

PENENTUAN RISIKO TOTAL BENCANA ALAM DI INDONESIA YANG DIKELOMPOKKAN DENGAN METODE *K-MEANS CLUSTERING*

YOSUA PARDAMEAN SIPANGKAR



**PROGRAM STUDI AKTUARIA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2025**



PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “Penentuan Risiko Total Bencana Alam di Indonesia yang Dikelompokkan dengan Metode *K-Means Clustering*” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Agustus 2025

Yosua Pardamean Sipangkar
G5402211044



ABSTRAK

YOSUA PARDAMEAN SIPANGKAR. Penentuan Risiko Total Bencana Alam di Indonesia yang Dikelompokkan dengan Metode *K-means clustering*. Dibimbing oleh BERLIAN SETIAWATY.

Indonesia merupakan negara kepulauan yang memiliki tingkat kerentanan tinggi terhadap bencana alam yang sering menimbulkan kerugian ekonomi signifikan. Penelitian ini bertujuan untuk menentukan risiko total bencana alam per tahun di Indonesia dengan mengelompokkan provinsi berdasarkan rata-rata banyak kejadian dan total kerugian menggunakan metode *K-means clustering*. Penelitian dilakukan menggunakan data kejadian bencana dan kerugian dari tahun 2018 hingga 2024 untuk 34 provinsi di Indonesia. Hasil klasterisasi menunjukkan tiga kelompok wilayah berdasarkan risiko bencana. Klaster pertama mencakup provinsi dengan frekuensi kejadian tinggi dan total kerugian sedang (risiko sedang), klaster kedua dengan frekuensi dan total kerugian rendah (risiko rendah), serta klaster ketiga dengan frekuensi sedang dan total kerugian tinggi (risiko tinggi). Setiap klaster dianalisis lebih lanjut untuk menentukan distribusi yang sesuai untuk banyak kejadian dan besar kerugian menggunakan metode *maximum likelihood estimation* dan besar kerugian menggunakan metode momen. Uji kecocokan model dilakukan dengan uji khi-kuadrat dan Anderson-Darling. Hasil menunjukkan bahwa distribusi binomial negatif cocok untuk memodelkan banyak kejadian, sementara distribusi *beta-prime* memberikan hasil yang sesuai untuk besar kerugian di seluruh klaster.

Kata kunci: *beta-prime*, binomial negatif, klaster, *K-means*, risiko total

ABSTRACT

YOSUA PARDAMEAN SIPANGKAR. Determining the Total Risk of Natural Disaster in Indonesia Using the K-Means Clustering Method. Supervised by BERLIAN SETIAWATY.

Indonesia is an archipelagic country with a high level of vulnerability to natural disasters, which often cause significant economic losses. This study aims to estimate the annual total risk of natural disasters in Indonesia by clustering provinces based on the average number of events and total losses using the K-means clustering method. The analysis utilizes disaster occurrence and loss data from 2018 to 2024 across 34 provinces in Indonesia. The clustering results reveal three groups of regions based on disaster risk levels. The first cluster consists of provinces with high event frequency and moderate total losses (medium risk), the second cluster includes provinces with both low frequency and low total losses (low risk), and the third cluster comprises provinces with moderate frequency and high total losses (high risk). Each cluster is further analyzed to identify appropriate probability distributions for the number of events and the size of losses. The number of events is modeled using the Maximum Likelihood Estimation method, while the size of losses is estimated using the method of moments. Goodness-of-fit tests, including the chi-square and Anderson-Darling tests, are applied to evaluate the models. The results indicate that the Negative Binomial distribution is suitable for modeling the number of events, while the Beta Prime distribution provides a good fit for modeling the size of losses across all clusters.

Keywords: beta-prime, clustering, K-means, negative binomial, total risk.



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2025
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.

PENENTUAN RISIKO TOTAL BENCANA ALAM DI INDONESIA YANG DIKELOMPOKKAN DENGAN METODE *K-MEANS CLUSTERING*

YOSUA PARDAMEAN SIPANGKAR

Skripsi
sebagai salah satu syarat untuk memperoleh gelar
Sarjana Aktuaria pada
Program Studi Aktuaria

**PROGRAM STUDI AKTUARIA
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2025**



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Tim Penguji pada Ujian Skripsi:
1. Ruhiyat, M.Si., M.Act.Sc.
2. Nur Agustiani, M.Si.



Judul Skripsi : Penentuan Risiko Total Bencana Alam di Indonesia yang
Dikelompokkan dengan Metode *K-Means Clustering*

Nama : Yosua Pardamean Sipangkar
NIM : G5402211044

Disetujui oleh

Pembimbing 1:
Dr. Berlian Setiawaty, M.S.



Diketahui oleh

Ketua Program Studi:
Dr. Ir. I Gusti Putu Purnaba, DEA.
NIP.196512181990021001



Tanggal Ujian: 31 Juli 2025

Tanggal Lulus:



PRAKATA

Puji dan syukur penulis ucapkan kepada Tuhan Yesus Kristus atas segala kasih karunia-Nya sehingga skripsi dengan judul “ Penentuan Risiko Total Asuransi Bencana Alam di Indonesia yang Dikelompokkan dengan Metode *K-Means Clustering*” ini dapat diselesaikan.

Skripsi ini dapat diselesaikan dengan baik karena doa serta dukungan dari berbagai pihak, sehingga penulis menyampaikan terima kasih kepada:

1. Ibu Dr. Dra. Berlian Setiawaty, M.S. selaku dosen pembimbing penulis yang dengan penuh kesabaran telah meluangkan waktu, berbagi ilmu, memberikan motivasi, serta menyampaikan saran dan masukan yang berharga dalam penyusunan tugas akhir ini.
2. Seluruh dosen dan staf Program Studi Aktuaria IPB yang telah memfasilitasi, membimbing, mendoakan, dan memberikan ilmunya selama masa studi penulis.
3. Orang tua saya, Bapak Josten Sipangkar dan Ibu Posmaida Aritoang, Rizky Nuari Sipangkar, Daniel Sipangkar, Ruth Septriana Sipangkar serta seluruh keluarga besar yang selalu memberikan semangat, dukungan, dan doa tanpa henti.
4. Eunike, Rindu, Ryan, Chanra, Wyliam, Thesa, Kezia, Nata, dan Jesika yang selalu setia menemani selama perkuliahan, memberikan dukungan, serta menjadi tempat berbagi cerita dan keluh kesah.
5. Keluarga Besar Persekutuan Mahasiswa Kristen IPB, Komisi Pelayanan Anak PMK IPB, ACD ASSA FMIPA IPB, dan Paguyuban KSE IPB.
6. Teman-teman Sekretariat KPS IPB, Chandra, Martin, Tulus, Andre, Valdo, Ferdinand, dan Raynaldi yang menjadi teman serumah selama masa perkuliahan.

Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan.

Bogor, Agustus 2025

Yosua Pardamean Sipangkar

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

DAFTAR ISI

DAFTAR TABEL	x
DAFTAR GAMBAR	x
DAFTAR LAMPIRAN	x
I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan	2
1.4 Manfaat	2
II TINJAUAN PUSTAKA	3
2.1 Asuransi, Polis Asuransi, Premi, dan Kejadian	3
2.2 Teori Peluang	3
2.3 Sebaran-Sebaran Khusus	5
2.4 <i>Maximum Likelihood Estimation</i>	6
2.5 Pendugaan Momen	7
2.6 Risiko Total, Nilai Harapan, dan Ragam Risiko Total	8
2.7 Uji Kebaikan Suai	9
2.8 Pengelompokan Data	11
III METODE	13
3.1 Sumber Data	13
3.2 Perangkat Lunak	13
3.3 Langkah-langkah Penelitian	13
IV HASIL DAN PEMBAHASAN	14
4.1 Eksplorasi Data	14
4.2 Klasterisasi <i>K-Means</i>	15
4.3 Sebaran Banyak Kejadian	19
4.4 Sebaran Besar Kerugian	22
4.5 Risiko Total	25
V SIMPULAN DAN SARAN	27
5.1 Simpulan	27
5.2 Saran	27
DAFTAR PUSTAKA	28
LAMPIRAN	30
RIWAYAT HIDUP	47



DAFTAR TABEL

1	Rekapitulasi banyak kejadian dan total kerugian akibat bencana per tahun di Indonesia	14
2	Hasil <i>z-score</i> rata-rata kejadian dan total kerugian per tahun	15
3	Nilai <i>sum of square error</i> untuk berbagai jumlah klaster	16
4	Hasil klasterisasi <i>K-means</i>	18
5	Statistik deskriptif banyak kejadian tiap klaster	19
6	Nilai dugaan parameter sebaran banyak kejadian	20
7	Tabel uji khi-kuadrat klaster 1	21
8	Tabel uji khi-kuadrat klaster 2	21
9	Tabel uji khi kuadrat klaster 3	21
10	Perbandingan nilai harapan dan ragam tiap klaster	21
11	Statistik deskriptif data besar kerugian tiap klaster	22
12	Parameter duga sebaran besar kerugian tiap klaster	24
13	Perbandingan nilai harapan dan ragam sebaran <i>beta-prime</i> dan data empirik (dalam ribu rupiah)	25
14	Perbandingan nilai raran dan ragam risiko total tiap klaster (dalam ribu rupiah)	25

DAFTAR GAMBAR

1	Metode <i>elbow</i> untuk <i>k</i> yang optimal	17
2	Visualisasi hasil klasterisasi	17
3	Histogram banyak kejadian tiap klaster	19
4	Perbandingan peluang empirik dan fungsi massa peluang sebaran binomial negatif tiap klaster	20
5	Histogram data besar kerugian tiap klaster	22
6	Plot <i>Cullen and Frey graph</i> tiap klaster	23
7	Plot besarnya klaim dengan fungsi kepekatan peluang sebaran <i>beta-prime</i> tiap klaster	24

DAFTAR LAMPIRAN

1	Data penelitian	31
2	Pembuktian nilai harapan dan ragam sebaran <i>beta-prime</i>	41
3	Pembuktian pendugaan parameter sebaran <i>beta-prime</i> dengan metode momen	42
4	Program <i>K-means clustering</i>	43
5	Program pendugaan sebaran banyak kejadian	44
6	Program pendugaan sebaran besar kerugian	46