

PEMETAAN POLA DISTRIBUSI WISATAWAN MENGGUNAKAN ANALISIS PENGELOMPOKAN DATA MEDIA SOSIAL DI PROVINSI BALI TAHUN 2013 – 2024

Naurin Tazkiya Tabarani

21/480320/GE/09671

INTISARI

Peningkatan aktivitas wisata di Provinsi Bali pascapandemi COVID-19 menunjukkan dinamika yang signifikan, terutama pada wilayah dengan konsentrasi wisatawan tinggi. Aktivitas wisata tersebut tidak terdistribusi secara merata, sehingga diperlukan pemahaman yang lebih mendalam mengenai pola distribusi spasial aktivitas wisatawan. Data resmi kunjungan wisatawan umumnya masih disajikan dalam bentuk statistik agregat, sehingga belum mampu menggambarkan persebaran aktivitas wisatawan secara spasial. Oleh karena itu, penelitian ini bertujuan untuk menyusun basis data spasial aktivitas wisatawan di Provinsi Bali periode 2013–2024 menggunakan data geotagging foto dari media sosial Flickr, menganalisis pola persebarannya melalui metode klusterisasi DBSCAN, serta mengkaji keterkaitannya dengan data kunjungan wisatawan resmi.

Data geotagging Flickr diperoleh melalui proses *data retrieval* dan selanjutnya dipra-pemrosesan untuk menghasilkan data yang layak dianalisis sebagai representasi lokasi aktivitas wisatawan. Analisis klusterisasi DBSCAN diterapkan untuk mengidentifikasi konsentrasi dan pola spasial aktivitas wisatawan, sementara analisis korelasi dan regresi digunakan untuk menguji hubungan antara data Flickr dan data kunjungan wisatawan resmi dari Dinas Pariwisata Provinsi Bali.

Hasil penelitian menunjukkan bahwa dari total 117.020 unggahan, sebanyak 20.373 data geotagging memenuhi kriteria analisis. Penerapan DBSCAN menghasilkan 181 kluster dari 87 lokasi, dengan dominasi kategori wisata alam. Distribusi kluster menunjukkan konsentrasi aktivitas yang tinggi di wilayah Bali Selatan, khususnya di Pantai Kuta dan Pantai Seminyak. Hasil analisis korelasi dan regresi memperlihatkan hubungan positif dan konsisten antara data Flickr dan data kunjungan wisatawan resmi. Hasil penelitian ini menunjukkan bahwa data geotagging Flickr berpotensi sebagai sumber informasi pelengkap dalam pemetaan dan pemantauan distribusi spasial aktivitas wisatawan di Provinsi Bali.

Kata kunci: DBSCAN, Geotagging, Flickr, Pola Spasial, Pariwisata Bali

***A CLUSTERING-BASED ANALYSIS OF SOCIAL MEDIA DATA FOR
MAPPING TOURIST DISTRIBUTION PATTERNS IN BALI PROVINCE
(2013 – 2024)***

Naurin Tazkiya Tabarani

21/480320/GE/09671

ABSTRACT

The increase in tourism activity in Bali Province following the COVID-19 pandemic has exhibited significant dynamics, particularly in areas with high tourist concentrations. These tourism activities are not evenly distributed, highlighting the need for a deeper understanding of the spatial distribution patterns of tourist activities. Official tourist visitation data are generally presented in aggregated statistical form and therefore remain insufficient to represent the spatial distribution of tourist activities. Accordingly, this study aims to develop a spatial database of tourist activities in Bali Province for the period 2013–2024 using geotagged photo data from the social media platform Flickr, to analyze spatial distribution patterns using the DBSCAN clustering method, and to examine their relationship with official tourist visitation data.

Geotagged Flickr data were obtained through a data retrieval process and subsequently preprocessed to generate datasets suitable for representing tourist activity locations. DBSCAN clustering analysis was applied to identify concentrations and spatial patterns of tourist activities, while correlation and regression analyses were employed to examine the relationship between Flickr data and official tourist visitation data from the Bali Provincial Tourism Office.

The results show that out of a total of 117,020 uploaded photos, 20,373 geotagged records met the criteria for analysis. The application of DBSCAN produced 181 clusters across 87 locations, with nature-based tourism dominating the cluster categories. The spatial distribution of clusters indicates a high concentration of tourist activities in southern Bali, particularly at Kuta Beach and Seminyak Beach. The correlation and regression analyses reveal a positive and consistent relationship between Flickr data and official tourist visitation statistics. These findings demonstrate that Flickr geotagged data have strong potential as a complementary information source for mapping and monitoring the spatial distribution of tourist activities in Bali Province.

Keywords: *DBSCAN, Geotagging, Flickr, Spatial Pattern, Bali Tourism*