

# Tugas Besar Komputasi Statistik

**Pengembangan Fungsi R untuk  
Analisis Sebaran Spasial-  
Temporal dan Clustering  
Hotspot Karhutla di Provinsi  
Sumatera Selatan (Januari - Juli  
2025)**

Kelompok 3 RB - Pemrograman Fungsi R

01

# Anggota Kelompok



Luthfia Laila Ramadhani -  
123450004



Feby Angelina -  
123450039



Ginda Fajar Riadi Marpaung -  
123450103



Muhammad Dzikra -  
123450124

# Pendahuluan

## Latar Belakang

- Permasalahan Karhutla di Indonesia sangat mengancam ekosistem dan kesehatan masyarakat. Daerah Sumatera Selatan termasuk wilayah paling rawan terhadap Karhutla
- Data hotspot dari satelit VIIRS dapat digunakan untuk memantau dan menganalisis pola sebaran titik api dan perlu diolah agar menghasilkan informasi spasial-temporal yang berguna
- Diperlukan fungsi otomatis R untuk memproses, memetakan, dan menganalisis hotspot secara cepat dan akurat

## Rumusan Masalah

- Bagaimana merancang fungsi R untuk data wrangling & analisis spasial-temporal?
- Bagaimana implementasi K-Means untuk zonasi hotspot?
- Bagaimana memvalidasi stabilitas cluster dengan bootstrap?
- Bagaimana menghasilkan visualisasi & output yang informatif?

## Tujuan

- Membangun fungsi R untuk analisis hotspot otomatis
- Menerapkan K-Means clustering
- Mengintegrasikan bootstrap validation
- Menghasilkan grafik, peta, dan ringkasan statistik lengkap

03

# Metode Penelitian

## Sumber Data

Dataset yang digunakan dalam penelitian ini bersumber dari NASA FIRMS, sebuah platform penyedia data titik panas (hotspot) secara near real-time yang dioperasikan oleh NASA. Data hotspot diperoleh dari sensor VIIRS yang terpasang pada satelit NOAA-20 (2 Januari–31 Juli 2025) dalam format CSV dan Shapefile, berisi informasi lengkap seperti koordinat, brightness, FRP, confidence, serta waktu akuisisi. Sensor VIIRS beresolusi 375 meter dengan dua lintasan harian dan latensi kurang dari 3 jam.

## Teknik Analisis

Data preprocessing → Konversi ke objek spasial (sf) → Analisis bulanan & per jam → K-Means clustering ( $k=3$ ) → Bootstrap 100 iterasi → Visualisasi

# Dataset

Variable	Mean	Min	Max	Std Dev
Latitude	-2.721	-4.499	-1.535	0.839
Longitude	104.236	102.001	106.49	1.445
Brightness (K)	334.15	295.30	367.00	12.48
FRP (MW)	7.46	1.00	73.70	6.92

## Statistik Deskriptif Variabel Utama

# Fungsi R yang Dikembangkan

## 1. Data Wrangling Functions

- `load_and_clean_hotspot()` → membersihkan, filter FRP, mengolah waktu
- `convert_to_spatial()` → mengubah dataframe menjadi objek spasial (sf)
- `join_with_shapefile()` → menggabungkan hotspot dengan shapefile wilayah
- `create_spatial_grid()` → membuat grid dan menghitung statistik per grid
- `get_data_summary()` → menghasilkan ringkasan statistik dataset

## 2. Temporal Analysis Functions

- `analyze_monthly_distribution()` → analisis hotspot bulanan
- `analyze_hourly_pattern()` → analisis pola hotspot berdasarkan jam

## 3. Clustering Functions

- `perform_kmeans_clustering()` → mengelompokkan hotspot menggunakan K-Means
- `find_optimal_k()` → menentukan jumlah cluster optimal (Elbow Method)

# Fungsi R yang Dikembangkan

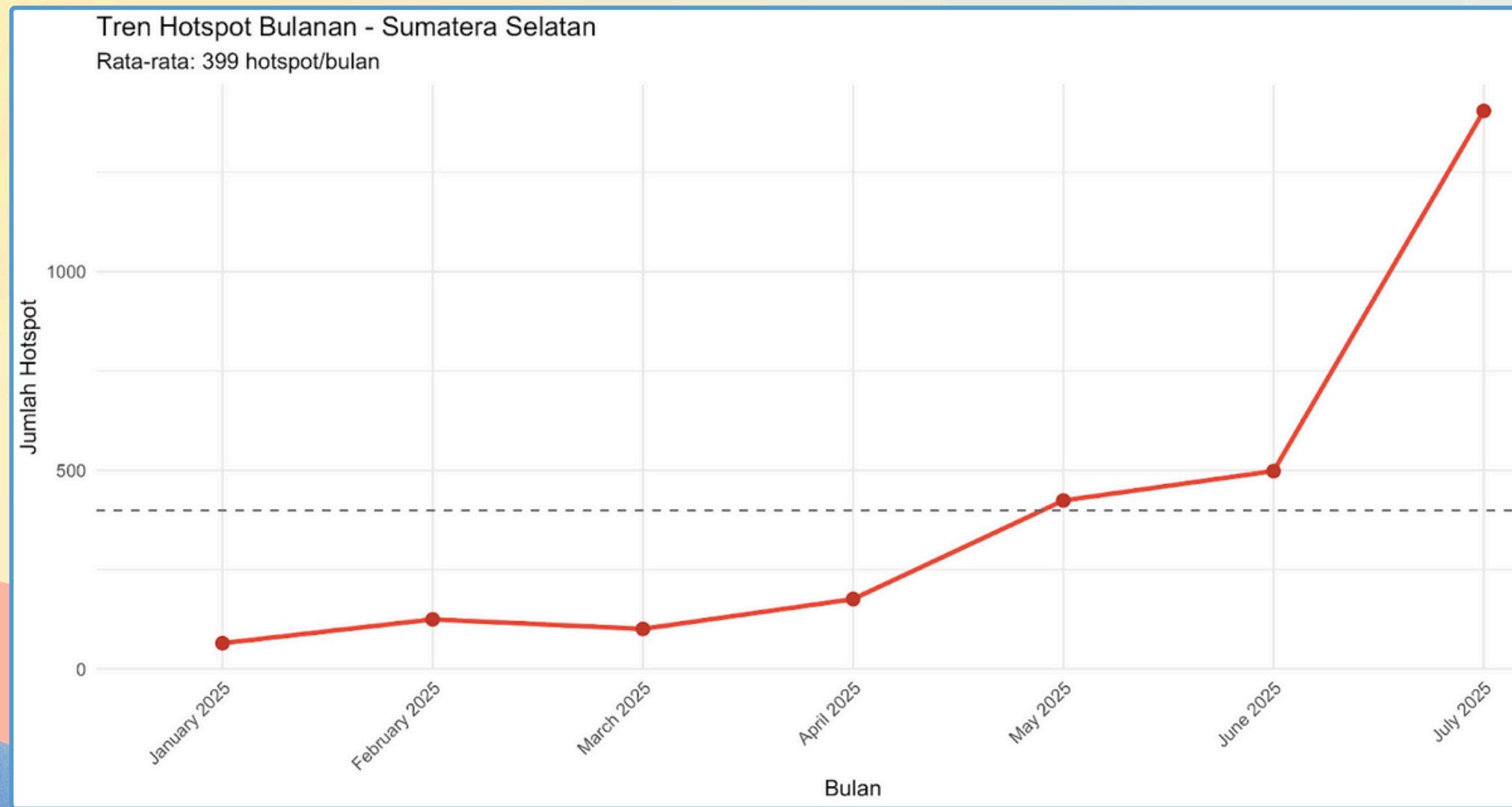
## 4. Bootstrap Validation

- `bootstrap_cluster_validation()` – menilai stabilitas cluster (100 iterasi)

## 5. Visualization Functions

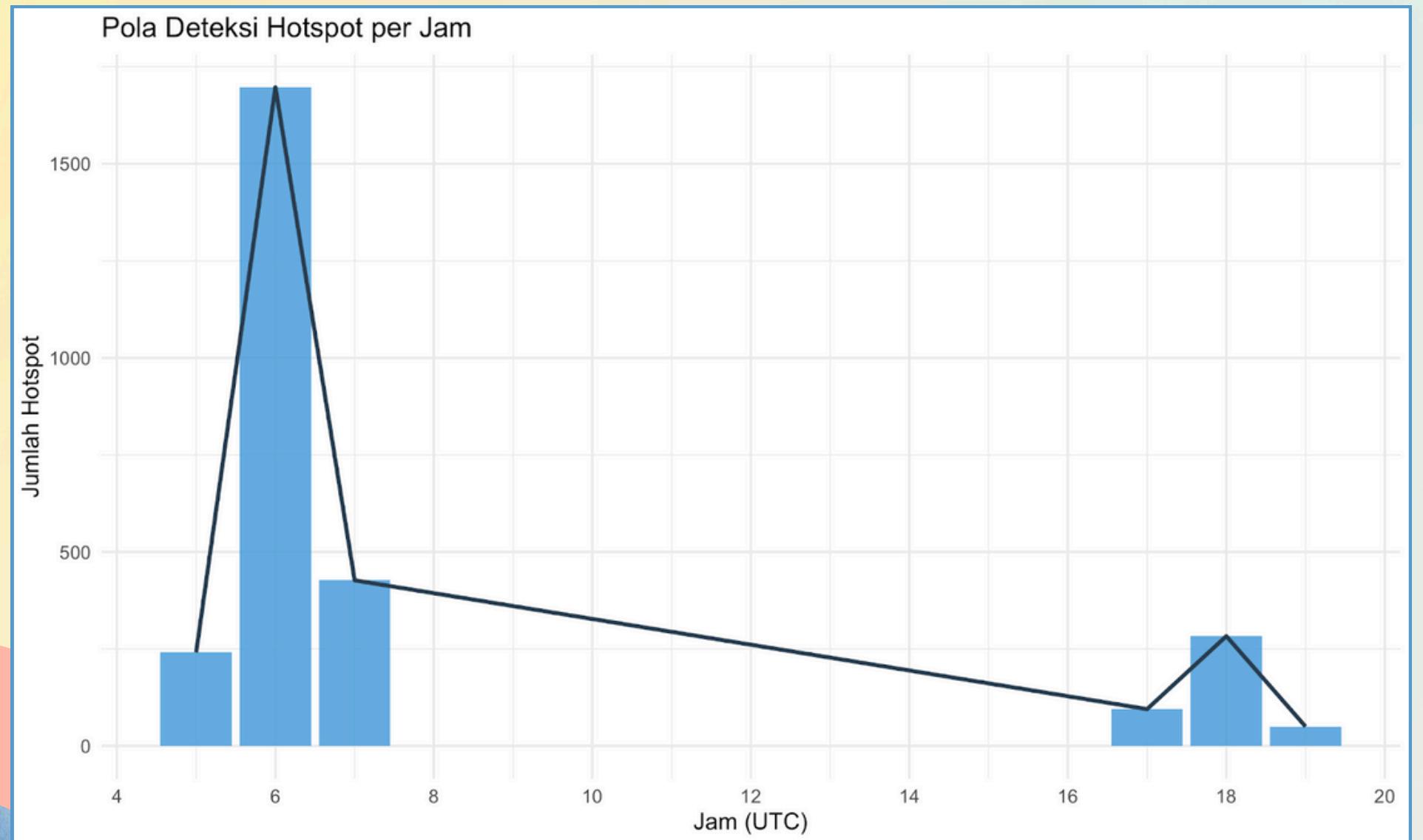
- `create_interactive_map()` → peta interaktif leaflet
- `plot_monthly_trend()` → grafik tren bulanan
- `plot_hourly_pattern()` → grafik pola per jam
- `plot_cluster_spatial()` → visualisasi cluster spasial
- `plot_bootstrap_stability()` → grafik stabilitas bootstrap
- `plot_elbow_method()` → grafik metode elbow
- `plotRegionalDistribution()` → distribusi hotspot per wilayah

# Analisis Trend Bulanan



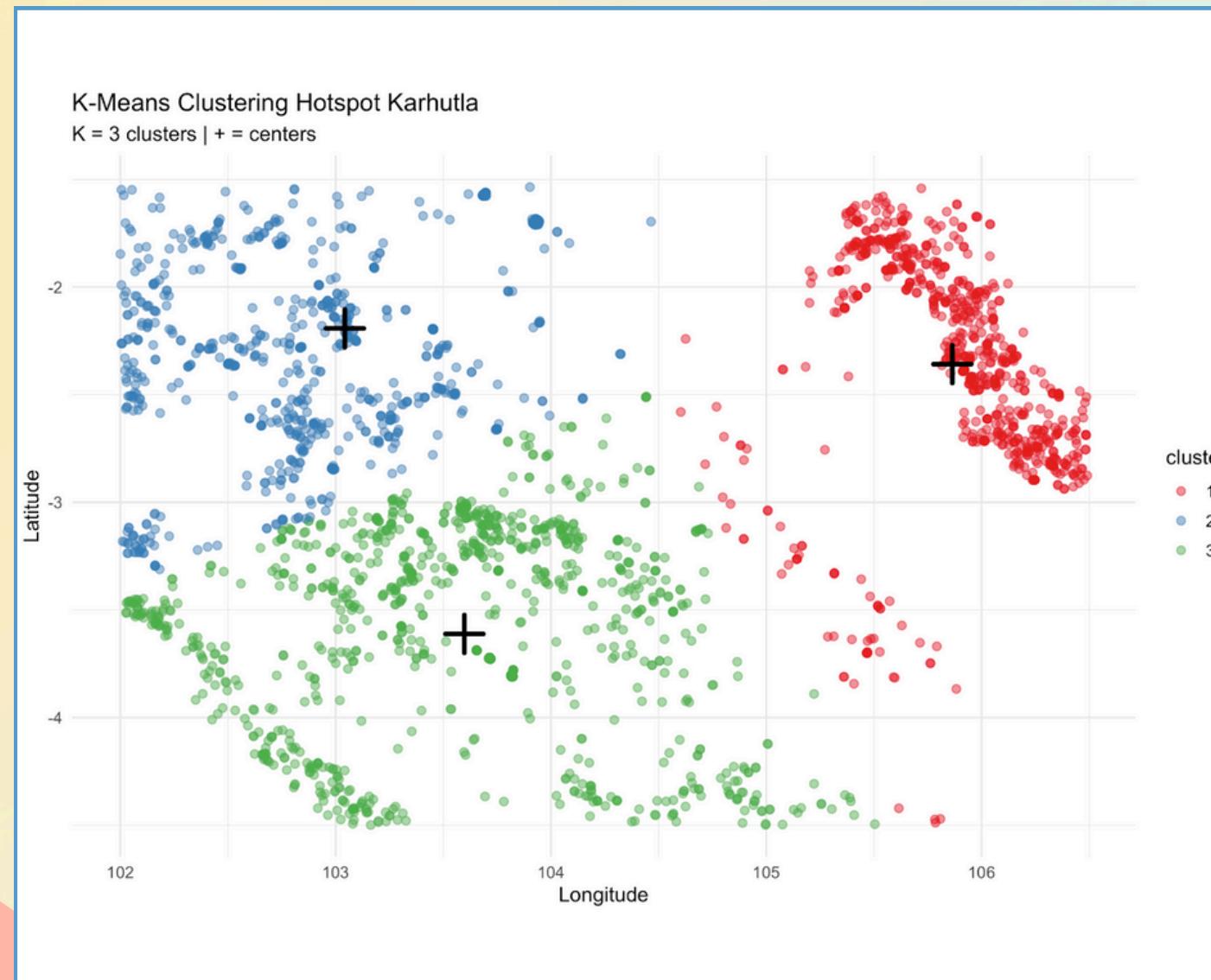
Tren hotspot Januari–Juli 2025 menunjukkan peningkatan eksponensial, dengan puncak pada Juli sebesar 1.404 hotspot atau 3,5 kali rata-rata bulanan. Pola ini mencerminkan masuknya musim kemarau yang semakin intens pada pertengahan tahun. Kenaikan bertahap dari Januari hingga April menunjukkan transisi dari musim hujan menuju kondisi kering, lalu meningkat tajam mulai Mei dan Juni. Intensitas kebakaran juga meningkat, terlihat dari kenaikan FRP rata-rata dari 5,31 MW pada Februari menjadi 8,29 MW pada Juli.

# Pola Deteksi Per Jam



Distribusi temporal hotspot menunjukkan puncak utama pada 06:00 UTC dengan lebih dari 1.600 deteksi, serta puncak sekunder pada 17:00–18:00 UTC. Aktivitas hotspot sangat rendah pada rentang 08:00–16:00 UTC, membentuk pola bimodal yang jelas. Puncak pagi hari berkaitan dengan waktu lintasan utama satelit NOAA-20 serta kondisi atmosfer yang optimal untuk deteksi thermal. Pola ini penting untuk mengoptimalkan jadwal monitoring dan mendukung respons pemadaman kebakaran.

# Statistik Clusterin



## Cluster 1 – Zona Timur/Pesisir warna Merah

- 722 titik
- FRP tertinggi (8.73 MW)
- Dominan lahan gambut

## Cluster 2 – Zona Barat warna Biru

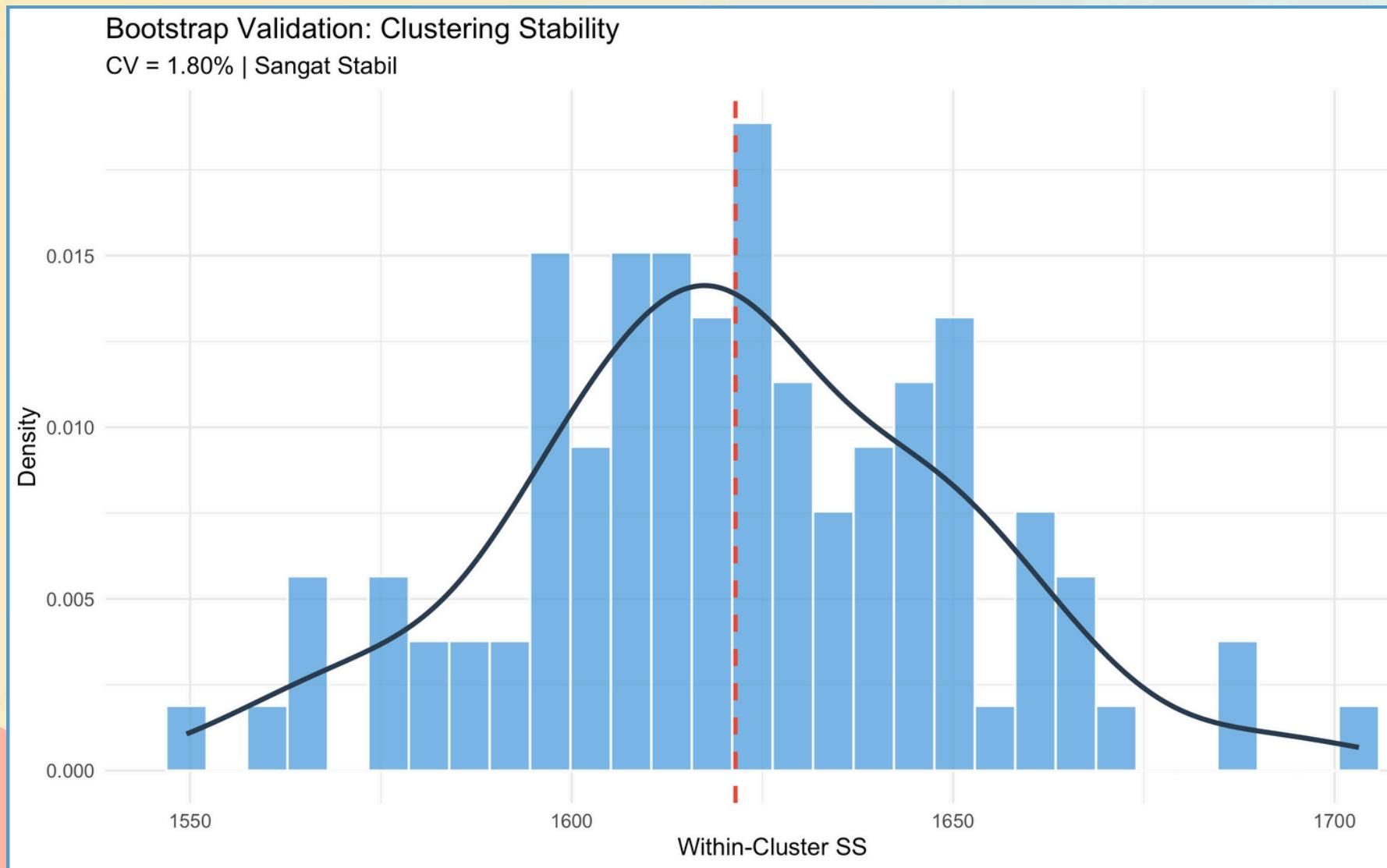
- 894 titik
- FRP sedang (6.87 MW)
- Wilayah transisi menuju Bukit Barisan

## Cluster 3 – Tengah-Selatan warna Hijau

- 1.177 titik
- Frekuensi tertinggi
- Area campuran pertanian–pemukiman

- Between-cluster SS bernilai 74.5% yang artinya pemisahan sangat baik
- Total WSS = 1.617,9
- Cluster jelas dan tidak saling tumpang tindih

# Bootstrap Validation



Histogram menampilkan sebaran WSS yang relatif rapat dengan pola yang mendekati distribusi normal, sementara garis merah putus-putus menandai nilai rata-rata WSS. Kepadatan distribusi yang sempit serta nilai CV = 1.80% menunjukkan bahwa hasil clustering sangat stabil dan konsisten terhadap variasi data. Grafik ini memberikan bukti kuat bahwa struktur klaster yang diperoleh bersifat robust dan tidak sensitif terhadap perubahan sampel.

# Penutupan

## Kesimpulan

1. Berhasil dibuat fungsi R lengkap untuk analisis hotspot
2. Tren hotspot menunjukkan kenaikan eksponensial, puncak Juli
3. Pola harian dominan pukul 06:00 UTC
4. K-Means berhasil membentuk 3 cluster geografis yang jelas
5. Bootstrap menunjukkan stabilitas sangat tinggi (CV 1.80%)

## Saran

Pengembangan fungsi R:

1. Tambahkan penanda wilayah administratif otomatis
2. Bandingkan K-Means dengan DBSCAN untuk data noisy

Kebijakan:

1. Prioritaskan zona pesisir timur (FRP tinggi)
2. Intensifkan pemantauan pada 12–14 WIB
3. Lakukan mitigasi sebelum Mei karena tren mulai naik

**TERIMA  
KASIH**