

**PENYELEKSIAN KELOMPOK PEUBAH  
YANG MEMENGARUHI SKOR SAINS PISA  
INDONESIA MENGGUNAKAN *GROUP LASSO***

**Skripsi**

**Disusun untuk memenuhi salah satu syarat memperoleh gelar  
Sarjana Statistika**



**Rizkha Hayati  
1314617012**

**PROGRAM STUDI STATISTIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS NEGERI JAKARTA  
2022**

## ABSTRAK

**RIZKHA HAYATI.** Penyeleksian Kelompok Peubah yang Memengaruhi Skor Sains PISA Indonesia menggunakan *Group Lasso*. Skripsi, Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta. Januari 2022.

Literasi sains merupakan kemampuan seseorang menerapkan pengetahuan dan mengembangkan pola pikir secara reflektif sehingga dapat berpartisipasi dalam mengatasi isu-isu dan gagasan terkait sains. Selain itu, literasi sains juga merupakan salah satu pengetahuan yang harus dimiliki oleh seorang pelajar. Kualitas literasi sains Indonesia berdasarkan hasil *Programme for International Students Assessment* (PISA) yang diikuti sejak tahun 2000 hingga tahun 2018 masih relative rendah. Indonesia menduduki peringkat ke-71 dari 79 negara dengan rata-rata skor sains sebesar 396 pada PISA 2018, dimana hasil tersebut masih jauh di bawah rata-rata OECD. Evaluasi terkait hasil PISA sangat diperlukan untuk mendapatkan skor yang lebih baik. Data PISA memiliki kompleksitas data yang tinggi dengan peubah bebas berkorelasi kuat, sehingga memerlukan metode yang tepat untuk menganalisis. Metode *group LASSO* merupakan salah satu metode yang dapat digunakan untuk menyeleksi kelompok peubah yang memengaruhi literasi sains Indonesia berdasarkan data PISA. Penelitian ini menganalisis hasil survey diagnostik PISA 2018 dari 6174 siswa berumur 15 tahun. Hasil analisis menunjukkan terdapat 11 kelompok peubah penjelas yang memengaruhi skor sains PISA Indonesia dengan nilai  $R^2$  sebesar 0,31. Hal tersebut berarti hanya 31% keragaman dari rata rata skor sains yang dapat dijelaskan oleh peubah penjelas, sedangkan sisanya dijelaskan oleh faktor lain diluar penelitian.

**Kata kunci.** PISA, *group LASSO*, multikolinearitas, literasi sains

## ABSTRACT

**RIZKHA HAYATI.** Selection in Groups of Variables that Affect Indonesian's PISA Science Scores Using The Group Lasso. Undergraduate Thesis, Statistics Major, Faculty of Mathematics and Natural Sciences, State University of Jakarta. January 2022.




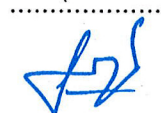

Scientific literacy is a person's ability to apply knowledge and develop a reflective mindset so that they can participate in overcoming issues and ideas related to science. In addition, scientific literacy is also one of the knowledge that must be possessed by a student. The quality of Indonesian scientific literacy based on the results of the Program for International Students Assessment (PISA) which was followed from 2000 to 2018 is still relatively low. Indonesia is ranked 71st out of 79 countries with an average science score of 396 at PISA 2018, which is still far below the OECD average. Evaluation related to PISA results is needed to get a better score. PISA data has high data complexity with strongly correlated independent variables, so it requires an appropriate method to analyze. The group LASSO method is one method that can be used to select groups of variables that affect Indonesian scientific literacy based on PISA data. This study analyzes the results of the 2018 PISA diagnostic survey of 6174 15-year-old students. The results of the analysis show that there are 11 groups of explanatory variables that affect the Indonesian PISA science score with an  $R^2$  value of 0.31. This means that only 31% of the diversity of the average science score can be explained by the explanatory variable, while the rest is explained by other factors outside of research.

**Keywords.** PISA, group lasso, multicollinearity, Scientific literacy

## LEMBAR PERSETUJUAN SIDANG SKRIPSI

### PENYELEKSIAN KELOMPOK PEUBAH YANG MEMENGARUHI SKOR SAINS PISA INDONESIA MENGUNAKAN *GROUP LASSO*

Nama : Rizkha Hayati  
NIM : 1314617012

	Nama	Tanda Tangan	Tanggal
<b>Penanggung Jawab :</b>			
Dekan	: <u>Prof. Dr. Muktiningsih N. M. Si.</u> NIP 196405111989032001	 .....	24/02/2022
<b>Wakil Penanggung Jawab :</b>			
Pembantu Dekan I	: <u>Dr. Esmar Budi, S. Si., MT.</u> NIP 197207281999032001	 .....	23/02/2022
Ketua Penguji	: <u>Dra. Widyanti Rahayu, M. Si.</u> NIP 196611032001122001	<u>Wid RH</u> .....	21/02/2022
Sekretaris	: <u>Dania Siregar, S. Stat., M. Si.</u> NIDN 8840600016	 .....	21/02/2022
<b>Anggota:</b>			
Pembimbing I	: <u>Vera Maya Santi, M. Si.</u> NIP.197905312005012006	<u>Vera Maya S.</u> .....	22/02/2022
Pembimbing II	: <u>Dr., Ir. Bagus Sumargo, M. Si.</u> NIP 196309221986011001	 .....	22/02/2022
Penguji Ahli	: <u>Dian Handayani, M. Si.</u> NIP 197404151998032001	 .....	22/02/2022

Dinyatakan lulus ujian skripsi tanggal ; 8 Februari 2022



## LEMBAR PERNYATAAN KEASLIAN SKRIPSI

Saya menyatakan dengan sesungguhnya bahwa skripsi dengan judul **“Penyeleksian Kelompok Peubah yang Memengaruhi Skor Sains PISA Indonesia Menggunakan *Group Lasso*”** yang disusun sebagai syarat untuk memperoleh gelar Sarjana Statistika dari Program Studi Statistika Universitas Negeri Jakarta adalah karya ilmiah saya dengan arahan dari dosen pembimbing.

Sumber informasi yang diperoleh dari penulis lain yang telah dipublikasikan yang disebutkan dalam teks skripsi ini, telah dicantumkan dalam Daftar Pustaka sesuai dengan norma, kaidah, dan etika penulisan ilmiah.

Jika dikemudian hari ditemukan sebagian besar skripsi ini bukan hasil karya saya sendiri dalam bagian bagian tertentu, saya bersedia menerima sanksi pencabutan gelar akademik yang saya sanding dan sanksi-sanksi lainnya sesuai dengan peraturan perundang-undangan yang berlaku.

Jakarta, 27-Januari-2022

  
SEPUJUH RIBU-RIBU  
METERAI  
TEMPEL  
E00D7AJX739167235  
Rizkha Hayau



KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN  
UNIVERSITAS NEGERI JAKARTA  
UPT PERPUSTAKAAN

Jalan Rawamangun Muka Jakarta 13220  
Telepon/Faksimili: 021-4894221  
Laman: [lib.unj.ac.id](http://lib.unj.ac.id)

**LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI  
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademika Universitas Negeri Jakarta, yang bertanda tangan di bawah ini, saya:

Nama : Rizkha Hayati  
NIM : 1314617012  
Fakultas/Prodi : FMIPA/Statistika  
Alamat email : rhayati1811@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada UPT Perpustakaan Universitas Negeri Jakarta, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah:

Skripsi     Tesis     Disertasi     Lain-lain (.....)

yang berjudul :

Penyeleksian Kelompok Peubah yang Memengaruhi Skor Sains PISA Indonesia Menggunakan Group LASSO

Dengan Hak Bebas Royalti Non-Eksklusif ini UPT Perpustakaan Universitas Negeri Jakarta berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya, dan menampilkan/mempublikasikannya di internet atau media lain secara *fulltext* untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan Universitas Negeri Jakarta, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 25 Februari 2022

Penulis

( Rizkha Hayati )  
*nama dan tanda tangan*

## KATA PENGANTAR

Alhamdulillah Rabbil ‘Alamin, segala puji dan syukur penulis panjatkan kehadirat Allah SWT yang telah melimpahkan rahmat dan nikmat-Nya sehingga akhirnya penulis dapat menyelesaikan skripsi ini sebagai salah satu syarat memperoleh gelar Sarjana Statistika. Jenis penelitian yang dipilih adalah penelitian kualitatif yang berjudul “Penyeleksian Kelompok Peubah yang Memengaruhi Skor Sains Pisa Indonesia Menggunakan *Group Lasso*”.

Selama menyelesaikan skripsi, tentunya tidak mudah dan melibatkan banyak pihak. Oleh karena itu, dalam kesempatan kali ini penulis ingin memberikan ucapan terima kasih kepada :

1. Ibu Vera Maya Santi, M.Si selaku dosen pembimbing I yang telah membimbing, mengarahkan, serta memberikan dorongan semangat sehingga skripsi ini dapat terselesaikan
2. Bapak Dr., Ir. Bagus Sumargo, M.Si selaku dosen pembimbing II serta Koordinator Program Studi Statistika yang telah membimbing penulis secara akademik dan memberikan dukungan moril selama penulis mengenyam pendidikan di Program Studi Statistika UNJ.
3. Seluruh Bapak/Ibu dosen pengajar yang telah memberikan ilmu dan pelajaran yang tak ternilai selama penulis menempuh pendidikan di Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Jakarta.
4. Wakil Dekan Bidang Akademik yaitu Bapak Dr. Esmar Budi, S. Si., MT. dan Dekan FMIPA UNJ Ibu Dr. Muktiningsih N M.Si. yang telah membantu penulis selama penyelesaian studi.
5. Keluarga penulis, orang tua yang telah memberikan semangat, dukungan, serta doa dan kasih sayang kepada penulis.
6. Teman–teman seperjuangan Program Studi Statistika UNJ angkatan 2017, sahabat–sahabat, dan semua pihak yang tidak dapat disebutkan satu persatu yang telah memberikan motivasi, bantuan, do’a serta dukungannya kepada penulis selama menyelesaikan skripsi ini.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna, masih banyak kekurangan dari berbagai sisi. Oleh karena itu, kritik dan saran yang membangun

dari semua pihak akan sangat berarti. Semoga skripsi ini dapat bermanfaat dan membantu banyak pihak, baik di luar maupun di dalam lingkungan akademik tempat penulis menuntut ilmu.

Jakarta, 27 Januari 2022



Rizkha Hayati



## DAFTAR ISI

<b>ABSTRAK .....</b>	<b>i</b>
<b>ABSTRACT .....</b>	<b>ii</b>
<b>KATA PENGANTAR.....</b>	<b>iii</b>
<b>DAFTAR TABEL .....</b>	<b>vii</b>
<b>DAFTAR GAMBAR.....</b>	<b>vii</b>
<b>DAFTAR LAMPIRAN .....</b>	<b>vii</b>
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Perumusan Masalah.....	4
1.3 Tujuan Penelitian.....	5
1.4 Manfaat Penelitian.....	5
<b>BAB II KAJIAN PUSTAKA .....</b>	<b>6</b>
2.1 <i>Programme for International Students Assessment (PISA)</i> .....	6
2.2 Analisis Regresi Linear .....	8
2.3 Metode Kuadrat Terkecil.....	9
2.4 Multikolinearitas .....	11
2.5 <i>LASSO (Least Shrinkage and Selection Operator)</i> .....	13
2.6 <i>Group LASSO</i> .....	15
2.7 Validasi Silang .....	17
<b>BAB III METODOLOGI PENELITIAN .....</b>	<b>19</b>
3.1 Data .....	19
3.2 Proses Analisis Data.....	19
3.3 Diagram Alir.....	20
<b>BAB VI HASIL PENELITIAN DAN PEMBAHASAN.....</b>	<b>22</b>
4.1 Eksplorasi Data.....	22



4.2 Analisis <i>Group LASSO</i> .....	24
<b>BAB V KESIMPULAN DAN SARAN</b> .....	<b>33</b>
5.1 Kesimpulan.....	33
5.2 Saran.....	33
<b>DAFTAR PUSTAKA</b> .....	<b>34</b>
<b>LAMPIRAN</b> .....	<b>38</b>
<b>RIWAYAT HIDUP</b> .....	<b>49</b>



## DAFTAR TABEL

Tabel 4.1 Statistik 5 serangkai dan rata-rata skor sains PISA Indonesia .....	22
Tabel 4.2 Peubah penjelas dengan nilai koefisien korelasi tinggi .....	24
Tabel 4.3 Hasil estimasi parameter berdasarkan $\lambda_{min}$ dan $\lambda_{1se}$ .....	25
Tabel 4.4. Hasil estimasi parameter yang masuk ke dalam model .....	26

## DAFTAR GAMBAR

Gambar 1.0.1 Grafik Skor Sains PISA Indonesia Tahun 2000 - 2018 .....	2
Gambar 2.0.1 Pendugaan dengan Lasso (kiri) dan Pendugaan dengan Regresi Ridge (kanan) .....	15
Gambar 2.0.2 Perbedaan LASSO dan <i>group</i> LASSO dalam pemilihan variabel. 16	
Gambar 3.0.1 Diagram Alir Penelitian .....	21
Gambar 4.0.1 Boxplot skor sains PISA Indonesia .....	22
Gambar 4.2 Normal <i>Q-Q Plot</i> skor sains PISA Indonesia .....	23
Gambar 4.3 Kurva nilai $\lambda$ berdasarkan validasi silang 10-fold .....	24

## DAFTAR LAMPIRAN

Lampiran 1 Tabel Peubah Penjelas dalam Penelitian .....	39
Lampiran 2 Tabel Korelasi Peubah Penjelas .....	44
Lampiran 3 Nilai Lambda dan Jumlah Peubah Penjelas yang Masuk ke dalam Model .....	48