



**TUGAS AKHIR - SS141501**

**PEMODELAN FAKTOR-FAKTOR YANG  
MEMPENGARUHI TINGKAT PENGANGGURAN  
TERBUKA DI PROVINSI ACEH MENGGUNAKAN  
PENDEKATAN REGRESI NONPARAMETRIK  
SPLINE**

**LEISA NOVIANA SANI  
NRP 1311 100 112**

**Dosen Pembimbing  
Prof. Dr. Drs. I Nyoman Budiantara, M.Si**

**Program Studi S1 Statistika  
Fakultas Matematika dan Ilmu Pengetahuan Alam  
Institut Teknologi Sepuluh Nopember  
Surabaya 2015**



**FINAL PROJECT - SS141501**

**MODELLING OF FACTORS THAT AFFECT OPEN  
UNEMPLOYMENT IN ACEH USING  
NONPARAMETRIC SPLINE REGRESSION  
APPROACH**

**LEISA NOVIANA SANI  
NRP 1311 100 112**

**Supervisor  
Prof. Dr. Drs. I Nyoman Budiantara, M.Si**

**Undergraduate Programme of Statistics  
Faculty of Mathematics and Natural Science  
Institut Teknologi Sepuluh Nopember  
Surabaya 2015**

**LEMBAR PENGESAHAN**

**PEMODELAN FAKTOR-FAKTOR YANG  
MEMPENGARUHI TINGKAT PENGANGGURAN  
TERBUKA DI PROVINSI ACEH MENGGUNAKAN  
PENDEKATAN REGRESI NONPARAMETRIK  
SPLINE**

**TUGAS AKHIR**

Diajukan untuk Memenuhi Salah Satu Syarat  
Memperoleh Gelar Sarjana Sains  
pada

Program Studi S-1 Jurusan Statistika  
Fakultas Matematika dan Ilmu Pengetahuan Alam  
Institut Teknologi Sepuluh Nopember

Oleh:

**LEISA NOVIANA SANI**

**NRP. 1311 100 112**

Disetujui oleh Dosen Pembimbing Tugas Akhir :

**Prof. Dr. Drs. INyoman Budiantara, M.Si**

**NIP. 19650603 198903 1 003**

Mengetahui

Ketua Jurusan Statistika FMIPA-ITS

**Dr. Muhammad Mashuri, MT**

**NIP. 19620408 198701 1 001**

**SURABAYA, JULI 2015**

**PEMODELAN FAKTOR-FAKTOR YANG  
MEMPENGARUHI TINGKAT PENGANGGURAN  
TERBUKA DI PROVINSI ACEH MENGGUNAKAN  
PENDEKATAN REGRESI NONPARAMETRIK SPLINE**

**Nama Mahasiswa : Leisa Noviana Sani**  
**NRP : 1311 100 112**  
**Jurusan : Statistika, FMIPA-ITS**  
**Pembimbing : Prof.Dr.Drs. I Nyoman  
Budiantara, M.Si**

**ABSTRAK**

*Pengangguran merupakan salah satu faktor yang dapat menghambat tercapainya kesejahteraan. Aceh termasuk salah satu provinsi dengan Tingkat Pengangguran Terbuka (TPT) tertinggi di Indonesia. Pada tahun 2013, Aceh merupakan provinsi dengan TPT tertinggi di Indonesia dan selama tiga tahun berturut-turut TPT Aceh berada di atas TPT nasional. Selain itu, dalam satu tahun fluktuasi TPT Aceh cukup besar dibandingkan provinsi lain. Untuk mengurangi TPT, perlu diselidiki faktor-faktor yang mempengaruhinya supaya dapat dilakukan tindakan terhadap faktor-faktor tersebut. Penelitian ini bertujuan menganalisis faktor-faktor yang mempengaruhi TPT dengan menggunakan pendekatan regresi nonparametrik spline. Model terbaik yang digunakan adalah model dengan nilai GCV terkecil, yaitu menggunakan kombinasi banyak titik knot. Pemodelan dilakukan dengan tiga variabel prediktor signifikan, yaitu kepadatan penduduk, Angka Melek Huruf (AMH), dan Angka Partisipasi Kasar (APK) jenjang SLTA dengan koefisien determinasi sebesar 88,55 persen.*

**Kata Kunci :** *Tingkat Pengangguran Terbuka, Regresi Nonparametrik Spline, GCV*

*(Halaman Ini Sengaja Dikosongkan)*

# MODELLING OF FACTORS THAT AFFECT OPEN UNEMPLOYMENT IN ACEH USING NONPARAMETRIC SPLINE REGRESSION APPROACH

**Name** : Leisa Noviana Sani  
**NRP** : 1311 100 112  
**Department** : Statistics, FMIPA-ITS  
**Supervisor** : Prof.Dr.Drs. I Nyoman  
Budiantara, M.Si

## ABSTRACT

*Unemployment is one of some factors that detains Indonesia from public welfare. Aceh is one of the provinces in Indonesia with the highest Open Unemployment. In 2013, Aceh had the highest Open Unemployment Rate in Indonesia and for three years in a row Aceh's Open Unemployment Rate was higher than the nation's. Moreover, it fluctuated more than other provinces' Open Unemployment Rate. The purpose of this study is to analyze the factors that affect Open Unemployment Rate using nonparametric spline regression approach. The best model is the one with the smallest GCV, which is knots combination. The modelling of Open Unemployment is formed using three predictors that significant. The variables significant are population density, literacy rate, and labor age of high school rate. This model's coefficient of determination is 88,55 percent.*

**Keywords** : *Open Unemployment, nonparametrik spline regression, GCV.*

*(Halaman Ini Sengaja Dikosongkan)*

## **KATA PENGANTAR**

**Assalamualaikum Wr. Wb.**

Puji syukur dipanjatkan kehadirat Allah SWT atas berkat dan rahmat-Nya sehingga penulis dapat menyelesaikan dengan baik Tugas Akhir dengan judul

**“PEMODELAN FAKTOR-FAKTOR YANG  
MEMPENGARUHI TINGKAT PENGANGGURAN  
TERBUKA DI PROVINSI ACEH MENGGUNAKAN  
PENDEKATAN REGRESI NONPARAMETRIK SPLINE”**

serta tak lupa sholawat serta salam pada junjungan Nabi Muhammad SAW. Tugas Akhir ini tidak akan terselesaikan dengan baik tanpa dukungan dan bantuan dari banyak pihak selama proses pengerjaannya, oleh karena itu penulis ingin memberikan penghormatan dan ucapan terima kasih yang sebesar-besarnya kepada :

1. Bapak Prof. Dr. Drs. I Nyoman Budiantara, M.Si selaku dosen pembimbing yang telah bersedia meluangkan waktu serta memberikan ilmu, kesabaran, semangat, dan saran kepada penulis.
2. Bapak Dr. Muhammad Mashuri, M.T. selaku Ketua Jurusan Statistika ITS dan Ibu Dra. Lucia Aridinanti, M.T. selaku koordinator Tugas Akhir Jurusan Statistika ITS.
3. Kedua orang tua tercinta, bapak (Haswan Sani) dan ibu (Vini Kusvini), atas segala doa dan dukungan yang dicurahkan sampai saat ini.
4. Ibu Dra. Madu Ratna, M.Si dan Ibu Erma Oktania Permatasari, M.Si selaku dosen penguji atas saran dan masukan demi perbaikan Tugas Akhir ini.

5. Dr. Sutikno, M.Si selaku dosen wali yang telah memberikan nasehat, motivasi, serta bimbingan selama perkuliahan.
6. Seluruh dosen Jurusan Statistika ITS atas ilmu dan pengalaman yang dibagikan pada penulis.
7. Kakak (Risa Desalvina Sani, M.M) yang membantu serta memberikan doa dan dukungan selama pembuatan Tugas Akhir ini.
8. Sahabat-sahabat (Feby, Riska P., dan Saidah), atas bantuan, dukungan, saran, dan persahabatan dalam suka-duka selama menempuh pendidikan S1 di Statistika ITS.
9. Teman-teman satu bimbingan atas ilmu yang dibagikan, semangat dan bantuan yang diberikan selama proses pengerjaan Tugas Akhir.
10. Teman-teman satu angkatan S1 2011 yang telah sama-sama berjuang di Statistika ITS.

Penulis menyadari akan ketidaksempurnaan dan masih adanya kekurangan dalam Tugas Akhir ini, oleh karena itu saran dan masukan sangatlah diharapkan. Semoga Tugas Akhir ini dapat bermanfaat bagi pembaca dan semua pihak. Amin.

**Wassalamualaikum Wr. Wb.**

Surabaya, Juni 2015

Penulis

# DAFTAR ISI

	Halaman
<b>HALAMAN PENGESAHAN</b> .....	<b>v</b>
<b>ABSTRAK</b> .....	<b>vii</b>
<b>ABSTRACT</b> .....	<b>ix</b>
<b>KATA PENGANTAR</b> .....	<b>xi</b>
<b>DAFTAR ISI</b> .....	<b>xiii</b>
<b>DAFTAR TABEL</b> .....	<b>xvi</b>
<b>DAFTAR GAMBAR</b> .....	<b>xix</b>
<b>DAFTAR LAMPIRAN</b> .....	<b>xxi</b>
<b>BAB I PENDAHULUAN</b> .....	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	4
1.3 Tujuan Penelitian .....	4
1.4 Manfaat Penelitian .....	5
1.5 Batasan Penelitian .....	5
<b>BAB II TINJAUAN PUSTAKA</b> .....	<b>7</b>
2.1 Statistik Deskriptif .....	7
2.2 Analisis Regresi .....	8
2.3 Regresi Parametrik .....	8
2.4 Regresi Nonparametrik Spline .....	9
2.5 Pemilihan Titik Knot Optimal .....	10
2.6 Pengujian Parameter Model .....	10
2.7 Asumsi Residual Model .....	13
2.7.1 Asumsi Residual Independen .....	13
2.7.2 Asumsi Residual Identik .....	14
2.7.3 Asumsi Normalitas Residual .....	14
2.8 Pengangguran Terbuka .....	15
2.9 Tinjauan Pengangguran Terbuka .....	15
<b>BAB III METODOLOGI PENELITIAN</b> .....	<b>19</b>
3.1 Sumber Data dan Variabel Penelitian .....	19

3.2	Langkah Penelitian .....	20
<b>BAB IV</b>	<b>ANALISIS DAN PEMBAHASAN .....</b>	<b>23</b>
4.1	Karakteristik Tingkat Pengangguran Terbuka (TPT) di Aceh dan Faktor-Faktor yang Diduga Mempengaruhinya .....	23
4.2	Pola Hubungan Antara TPT dengan Faktor-Faktor yang Diduga Mempengaruhinya .....	27
4.3	Model Regresi Nonparametrik Spline .....	28
4.4	Pemilihan Titik Knot Optimal .....	29
4.4.1	Pemilihan Titik Knot Optimal dengan Satu Titik Knot .....	29
4.4.2	Pemilihan Titik Knot Optimal dengan Dua Titik Knot .....	30
4.4.3	Pemilihan Titik Knot Optimal dengan Tiga Titik Knot .....	32
4.4.4	Pemilihan Titik Knot Optimal dengan Kombinasi Knot .....	34
4.5	Pemilihan Model Terbaik .....	36
4.6	Penaksiran Parameter Model Regresi Nonparametrik Spline .....	36
4.7	Pengujian Signifikansi Parameter .....	37
4.8	Pemilihan Titik Knot Optimal dengan Tiga Variabel Prediktor .....	39
4.8.1	Pemilihan Titik Knot Optimal dengan Satu Titik Knot .....	39
4.8.2	Pemilihan Titik Knot Optimal dengan Dua Titik Knot .....	40
4.8.3	Pemilihan Titik Knot Optimal dengan Tiga Titik Knot .....	42
4.8.4	Pemilihan Titik Knot Optimal dengan Kombinasi Knot .....	44

4.9	Pemilihan Model Terbaik dengan Tiga Variabel Prediktor .....	45
4.10	Penaksiran Parameter Model Regresi Nonparametrik Spline dengan Tiga Variabel.....	46
4.11	Pengujian Signifikansi Parameter dengan Tiga Variabel Prediktor .....	47
4.12	Pemeriksaan Asumsi Residual .....	48
4.12.1	Pemeriksaan Asumsi Independen Residual.....	49
4.12.2	Pengujian Asumsi Identik Residual.....	49
4.12.3	Pengujian Asumsi Normalitas Residual.....	50
4.13	Interpretasi Model Regresi Nonparametrik Spline .....	50
	<b>BAB V KESIMPULAN DAN SARAN.....</b>	<b>55</b>
5.1	Kesimpulan.....	55
5.2	Saran.....	57
	<b>DAFTAR PUSTAKA.....</b>	<b>59</b>
	<b>LAMPIRAN.....</b>	<b>61</b>

*(Halaman Ini Sengaja Dikosongkan)*

## DAFTAR TABEL

	Halaman
Tabel 2.1 <i>Analysis of Variance</i> (ANOVA).....	11
Tabel 3.1 Variabel Penelitian.....	19
Tabel 4.1 Statistik Deskriptif TPT di Aceh dan Lima Variabel yang Diduga Mempengaruhinya.....	23
Tabel 4.2 Nilai GCV Regresi Nonparametrik Spline dengan Satu Titik Knot.....	29
Tabel 4.3 Nilai GCV Regresi Nonparametrik Spline dengan Dua Titik Knot.....	30
Tabel 4.4 Nilai GCV Regresi Nonparametrik Spline dengan Tiga Titik Knot.....	32
Tabel 4.5 Nilai GCV Regresi Nonparametrik Spline dengan Kombinasi Knot.....	34
Tabel 4.6 Nilai GCV Terkecil dari Pemodelan Tiap Jumlah Knot.....	36
Tabel 4.7 Tabel ANOVA.....	38
Tabel 4.8 Hasil Uji Signifikansi Parameter Secara Parsial.....	38
Tabel 4.9 Nilai GCV Regresi Nonparametrik Spline Tiga Variabel dengan Satu Titik Knot.....	40
Tabel 4.10 Nilai GCV Regresi Nonparametrik Spline Dua Variabel dengan Dua Titik Knot.....	41
Tabel 4.11 Nilai GCV Regresi Nonparametrik Spline Tiga Variabel dengan Tiga Titik Knot.....	42
Tabel 4.12 Nilai GCV Regresi Nonparametrik Spline Tiga Variabel dengan Kombinasi Knot.....	44
Tabel 4.13 Nilai GCV Terkecil dari Pemodelan Tiap Jumlah Knot dengan Tiga Variabel Prediktor.....	46
Tabel 4.14 Tabel ANOVA Pemodelan dengan Tiga Variabel Prediktor.....	47
Tabel 4.15 Hasil Uji Signifikansi Parameter Secara Parsial	

Model dengan Tiga Variabel Prediktor.....	48
Tabel 4.16 ANOVA Uji Glejser.....	49

## DAFTAR GAMBAR

	Halaman
Gambar 3.1	Diagram Alir Penelitian..... 22
Gambar 4.1	TPT di Aceh Berdasarkan Kabupaten/Kota pada Tahun 2013..... 24
Gambar 4.2	<i>Scatterplot</i> Antara TPT di Aceh dan Variabel- Variabel yang Diduga Mempengaruhinya..... 27
Gambar 4.3	Plot ACF Residual..... 49
Gambar 4.4	Plot Distribusi Normal Residual..... 50

*(Halaman Ini Sengaja Dikosongkan)*

## DAFTAR LAMPIRAN

	Halaman
Lampiran 1	Data TPT di Aceh dan Variabel-Variabel yang Diduga Berpengaruh.....61
Lampiran 2	Program Regresi Nonparametrik Spline Linier Menggunakan <i>Software R</i> ..... 62
Lampiran 3	Program Uji Glejser Menggunakan <i>Software R</i> ...94
Lampiran 4	<i>Output</i> Uji Signifikansi Parameter 5 Prediktor.....97
Lampiran 5	<i>Output</i> Uji Signifikansi Parameter 3 Prediktor...100
Lampiran 6	<i>Output</i> Uji Glejser.....103

*(Halaman Ini Sengaja Dikosongkan)*

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Pembangunan ekonomi yang dilakukan dalam suatu negara bertujuan antara lain untuk memberikan kesejahteraan bagi masyarakat. Secara potensial Indonesia memiliki kemampuan sumber daya manusia yang cukup untuk dikembangkan. Namun, untuk mencapai kesejahteraan tidaklah mudah. Banyaknya pengangguran merupakan salah satu faktor yang dapat menghambat tercapainya kesejahteraan tersebut. Munculnya masalah ketenagakerjaan ini disebabkan karena ketidakseimbangan antara lapangan kerja dan jumlah penduduk, kompetensi pencari kerja yang tidak memenuhi kebutuhan pasar kerja, dan perekonomian di wilayah itu sendiri. Tingkat pengangguran di Indonesia pada Agustus 2014 adalah sebesar 5,94 persen mengalami peningkatan dibandingkan pada Februari 2014 yaitu 5,7 persen, dan mengalami penurunan dibandingkan pada Agustus 2013 yaitu 6,17 persen (BPS, 2014). Pemerintah sebaiknya menekan angka pengangguran seminimal mungkin karena pengangguran lebih berpotensi menyebabkan permasalahan sosial seperti kriminalitas dan kemiskinan. Salah satu upaya pemerintah dalam memperbaiki keadaan ketenagakerjaan di Indonesia adalah dengan meningkatkan kompetensi pencari kerja melalui pelatihan kewirausahaan dan pelatihan calon tenaga kerja.

Aceh termasuk salah satu provinsi dengan tingkat pengangguran terbuka (TPT) tertinggi di Indonesia. Pada Mei 2014 tercatat bahwa TPT Aceh sebesar 7,15 persen, menjadikan Aceh provinsi dengan TPT tertinggi kelima setelah Banten, DKI Jakarta, Kalimantan Timur, dan Jawa Barat (BPS, 2014). Jumlah penganggur di Aceh pada Agustus 2014 dibandingkan dengan Agustus 2013 menurun sebanyak 16 ribu orang, tetapi meningkat sebanyak 44 ribu orang dibandingkan pada Februari 2014 (BPS,

2014). Hal tersebut menunjukkan bahwa sempat terjadi penurunan jumlah penganggur pada Februari 2014 dibandingkan pada Agustus 2013 tetapi pada Agustus 2014 jumlah penganggur meningkat kembali. Apabila dibandingkan dengan provinsi lain, fluktuasi TPT di Aceh cukup besar. Oleh karena itu perlu diselidiki penyebab perubahan banyaknya penganggur di Aceh.

Kondisi ketenagakerjaan di Aceh dapat dipengaruhi berbagai faktor, seperti halnya tingkat pendidikan dan banyaknya perusahaan. Selain itu, faktor kependudukan juga dapat berpengaruh. Semakin banyak penduduk di suatu wilayah, ada kemungkinan kesempatan untuk mendapat pekerjaan akan berkurang. Artinya, akibat dari pertumbuhan penduduk yang tinggi, kebutuhan akan penciptaan lapangan kerja lebih besar. Oleh karena itu, akan lebih mudah bagi pemerintah Aceh dalam menangani masalah pengangguran apabila diketahui hal-hal apa saja yang mempengaruhinya.

Beberapa penelitian sebelumnya yang berkaitan dengan masalah pengangguran telah dilakukan oleh Sari (2012), Ariane (2012), Putra (2013), dan Hajji (2013). Penelitian yang dilakukan oleh Sari (2012) adalah melakukan pemodelan faktor-faktor yang mempengaruhi TPT di Jawa Timur menggunakan pendekatan regresi spline multivariabel. Hasil penelitian menunjukkan bahwa dari tujuh variabel yang digunakan, tiga variabel yang signifikan mempengaruhi TPT di Jawa Timur adalah persentase penduduk usia kerja dengan pendidikan tertinggi SLTA, Angka Partisipasi Kasar (APK), dan tingkat investasi. Ariane (2012) meneliti tentang faktor-faktor yang berpengaruh terhadap TPT di Jawa Timur dan Jawa Tengah menggunakan metode regresi *ridge*. Hasil penelitian menunjukkan bahwa variabel yang mempengaruhi TPT di Jawa Timur adalah pertumbuhan penduduk, Tingkat Partisipasi Angkatan Kerja (TPAK), dan rata-rata lama sekolah, sedangkan variabel yang mempengaruhi TPT di Jawa

Tengah hanya persentase penduduk yang tinggal di perkotaan. Penelitian lain oleh Putra (2013) mendapatkan hasil bahwa pertumbuhan penduduk dan upah berpengaruh signifikan terhadap rasio pengangguran di Indonesia sedangkan inflasi tidak berpengaruh secara signifikan. Hajji (2013) juga meneliti tentang faktor yang mempengaruhi TPT di Indonesia. Hasil penelitian menyatakan bahwa Upah Minimum Provinsi (UMP) dan Angka Melek Huruf (AMH) mempengaruhi TPT di Indonesia.

Analisis regresi merupakan metode yang dapat menjelaskan hubungan antara variabel prediktor dan variabel responnya. Terdapat banyak pendekatan analisis regresi yang penggunaannya disesuaikan dengan bentuk kurva regresi. Beberapa diantaranya adalah regresi linier, kuadratik, dan kubik untuk regresi parametrik, serta regresi spline, kernel, dan wavelets untuk regresi nonparametrik. Pendekatan regresi nonparametrik digunakan apabila bentuk kurva regresi tidak diketahui. Regresi spline memiliki beberapa kelebihan, salah satunya adalah cenderung mencari sendiri estimasi data dari pola data yang terbentuk (Budiantara, 2009). Beberapa penelitian terdahulu yang menggunakan regresi spline adalah Mubarak (2012) yang menggunakan regresi spline multivariabel untuk pemodelan kematian penderita Demam Berdarah Dengue (DBD) di Jawa Timur. Selain itu, Dewi (2012) menganalisis faktor-faktor yang mempengaruhi angka gizi buruk di Jawa Timur menggunakan pendekatan regresi spline. Merdekawati (2013) melakukan pemodelan regresi spline *truncated* multivariabel pada faktor-faktor yang mempengaruhi kemiskinan di kabupaten/kota Provinsi Jawa Tengah. Marina (2013) menggunakan regresi semiparametrik spline untuk pemodelan faktor-faktor yang mempengaruhi persentase kriminalitas di Jawa Timur.

Pada penelitian ini akan dilakukan pemodelan tingkat pengangguran terbuka di Aceh dengan menggunakan pendekatan

regresi spline untuk mengetahui variabel apa saja yang mempengaruhinya. Metode analisis regresi spline digunakan karena lebih sederhana apabila dibandingkan dengan metode regresi nonparametrik yang lain. Selain itu, koefisien determinasi yang dihasilkan juga cenderung lebih tinggi. Hasil penelitian diharapkan dapat dijadikan salah satu bahan masukan untuk pemerintah dalam penentuan kebijakan terkait masalah pengangguran.

## **1.2 Rumusan Masalah**

Aceh termasuk salah satu provinsi dengan TPT tertinggi di Indonesia. Pada tahun 2013, Aceh merupakan provinsi dengan TPT tertinggi di Indonesia dan selama tiga tahun berturut-turut TPT Aceh lebih tinggi dari TPT nasional. Selain itu, dalam satu tahun fluktuasi TPT Aceh cukup besar dibandingkan provinsi lain. Untuk mengurangi TPT, perlu diselidiki faktor-faktor yang mempengaruhinya supaya dapat dilakukan tindakan terhadap faktor-faktor tersebut. Penelitian ini bertujuan menganalisis faktor-faktor yang mempengaruhi TPT. Oleh karena itu rumusan masalahnya adalah sebagai berikut:

1. Bagaimana karakteristik tingkat pengangguran terbuka di Aceh beserta karakteristik faktor-faktor yang diduga mempengaruhinya?
2. Bagaimana pemodelan tingkat pengangguran terbuka di Aceh dan variabel yang berpengaruh secara signifikan?

## **1.3 Tujuan Penelitian**

Berdasarkan perumusan masalah yang telah diuraikan, maka tujuan penelitian yang ingin dicapai adalah :

1. Mendeskripsikan karakteristik tingkat pengangguran terbuka di Aceh beserta karakteristik faktor-faktor yang diduga mempengaruhinya menggunakan statistik deskriptif.

2. Melakukan pemodelan tingkat pengangguran terbuka di Aceh dan menganalisis variabel yang berpengaruh secara signifikan.

#### **1.4 Manfaat Penelitian**

Manfaat yang akan diperoleh dari penelitian ini adalah :

1. Memberikan wawasan keilmuan yang berkaitan dengan regresi nonparametrik dan aplikasinya di bidang sosial.
2. Memberikan informasi mengenai pengangguran terbuka di Aceh dan variabel-variabel yang memengaruhinya yang dapat dijadikan bahan masukan kepada pemerintah Aceh dalam penentuan kebijakan terkait masalah pengangguran.

#### **1.5 Batasan Penelitian**

Pada penelitian ini terdapat beberapa batasan masalah yaitu :

1. Data yang digunakan adalah data indikator tingkat pengangguran terbuka Provinsi Aceh tahun 2013.
2. Faktor banyak lapangan kerja yang tersedia adalah banyak perusahaan di tiap kabupaten/kota.
3. Metode yang digunakan untuk memilih titik knot optimal adalah *Generalized Cross Validation (GCV)*.
4. Titik knot yang digunakan sebanyak 1, 2, 3 knot, dan kombinasi knot.

*(Halaman Ini Sengaja Dikosongkan)*

## BAB II TINJAUAN PUSTAKA

Terdapat beberapa pustaka yang digunakan dalam penyelesaian permasalahan pada penelitian ini. Pada bab ini akan dibahas materi mengenai statistik deskriptif, analisis regresi, regresi parametrik, regresi nonparametrik spline, asumsi residual yang harus terpenuhi, dan tinjauan mengenai pengangguran terbuka.

### 2.1 Statistik Deskriptif

Statistik deskriptif adalah metode-metode yang berkaitan dengan pengumpulan dan penyajian suatu gugus data sehingga memberikan informasi yang berguna. Statistik deskriptif memberikan informasi hanya mengenai data yang dipunyai dan sama sekali tidak menarik inferensia atau kesimpulan apapun tentang gugus data induknya yang lebih besar (Walpole, 1995). Gambaran umum dari data pada suatu kasus didapatkan melalui pengukuran pemusatan data, penyebaran data, dan menampilkan grafik atau diagram. Pengukuran pemusatan data dilakukan dengan menghitung nilai rata-rata (*mean*) dan pengukuran penyebaran data dilakukan dengan menghitung nilai deviasi standar.

*Mean* didapatkan dari membagi jumlah nilai pada data dengan banyaknya data tersebut. Rumus yang digunakan untuk menghitung *mean* data adalah :

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (2.1)$$

serta deviasi standar diberikan oleh :

$$S = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} \quad (2.2)$$

dengan  $x_i$  adalah data ke- $i$ ,  $\bar{x}$  adalah rata-rata, dan  $n$  adalah banyak data. Dalam memberi gambaran umum dari data, dapat disajikan pula menggunakan histogram.

## 2.2 Analisis Regresi

Analisis regresi merupakan analisis statistika yang digunakan untuk menyelidiki pola hubungan fungsional antara variabel respon dan variabel prediktor (Budiantara, 2009). Selain untuk menyelidiki pola hubungan, regresi juga dapat digunakan untuk peramalan variabel respon dengan nilai dari variabel prediktor yang telah diketahui. Dalam analisis regresi terdapat tiga pendekatan yaitu regresi parametrik, regresi nonparametrik, dan regresi semiparametrik. Penentuan pendekatan yang tepat dapat dilakukan dengan mempelajari pola hubungan antar variabel respon dan variabel prediktor menggunakan *scatter plot*. Apabila bentuk kurva yang dihasilkan diketahui, maka pendekatan yang tepat adalah regresi parametrik. Sedangkan apabila kurva tidak dapat ditentukan bentuk polanya, maka pendekatan yang lebih sesuai adalah regresi nonparametrik.

## 2.3 Regresi Parametrik

Regresi parametrik merupakan pendekatan yang digunakan apabila bentuk kurva regresi diketahui. Pendekatan model regresi parametrik sangat populer dan banyak disukai berbagai kalangan. Hal tersebut disebabkan oleh sifat regresi parametrik yang sangat baik dari pandangan statistika inferensi, seperti parsimoni, mudah diinterpretasikan, efisien, konsisten, dan estimatornya tidak bias (Budiantara, 2009). Secara umum model regresi parametrik linier dapat dituliskan sebagai berikut.

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_q x_{qi} + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (2.3)$$

dengan  $q$  adalah banyaknya variabel prediktor,  $n$  adalah banyaknya data, dan *error* random  $\varepsilon_i$  diasumsikan independen, identik, dan

mengikuti distribusi normal dengan *mean* nol dan varian  $\zeta^2$  (Eubank, 1999).

## 2.4 Regresi Nonparametrik Spline

Regresi nonparametrik merupakan metode yang dapat digunakan apabila pola data yang diketahui bentuk kurva regresinya atau tidak terdapat informasi masa lalu yang lengkap tentang bentuk pola data (Eubank, 1988). Menurut Budiantara (2009), biarkan data mencari bentuk estimasi dari kurva regresinya sendiri tanpa dipengaruhi oleh faktor subyektifitas peneliti. Spline memiliki fleksibilitas yang tinggi, mampu menangani data atau fungsi yang mulus (*smooth*), dan memiliki kemampuan yang sangat baik dalam menangani data yang perilakunya berubah-ubah pada sub interval tertentu. Oleh karena itu, spline digunakan dalam regresi nonparametrik. Bentuk model regresi nonparametrik secara umum disajikan sebagai berikut.

$$y_i = f(x_i) + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (2.4)$$

dengan  $y_i$  adalah variabel respon, fungsi  $f$  adalah kurva regresi yang tidak diketahui bentuknya,  $x_i$  adalah variabel prediktor ke- $i$ , dan  $\varepsilon_i$  adalah *error* random ke- $i$  (Eubank, 1999). Fungsi  $f(x_i)$  pada model regresi nonparametrik spline dihampiri dengan fungsi spline berorde  $p$  dan titik knot  $K_1, K_2, \dots, K_r$  yang dapat diberikan oleh persamaan :

$$f(x_i) = \sum_{j=0}^p \delta_j x_i^j + \sum_{k=1}^r \beta_k (x_i - K_k)_+^p \quad (2.5)$$

$$\text{dimana } (x_i - K_k)_+^p = \begin{cases} (x_i - K_k)^p, & x_i \geq K_k \\ 0, & x_i < K_k \end{cases} \quad (2.6)$$

Parameter  $\delta_j$  dan  $\beta_k$  merupakan parameter dari fungsi spline dengan  $j = 1, 2, \dots, p$  dan  $k = 1, 2, \dots, r$ , sedangkan  $K_k$  adalah titik knot. Titik knot adalah suatu titik pertemuan dua pola data

yang berbeda. Persamaan (2.4) dapat disajikan dalam bentuk matriks sebagai berikut.

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2.7)$$

dengan  $\mathbf{y}$  merupakan vektor variabel respon berukuran  $n \times 1$ ,  $\mathbf{X}$  merupakan matriks berukuran  $n \times (p + r + 1)$  yang terdiri dari  $p + r$  kolom variabel prediktor dan 1 kolom angka 1,  $\boldsymbol{\beta}$  merupakan parameter dari fungsi spline, dan  $\boldsymbol{\varepsilon}$  merupakan vektor *error* random berukuran  $n \times 1$ . Estimasi dari  $\boldsymbol{\beta}$  didapatkan melalui persamaan berikut.

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \quad (2.8)$$

## 2.5 Pemilihan Titik Knot Optimal

Sesuai dengan tujuan dari pendekatan regresi nonparametrik yaitu ingin mendapatkan kurva yang mulus dan mempunyai knot optimal menggunakan data amatan sebanyak  $n$ , maka diperlukan ukuran kinerja atas estimator yang dapat diterima secara universal (Eubank, 1988). Fungsi spline terbaik akan didapatkan jika titik knot optimal sudah diperoleh. Salah satu metode pemilihan titik knot optimal adalah menggunakan *Generalized Cross Validation* atau GCV (Budiantara, 2000). Titik knot yang optimal didapatkan dari nilai GCV yang paling kecil. Fungsi GCV dapat dituliskan sebagai berikut.

$$GCV(K_1, K_2, \dots, K_r) = \frac{MSE(K_1, K_2, \dots, K_r)}{(n^{-1} \text{trace}[I - \mathbf{A}(K_1, K_2, \dots, K_r)])^2} \quad (2.9)$$

dimana  $MSE(K_1, K_2, \dots, K_r) = n^{-1} \sum_{i=1}^n (y_i - \hat{y}_i)^2$  dan

$$\mathbf{A}(K_1, K_2, \dots, K_r) = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' .$$

## 2.6 Pengujian Parameter Model

Uji parameter model dilakukan untuk menentukan variabel prediktor yang berpengaruh terhadap variabel respon. Pada regresi spline, uji parameter model dilakukan setelah mendapatkan model regresi dengan titik knot optimal berdasarkan nilai GCV yang paling minimum. Terdapat dua tahap pengujian parameter yaitu pengujian secara serentak dan secara parsial.

### a. Uji Serentak

Uji serentak merupakan pengujian signifikansi seluruh parameter yang terdapat dalam model secara keseluruhan. Hipotesis yang digunakan untuk uji serentak sesuai dengan parameter fungsi regresi pada persamaan (2.5) adalah sebagai berikut.

$$H_0 : \delta_1 = \delta_2 = \dots = \delta_p = \beta_1 = \beta_2 = \dots = \beta_r = 0$$

$$H_1 : \text{Minimal ada satu } \delta_j \neq 0 \text{ atau } \beta_k \neq 0,$$

$$j = 1, 2, \dots, p \text{ dan } k = 1, 2, \dots, r$$

serta menggunakan nilai signifikansi  $\alpha$ . Statistik uji yang digunakan adalah :

$$F_{hitung} = \frac{MS_{regresi}}{MS_{error}} \quad (2.10)$$

Berikut adalah *Analysis of Variance* (ANOVA) untuk mendapatkan nilai  $MS_{regresi}$  dan  $MS_{error}$ .

**Tabel 2.1** *Analysis of Variance* (ANOVA)

Sumber Variasi	Derajat Bebas (df)	Jumlah Kuadrat (SS)	Rataan Kuadrat (MS)	$F_{hitung}$
Regresi	$p + r$	$\hat{\beta}'\mathbf{X}'\mathbf{y} - n\bar{y}^2$	$\frac{SS_{regresi}}{df_{regresi}}$	$MS_{regresi}$
Error	$n - (p + r + 1)$	$\mathbf{y}'\mathbf{y} - \hat{\beta}'\mathbf{X}'\mathbf{y}$	$\frac{SS_{error}}{df_{error}}$	$MS_{error}$
Total	$n - 1$	$\mathbf{y}'\mathbf{y} - n\bar{y}^2$	-	

Daerah penolakannya adalah tolak  $H_0$  apabila  $F_{hitung}$  lebih besar dari nilai  $F_{\alpha, (p+r, n-p-r-1)}$ . Jika  $H_0$  ditolak artinya terdapat minimal satu parameter pada model yang signifikan. Nilai koefisien determinasi ( $R^2$ ) didapatkan dari persamaan berikut.

$$R^2 = \frac{SS_{regresi}}{SS_{total}} \times 100\% \quad (2.11)$$

### b. Uji Parsial

Uji parsial dilakukan untuk mengetahui parameter yang berpengaruh signifikan terhadap model secara parsial. Pada pengujian ini dilakukan pengujian parameter polinomial ( $\delta$ ) dan pengujian parameter potongan polinomial ( $\beta$ ). Nilai signifikansi yang digunakan adalah  $\alpha$ . Hipotesis yang digunakan untuk pengujian parameter polinomial adalah sebagai berikut.

$$H_0 : \delta_j = 0$$

$$H_1 : \delta_j \neq 0, j = 1, 2, \dots, p$$

Statistik uji yang digunakan adalah :

$$t_{hitung} = \frac{\delta_j}{SE(\delta_j)}, j = 1, 2, \dots, p \quad (2.12)$$

Hipotesis yang digunakan untuk pengujian parameter potongan polinomial adalah sebagai berikut.

$$H_0 : \beta_k = 0$$

$$H_1 : \beta_k \neq 0, k = 1, 2, \dots, r$$

Statistik uji yang digunakan adalah :

$$t_{hitung} = \frac{\hat{\beta}_k}{SE(\hat{\beta}_k)}, k = 1, 2, \dots, r \quad (2.13)$$

Nilai standar *error* didapatkan dari akar varian  $\hat{\beta}$  sebagai berikut.

$$\begin{aligned} \text{var}(\hat{\beta}) &= \text{var}((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}) \\ &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\text{var}(\mathbf{y})(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \end{aligned}$$

$$\begin{aligned}
&= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\sigma^2((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')' \\
&= \sigma^2(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \\
&= \sigma^2(\mathbf{X}'\mathbf{X})^{-1}
\end{aligned}$$

serta nilai  $\sigma^2 = \text{MSE}$ .

Daerah penolakannya adalah tolak  $H_0$  apabila nilai  $|t_{hitung}|$  lebih besar dari nilai  $t_{(\alpha/2, n-(p+r)-1)}$ .

## 2.7 Asumsi Residual Model

Setelah mendapatkan model terbaik dari regresi spline, perlu dilakukan pengujian asumsi residual untuk mengetahui asumsi residual telah terpenuhi atau belum. Berikut adalah asumsi-asumsi yang harus dipenuhi.

### 2.7.1 Asumsi Residual Independen

Pemeriksaan asumsi residual independen digunakan untuk mengetahui adanya atau tidaknya autokorelasi antar residual. Salah satu cara untuk pemeriksaan asumsi residual independen adalah dengan menggunakan plot *Autocorrelation Function* (ACF). Persamaan yang digunakan untuk menghitung ACF adalah sebagai berikut (Wei, 2006).

$$\rho_h = \frac{\gamma_h}{\gamma_0} = \frac{\sum_{i=1}^{n-h} (e_i - \bar{e})(e_{i+h} - \bar{e})}{\sum_{i=1}^n (e_i - \bar{e})^2} \quad (2.14)$$

Interval konfidensi  $\rho_h$  diberikan oleh :

$$-t_{(\frac{\alpha}{2}, n-1)}SE(\hat{\rho}_h) \leq \rho_h \leq t_{(\frac{\alpha}{2}, n-1)}SE(\hat{\rho}_h) \quad (2.15)$$

dengan

$$SE(\rho_h) = \sqrt{\frac{1 + 2 \sum_{m=1}^{h-1} (\hat{\rho}_m)^2}{n}} \quad (2.16)$$

Apabila terdapat ACF di luar interval konfidensi, maka terdapat autokorelasi antar residual.

### 2.7.2 Asumsi Residual Identik

Uji asumsi identik digunakan untuk mengetahui homogenitas variansi residual. Jika asumsi ini tidak terpenuhi artinya terdapat heteroskedastisitas yang mengakibatkan kerugian bagi efisiensi estimator (Eubank, 1988). Salah satu cara untuk mendeteksi adanya heteroskedastisitas adalah menggunakan uji *glejser*. Uji *glejser* dilakukan dengan meregresikan absolut dari residual dengan variabel prediktornya (Gujarati, 2003). Hipotesis yang digunakan untuk uji *glejser* adalah sebagai berikut.

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$$

$$H_1 : \text{minimal ada satu } \sigma_i^2 \neq \sigma^2; i = 1, 2, \dots, n$$

serta menggunakan nilai signifikansi  $\alpha$ . Statistik uji yang digunakan adalah :

$$F_{hitung} = \frac{\left[ \sum_{i=1}^n (|\hat{e}_i| - |\bar{e}|)^2 \right] / (s-1)}{\left[ \sum_{i=1}^n (|e_i| - |\hat{e}_i|)^2 \right] / (n-s)} \quad (2.17)$$

Diambil keputusan tolak  $H_0$  apabila  $F_{hitung}$  lebih besar dari nilai  $F_{\alpha; (s-1, n-s)}$ . Nilai  $s$  merupakan banyaknya parameter dalam model. Apabila  $H_0$  ditolak, maka artinya terjadi kasus heteroskedastisitas, sehingga asumsi residual identik tidak terpenuhi.

### 2.7.3 Asumsi Normalitas Residual

Uji normalitas residual dilakukan untuk mengetahui residual mengikuti distribusi normal atau tidak. Pengujian ini menggunakan uji *Kolmogorov-Smirnov* dengan hipotesis sebagai berikut.

$H_0 : F_n(x) = F_0(x)$  (Residual mengikuti distribusi normal)

$H_1 : F_n(x) \neq F_0(x)$  (Residual tidak mengikuti distribusi normal)

serta menggunakan nilai signifikansi  $\alpha$ . Statistik uji yang digunakan adalah :

$$D = \text{maksimal}|F_n(x) - F_0(x)| \quad (2.18)$$

dimana  $F_n(x)$  adalah nilai distribusi kumulatif,  $F_0(x)$  adalah nilai distribusi kumulatif di bawah  $p(Z < Z_i)$ . Diambil keputusan tolak  $H_0$  jika  $D > D_{tabel}$  dengan nilai  $D_{tabel}$  didapatkan dari Tabel *Kolmogorov-Smirnov*. Selain dengan perhitungan nilai  $D$ , normalitas residual juga dapat ditunjukkan secara visual dari *probability plot*. Residual normal apabila titik-titik pada *probability plot* mengikuti garis distribusi normal.

## 2.8 Pengangguran Terbuka

Menurut BPS, bekerja merupakan kegiatan ekonomi yang dilakukan oleh seseorang dengan maksud memperoleh atau membantu memperoleh pendapatan atau keuntungan, paling sedikit satu jam (tidak terputus) dalam seminggu yang lalu. Pengangguran adalah suatu keadaan seseorang yang sudah termasuk angkatan kerja (15 tahun ke atas) dan ingin memperoleh pekerjaan tetapi belum mendapatkan pekerjaan. Terdapat beberapa jenis pengangguran, salah satunya adalah pengangguran terbuka. Penduduk yang termasuk dalam kategori pengangguran terbuka adalah penduduk yang termasuk angkatan kerja yang mencari pekerjaan, mempersiapkan usaha, tidak mencari pekerjaan karena merasa tidak mungkin mendapat pekerjaan, dan sudah punya pekerjaan tetapi belum mulai bekerja.

## 2.9 Tinjauan Pengangguran Terbuka

Di Indonesia masih banyak masyarakat yang masih hidup di bawah garis kemiskinan. Salah satu penyebab kemiskinan tersebut adalah tidak bekerjanya seseorang pada usia kerja.

Penduduk yang termasuk usia kerja sebenarnya adalah kelompok usia yang konsumtif dan belum produktif. Hal tersebut akan menghambat pertumbuhan karena sebagian besar pertambahan pendapatan akan habis dikonsumsi oleh penganggur dan perekonomian dapat mengalami kemunduran. Banyaknya pengangguran dapat disebabkan oleh kurangnya lapangan kerja sedangkan pencari kerja sangat banyak. Terkadang lapangan kerja sudah tersedia tetapi pendidikan pencari kerja tidak sesuai dengan yang diinginkan oleh lapangan kerja. Selain itu, Sudradjad (2000) mengungkapkan bahwa terdapat lima penyebab pengangguran. Penyebab yang pertama adalah warisan sifat feodal. Pada umumnya para penganggur kurang tertarik dengan pekerjaan swasta dengan alasan takut akan turunnya martabat di mata masyarakat. Penyebab lainnya adalah tidak ada motivasi untuk bekerja, lapangan kerja yang tersedia memerlukan kemampuan khusus, hilangnya kepercayaan diri setelah tidak diterima bekerja, dan pertumbuhan ekonomi. Krisis ekonomi berkepanjangan memberikan pengaruh pertumbuhan ekonomi yang kurang menguntungkan disertai dengan pertumbuhan penduduk yang cukup tinggi. Hal tersebut mengakibatkan beratnya pertumbuhan ekonomi sehingga tingkat investasi yang dilakukan pemerintah maupun swasta juga melambat. Besarnya lapangan kerja yang diciptakan dari investasi relatif kecil dibandingkan dengan jumlah pencari kerja.

Menurut Badan Pusat Statistik (2014), wilayah dengan jumlah penduduk yang besar akan dihadapkan pada persoalan meningkatnya jumlah pengangguran karena tidak memadainya penyediaan lapangan pekerjaan, permasalahan kebutuhan lahan untuk pemukiman, serta tidak memadainya akses fasilitas pendidikan dan kesehatan serta masalah-masalah sosial lainnya. Keterbatasan lapangan kerja yang tersediabagi penduduk yang mempunyai ijazah tinggi menyebabkan penduduk dengan ijazah tinggi tidak terserap pada lapangan usaha tersebut. Penduduk

yang merupakan lulusan pendidikan tinggi enggan menerima pekerjaan yang tidak sesuai dengan jenis keahliannya atau dengan jenjang pendidikan yang telah ditamatkan.

Terdapat beberapa penelitian mengenai faktor-faktor yang mempengaruhi pengangguran terbuka. Berdasarkan penelitian-penelitian terdahulu yang dilakukan oleh Sari (2012), Ariane (2012), Putra (2013), dan Hajji (2013), diketahui bahwa faktor yang berpengaruh terhadap pengangguran adalah kependudukan, pendidikan, ekonomi daerah, dan keadaan industri daerah.

*(Halaman Ini Sengaja Dikosongkan)*

## BAB III METODOLOGI PENELITIAN

### 3.1 Sumber Data dan Variabel Penelitian

Data yang digunakan dalam penelitian ini adalah data sekunder yang diambil dari Badan Pusat Statistik (BPS) Provinsi Aceh dan Susenas Kor 2013. Unit observasi yang digunakan adalah 23 kabupaten/kota yang ada di provinsi Aceh. Variabel-variabel yang digunakan dalam penelitian ini teringkas dalam Tabel 3.1.

**Tabel 3.1** Variabel Penelitian

Variabel	Keterangan
$y$	Tingkat Pengangguran Terbuka (TPT)
$x_1$	Kepadatan Penduduk
$x_2$	Angka Melek Huruf (AMH) Penduduk Usia Kerja
$x_3$	Rata-rata Lama Sekolah Penduduk Usia Kerja
$x_4$	Angka Partisipasi Kasar (APK) SLTA
$x_5$	Banyak Industri (Perusahaan)

Berikut adalah definisi operasional dari variabel penelitian.

- a. Variabel  $y$  merupakan variabel respon yang menunjukkan persentase pengangguran terbuka terhadap jumlah penduduk yang termasuk dalam angkatan kerja (TPT). Pengangguran terbuka merupakan kondisi penduduk yang sedang mencari kerja, atau sedang mempersiapkan usaha, atau merasa tidak mungkin mendapat pekerjaan, atau sudah punya pekerjaan tetapi belum mulai bekerja. Nilai TPT didapatkan dari :

$$TPT = \frac{\text{jumlah pengangguran terbuka}}{\text{jumlah angkatan kerja}} \times 100$$

- b. Variabel  $x_1$  merupakan variabel yang menyatakan kepadatan penduduk di setiap kabupaten/kota di Aceh. Nilai ini di-

dapatkan dengan membagi jumlah penduduk dengan luas wilayah.

$$\text{Kepadatan Penduduk} = \frac{\text{jumlah penduduk(jiwa)}}{\text{jumlah wilayah (km}^2\text{)}}$$

- c. Variabel  $x_2$  merupakan variabel yang menyatakan persentase jumlah penduduk yang mampu membaca dan menulis. Nilai ini didapatkan dari membagi jumlah penduduk usia kerja yang mampu membaca dan menulis dengan jumlah penduduk usia kerja kemudian dikalikan 100.

$$AMH = \frac{\text{jumlah penduduk usia kerja mampu baca tulis}}{\text{jumlah penduduk usia kerja}} \times 100$$

- d. Variabel  $x_3$  merupakan variabel yang menyatakan rata-rata lama sekolah penduduk usia kerja di suatu kabupaten/kota di Aceh.
- e. Variabel  $x_4$  merupakan variabel yang menunjukkan tingkat partisipasi penduduk di jenjang SLTA yang didapatkan dari membagi jumlah penduduk yang sedang bersekolah di jenjang SLTA dengan jumlah penduduk kelompok usia SLTA (16-18 tahun).

$$APK(SLTA) = \frac{\text{jumlah penduduk sedang menempuh SLTA}}{\text{jumlah penduduk usia SLTA}} \times 100$$

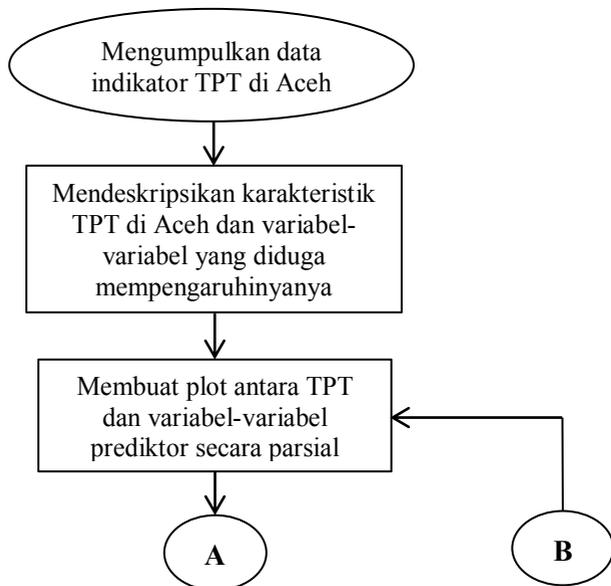
- f. Variabel  $x_5$  merupakan variabel yang menyatakan banyak unit perusahaan yang ada di suatu kabupaten/kota di Aceh.

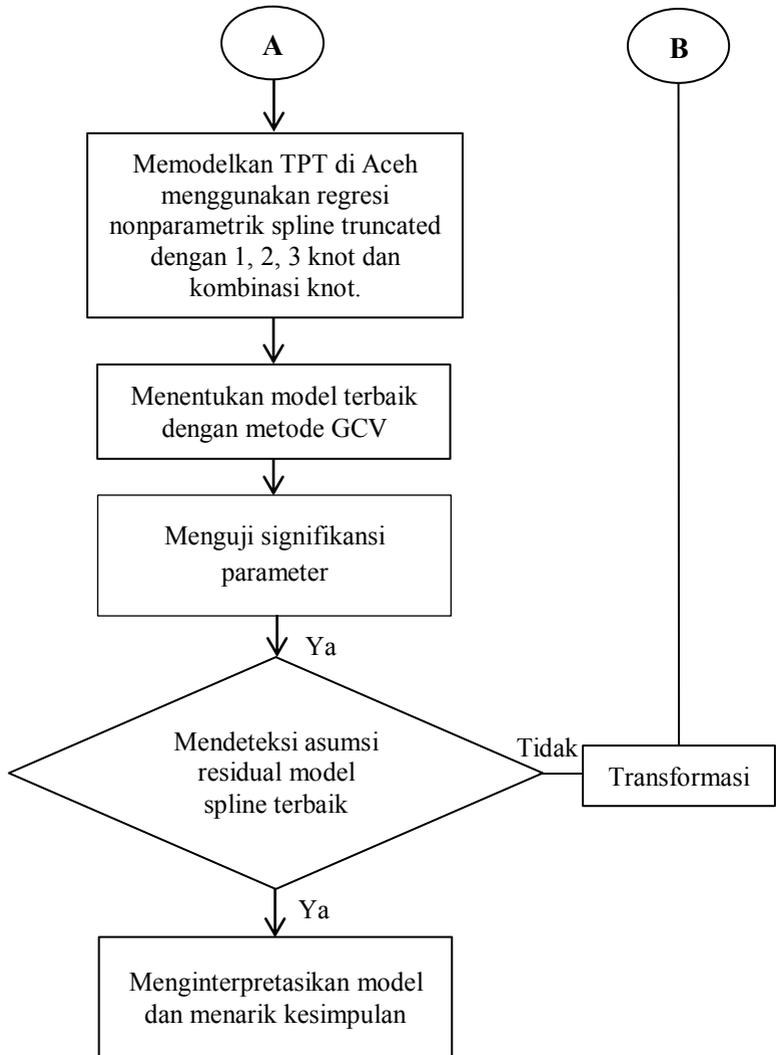
### 3.2 Langkah Penelitian

Langkah awal yang dilakukan adalah mendeskripsikan karakteristik tingkat pengangguran terbuka dan faktor-faktor yang diduga berpengaruh berdasarkan kabupaten/kota di Aceh berupa grafik atau diagram. Selanjutnya melakukan analisis untuk mengetahui pengaruh variabel-variabel prediktor terhadap variabel respon menggunakan regresi nonparametrik spline. Rincian tahapan analisis adalah sebagai berikut.

1. Membuat plot antara tingkat pengangguran terbuka dan variabel-variabel prediktor.
2. Memodelkan tingkat pengangguran terbuka di Aceh menggunakan regresi nonparametrik spline dengan 1, 2, 3 knot dan kombinasi knot.
3. Menentukan model terbaik dengan menggunakan metode GCV.
4. Melakukan uji signifikansi parameter secara serentak dan parsial.
5. Mendeteksi asumsi residual model spline terbaik. Apabila residual model spline terbaik tidak memenuhi asumsi, maka dilakukan transformasi pada data. Kemudian melakukan kembali mulai langkah (2).
6. Menginterpretasikan model dan menarik kesimpulan.

Langkah analisis dapat disajikan secara ringkas dalam diagram alir sebagai berikut.





**Gambar 3.1** Diagram Alir Penelitian



## BAB IV

### ANALISIS DAN PEMBAHASAN

Pada bab ini akan dibahas analisa yang dilakukan mengenai Tingkat Pengangguran Terbuka (TPT) di Aceh beserta faktor-faktor yang diduga mempengaruhinya. Analisa yang dilakukan adalah menjelaskan karakteristik TPT di Aceh dan faktor-faktor yang diduga berpengaruh menggunakan statistik deskriptif serta memodelkannya dengan pendekatan regresi nonparametrik spline.

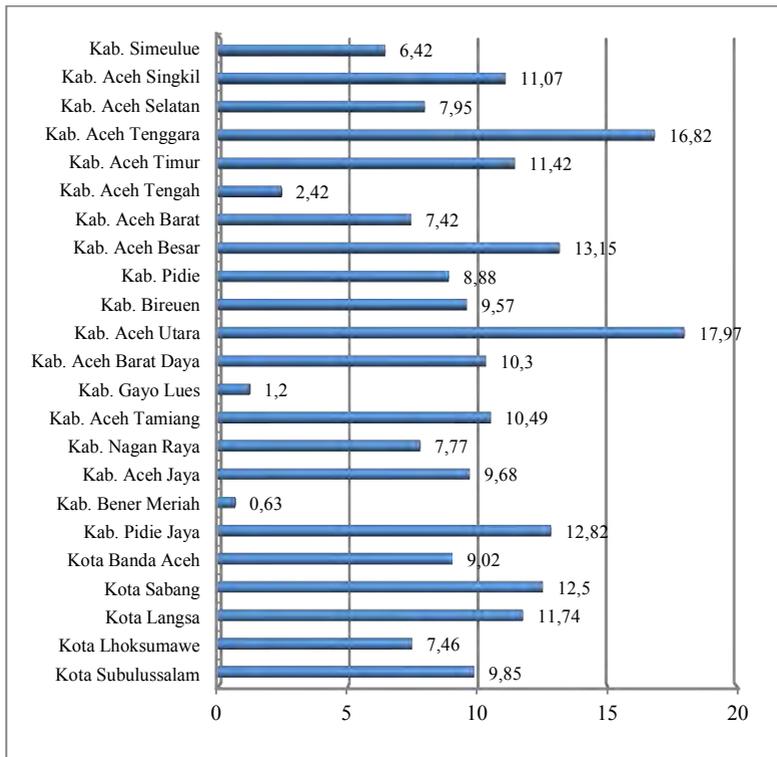
#### 4.1 Karakteristik TPT di Aceh dan Faktor-Faktor yang Diduga Mempengaruhinya

Provinsi Aceh merupakan provinsi dengan TPT tertinggi di Indonesia pada tahun 2013. Kemudian pada tahun 2014 TPT di Aceh menempati peringkat kelima di Indonesia. Walaupun tidak berada di peringkat pertama, tetapi TPT di Aceh masih cenderung tinggi. Selama 3 tahun berturut-turut TPT di Aceh berada di atas garis TPT nasional . Berdasarkan data tersebut diketahui bahwa TPT merupakan permasalahan sosial yang cukup memprihatinkan di Aceh. Dalam penelitian ini, digunakan lima variabel yang diduga mempengaruhi TPT di Aceh. Karakteristik TPT dan lima variabel yang diduga mempengaruhi terdapat pada Tabel 4.1.

**Tabel 4.1** Statistik Deskriptif TPT di Aceh dan Lima Variabel yang Diduga Mempengaruhinya

Variabel	Rata-Rata	Standar Deviasi	Maksimum	Minimum
$y$	9,42	4,23	17,97	0,63
$x_1$	361,26	931,64	4451	15
$x_2$	96,42	2,69	99,79	90,37
$x_3$	9,21	1,04	12,27	7.66
$x_4$	78,512	11,62	98,25	59,84
$x_5$	1,91	2,56	12	0

Diketahui rata-rata TPT di Aceh sebesar 9,42 persen. Nilai tersebut dapat dikatakan tinggi karena dengan kata lain untuk setiap 10000 penduduk Aceh diperkirakan terdapat 942 orang yang termasuk pengangguran terbuka. Selain itu diketahui pula bahwa standar deviasi TPT di Aceh sebesar 4,23. Artinya keragaman data cukup besar. Keragaman tersebut dapat terlihat secara visual pada Gambar 4.1.



**Gambar 4.1** TPT di Aceh Berdasarkan Kabupaten/Kota pada Tahun 2013

Berdasarkan Gambar 4.1, Kabupaten Aceh Utara merupakan kabupaten/kota dengan TPT tertinggi di Aceh yaitu sebesar 17,97 persen dengan kata lain untuk setiap 10000 penduduk, terdapat 1797 orang termasuk pengangguran terbuka. Setelah Kabupaten

Aceh Utara, TPT tertinggi kedua ditempati oleh Kabupaten Aceh Tenggara sebesar 16,82 persen yang artinya untuk setiap 10000 penduduk, terdapat 1682 orang yang termasuk pengangguran terbuka. Kabupaten/kota dengan TPT paling rendah adalah Kabupaten Bener Meriah dengan TPT sebesar 0,63 persen, artinya untuk setiap 10000 penduduk terdapat 63 orang yang termasuk pengangguran terbuka. Dapat terlihat keragaman TPT yang besar di Aceh. Hal ini dapat disebabkan oleh keadaan di masing-masing wilayah.

Menurut teori, semakin padat penduduk di suatu wilayah maka semakin tinggi pula TPT di wilayah tersebut. Rata-rata kepadatan penduduk ( $x_1$ ) di Aceh adalah 361,26. Hal tersebut menunjukkan bahwa rata-rata perkilometer persegi wilayah di kabupaten/kota di Aceh dihuni oleh 362 orang. Standar deviasi dari data kepadatan penduduk kabupaten/kota di Aceh adalah sebesar 931,64 yang berarti keragaman data sangat besar. Wilayah dengan kepadatan penduduk tertinggi adalah Kota Banda Aceh yaitu 4451 orang perkilometer persegi wilayahnya. Sementara itu kepadatan penduduk paling rendah ditempati oleh Kabupaten Gayo Lues dengan 15 orang perkilometer persegi wilayahnya. Kepadatan penduduk di kota cenderung lebih tinggi dibandingkan dengan di kabupaten. Kebanyakan orang lebih memilih untuk tinggal di kota dibandingkan di kabupaten. Hal ini terjadi karena perkembangan di kota lebih pesat daripada kabupaten.

Variabel lain yang diduga mempengaruhi TPT di Aceh adalah Angka Melek Huruf ( $x_2$ ). Rata-rata 96,42 persen dari total penduduk di tiap kabupaten/kota di Aceh sudah melek huruf. Standar deviasi data Angka Melek Huruf (AMH) di Aceh cukup kecil yaitu sebesar 2,69. AMH tertinggi ditempati oleh Kabupaten Simeulue dengan AMH sebesar 99,79 persen. Artinya 99,42 persen dari total penduduk di Kabupaten Simeulue sudah melek

huruf. Sementara itu AMH paling rendah ditempati oleh Kabupaten Gayo Lues dengan AMH sebesar 90,37 persen yang artinya 90,37 persen dari total penduduk di Kabupaten Gayo Lues sudah melek huruf.

Rata-rata lama sekolah ( $x_3$ ) juga diduga mempengaruhi TPT di Aceh. Variabel rata-rata lama sekolah untuk 23 kabupaten/kota di Aceh memiliki rata-rata 9,21 tahun. Artinya untuk setiap kabupaten/kota rata-rata penduduk wilayah tersebut menjalani sekolah selama 9 sampai 10 tahun. Wilayah dengan rata-rata lama sekolah paling besar adalah Kota Banda Aceh sedangkan wilayah dengan rata-rata lama sekolah paling kecil adalah Kota Subulussalam.

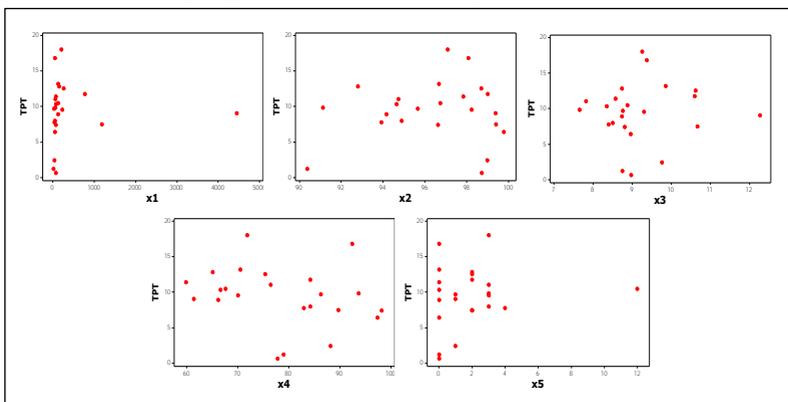
Selain itu, Angka Partisipasi Kasar (APK) jenjang SLTA ( $x_4$ ) juga diduga mempengaruhi TPT di Aceh. Diketahui bahwa pada setiap kabupaten/kota terdapat rata-rata 78,52 persen dari total penduduk yang tergolong umur SLTA (16-18 tahun) wilayah tersebut menjalani pendidikan jenjang SLTA. Standar deviasi data APK SLTA cukup besar yaitu 11,62. Wilayah dengan APK SLTA tertinggi adalah Kabupaten Aceh Barat yaitu 98,25 persen sedangkan wilayah dengan APK SLTA paling rendah adalah Kabupaten Aceh Timur yaitu 59,84 persen.

Selain variabel-variabel yang telah disebutkan, banyaknya perusahaan pada suatu wilayah juga diduga mempengaruhi TPT di Aceh. Rata-rata terdapat 2 perusahaan di setiap wilayah. Tetapi standar deviasi data banyak perusahaan sangat besar yaitu 2,56. Wilayah dengan perusahaan terbanyak adalah Kabupaten Aceh Tamiang yaitu sebanyak 12 perusahaan, jauh lebih banyak dibandingkan dengan wilayah-wilayah lainnya. Wilayah dengan banyak perusahaan terbanyak kedua adalah Kabupaten Nagan Raya yaitu sebanyak 4 perusahaan, berbeda cukup jauh dengan Kabupaten Aceh Tamiang. Terdapat 8 wilayah yang tidak terdapat perusahaan sama sekali yaitu Kabupaten Simeulue,

Kabupaten Aceh Tenggara, Kabupaten Aceh Timur, Kabupaten Aceh Besar, Kabupaten Pidie, Kabupaten Aceh Barat Daya, Kabupaten Gayo Lues, dan Kabupaten Bener Meriah.

#### 4.2 Pola Hubungan Antara TPT dengan Faktor-Faktor yang Diduga Mempengaruhinya

Terdapat tiga pendekatan dalam analisis regresi yaitu parametrik, nonparametrik, dan semiparametrik. Pendekatan yang tepat digunakan dapat ditentukan dengan melihat pola hubungan antara variabel respon dan variabel-variabel prediktornya. Maka untuk mengetahui pendekatan yang tepat untuk penelitian ini, sebaiknya dicari tahu terlebih dahulu pola hubungan antara TPT dengan variabel-variabel yang diduga mempengaruhinya. Pola hubungan tersebut dapat diketahui dengan menggunakan plot sebaran data (*scatterplot*). Apabila sebaran data membentuk pola tertentu, maka pendekatan yang digunakan adalah parametrik. Apabila sebaran data tidak membentuk pola tertentu, maka pendekatan yang digunakan adalah nonparametrik. Apabila sebaran data beberapa variabel pada suatu penelitian membentuk suatu pola dan beberapa variabel tidak membentuk pola, maka pendekatan yang digunakan adalah semiparametrik.



**Gambar 4.2** Scatterplot Antara TPT di Aceh dan Variabel-Variabel yang Diduga Mempengaruhinya

Gambar 4.2 memperlihatkan plot sebaran data antara TPT di Aceh dengan variabel-variabel yang diduga mempengaruhinya. Diketahui bahwa hubungan antara TPT di Aceh dengan lima variabel yang diduga mempengaruhinya yaitu kepadatan penduduk ( $x_1$ ), AMH ( $x_2$ ), rata-rata lama sekolah ( $x_3$ ), APK SLTA ( $x_4$ ), dan banyak perusahaan ( $x_5$ ) tidak membentuk pola tertentu. Maka dari itu akan digunakan pendekatan nonparametrik.

### 4.3 Model Regresi Nonparametrik Spline

Berdasarkan *scatterplot* antara TPT di Aceh dengan variabel-variabel yang diduga mempengaruhinya, seluruh *scatterplot* tidak membentuk pola tertentu. Maka dapat diketahui bahwa pendekatan regresi yang lebih sesuai adalah pendekatan nonparametrik. Pada penelitian ini, analisis dilakukan menggunakan regresi nonparametrik spline linier. Model regresi nonparametrik spline linier TPT di Aceh dan lima variabel yang diduga mempengaruhinya adalah sebagai berikut.

$$\begin{aligned}
 y_i &= \alpha_0 + f(x_{i1}) + f(x_{i2}) + f(x_{i3}) + f(x_{i4}) + f(x_{i5}) + \varepsilon_i \\
 &= \alpha_0 + \sum_{j=0}^p \alpha_j x_{i1}^j + \sum_{k=1}^r \beta_k (x_{i1} - k_k)_+^p + \sum_{j=0}^p \alpha_j x_{i2}^j + \\
 &\quad \sum_{k=1}^r \beta_k (x_{i2} - k_k)_+^p + \sum_{j=0}^p \alpha_j x_{i3}^j + \sum_{k=1}^r \beta_k (x_{i3} - k_k)_+^p + \\
 &\quad \sum_{j=0}^p \alpha_j x_{i4}^j + \sum_{k=1}^r \beta_k (x_{i4} - k_k)_+^p + \sum_{j=0}^p \alpha_j x_{i5}^j + \\
 &\quad \sum_{k=1}^r \beta_k (x_{i5} - k_k)_+^p + \varepsilon_i
 \end{aligned}$$

Model yang digunakan adalah regresi nonparametrik spline linier dengan  $p = 1$ . Banyak knot yang dicobakan adalah satu titik knot, dua titik knot, dan tiga titik knot ( $r = 1, 2, 3$ ).

#### 4.4 Pemilihan Titik Knot Optimal

Titik knot adalah suatu titik pertemuan dua pola data yang berbeda. Pemilihan titik knot optimal dilakukan dengan melihat nilai GCV yang paling kecil. Pada penelitian ini digunakan beberapa kemungkinan titik knot yaitu satu titik knot, dua titik knot, tiga titik knot, dan kombinasi ketiganya.

##### 4.4.1 Pemilihan Titik Knot Optimal dengan Satu Titik Knot

Suatu analisis sebaiknya mengutamakan model yang paling sederhana. Kelima variabel yang diduga mempengaruhi TPT di Aceh menggunakan regresi nonparametrik spline dicoba dimodelkan dengan satu titik knot terlebih dahulu. Berikut adalah model regresi nonparametrik spline dengan satu titik knot.

$$y_i = \alpha_0 + \alpha_1 x_{i1} + \beta_1 (x_{i1} - k_1)_+ + \alpha_2 x_{i2} + \beta_2 (x_{i2} - k_1)_+ + \alpha_3 x_{i3} + \beta_3 (x_{i3} - k_1)_+ + \alpha_4 x_{i4} + \beta_4 (x_{i4} - k_1)_+ + \alpha_5 x_{i5} + \beta_5 (x_{i5} - k_1)_+ + \varepsilon_i$$

Tabel 4.2 berikut menyajikan nilai GCV yang dihasilkan oleh regresi nonparametrik spline dengan satu titik knot.

**Tabel 4.2** Nilai GCV Regresi Nonparametrik Spline dengan Satu Titik Knot

Knot					GCV
$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	
2821,449	96,330	10,577	84,140	7,592	34,960
2911,980	96,522	10,671	84,924	7,837	30,591
3002,510	96,714	10,765	85,708	8,082	30,598
3093,041	96,906	10,859	86,492	8,327	29,916
3183,571	97,099	10,953	87,276	8,571	29,050
<b>3274,102</b>	<b>97,291</b>	<b>11,047</b>	<b>88,060</b>	<b>8,816</b>	<b>28,793</b>
3364,633	97,483	11,141	88,843	9,061	28,910
3455,163	97,675	11,235	89,627	9,306	29,287
3545,694	97,868	11,329	90,411	9,551	30,141
3636,224	98,060	11,423	91,195	9,796	31,212

Berdasarkan Tabel 4.2, nilai GCV terkecil yang diperoleh adalah 28,793. Nilai tersebut dihasilkan pada titik knot 3274,102 untuk variabel kepadatan penduduk ( $x_1$ ), titik knot 97,291 untuk variabel AMH ( $x_2$ ), titik knot 11,047 untuk variabel rata-rata lama sekolah ( $x_3$ ), titik knot 88,06 untuk variabel APK SLTA ( $x_4$ ), dan titik knot 8,816 untuk variabel banyak perusahaan ( $x_5$ ).

#### 4.4.2 Pemilihan Titik Knot Optimal dengan Dua Titik Knot

Pemodelan menggunakan dua titik knot juga dilakukan untuk perbandingan model yang lebih baik. Pemodelan variabel-variabel yang diduga mempengaruhi TPT di Aceh menggunakan regresi nonparametrik spline dengan dua titik knot adalah sebagai berikut.

$$y_i = \alpha_0 + \alpha_1 x_{i1} + \beta_1 (x_{i1} - k_1)_+ + \beta_2 (x_{i1} - k_2)_+ + \alpha_2 x_{i2} + \beta_3 (x_{i2} - k_1)_+ + \beta_4 (x_{i2} - k_2)_+ + \alpha_3 x_{i3} + \beta_5 (x_{i3} - k_1)_+ + \beta_6 (x_{i3} - k_2)_+ + \alpha_4 x_{i4} + \beta_7 (x_{i4} - k_1)_+ + \beta_8 (x_{i4} - k_2)_+ + \alpha_5 x_{i5} + \beta_9 (x_{i5} - k_1)_+ + \beta_{10} (x_{i5} - k_2)_+ + \varepsilon_i$$

Nilai GCV yang dihasilkan oleh regresi nonparametrik spline dengan dua titik knot disajikan dalam Tabel 4.3 berikut.

**Tabel 4.3** Nilai GCV yang Regresi Nonparametrik Spline dengan Dua Titik Knot

Knot					GCV
$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	
3455,163	97,675	11,235	89,627	9,306	11,221
3636,224	98,060	11,423	91,195	9,796	
<b>3455,163</b>	<b>97,675</b>	<b>11,235</b>	<b>89,627</b>	<b>9,306</b>	<b>9,508</b>
<b>3726,755</b>	<b>98,252</b>	<b>11,517</b>	<b>91,979</b>	<b>10,041</b>	
3455,163	97,675	11,235	89,627	9,306	9,837
3817,286	98,444	11,611	92,763	10,286	

**Tabel 4.3** Nilai GCV Regresi Nonparametrik Spline dengan Dua Titik Knot (lanjutan)

Knot					GCV
$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	
3455,163	97,675	11,235	89,627	9,306	13,174
3907,816	98,637	11,706	93,547	10,531	
3455,163	97,675	11,235	89,627	9,306	14,196
3998,347	98,829	11,800	94,331	10,776	
3455,163	97,675	11,235	89,627	9,306	15,578
4088,878	99,021	11,894	95,114	11,020	
3455,163	97,675	11,235	89,627	9,306	13,443
4179,408	99,213	11,988	95,898	11,265	
3455,163	97,675	11,235	89,627	9,306	11,102
4269,939	99,406	12,082	96,682	11,510	
3455,163	97,675	11,235	89,627	9,306	11,058
4360,469	99,598	12,176	97,466	11,755	
3455,163	97,675	11,235	89,627	9,306	29,287
4451,000	99,790	12,270	98,250	12	

Berdasarkan Tabel 4.3, nilai GCV terkecil yang diperoleh dari pemodelan regresi nonparametrik spline dengan dua titik knot adalah 9,508. Nilai GCV tersebut diperoleh pada titik knot 3455,163 dan 3726,755 untuk variabel  $x_1$ , titik knot 97,675 dan 98,252 untuk variabel  $x_2$ , titik knot 11,235 dan 11,517 untuk variabel  $x_3$ , titik knot 89,627 dan 91,979 untuk variabel  $x_4$ , dan titik knot 9,306 dan 10,041 untuk variabel  $x_5$ . Apabila dibandingkan dengan nilai GCV terkecil yang dihasilkan oleh satu knot, nilai GCV terkecil yang dihasilkan oleh dua knot lebih kecil. Maka regresi nonparametrik spline dengan dua titik knot lebih baik daripada dengan satu titik knot saja.

#### 4.4.3 Pemilihan Titik Knot Optimal dengan Tiga Titik Knot

Selain memodelkan data dengan titik knot sebanyak satu dan dua, data juga dimodelkan dengan menggunakan tiga titik knot. Pemodelan variabel-variabel yang diduga mempengaruhi TPT di Aceh menggunakan regresi nonparametrik spline dengan tiga titik knot adalah sebagai berikut.

$$\begin{aligned}
 y_i = & \alpha_0 + \alpha_1 x_{i1} + \beta_1 (x_{i1} - k_1)_+ + \beta_2 (x_{i1} - k_2)_+ + \beta_3 (x_{i1} - k_3)_+ + \\
 & \alpha_2 x_{i2} + \beta_4 (x_{i2} - k_1)_+ + \beta_5 (x_{i2} - k_2)_+ + \beta_6 (x_{i1} - k_3)_+ + \alpha_3 x_{i3} + \\
 & \beta_7 (x_{i3} - k_1)_+ + \beta_8 (x_{i3} - k_2)_+ + \beta_9 (x_{i1} - k_3)_+ + \alpha_4 x_{i4} + \\
 & \beta_{10} (x_{i4} - k_1)_+ + \beta_{11} (x_{i4} - k_2)_+ + \beta_{12} (x_{i1} - k_3)_+ + \alpha_5 x_{i5} + \\
 & \beta_{13} (x_{i5} - k_1)_+ + \beta_{14} (x_{i5} - k_2)_+ + \beta_{15} (x_{i1} - k_3)_+ + \varepsilon_i
 \end{aligned}$$

Nilai GCV yang dihasilkan oleh regresi nonparametrik spline dengan tiga titik knot disajikan dalam Tabel 4.4 berikut.

**Tabel 4.4** Nilai GCV Regresi Nonparametrik Spline dengan Tiga Titik Knot

Knot					GCV
$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	
377,122	91,139	8,036	62,976	0,980	25,865
3545,694	97,868	11,329	90,411	9,551	
3907,816	98,637	11,706	93,547	10,531	
377,122	91,139	8,036	62,976	0,980	29,251
3545,694	97,868	11,329	90,411	9,551	
3998,347	98,829	11,800	94,331	10,776	
377,122	91,139	8,036	62,976	0,980	34,434
3545,694	97,868	11,329	90,411	9,551	
4088,878	99,021	11,894	95,114	11,020	

**Tabel 4.4** Nilai GCV Regresi Nonparametrik Spline dengan Tiga Titik Knot (lanjutan)

Knot					GCV
$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	
377,122	91,139	8,036	62,976	0,980	29,811
3545,694	97,868	11,329	90,411	9,551	
4179,408	99,213	11,988	95,898	11,265	
377,122	91,139	8,036	62,976	0,980	24,075
3545,694	97,868	11,329	90,411	9,551	
4269,939	99,406	12,082	96,682	11,510	
377,122	91,139	8,036	62,976	0,980	23,948
3545,694	97,868	11,329	90,411	9,551	
4360,469	99,598	12,176	97,466	11,755	
<b>377,122</b>	<b>91,139</b>	<b>8,036</b>	<b>62,976</b>	<b>0,980</b>	<b>6,846</b>
<b>3636,224</b>	<b>98,060</b>	<b>11,423</b>	<b>91,195</b>	<b>9,796</b>	
<b>3726,755</b>	<b>98,252</b>	<b>11,517</b>	<b>91,979</b>	<b>10,041</b>	
377,122	91,139	8,036	62,976	0,980	15,026
3636,224	98,060	11,423	91,195	9,796	
3817,286	98,444	11,611	92,763	10,286	
377,122	91,139	8,036	62,976	0,980	30,674
3636,224	98,060	11,423	91,195	9,796	
3907,816	98,637	11,706	93,547	10,531	
377,122	91,139	8,036	62,976	0,980	33,941
3636,224	98,060	11,423	91,195	9,796	
3998,347	98,829	11,800	94,331	10,776	

Berdasarkan Tabel 4.4, dapat diketahui bahwa nilai GCV terkecil yang diperoleh dari pemodelan regresi nonparametrik spline dengan tiga titik knot adalah 6,846. Nilai GCV tersebut diperoleh pada titik knot 377,122 , 3636,224 dan 3726, 755 untuk variabel  $x_1$ , titik knot 91,139 , 98,060 dan 98,252 untuk variabel  $x_2$ , titik knot 8,036 , 11,423 dan 11,517 untuk variabel  $x_3$ , titik knot

62,976 , 91,195 dan 91,979 untuk variabel  $x_4$ , dan titik knot 0,98 , 9,796 dan 10,041 untuk variabel  $x_5$ . Nilai GCV yang dihasilkan dengan menggunakan tiga titik knot lebih kecil daripada nilai GCV yang dihasilkan menggunakan satu dan dua titik knot.

#### 4.4.4 Pemilihan Titik Knot Optimal dengan Kombinasi Knot

Telah didapatkan titik knot optimum dari pemodelan dengan satu titik knot, dua titik knot, dan tiga titik knot. Tetapi untuk setiap variabel tidak diharuskan memiliki jumlah titik knot yang sama. Oleh karena itu data dimodelkan dengan mengkombinasikan jumlah knot tersebut. Sehingga dengan mengkombinasikan jumlah titik knot diharapkan menghasilkan model yang terbaik. Nilai GCV dengan kombinasi knot disajikan dalam Tabel 4.5.

**Tabel 4.5** Nilai GCV Regresi Nonparametrik Spline dengan Kombinasi Knot

Knot					GCV
$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	
377,122	91,139				
3636,224	98,060	11,047	89,627	9,306	5,455
3726,755	98,252		91,979	10,041	
<b>377,122</b>	<b>91,139</b>				
<b>3636,224</b>	<b>98,060</b>	<b>11,047</b>	<b>89,627</b>	<b>9,796</b>	<b>5,265</b>
<b>3726,755</b>	<b>98,252</b>		<b>91,979</b>	<b>10,041</b>	
377,122	91,139		62,976		
3636,224	98,060	11,047	91,195	8,816	6,266
3726,755	98,252		91,979		
377,122	91,139		62,976		
3636,224	98,060	11,047	91,195	9,306	6,266
3726,755	98,252		91,979	10,041	
3455,163	97,675		89,627		
3726,755	98,252	11,047	91,979	8,816	9,508

**Tabel 4.5** Nilai GCV Regresi Nonparametrik Spline dengan Kombinasi Knot (lanjutan)

Knot					GCV
$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	
377,122	91,139	11,235	88,060	8,816	16,833
3636,224	98,060	11,517			
3726,755	98,252				
377,122	91,139	11,235	88,060	9,306	16,833
3636,224	98,060	11,517		10,041	
3726,755	98,252				
377,122	91,139	11,235	88,060	0,980	20,334
3636,224	98,060	11,517		9,796	
3726,755	98,252			10,041	
377,122	91,139	11,235	89,627	8,816	5,455
3636,224	98,060	11,517	91,979		
3726,755	98,252				
377,122	97,675	11,047	89,627	8,816	8,434
3636,224	98,252		91,979		
3726,755					

Berdasarkan Tabel 4.5, nilai GCV terkecil dengan kombinasi knot adalah 5,265. Nilai GCV tersebut dihasilkan apabila digunakan tiga titik knot untuk variabel  $x_1$ ,  $x_2$ , dan  $x_5$ , satu titik knot untuk variabel  $x_3$ , serta dua titik knot untuk variabel  $x_4$ . Lokasi titik knot untuk variabel  $x_1$  adalah  $k_1 = 377,122$ ,  $k_2 = 3636,224$ , dan  $k_3 = 3726,755$ . Lokasi titik knot untuk variabel  $x_2$  adalah  $k_1 = 91,139$ ,  $k_2 = 98,06$ ,  $k_3 = 98,252$ . Lokasi titik knot untuk variabel  $x_3$  adalah 11,047. Lokasi titik knot untuk variabel  $x_4$  adalah  $k_1 = 89,627$  dan  $k_2 = 91,979$ . Lokasi titik knot untuk variabel  $x_5$  adalah  $k_1 = 0,98$ ,  $k_2 = 9,796$ ,  $k_3 = 10,041$ .

#### 4.5 Pemilihan Model Terbaik

Telah didapatkan knot-knot optimal dengan menggunakan satu knot, dua knot, tiga knot, serta kombinasi knot. Selanjutnya akan dipilih model terbaik dari keempat model optimal yang sudah didapat dari tiap jumlah knot. Pemilihan model terbaik dilakukan dengan membandingkan nilai GCV terkecil dari masing-masing model. Tabel 4.6 menunjukkan nilai GCV terkecil dari masing-masing model.

**Tabel 4.6** Nilai GCV Terkecil dari Pemodelan Tiap Jumlah Knot

Jumlah Knot	GCV	Jumlah Parameter
1 Titik Knot	28,79309	11
2 Titik Knot	9,508306	16
3 Titik Knot	6,845799	21
Kombinasi (33123)	5,264601	18

Pada Tabel 4.6 dapat diketahui bahwa pemodelan yang menghasilkan nilai GCV yang paling kecil adalah dengan menggunakan kombinasi jumlah knot yaitu tiga titik knot untuk variabel  $x_1$ ,  $x_2$ , dan  $x_5$ , satu titik knot untuk variabel  $x_3$ , serta dua titik knot untuk variabel  $x_4$ . Pemodelan menggunakan tiga titik knot menghasilkan nilai GCV yang kecil pula dan tidak berbeda jauh dengan kombinasi knot. Namun, jumlah parameter yang digunakan dalam model kombinasi knot lebih sedikit sehingga model lebih sederhana. Oleh karena itu diputuskan bahwa model terbaik yang dipilih adalah model regresi nonparametrik spline dengan menggunakan kombinasi jumlah knot (33123).

#### 4.6 Penaksiran Parameter Model Regresi Nonparametrik Spline

Berdasarkan kriteria pemilihan model terbaik, dipilih model regresi nonparametrik spline dengan tiga titik knot untuk variabel  $x_1$ ,  $x_2$ , dan  $x_5$ , satu titik knot untuk variabel  $x_3$ , serta dua titik knot

untuk variabel  $x_4$ . Estimasi parameter model didapatkan menggunakan metode OLS (*Ordinary Least Square*). Berikut adalah model regresi nonparametrik spline terbaik.

$$\begin{aligned}
 y = & -1,7303 + 0,0457x_1 - 0,0635(x_1 - 377,122)_+ + \\
 & 0,0404(x_1 - 3636,224)_+ + 0,0359(x_1 - 3726,755)_+ + \\
 & 0,126x_2 + 0,9869(x_2 - 91,139)_+ - 68,2952(x_2 - 98,06)_+ + \\
 & 74,2105(x_2 - 98,252)_+ - 0,4509x_3 + 0,00006(x_3 - 11,047)_+ - \\
 & 0,0678x_4 + 4,0068(x_4 - 89,627)_+ - 5,9449(x_4 - 91,979)_+ + \\
 & 2,4777x_5 - 2,6072(x_5 - 0,98)_+ - 0,9193(x_5 - 9,796)_+ - \\
 & 0,8171(x_5 - 10,041)_+
 \end{aligned}$$

Koefisien determinasi yang dihasilkan oleh model tersebut adalah sebesar 95,28 persen. Artinya variabilitas TPT di Aceh dapat dijelaskan oleh model ini sebesar 95,28 persen sedangkan sisanya dijelaskan oleh variabel-variabel yang tidak termasuk dalam model tersebut.

#### 4.7 Pengujian Signifikansi Parameter

Setelah mendapatkan model terbaik dan estimasi parameter-nya, dilakukan pengujian terhadap parameter tersebut untuk mengetahui variabel apa saja yang mempengaruhi TPT di Aceh secara signifikan. Parameter diuji signifikansi secara serentak terlebih dahulu. Apabila hasil uji signifikansi secara serentak menunjukkan bahwa terdapat minimal satu parameter yang signifikan, maka pengujian dilanjutkan secara parsial.

##### a. Uji Serentak

Tujuan dari uji serentak adalah menguji signifikansi seluruh parameter yang terdapat dalam model secara keseluruhan. Nilai  $\alpha$  yang digunakan adalah 0,05. Hasil uji serentak disajikan dalam Tabel 4.7 berikut.

**Tabel 4.7** Tabel ANOVA

<b>Sumber Variasi</b>	<b>df</b>	<b>SS</b>	<b>MS</b>	<b>F</b>	<b>P-value</b>
Regresi	17	374,6102	22,0359		
<i>Error</i>	5	18,5406	3,7081	5,9426	0,0291
Total	22	393,1507			

Berdasarkan Tabel 4.7 di atas, uji signifikansi parameter secara serentak menghasilkan statistik uji F sebesar 5,9426. Nilai  $F_{tabel}$  untuk uji ini adalah 4,596, lebih kecil dari  $F$  sehingga keputusan yang diambil adalah tolak  $H_0$ . Hal tersebut menunjukkan bahwa terdapat minimal satu parameter yang signifikan.

### **b. Uji Parsial**

Uji serentak yang telah dilakukan menunjukkan bahwa terdapat minimal satu parameter yang signifikan. Maka pengujian signifikansi parameter model dilanjutkan pengujian secara parsial. Tabel 4.8 menunjukkan hasil pengujian parameter secara parsial.

**Tabel 4.8** Hasil Uji Signifikansi Parameter Secara Parsial

<b>Variabel</b>	<b>Parameter</b>	<b>Estimator</b>	<b>t</b>	<b>P-value</b>
Konstan	$\delta_0$	-1,7303	-5,8035	0,0021*
$x_1$	$\delta_1$	0,0457	2,7384	0,0409*
	$\beta_1$	-0,0635	-2,8902	0,0342*
	$\beta_2$	0,0404	2,3719	0,0638
	$\beta_3$	0,0359	2,3719	0,0638
	$\delta_2$	0,126	1,187	0,2885
$x_2$	$\beta_4$	0,9869	3,4148	0,0189*
	$\beta_5$	-68,2952	-5,8906	0,002*
	$\beta_6$	74,2105	5,5866	0,0025*
$x_3$	$\delta_3$	-0,4509	-0,2854	0,7868
	$\beta_7$	0,00006	2,3719	0,0638

**Tabel 4.8** Hasil Uji Signifikansi Parameter Secara Parsial (lanjutan)

Variabel	Parameter	Estimator	$t$	$P$ -value
$x_4$	$\delta_4$	-0,0678	-0,5115	0,6308
	$\beta_8$	4,0068	3,2789	0,022*
	$\beta_9$	-5,9449	-4,0643	0,0097*
$x_5$	$\delta_5$	2,4777	1,4244	0,2136
	$\beta_{10}$	-2,6072	-1,1815	0,2905
	$\beta_{11}$	-0,9193	-0,4123	0,6972
	$\beta_{12}$	-0,8171	-0,4123	0,6972

\*signifikan pada  $\alpha = 0,05$  dan  $|t| > t_{tabel} = 2,57058$

Berdasarkan Tabel 4.8, terdapat tiga variabel yang berpengaruh secara signifikan terhadap model. Ketiga variabel tersebut adalah kepadatan penduduk ( $x_1$ ), AMH ( $x_2$ ), dan APK SLTA ( $x_4$ ). Dua variabel lainnya, yaitu rata-rata lama sekolah ( $x_3$ ) dan banyak perusahaan ( $x_5$ ) tidak berpengaruh signifikan terhadap TPT di Aceh.

#### 4.8 Pemilihan Titik Knot Optimal dengan Tiga Variabel Prediktor

Pada pengujian signifikansi parameter telah diketahui bahwa hanya tiga variabel yang berpengaruh secara signifikan terhadap model, yaitu  $x_1$ ,  $x_2$ , dan  $x_4$ . Maka dilakukan pemodelan kembali dengan menggunakan tiga variabel tersebut saja. Pemodelan dengan tiga variabel juga dilakukan dengan menggunakan satu titik knot, dua titik knot, tiga titik knot, dan kombinasi titik knot.

##### 4.8.1 Pemilihan Titik Knot Optimal dengan Satu Titik Knot

Seperti halnya pemodelan dengan lima variabel prediktor, pemodelan dengan tiga variabel prediktor dicobakan terlebih dahulu dengan menggunakan satu titik knot. Berikut adalah model regresi nonparamtrik spline dengan tiga variabel prediktor menggunakan satu titik knot.

$$y_i = \alpha_0 + \alpha_1 x_{i1} + \beta_1 (x_{i1} - k_1)_+ + \alpha_2 x_{i2} + \beta_2 (x_{i2} - k_1)_+ + \alpha_3 x_{i4} + \beta_3 (x_{i4} - k_1)_+ + \varepsilon_i$$

dengan nilai GCV yang dihasilkan sebagai berikut.

**Tabel 4.9** Nilai GCV Regresi Nonparametrik Spline Tiga Variabel dengan Satu Titik Knot

Knot			GCV
$x_1$	$x_2$	$x_4$	
2821,449	96,33	84,140	24,221
2911,98	96,523	84,924	24,138
3002,51	96,714	85,708	23,883
3093,041	96,906	86,492	23,039
3183,571	97,099	87,276	22,062
3274,102	97,291	88,06	21,675
3364,633	97,483	88,843	21,435
<b>3455,163</b>	<b>97,675</b>	<b>89,627</b>	<b>21,347</b>
3545,694	97,868	90,411	21,601
3636,224	98,06	91,195	21,872

Berdasarkan Tabel 4.9 diketahui bahwa nilai GCV terkecil adalah 21,347. Nilai tersebut didapatkan pada titik knot 3455,163 untuk variabel  $x_1$ , 97,675 untuk variabel  $x_2$ , dan 89,627 untuk variabel  $x_4$ .

#### 4.8.2 Pemilihan Titik Knot Optimal dengan Dua Titik Knot

Selanjutnya adalah memodelkan data dengan tiga variabel prediktor menggunakan dua titik knot. Pemodelan dengan tiga variabel prediktor yang diduga mempengaruhi TPT di Aceh menggunakan regresi nonparametrik spline dengan dua titik knot adalah sebagai berikut.

$$y_i = \alpha_0 + \alpha_1 x_{i1} + \beta_1 (x_{i1} - k_1)_+ + \beta_2 (x_{i1} - k_2)_+ + \alpha_2 x_{i2} + \beta_3 (x_{i2} - k_1)_+ + \beta_4 (x_{i2} - k_2)_+ + \alpha_3 x_{i4} + \beta_5 (x_{i4} - k_1)_+ + \beta_6 (x_{i4} - k_2)_+ + \varepsilon_i$$

Tabel 4. 10 berikut menyajikan nilai GCV yang dihasilkan.

**Tabel 4.10** Nilai GCV Regresi Nonparametrik Spline Tiga Variabel dengan Dua Titik Knot

Knot			GCV
$x_1$	$x_2$	$x_4$	
3455,163	97,675	89,627	17.450
3907,816	98,637	93,547	
3455,163	97,675	89,627	17.435
3998,347	98,829	94,331	
3455,163	97,675	89,627	17.339
4088,878	99,021	95,114	
3455,163	97,675	89,627	16.045
4179,408	99,213	95,898	
3455,163	97,675	89,627	14.084
4269,939	99,406	96,682	
<b>3455,163</b>	<b>97,675</b>	<b>89,627</b>	<b>14.032</b>
<b>4360,469</b>	<b>99,598</b>	<b>97,466</b>	
3455,163	97,675	89,627	21.347
4451,000	99,790	98,250	
3545,694	97,868	90,411	17.823
3636,224	98,060	91,195	
3545,694	97,868	90,411	17.565
3726,755	98,252	91,979	
3545,694	97,868	90,411	17.531
3817,286	98,444	92,763	

Berdasarkan Tabel 4.9 diketahui bahwa nilai GCV terkecil adalah 14,032. Nilai GCV tersebut diperoleh pada titik knot 3455,163 dan 4360,469 untuk variabel  $x_1$ , titik knot 97,675 dan 99,598 untuk variabel  $x_2$ , dan titik knot 89,627 dan 97,466 untuk variabel  $x_4$ . Seperti halnya pemodelan dengan lima variabel prediktor, nilai GCV terkecil yang dihasilkan oleh dua knot lebih kecil dari nilai GCV terkecil yang dihasilkan oleh satu knot. Maka regresi nonparametrik spline dengan dua titik knot lebih baik daripada dengan satu titik knot saja.

### 4.8.3 Pemilihan Titik Knot Optimal dengan Tiga Titik Knot

Selanjutnya memodelkan data dengan titik knot sebanyak tiga. Pemodelan TPT di Aceh dengan tiga variabel prediktor menggunakan regresi nonparametrik spline dengan tiga titik knot adalah sebagai berikut.

$$y_i = \alpha_0 + \alpha_1 x_{i1} + \beta_1 (x_{i1} - k_1)_+ + \beta_2 (x_{i1} - k_2)_+ + \beta_3 (x_{i1} - k_3)_+ + \alpha_2 x_{i2} + \beta_4 (x_{i2} - k_1)_+ + \beta_5 (x_{i2} - k_2)_+ + \beta_6 (x_{i1} - k_3)_+ + \alpha_4 x_{i4} + \beta_{10} (x_{i4} - k_1)_+ + \beta_{11} (x_{i4} - k_2)_+ + \beta_{12} (x_{i1} - k_3)_+ + \varepsilon_i$$

Nilai GCV yang dihasilkan oleh regresi nonparametrik spline dengan tiga titik knot disajikan dalam Tabel 4.11 berikut.

**Tabel 4.11** Nilai GCV Regresi Nonparametrik Spline Tiga Variabel dengan Tiga Titik Knot

Knot			GCV
$x_1$	$x_2$	$x_4$	
<b>286,592</b>	<b>90,947</b>	<b>62,192</b>	<b>6,567</b>
<b>3636,224</b>	<b>98,06</b>	<b>91,195</b>	
<b>3726,755</b>	<b>98,252</b>	<b>91,979</b>	
286,592	90,947	62,192	8,507
3636,224	98,06	91,195	
3817,286	98,444	92,763	

**Tabel 4.11** Nilai GCV Regresi Nonparametrik Spline Tiga Variabel dengan Tiga Titik Knot (lanjutan)

<b>Knot</b>			<b>GCV</b>
$x_1$	$x_2$	$x_4$	
286,592	90,947	62,192	13,388
3636,224	98,06	91,195	
3907,816	98,637	93,547	
377,122	91,139	62,976	8,150
3455,163	97,675	89,627	
3817,286	98,444	92,763	
377,122	91,139	62,976	11,070
3455,163	97,675	89,627	
3907,816	98,637	93,547	
377,122	91,139	62,975	11,903
3455,163	97,675	89,627	
3998,347	98,829	94,331	
377,122	91,139	62,976	13,350
3455,163	97,675	89,627	
4088,878	99,021	95,114	
377,122	91,139	62,976	12,670
3455,163	97,675	89,627	
4179,408	99,213	95,898	
377,122	91,139	62,976	12,238
3455,163	97,675	89,627	
4269,939	99,406	96,682	
377,122	91,139	62,976	12,237
3455,163	97,675	89,627	
4360,469	99,598	97,466	

Berdasarkan Tabel 4.11, nilai GCV terkecil dengan tiga titik knot adalah 6,567. Nilai GCV tersebut dihasilkan pada titik knot untuk variabel  $x_1$  adalah  $k_1 = 286,592$ ,  $k_2 = 3636,224$ ,  $k_3 = 3726,755$ . Titik knot untuk variabel  $x_2$  adalah  $k_1 = 90,947$ ,  $k_2 = 98,06$ ,  $k_3 = 98,252$ . Titik knot untuk variabel  $x_4$  adalah  $k_1 = 62,192$ ,  $k_2 = 91,195$ ,  $k_3 = 91,979$ .

#### 4.8.4 Pemilihan Titik Knot Optimal dengan Kombinasi Knot

Telah didapatkan titik knot optimum dari pemodelan dengan satu titik knot, dua titik knot, dan tiga titik knot. Kemudian data dengan tiga variabel prediktor dimodelkan dengan mengkombinasikan jumlah knot tersebut. Nilai GCV dengan kombinasi knot disajikan dalam Tabel 4.12.

**Tabel 4.12** Nilai GCV Regresi Nonparametrik Spline Tiga Variabel dengan Kombinasi Knot

Knot			GCV
$x_1$	$x_2$	$x_4$	
3455,163	97,675	89,627 97,466	12,380
3455,163 4360,469	97,675 99,598	89,627 97,466	12,380
3455,163 4360,469	97,675 99,598	62,192 91,195 91,979	15,485
286,592 3636,224 3726,755	97,675	89,627 97,466	9,165
286,592 3636,224 3726,755	97,675 99,598	89,627	17,608
286,592 3636,224	97,675 99,598	89,627 97,466	10,628

**Tabel 4.12** Nilai GCV Regresi Nonparametrik Spline Tiga Variabel dengan Kombinasi Knot (lanjutan)

<b>Knot</b>			<b>GCV</b>
$x_1$	$x_2$	$x_4$	
3726,755			
286,592		62,192	
3636,224	97,675	91,195	8,246
3726,755	99,598	91,979	
286,592	90,947		
3636,224	98,060	89,627	11,594
3726,755	98,252		
<b>286,592</b>	<b>90,947</b>	<b>89,627</b>	<b>6,126</b>
<b>3636,224</b>	<b>98,060</b>	<b>97,466</b>	
<b>3726,755</b>	<b>98,252</b>		
286,592	90,947	62,192	
3636,224	98,060	91,195	6,567
3726,755	98,252	91,979	

Berdasarkan Tabel 4.12, nilai GCV terkecil dengan kombinasi knot adalah 6,126. Nilai GCV tersebut dihasilkan apabila digunakan tiga titik knot untuk variabel  $x_1$  dan  $x_2$ , dan  $x_3$ , serta dua titik knot untuk variabel  $x_4$ . Lokasi titik knot untuk variabel  $x_1$  adalah  $k_1 = 286,592$ ,  $k_2 = 3636,224$ ,  $k_3 = 3726,755$ . Lokasi titik knot untuk variabel  $x_2$  adalah  $k_1 = 90,947$ ,  $k_2 = 98,06$ ,  $k_3 = 98,252$ . Lokasi titik knot untuk variabel  $x_4$  adalah  $k_1 = 89,627$  dan  $k_2 = 97,466$ .

#### 4.9 Pemilihan Model Terbaik dengan Tiga Variabel Prediktor

Telah didapatkan knot-knot optimal dari pemodelan TPT di Aceh dengan tiga variabel prediktor menggunakan satu knot, dua knot, tiga knot, serta kombinasi knot. Selanjutnya membandingkan nilai GCV terkecil dari masing-masing model untuk memilih

model yang terbaik. Nilai GCV terkecil dari masing-masing model disajikan dalam Tabel 4.13.

**Tabel 4.13** Nilai GCV Terkecil dari Pemodelan Tiap Jumlah Knot dengan Tiga Variabel Prediktor

<b>Jumlah Knot</b>	<b>GCV</b>	<b>Jumlah Parameter</b>
1 Titik Knot	28,793	7
2 Titik Knot	9,508	10
3 Titik Knot	6,846	13
Kombinasi (332)	6,125	12

Pada Tabel 4.13 diketahui bahwa pemodelan yang menghasilkan nilai GCV yang paling kecil adalah dengan menggunakan kombinasi jumlah knot yaitu tiga titik knot untuk variabel  $x_1$  dan  $x_2$ , serta dua titik knot untuk variabel  $x_4$ . Selain itu, banyak parameter yang digunakan dalam pemodelan kombinasi titik knot cukup sederhana. Oleh karena itu diputuskan bahwa model terbaik yang dipilih untuk pemodelan dengan tiga variabel prediktor adalah model regresi nonparametrik spline dengan menggunakan kombinasi jumlah knot (332).

#### **4.10 Penaksiran Parameter Model Regresi Nonparametrik Spline dengan Tiga Variabel**

Telah didapatkan model terbaik yaitu dengan tiga titik knot untuk variabel  $x_1$  dan  $x_2$ , serta dua titik knot untuk variabel  $x_4$ . Seperti halnya pemodelan dengan lima variabel prediktor, estimasi parameter model dengan tiga variabel prediktor juga didapatkan menggunakan metode OLS. Berikut adalah model regresi nonparametrik spline terbaik.

$$\begin{aligned}
 y = & -0,6362 + 0,0517x_1 - 0,0575(x_1 - 286,592)_+ + \\
 & 0,0163(x_1 - 3636,224)_+ + 0,0145(x_1 - 3726,755)_+ - \\
 & 0,0845x_2 + 1,2135(x_2 - 90,947)_+ - 55,0797(x_2 - 98,06)_+ + \\
 & 51,8247(x_2 - 98,252)_+ + 0,1291x_4 + 0,7487(x_4 - 89,627)_+ - \\
 & 17,0144(x_4 - 97,466)_+
 \end{aligned}$$

Koefisien determinasi yang dihasilkan oleh model tersebut adalah sebesar 88,55 persen. Nilai tersebut menunjukkan bahwa variabilitas TPT di Aceh dapat dijelaskan oleh model dengan menggunakan tiga variabel prediktor sebesar 88,55 persen sedangkan sisanya dijelaskan oleh variabel-variabel yang tidak termasuk dalam model tersebut.

#### 4.11 Pengujian Signifikansi Parameter dengan Tiga Variabel Prediktor

Selanjutnya adalah menguji parameter yang telah didapat untuk mengetahui variabel apa saja yang mempengaruhi TPT di Aceh secara signifikan. Apabila hasil uji signifikansi secara serentak menunjukkan bahwa terdapat minimal satu parameter yang signifikan, maka pengujian dilanjutkan secara parsial.

##### a. Uji Serentak

Nilai  $\alpha$  yang digunakan dalam pengujian ini adalah 0,05. Hasil uji serentak disajikan dalam Tabel 4.14 di bawah.

**Tabel 4.14** Tabel ANOVA Pemodelan dengan Tiga Variabel Prediktor

Sumber Variasi	df	SS	MS	F	P-value
Regresi	11	347,9388	31,6308		
Error	11	45,01106	4,0919	7,73	0,001
Total	22	392,9498			

Tabel 4.14 menunjukkan bahwa uji signifikansi parameter secara serentak menghasilkan statistik uji F sebesar 7,73. Nilai  $F_{tabel}$

untuk uji ini adalah 2,82, lebih kecil dari  $F$  sehingga keputusan yang diambil adalah tolak  $H_0$ . Artinya terdapat minimal satu parameter yang signifikan.

### b. Uji Parsial

Uji serentak yang telah dilakukan menunjukkan bahwa terdapat minimal satu parameter yang signifikan. Maka pengujian signifikansi parameter model dilanjutkan pengujian secara parsial. Tabel 4.15 menunjukkan hasil pengujian parameter secara parsial.

**Tabel 4.15** Hasil Uji Signifikansi Parameter Secara Parsial Model dengan Tiga Variabel Prediktor

Variabel	Parameter	Estimator	$t$	$P$ -value
Konstan	$\delta_0$	-0,6362	-4,4221	0,001*
$x_1$	$\delta_1$	0,0517	5,3852	0,0002*
	$\beta_1$	-0,0575	-4,3667	0,0011*
	$\beta_2$	0,0163	1,4014	0,1887
	$\beta_3$	0,0145	1,4014	0,1887
	$\delta_2$	-0,0845	-1,247	0,2383
$x_2$	$\beta_4$	1,2135	4,2242	0,0014*
	$\beta_5$	-55,0797	-4,5343	0,0009*
	$\beta_6$	51,8247	3,6071	0,0041*
$x_4$	$\delta_4$	0,1291	1,7338	0,1108
	$\beta_8$	0,7487	1,7179	0,1138
	$\beta_9$	-17,0144	-3,626	0,004*

\*signifikan pada  $\alpha = 0,05$  dan  $|t| > t_{tabel} = 2,20099$

Berdasarkan Tabel 4.15, ketiga variabel prediktor yang digunakan yaitu kepadatan penduduk ( $x_1$ ), AMH ( $x_2$ ), dan APK SLTA ( $x_4$ ) berpengaruh secara signifikan terhadap TPT di Aceh.

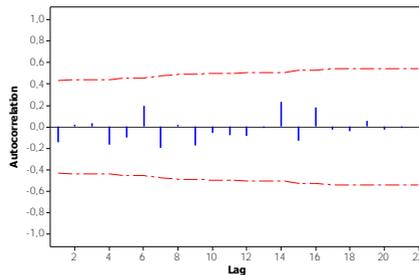
### 4.12 Pemeriksaan Asumsi Residual

Residual dari model regresi nonparametrik spline terbaik yang telah didapatkan harus memenuhi asumsi identik,

independen, dan berdistribusi normal. Oleh karena itu dilakukan pemeriksaan apakah asumsi-asumsi tersebut telah terpenuhi.

#### 4.12.1 Pemeriksaan Asumsi Independen Residual

Salah satu asumsi yang harus dipenuhi adalah residual yang independen. Yang dimaksud dengan residual yang independen adalah tidak terjadi autokorelasi antar residual. Plot *Auto-correlation Function* (ACF) digunakan untuk mengetahui adanya autokorelasi antar residual. Pada Gambar 4.3 terlihat bahwa seluruh autokorelasi berada di dalam interval konfidensi, artinya tidak terjadi autokorelasi, dengan kata lain residual independen dan asumsi independen residual terpenuhi.



Gambar 4.3 Plot ACF Residual

#### 4.12.2 Pengujian Asumsi Identik Residual

Asumsi lain yang harus dipenuhi adalah residual dari model yang identik. Maksud dari identik adalah varian residual yang homogen, dengan kata lain tidak terjadi heteroskedastisitas. Untuk mendeteksi terjadinya heteroskedastisitas, dilakukan uji glejser. Berikut adalah hasil uji glejser dari residual model.

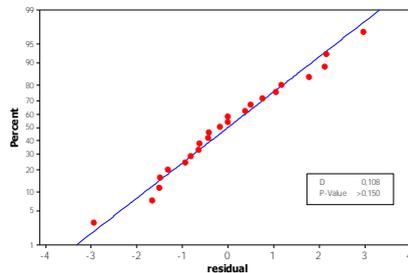
Tabel 4.16 ANOVA Uji Glejser

Sumber Variasi	df	SS	MS	F	P-value
Regresi	11	5,5583	0,5053		
Error	11	10,4	0,9454	0,5345	0,8432
Total	22	15,9583			

Dalam pengujian ini,  $\alpha$  yang digunakan adalah 0,05. Nilai  $F_{tabel}$  untuk uji ini adalah 2,82, lebih besar dari nilai  $F$  sehingga  $H_0$  gagal ditolak. Hal tersebut menunjukkan bahwa tidak terjadi heteroskedastisitas, artinya asumsi residual yang identik terpenuhi.

#### 4.12.3 Pengujian Asumsi Normalitas Residual

Asumsi terakhir yang harus dipenuhi adalah residual yang berdistribusi normal. Salah satu uji yang dapat digunakan untuk pengujian distribusi normal adalah dengan uji Kolmogorov-Smirnov. Berikut adalah *probability plot* residual model.



**Gambar 4.4** Plot Distribusi Normal Residual

Secara visual, Gambar 4.4 menunjukkan bahwa sebaran titik mengikuti garis distribusi normal yang artinya residual mengikuti distribusi normal. Selain itu, hal tersebut ditunjukkan pula oleh nilai uji Kolmogorov-Smirnov  $D$  dan  $p$ -value. Nilai  $D$  tersebut lebih kecil dibandingkan dengan nilai  $D_{tabel}$  yaitu 0,275 dengan  $\alpha = 0,05$  dan  $p$ -value lebih besar dari 0,15 yang artinya lebih dari  $\alpha$ . Kedua hal tersebut menunjukkan bahwa residual mengikuti distribusi normal.

#### 4.13 Interpretasi Model Regresi Nonparametrik Spline

Model regresi nonparametrik spline yang didapat telah memenuhi asumsi IIDN( $0, \sigma^2$ ), maka model tersebut dapat

diinterpretasikan. Model regresi nonparametrik spline linier terbaik untuk data TPT di Aceh adalah sebagai berikut.

$$y = -0,6362 + 0,0517x_1 - 0,0575(x_1 - 286,592)_+ + 0,0163(x_1 - 3636,224)_+ + 0,0145(x_1 - 3726,755)_+ - 0,0845x_2 + 1,2135(x_2 - 90,947)_+ - 55,0797(x_2 - 98,06)_+ + 51,8247(x_2 - 98,252)_+ + 0,1291x_4 + 0,7487(x_4 - 89,627)_+ - 17,0144(x_4 - 97,466)_+$$

Berikut adalah interpretasi model terhadap masing-masing variabel yang berpengaruh secara signifikan terhadap TPT di Aceh.

1. Interpretasi model terhadap variabel kepadatan penduduk ( $x_1$ ) adalah sebagai berikut.

$$y = \begin{cases} 0,0517x_1 & ; x_1 < 286,592 \\ -0,0058x_1 + 16,479 & ; x_1 \geq 286,592 \end{cases}$$

Untuk wilayah dengan kepadatan penduduk kurang dari 286,592 jiwa/km<sup>2</sup>, apabila kepadatan penduduk naik 1 jiwa/km<sup>2</sup> maka TPT di wilayah tersebut akan meningkat sebesar 0,0517 persen. Wilayah yang termasuk interval ini adalah seluruh kabupaten/kota kecuali Kota Banda Aceh, Kota Langsa, dan Kota Lhoksumawe. Untuk wilayah dengan kepadatan penduduk sebesar 286,592 jiwa/km<sup>2</sup> atau lebih, apabila kepadatan penduduk naik 1 jiwa/km<sup>2</sup> maka TPT di wilayah tersebut akan menurun sangat rendah yaitu sebesar 0,0058 persen. Wilayah yang termasuk interval ini adalah Kota Banda Aceh, Kota Langsa, dan Kota Lhoksumawe.

2. Interpretasi model terhadap variabel AMH ( $x_2$ ) adalah sebagai berikut.

$$y = \begin{cases} -0,0845x_2 & ; x_2 < 90,947 \\ 1,129x_2 - 110,364 & ; 90,947 \leq x_2 < 98,06 \\ -53,9507x_2 + 5290,751 & ; 98,06 \leq x_2 < 98,252 \\ -2,126x_2 + 198,871 & ; x_2 \geq 98,252 \end{cases}$$

Untuk wilayah dengan AMH kurang dari 90,947 persen, apabila AMH naik satu persen maka TPT di wilayah tersebut akan menurun sangat rendah yaitu sebesar 0,0845 persen. Wilayah yang termasuk interval ini adalah Kabupaten Gayo Lues. Untuk wilayah dengan AMH sebesar 90,947 persen atau lebih tetapi kurang dari 98,06 persen, apabila AMH naik satu persen maka TPT di wilayah tersebut akan meningkat sebesar 1,129 persen. Wilayah yang termasuk interval ini adalah Kabupaten Aceh Singkil, Kabupaten Aceh Selatan, Kabupaten Aceh Timur, Kabupaten Aceh Barat, Kabupaten Aceh Besar, Kabupaten Pidie, Kabupaten Aceh Utara, Kabupaten Aceh Barat Daya, Kabupaten Aceh Tamiang, Kabupaten Nagan Raya, Kabupaten Aceh Jaya, Kabupaten Pidie Jaya, dan Kota Subussalam. Untuk wilayah dengan AMH sebesar 98,06 persen atau lebih tetapi kurang dari 98,252 persen, apabila AMH naik satu persen maka TPT di wilayah tersebut akan menurun sebesar 53,95 persen. Wilayah yang termasuk interval ini adalah Kabupaten Aceh Tenggara, Kabupaten Aceh Tengah, dan Kabupaten Bireuen. Untuk wilayah dengan AMH sebesar 98,252 persen atau lebih, apabila AMH naik satu persen maka TPT di wilayah tersebut akan menurun sebesar 2,126 persen. Wilayah yang termasuk interval ini adalah Kabupaten Simeulue, Kabupaten Bener Meriah, Kota Banda Aceh, Kota Sabang, Kota Langsa, dan Kota Lhoksumawe.

3. Interpretasi model terhadap variabel APK jenjang SLTA ( $x_4$ ) adalah sebagai berikut.

$$y = \begin{cases} 0,8778x_4 - 67,1037 & ; x_4 < 97,466 \\ -16,1366x_4 + 1591,222 & ; x_4 \geq 97,466 \end{cases}$$

Untuk wilayah dengan APK SLTA kurang dari 97,466 persen, apabila APK SLTA naik satu persen maka TPT di wilayah tersebut akan meningkat sebesar 0,8778 persen. Wilayah yang termasuk interval ini adalah seluruh kabupaten/kota kecuali Kabupaten Simeulue dan Kabupaten Aceh Barat. Untuk wilayah dengan APK SLTA sebesar 97,466 persen atau lebih, apabila APK SLTA naik satu persen maka TPT di wilayah tersebut akan menurun sebesar 16,1366 persen. Wilayah yang termasuk interval ini adalah Kabupaten Simeulue dan Kabupaten Aceh Barat.

*(Halaman Ini Sengaja Dikosongkan)*

## **BAB V**

### **KESIMPULAN DAN SARAN**

Berdasarkan hasil analisis dan pembahasan pada bab sebelumnya, dapat diambil beberapa kesimpulan yang menjawab tujuan dari penelitian ini. Selain itu, terdapat beberapa hal yang disarankan untuk penelitian selanjutnya agar mendapatkan hasil yang lebih baik.

#### **5.1 Kesimpulan**

Berdasarkan analisis dan pembahasan yang telah dilakukan, dapat diambil dua kesimpulan dari penelitian ini.

1. Selama tiga tahun terakhir, TPT di Provinsi Aceh selalu berada di atas TPT nasional. Bahkan pada tahun 2013, Aceh menduduki peringkat pertama provinsi dengan TPT tertinggi di Indonesia. Pada tahun 2014, Aceh berada di peringkat kelima. Walaupun tidak lagi menduduki peringkat pertama, TPT di Aceh masih tinggi dan berada di atas TPT nasional. Berdasarkan data terbaru provinsi Aceh yaitu tahun 2013, rata-rata TPT provinsi adalah 9,42 persen, dengan kata lain untuk setiap 10000 penduduk di Aceh terdapat 942 orang yang termasuk pengangguran terbuka. Wilayah dengan TPT tertinggi adalah Kabupaten Aceh Utara yaitu sebesar 17,97 persen. TPT tertinggi kedua ditempati oleh Kabupaten Aceh Tenggara sebesar 16,82 persen. Sedangkan wilayah dengan TPT paling rendah adalah Kabupaten Bener Meriah dengan TPT sebesar 0,63 persen. Perbedaan TPT antar kabupaten/kota cukup besar, hal ini ditunjukkan oleh standar deviasi yang cukup besar yaitu 4,23. Rata-rata perkilometer persegi wilayah di kabupaten/kota di Provinsi Aceh dihuni oleh 362 orang, dengan standar deviasi yang sangat besar yaitu 931,64. Wilayah dengan kepadatan penduduk tertinggi adalah Kota Banda Aceh yaitu 4451 orang perkilometer persegi

wilayahnya sedangkan kepadatan penduduk Kabupaten Gayo Lues merupakan yang paing rendah yaitu 15 orang perkilometer persegi wilayahnya. Sebesar 96,42 persen penduduk dari setiap kabupaten/kota di Aceh telah melek huruf dan untuk setiap kabupaten/kota rata-rata penduduknya menjalani sekolah selama 9 sampai 10 tahun. Rata-rata APK jenjang SLTA di Aceh adalah sebesar 78,52. Artinya 78,52 persen dari total penduduk yang tergolong umur SLTA (16-18 tahun) di tiap kabupaten/kota menjalani pendidikan jenjang SLTA. Terdapat rata-rata dua perusahaan di setiap kabupaten/kota tetapi standar deviasinya cukup besar 2,56. Kabupaten Aceh Tamiang merupakan wilayah dengan perusahaan terbanyak yaitu sebanyak 12, jauh lebih banyak dari kabupaten/kota lainnya. Sementara itu terdapat 8 kabupaten/kota yang di wilayahnya tidak terdapat perusahaan sama sekali, yaitu Kabupaten Simeulue, Kabupaten Aceh Tenggara, Kabupaten Aceh Timur, Kabupaten Aceh Besar, Kabupaten Pidie, Kabupaten Aceh Barat Daya, Kabupaten Gayo Lues, dan Kabupaten Bener Meriah.

2. Model regresi nonparametrik spline terbaik yang dihasilkan dengan seluruh variabel prediktor adalah dengan menggunakan knot kombinasi. Nilai koefisien determinasi yang dihasilkan cukup tinggi yaitu 95,28 persen, tetapi terdapat dua variabel yang tidak berpengaruh secara signifikan terhadap TPT yaitu rata-rata lama sekolah ( $x_3$ ) dan banyak perusahaan ( $x_5$ ). Setelah kedua variabel tersebut dikeluarkan, model regresi nonparametrik spline terbaik dengan tiga variabel juga model dengan menggunakan kombinasi knot. Nilai GCV yang dihasilkan adalah 5,265 dan koefisien determinasi sebesar 88,55 persen. Model regresi nonparametrik spline yang didapatkan adalah sebagai berikut.

$$\begin{aligned}
y = & -0,6362 + 0,0517x_1 - 0,0575(x_1 - 286,592)_+ + \\
& 0,0163(x_1 - 3636,224)_+ + 0,0145(x_1 - 3726,755)_+ - \\
& 0,0845x_2 + 1,2135(x_2 - 90,947)_+ - 55,0797(x_2 - 98,06)_+ + \\
& 51,8247(x_2 - 98,252)_+ + 0,1291x_4 + 0,7487(x_4 - 89,627)_+ - \\
& 17,0144(x_4 - 97,466)_+
\end{aligned}$$

Uji signifikansi parameter menunjukkan bahwa ketiga variabel dalam model yaitu kepadatan penduduk ( $x_1$ ), AMH ( $x_2$ ), dan APK SLTA ( $x_4$ ) berpengaruh signifikan terhadap TPT di Aceh.

## 5.2 Saran

Pada penelitian ini, model yang digunakan adalah regresi nonparametrik spline linier dengan mencoba knot sebanyak satu, dua, tiga, dan kombinasi knot, serta variabel prediktor yang digunakan sebanyak lima. Untuk penelitian selanjutnya disarankan menambah variabel yang diduga mempengaruhi TPT. Selain itu, perlu dicoba pula pemodelan orde lainnya dan dengan banyak knot yang lebih banyak pula sehingga kombinasi knot yang dihasilkan lebih banyak pula.

*(Halaman Ini Sengaja Dikosongkan)*

## DAFTAR PUSTAKA

- Ariane, S. (2012). *Pendekatan Regresi Ridge Untuk Memodelkan Faktor-faktor yang Mempengaruhi Tingkat Pengangguran Terbuka di Provinsi Jawa Timur dan Jawa Tengah*. Surabaya.
- Badan Pusat Statistik. (2014). *Keadaan Ketenagakerjaan Agustus 2014*. Badan Pusat Statistik.
- Badan Pusat Statistik. (2014). *Keadaan Ketenagakerjaan Agustus 2014*. Badan Pusat Statistik Provinsi Aceh.
- Badan Pusat Statistik. (2014, Mei). *Publikasi BPS Indikator Pasar Tenaga Kerja Mei 2014*. Retrieved Januari 28, 2015, from BPS:  
[http://www.bps.go.id/hasil\\_publicasi/ind\\_pasar\\_tenaga\\_kerja\\_mei\\_2014/index3.php?pub=Indikator%20Pasar%20Tenaga%20Kerja%20Indonesia%20Mei%202013](http://www.bps.go.id/hasil_publicasi/ind_pasar_tenaga_kerja_mei_2014/index3.php?pub=Indikator%20Pasar%20Tenaga%20Kerja%20Indonesia%20Mei%202013)
- Budiantara, I. (2000). Metode U, GLM, CV, dan GCV dalam Regresi Nonparametrik Spline. *Majalah Ilmiah Himpunan Matematika Indonesia (MIHMI)*, 6,41-45.
- Budiantara, I. (2009). *Spline dalam Regresi Nonparametrik dan Semiparametrik : Sebuah Pemodelan Statistika Masa Kini dan Masa Mendatang*. Surabaya: ITS Press.
- Dewi, A. M. (2010). *Analisis Tingkat Penganggran dan Faktor-faktor yang Mempengaruhinya di Kota Semarang*. Semarang.
- Eubank, R. (1988). *Spline Smoothing and Nonparametric Regression*. New York: Marcel Dekker Inc.
- Eubank, R. (1999). *Nonparametric Regression and Spline Smoothing*. New York: Marcel Dekker Inc.

- Gujarati, D. N. (2003). *Basic Econometrics* (4 ed.). New York: McGraw-Hill Inc.
- Hajji, M. S. (2013). *Analisis PDRB, Inflasi, Upah Minimum Provinsi, dan Angka Melek Huruf Terhadap Tingkat Pengangguran Terbuka di Provinsi Jawa Tengah Tahun 1990-2011*. Semarang.
- Marina, S. M. (2013). *Pemodelan Faktor-faktor yang Mempengaruhi Persentase Kriminalitas di Jawa Timur dengan Pendekatan Regresi Semiparametrik Spline*. Surabaya.
- Merdekawati, I. P. (2013). *Pemodelan Regresi Spline Truncated Multivariabel pada Faktor-faktor yang Mempengaruhi Kemiskinan di Kabupaten/Kota Provinsi Jawa Tengah*. Surabaya.
- Mubarak, R. (2012). *Analisis Regresi Spline Multivariabel Untuk Pemodelan Kematian Penderita Demam Berdarah Dengue (DBD) di Jawa Timur*. Surabaya.
- Putra, P. P. (2013). *Analisis Faktor-faktor yang Mempengaruhi Pengangguran di Indonesia Tahun 1993-2012*. Surakarta.
- Sari, R. S. (2012). *Pemodelan Pengangguran Terbuka di Jawa Timur Dengan Menggunakan Pendekatan Regresi Spline Multivariabel*. Surabaya.
- Sudradjad, S. (2000). *Kiat Mengentaskan Pengangguran Melalui Wirausaha*. Jakarta: Bumi Aksara.
- Walpole, R. E. (1995). *Pengantar Statistika*. Jakarta: PT Gramedia Pusaka Utama.
- Wei, W. W. (2006). *Time Series Analysis Univariate and Multivariate Methods*. United States: Pearson Education, Inc.

## BIODATA PENULIS



Penulis yang bernama lengkap Leisa Noviana Sani merupakan anak terakhir dari pasangan Haswan Sani dan Vini Kusvini, serta lahir di Bandung pada tanggal 28 November 1992. Penulis menempuh jenjang pendidikan sejak TK sampai SMP di YLPI Al-Hikmah Surabaya, kemudian melanjutkan jenjang SMA di SMA Negeri 5 Surabaya, dan pada tahun 2011 telah diterima menjadi mahasiswa Jurusan Statistika. Mahasiswi dengan NRP 1311100112 ini bergabung dengan Unit Kegiatan Tari dan Karawitan (UKTK) ITS divisi *modern dance* sebagai anggota pada tahun 2011. Pada tahun kepengurusan 2012/2013, penulis menjabat sebagai *manager* divisi *modern dance* dan pada tahun kepengurusan 2013/2014 sebagai Ketua Bidang Kesenian UKTK-ITS. Penulis pernah mendapatkan dana hibah Dikti pada tahun ajaran 2014/2015 sebagai ketua untuk Program Kreativitas Mahasiswa (PKM) bidang Penelitian. Apabila pembaca ingin berdiskusi mengenai tugas akhir ini dan/atau mengenai materi lain yang berhubungan, serta kritik dan saran, dapat menghubungi penulis melalui email : [leisanoviana@gmail.com](mailto:leisanoviana@gmail.com).