedom
Multiple R-squared:  0.4696,    Adjusted R-squared:  0.4642
F-statistic: 86.78 on 1 and 98 DF,  p-value: 3.701e-15

Call:
ltsReg.formula(formula = y ~ x)
Residuals (from reweighted LS):
    Min       1Q   Median       3Q      Max
-1.8622 -0.5628  0.0000  0.5072  2.0547
Coefficients: Estimate Std. Error t value Pr(>|t|)
Intercept  -0.80974    0.18746  -4.32 3.94e-05 ***
x           1.06854    0.01544  69.20 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.8234 on 92 degrees of freedom
Multiple R-Squared:  0.9812,    Adjusted R-squared:  0.9809
F-statistic: 4789 on 1 and 92 DF,  p-value: < 2.2e-16

Call:
ltsReg.formula(formula = y1 ~ x)
Residuals (from reweighted LS):
    Min       1Q   Median       3Q      Max
-1.9220 -0.5325  0.0000  0.6187  2.2770
Coefficients: Estimate Std. Error t value Pr(>|t|)
Intercept  -0.80721    0.21179  -3.811 0.000253 ***
x           1.07180    0.01839  58.292 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.9163 on 90 degrees of freedom
Multiple R-Squared:  0.9742,    Adjusted R-squared:  0.9739
F-statistic: 3398 on 1 and 90 DF,  p-value: < 2.2e-16

Call:
lmwinsor(formula = y ~ x, data = datasim, trim = 0.1)
Residuals:

```

Lampiran 4. Output Analisis Data Simulasi (lanjutan)

```

      Min      1Q   Median      3Q      Max
-2.75335 -0.78352 -0.06705  0.76418  2.56965
Coefficients: Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.4490      0.3604  -1.246   0.216
x            1.0463      0.0307   34.087 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 1.107 on 98 degrees of freedom
Multiple R-squared:  0.9584,    Adjusted R-squared:  0.9673
F-statistic: 2259 on 1 and 98 DF,  p-value: < 2.2e-16

Call:
lmwinsor(formula = y1 ~ x, data = datasim, trim = 0.1)
Residuals:
      Min      1Q   Median      3Q      Max
-2.753 -0.784 -0.067  0.764  53.649
Coefficients: Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.4490      3.7741  -0.119  0.90554
x            1.0463      0.3215   3.255  0.00156 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 11.6 on 98 degrees of freedom
Multiple R-squared:  0.1736,    Adjusted R-squared:  0.3507
F-statistic: 20.59 on 1 and 98 DF,  p-value: 1.615e-05

#UJI ASUMSI NORMAL
data: residuals(ols1)
D = 0.063892, p-value = 0.8089
data: residuals(ols2)
D = 0.47668, p-value < 2.2e-16
data: residuals(lts1)
D = 0.076607, p-value = 0.6002
data: residuals(lts2)
D = 0.084271, p-value = 0.4765
data: residuals(lws1)
D = 0.066431, p-value = 0.7695
data: residuals(lws2)
D = 0.076886, p-value = 0.5955

#UJI ASUMSI IDENTIK
F-statistic: 1.363 on 1 and 98 DF,  p-value: 0.2459
F-statistic: 33.67 on 1 and 98 DF,  p-value: 8.015e-08
F-statistic: 0.8267 on 1 and 98 DF,  p-value: 0.3654
F-statistic: 14.15 on 1 and 98 DF,  p-value: 0.0002869
F-statistic: 1.137 on 1 and 98 DF,  p-value: 0.289
F-statistic: 13.41 on 1 and 98 DF,  p-value: 0.0004066

#NILAI MSE
> mseols1
[1] 0.9484461
> mseols2
[1] 104.3829
> mselts1
[1] 1.079013
> mselts2
[1] 135.929
> mselws1
[1] 1.22644
> mselws2
[1] 134.5348

```

Lampiran 5. *Output* Deteksi *Outlier*

3. Statistika

```

> data.dffits1[which(dffits1>cv)]
      48      50      59      68      89      90
-1.1459160 -0.8811852  0.9314997 -1.0097733 -1.0635966 -0.8931230
 0.8267339 -0.8072416
> data.dffits2[which(dffits2>cv)]
      38      40      71      80      89      100
-1.4088968 -0.8544738 -2.9892951 -1.1674524 -1.6189778 -0.8415726
 0.8415726
> data.dffits3[which(dffits3>cv)]
      18      19      38      44      47      71
-1.1509872 -1.1426151 -1.1777859  1.0097498 -0.9211276 -1.7779411
-1.8046827 -1.0098421

```

Output lebih lengkap di lampirkan pada CD

Lampiran 6. *Output* Pengujian Asumsi IIDN

3. Statistika

D = 0.31046,	p-value = 1.222e-10
D = 0.24186,	p-value = 1.266e-06
D = 0.28211,	p-value = 7.373e-09
F-statistic:	1.49 on 18 and 103 DF, p-value: 0.1088
F-statistic:	0.9562 on 18 and 103 DF, p-value: 0.5151
F-statistic:	1.487 on 18 and 103 DF, p-value: 0.1098
	GVIF Df GVIF ^{1/(2*Df)}
nf_verbal	1.224978 1 1.106787
nf_numerical	1.431402 1 1.196412
nf_figural	1.341455 1 1.158212
nf_matdas	2.008649 1 1.417268
nf_bind	1.252831 1 1.119299
nf_bing	1.359970 1 1.166178
nf_matipa	1.471184 1 1.212924
nf_fis	1.448227 1 1.203423
nf_kim	1.525563 1 1.235137
nf_bio	1.396233 1 1.181623
Asal_Daerah	1.560571 1 1.249228
Bidikmisi	1.103097 1 1.050284
Jenis_Kel	1.193739 1 1.092583
Jenis_Sek	1.257564 2 1.058967
Pilihan_Pro	1.483581 2 1.103641
PTN	1.275676 1 1.129458

Output lebih lengkap di lampirkan pada CD

Lampiran 7. Output Analisis OLS

3. Statistika

```

Call:
lm(formula = ips1 ~ nf_verbal + nf_numerical + nf_figural + nf_matdas +
    nf_bind + nf_bing + nf_matipa + nf_fis + nf_kim + nf_bio +
    Asal_Daerah + Bidikmisi + Jenis_Kej + Jenis_Sek + Pilihan_Pro
+ PTN, data = stattrain)
Residuals:
    Min       1Q   Median       3Q      Max
-1.05954 -0.14915  0.00601  0.20322  0.59325
Coefficients:
(Intercept)              1.535e+00  5.434e-01  2.824  0.0057 **
nf_verbal                5.803e-05  3.819e-04  0.152  0.8795
nf_numerical             1.242e-03  4.955e-04  2.506  0.0138 *
nf_figural              -9.070e-05  4.292e-04  -0.211  0.8331
nf_matdas                7.143e-05  2.289e-04  0.312  0.7557
nf_bind                 -3.441e-04  2.994e-04  -1.149  0.2531
nf_bing                 1.794e-04  2.852e-04  0.629  0.5306
nf_matipa               1.732e-04  2.904e-04  0.597  0.5521
nf_fis                  7.805e-04  3.030e-04  2.576  0.0114 *
nf_kim                  1.137e-03  2.625e-04  4.331  3.45e-05 ***
nf_bio                  -3.864e-04  3.519e-04  -1.098  0.2747
Asal_Daerahluar jawa    8.079e-02  6.994e-02  1.155  0.2507
Bidikmisiya             4.200e-02  7.581e-02  0.554  0.5807
Jenis_Kej               1.076e-01  6.302e-02  1.707  0.0909 .
Jenis_SekSMA           -1.539e-01  9.585e-02  -1.606  0.1114
Jenis_SekSMK           2.232e-01  2.479e-01  0.901  0.3699
Pilihan_Pro2           -5.207e-02  6.764e-02  -0.770  0.4431
Pilihan_Pro3           -1.172e-01  8.569e-02  -1.368  0.1744
PTNNON ITS             1.382e-01  7.624e-02  1.813  0.0727 .
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.3009 on 103 degrees of freedom
Multiple R-squared:  0.4233,    Adjusted R-squared:  0.3226
F-statistic: 4.201 on 18 and 103 DF,  p-value: 1.497e-06

Call:
lm(formula = ips2 ~ nf_verbal + nf_numerical + nf_figural + nf_matdas +
    nf_bind + nf_bing + nf_matipa + nf_fis + nf_kim + nf_bio +
    Asal_Daerah + Bidikmisi + Jenis_Kej + Jenis_Sek + Pilihan_Pro
+ PTN, data = stattrain)
Residuals:
    Min       1Q   Median       3Q      Max
-2.74121 -0.20845  0.06715  0.30103  0.70805
Coefficients:
(Intercept)              0.3579939  0.8982145  0.399  0.69104
nf_verbal                0.0009301  0.0006313  1.473  0.14373
nf_numerical             0.0014353  0.0008191  1.752  0.08271 .
nf_figural              -0.0005458  0.0007095  -0.769  0.44344
nf_matdas                0.0002624  0.0003784  0.693  0.48966
nf_bind                 -0.0001862  0.0004950  -0.376  0.70759
nf_bing                 -0.0003872  0.0004714  -0.821  0.41333
nf_matipa               0.0001786  0.0004800  0.372  0.71057
nf_fis                  0.0010837  0.0005008  2.164  0.03278 *
nf_kim                  0.0012349  0.0004338  2.846  0.00534 **
nf_bio                  -0.0003501  0.0005816  -0.602  0.54856
Asal_Daerahluar jawa    0.1667178  0.1156106  1.442  0.15232
Bidikmisiya            -0.0225387  0.1253060  -0.180  0.85761
Jenis_Kej              0.2865473  0.1041731  2.751  0.00703 **
Jenis_SekSMA           0.0089287  0.1584467  0.056  0.95517

```

Lampiran 7. Output Analisis OLS (lanjutan)

```

Jenis_SekSMK      -0.0625068  0.4097665  -0.153  0.87906
Pilihan_Pro2     -0.0426973  0.1118015  -0.382  0.70332
Pilihan_Pro3     -0.0096110  0.1416476  -0.068  0.94604
PTNNON ITS       0.3861176  0.1260229  3.064  0.00279 **
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.4975 on 103 degrees of freedom
Multiple R-squared:  0.347,    Adjusted R-squared:  0.2329
F-statistic: 3.041 on 18 and 103 DF,  p-value: 0.0002021

Call:
lm(formula = ipk ~ nf_verbal + nf_numerical + nf_figural + nf_matd
as +
    nf_bind + nf_bing + nf_matipa + nf_fis + nf_kim + nf_bio +
    Asal_Daerah + Bidikmisi + Jenis_kel + Jenis_Sek + Pilihan_Pro
+ PTN, data = stattrain)
Residuals:
    Min       1Q   Median       3Q      Max
-1.40170 -0.15046  0.04684  0.23426  0.89579
Coefficients:
(Intercept)      1.2152839  0.6809865  1.785 0.077271 .
nf_verbal        0.0001732  0.0004786  0.362 0.718122
nf_numerical    0.0008980  0.0006210  1.446 0.151207
nf_figural      -0.0004313  0.0005379  -0.802 0.424517
nf_matdas       0.0002471  0.0002869  0.861 0.391168
nf_bind        -0.0002431  0.0003753  -0.648 0.518503
nf_bing         0.0001846  0.0003574  0.517 0.606509
nf_matipa      0.0001905  0.0003639  0.523 0.601814
nf_fis         0.0011067  0.0003797  2.915 0.004364 **
nf_kim         0.0008220  0.0003289  2.499 0.014023 *
nf_bio         0.0001652  0.0004409  0.375 0.708742
Asal_Daerahluar jawa 0.1132585  0.0876508  1.292 0.199195
BidikmisiYa     0.0675243  0.0950014  0.711 0.478832
Jenis_Kelp      0.2913707  0.0789794  3.689 0.000361 ***
Jenis_SekSMA   -0.1673766  0.1201273  -1.393 0.166520
Jenis_SekSMK   -0.1431922  0.3106668  -0.461 0.645829
Pilihan_Pro2   -0.0778738  0.0847630  -0.919 0.360386
Pilihan_Pro3   -0.2446893  0.1073909  -2.278 0.024761 *
PTNNON ITS     0.0219492  0.0955450  0.230 0.818760
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.3771 on 103 degrees of freedom
Multiple R-squared:  0.4277,    Adjusted R-squared:  0.3277
F-statistic: 4.276 on 18 and 103 DF,  p-value: 1.095e-06

```

Output lebih lengkap di lampirkan pada CD

Lampiran 8. Output Analisis LTS

3. Statistika

```

Call:
ltsReg.formula(formula = ips1 ~ nf_verbal + nf_numerical + nf_figural +
  nf_matdas + nf_bind + nf_bing + nf_matipa + nf_fis + nf_kim +
  nf_bio + Asal_Daerah + Bidikmisi + Jenis_Kel + Jenis_Sek +
  Pilihan_Pro + PTN, data = stattrain)
Residuals (from reweighted LS):
  Min      1Q  Median      3Q      Max
-0.6065 -0.1045  0.0000  0.1238  0.4696
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
Intercept    1.826e+00  4.740e-01  3.852 0.000211 ***
nf_verbal    1.188e-04  3.276e-04  0.363 0.717648
nf_numerical 9.356e-04  4.392e-04  2.130 0.035706 *
nf_figural   3.700e-04  3.624e-04  1.021 0.309864
nf_matdas    1.607e-04  1.934e-04  0.831 0.408108
nf_bind      -2.384e-04  2.538e-04 -0.939 0.349872
nf_bing      -2.347e-04  2.433e-04 -0.964 0.337248
nf_matipa    5.186e-05  2.534e-04  0.205 0.838265
nf_fis       1.189e-03  2.633e-04  4.518 1.78e-05 ***
nf_kim       6.289e-04  2.303e-04  2.731 0.007507 **
nf_bio      -6.165e-04  3.006e-04 -2.051 0.042983 *
Asal_Daerahluar_jawa -4.380e-02  6.004e-02 -0.730 0.467434
Bidikmisiya  2.114e-02  6.760e-02  0.313 0.755215
Jenis_Kelp   5.734e-02  5.396e-02  1.063 0.290640
Jenis_SekSMA -6.026e-02  8.004e-02 -0.753 0.453402
Jenis_SekSMK 1.458e-01  2.075e-01  0.703 0.483848
Pilihan_Pro2 -9.668e-02  5.820e-02 -1.661 0.099962 .
Pilihan_Pro3 -9.265e-02  7.520e-02 -1.232 0.220962
PTNNON ITS   1.750e-01  6.408e-02  2.731 0.007514 **
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.2502 on 96 degrees of freedom
Multiple R-Squared: 0.4959,    Adjusted R-squared: 0.4014
F-statistic: 5.247 on 18 and 96 DF,  p-value: 3.255e-08

Call:
ltsReg.formula(formula = ips2 ~ nf_verbal + nf_numerical + nf_figural +
  nf_matdas + nf_bind + nf_bing + nf_matipa + nf_fis + nf_kim +
  nf_bio + Asal_Daerah + Bidikmisi + Jenis_Kel + Jenis_Sek +
  Pilihan_Pro + PTN, data = stattrain)
Residuals (from reweighted LS):
  Min      1Q  Median      3Q      Max
-0.712078 -0.170108  0.005956  0.198906  0.599803
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
Intercept    1.399e+00  6.584e-01  2.125 0.036175 *
nf_verbal    1.735e-04  4.299e-04  0.404 0.687472
nf_numerical 7.857e-04  5.967e-04  1.317 0.191087
nf_figural   -2.702e-04  4.952e-04 -0.546 0.586621
nf_matdas    3.337e-04  2.600e-04  1.283 0.202460
nf_bind      1.784e-04  3.426e-04  0.521 0.603793
nf_bing      -6.790e-05  3.417e-04 -0.199 0.842888
nf_matipa    9.765e-05  3.331e-04  0.293 0.770044
nf_fis       6.446e-04  3.526e-04  1.828 0.070636 .
nf_kim       8.232e-04  2.963e-04  2.778 0.006583 **
nf_bio      -5.111e-04  4.015e-04 -1.273 0.206085
Asal_Daerahluar_jawa 7.122e-02  7.781e-02  0.915 0.362300
Bidikmisiya -5.392e-02  8.627e-02 -0.625 0.533423
Jenis_Kelp   2.796e-01  7.337e-02  3.811 0.000244 ***
Jenis_SekSMA 2.781e-02  1.108e-01  0.251 0.802424

```

Lampiran 8. Output Analisis LTS (lanjutan)

```

Jenis_SekSMK      -1.623e-01  2.750e-01  -0.590  0.556425
Pilihan_Pro2     -6.066e-02  7.721e-02  -0.786  0.434019
Pilihan_Pro3      9.619e-02  9.762e-02  0.985  0.326925
PTNNON ITS       3.774e-01  8.671e-02  4.353  3.36e-05 ***
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.3324 on 96 degrees of freedom
Multiple R-Squared: 0.4402,    Adjusted R-squared: 0.3352
F-statistic: 4.194 on 18 and 96 DF,  p-value: 2.016e-06

Call:
ltsReg.formula(formula = ipk ~ nf_verbal + nf_numerical + nf_figur
al +
  nf_matdas + nf_bind + nf_bing + nf_matipa + nf_fis + nf_kim +
  nf_bio + Asal_Daerah + Bidikmisi + Jenis_Kel + Jenis_Sek +
  Pilihan_Pro + PTN, data = stattrain)
Residuals (from reweighted LS):
  Min      1Q  Median      3Q      Max
-0.6111 -0.1037  0.0000  0.1511  0.6432
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
Intercept    1.770e+00  4.710e-01  3.759 0.000297 ***
nf_verbal    2.740e-05  3.343e-04  0.082 0.934863
nf_numerical 1.016e-03  4.311e-04  2.357 0.020517 *
nf_figural   -4.706e-04  3.883e-04  -1.212 0.228663
nf_matdas    1.861e-04  1.924e-04  0.967 0.336078
nf_bind      -4.759e-04  2.592e-04  -1.836 0.069577 .
nf_bing      2.971e-05  2.507e-04  0.118 0.905930
nf_matipa    2.999e-04  2.527e-04  1.186 0.238497
nf_fis       8.314e-04  2.623e-04  3.170 0.002065 **
nf_kim       8.598e-04  2.254e-04  3.814 0.000246 ***
nf_bio      -1.206e-04  3.028e-04  -0.398 0.691320
Asal_Daerahluar jawa 1.376e-01  6.187e-02  2.224 0.028539 *
Bidikmisiya  -1.041e-02  6.303e-02  -0.165 0.869123
Jenis_Kelp    2.619e-01  5.407e-02  4.844 5.06e-06 ***
Jenis_SekSMA -9.228e-02  7.932e-02  -1.163 0.247657
Jenis_SekSMK -1.479e-01  2.031e-01  -0.728 0.468238
Pilihan_Pro2  -9.206e-02  5.761e-02  -1.598 0.113454
Pilihan_Pro3  -1.975e-01  7.722e-02  -2.557 0.012163 *
PTNNON ITS    4.013e-02  6.253e-02  0.642 0.522543
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.2454 on 93 degrees of freedom
Multiple R-Squared: 0.5573,    Adjusted R-squared: 0.4716
F-statistic: 6.504 on 18 and 93 DF,  p-value: 4.216e-10

```

Output lebih lengkap di lampirkan pada CD

Lampiran 9. Output Analisis LWS

3. Statistika

```

Call:
lmwinsor(formula = ips1 ~ nf_verbal + nf_numerical + nf_figural +
  nf_matdas + nf_bind + nf_bing + nf_matipa + nf_fis + nf_kim +
  nf_bio + Asal_Daerah + Bidikmisi + Jenis_Kel + Jenis_Sek +
  Pilihan_Pro + PTN, data = stattrain, trim = 0.1)
Residuals:
    Min       1Q   Median       3Q      Max
-1.07369 -0.13955  0.00803  0.20458  0.62104
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    1.116e+00  7.684e-01  1.453  0.1494
nf_verbal      1.433e-04  4.656e-04  0.308  0.7589
nf_numerical   1.060e-03  6.744e-04  1.572  0.1190
nf_figural     -8.876e-05  5.360e-04  -0.166  0.8688
nf_matdas      3.694e-05  2.821e-04  0.131  0.8961
nf_bind        -2.755e-04  3.593e-04  -0.767  0.4451
nf_bing         2.802e-04  3.504e-04  0.800  0.4258
nf_matipa      4.959e-04  3.720e-04  1.333  0.1855
nf_fis         7.447e-04  4.289e-04  1.736  0.0855
nf_kim         1.649e-03  3.608e-04  4.571  1.35e-05 ***
nf_bio         -3.977e-04  4.279e-04  -0.929  0.3549
Asal_Daerahluar_jawa  6.373e-02  7.523e-02  0.847  0.3989
BidikmisiYa    1.527e-02  7.911e-02  0.193  0.8473
Jenis_Kelp     1.112e-01  6.734e-02  1.652  0.1016
Jenis_SekSMA  -2.314e-01  1.093e-01  -2.118  0.0366 *
Jenis_SekSMK   9.806e-02  2.603e-01  0.377  0.7071
Pilihan_Pro2   -6.282e-02  7.206e-02  -0.872  0.3854
Pilihan_Pro3   -1.053e-01  8.891e-02  -1.184  0.2390
PTNON ITS      1.605e-01  8.029e-02  1.999  0.0483 *
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.3105 on 103 degrees of freedom
Multiple R-squared:  0.3927,    Adjusted R-squared:  0.316
F-statistic:  3.7 on 18 and 103 DF,  p-value: 1.221e-05

Call:
lmwinsor(formula = ips2 ~ nf_verbal + nf_numerical + nf_figural +
  nf_matdas + nf_bind + nf_bing + nf_matipa + nf_fis + nf_kim +
  nf_bio + Asal_Daerah + Bidikmisi + Jenis_Kel + Jenis_Sek +
  Pilihan_Pro + PTN, data = stattrain, trim = 0.1)
Residuals:
    Min       1Q   Median       3Q      Max
-2.66638 -0.22564  0.05049  0.28991  0.81000
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   -1.9787216  1.5818706  -1.251  0.21381
nf_verbal     0.0019943  0.0008340   2.391  0.01860 *
nf_numerical  0.0005166  0.0011665   0.443  0.65880
nf_figural    0.0001078  0.0009334   0.115  0.90832
nf_matdas     0.0001270  0.0004818   0.264  0.79268
nf_bind       0.0005496  0.0006562   0.838  0.40420
nf_bing       -0.0001912  0.0006003  -0.319  0.75074
nf_matipa     0.0008779  0.0006598   1.331  0.18624
nf_fis        0.0021511  0.0007435   2.893  0.00466 **
nf_kim        0.0018754  0.0006903   2.717  0.00773 **
nf_bio        -0.0005658  0.0007577  -0.747  0.45694
Asal_Daerahluar_jawa  0.3196445  0.1511740   2.114  0.03689 *
BidikmisiYa   0.0241346  0.1416661   0.170  0.86506
Jenis_Kelp    0.3554194  0.1223388   2.905  0.00449 **
Jenis_SekSMA  -0.0412379  0.1648812  -0.250  0.80300
Jenis_SekSMK  -0.2255967  0.4221175  -0.534  0.59419
Pilihan_Pro2  0.0101931  0.1314490   0.078  0.93834

```

Lampiran 9. Output Analisis LWS (lanjutan)

```

Pilihan_Pro3      0.0994003  0.1597312  0.622  0.53512
PTNNON ITS       0.3630339  0.1393145  2.606  0.01052 *
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.5008 on 103 degrees of freedom
Multiple R-squared:  0.3509,    Adjusted R-squared:  0.3446
F-statistic: 3.093 on 18 and 103 DF,  p-value: 0.0001615

Call:
lmwinsor(formula = ipk ~ nf_verbal + nf_numerical + nf_figural +
  nf_matdas + nf_bind + nf_bing + nf_matipa + nf_fis + nf_kim +
  nf_bio + Asal_Daerah + Bidikmisi + Jenis_Kel + Jenis_Sek +
  Pilihan_Pro + PTN, data = stattrain, trim = 0.1)
Residuals:
    Min       1Q   Median       3Q      Max
-1.39272 -0.17742  0.02521  0.20633  0.88141
Coefficients:
(Intercept)      0.5444572  0.9477527  0.574  0.566901
nf_verbal        0.0003465  0.0005716  0.606  0.545752
nf_numerical     0.0006957  0.0008475  0.821  0.413619
nf_figural      -0.0004669  0.0006580 -0.710  0.479534
nf_matdas       0.0001929  0.0003479  0.555  0.580402
nf_bind        -0.0001720  0.0004459 -0.386  0.700547
nf_bing         0.0003565  0.0004395  0.811  0.419160
nf_matipa       0.0005342  0.0004559  1.172  0.243952
nf_fis          0.0015357  0.0005369  2.860  0.005123 **
nf_kim          0.0009742  0.0004379  2.225  0.028270 *
nf_bio          0.0004585  0.0005329  0.860  0.391539
Asal_Daerahluar jawa 0.1172216  0.0938115  1.250  0.214298
BidikmisiYa     0.0539467  0.1039491  0.519  0.604893
Jenis_kelp      0.2921559  0.0827198  3.532  0.000619 ***
Jenis_SekSMA   -0.2555593  0.1527381 -1.673  0.097325 .
Jenis_SekSMK   -0.2808369  0.3273389 -0.858  0.392918
Pilihan_Pro2   -0.0971295  0.0902856 -1.076  0.284529
Pilihan_Pro3   -0.2382228  0.1107272 -2.151  0.033779 *
PTNNON ITS     0.0100104  0.1019241  0.098  0.921953
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.3799 on 103 degrees of freedom
Multiple R-squared:  0.4031,    Adjusted R-squared:  0.3451
F-statistic: 3.864 on 18 and 103 DF,  p-value: 6.107e-06

```

Output lebih lengkap di lampirkan pada CD

Lampiran 10. Surat Pernyataan Data

SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini, mahasiswa Departemen Statistika FMKSD ITS:

Nama : Khoirotus Nainiyah

NRP : 06211745000025

menyatakan bahwa data yang digunakan dalam Tugas Akhir/ Thesis ini merupakan data sekunder yang diambil dari penelitian/ buku/ Tugas Akhir/ Thesis/ publikasi lainnya yaitu:

Sumber : Panitia Pengkaji Jalur Ujian Tulis Perguruan Tinggi Negeri

Keterangan : Data yang digunakan adalah data nasional hasil ujian tulis masuk perguruan tinggi negeri di Indonesia tahun 2014 meliputi nilai mata uji dan kriteria calon mahasiswa serta data indeks prestasi mahasiswa

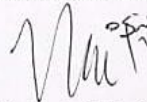
Surat Pernyataan ini dibuat dengan sebenarnya. Apabila terdapat pemalsuan data maka saya siap menerima sanksi sesuai aturan yang berlaku.

Mengetahui
Pembimbing Tugas Akhir



(Dr. Suhartono, M.Sc.)
NIP. 19710929 199512 1 001

Surabaya, 24 Juni 2019



(Khoirotus Nainiyah)
NRP. 06211745000025

*(coret yang tidak perlu)



BIODATA PENULIS

Penulis dalam kesehariannya biasa dipanggil 'Nainin', dilahirkan di Jombang pada 16 Oktober 1995 dengan nama lengkap Khoirotus Nainiyah. Penulis merupakan anak kedua dari tiga bersaudara. Riwayat pendidikan penulis dimulai dari RA Al-Islamiyah Sawiji (2000-2001), MIN Rejoso PP. Darul 'Ulum (2001-2007), SMPN 3 PP. Darul 'Ulum Peterongan Jombang (2007-2010). SMAN 2 Jombang (2010-2013).

Diploma III Statistika FMIPA ITS (2013-2016) dan Lintas Jalur Statistika FMKSD ITS (2017-2019). Selama kuliah di ITS penulis berkesempatan Kerja Praktik di Kantor Pelayanan Pajak Pratama Wonocolo dan PT. Pos Indonesia (Persero) Regional 7 Jawa Timur. Jika pembaca ingin memberikan kritik, saran atau ingin berdiskusi mengenai Tugas Akhir penulis, dapat menghubungi penulis melalui email: khoirotusnainiyah16@gmail.com ☺