

# BAB I

## PENDAHULUAN

### 1.1 Latar belakang

Pariwisata merupakan sebuah kegiatan yang memiliki beberapa bidang di mana diantaranya terdapat bidang ekonomi, bidang budaya, bidang sosial, bidang agama, bidang lingkungan dan beberapa bidang lainnya dijelaskan dalam Undang-Undang Nomor 10 Tahun 2009 yang membahas tentang kepariwisataan memaparkan bahwa sebuah kepariwisataan yang diselenggarakan memiliki tujuan supaya pendapatan nasional semakin meningkat sehingga dapat memajukan kesejahteraan serta kemakmuran rakyat, membuka lapangan pekerjaan dan kesempatan untuk berwirausaha, menyokong pemerataan pembangunan di daerah, memperluas informasi mengenai objek dan daya tarik pariwisata di Indonesia dengan tujuan menumbuhkan nasionalisme dan memperkuat hubungan antar bangsa (Batubara, 2020).

Indonesia merupakan sebuah Negara yang luas serta memiliki banyak pulau di mana Indonesia mencatat 17.504 pulau yang bahkan tidak sedikit warga Indonesia yang tidak mengetahui pulau apa saja yang dimiliki Indonesia, Kepulauan Riau menjadi Provinsi dengan pulau terbanyak yaitu 2.408 pulau, berikutnya terdapat Provinsi Papua Barat yang memiliki 1.945 pulau, wilayah Indonesia memiliki luas yang mencapai 1,91 juta km persegi (Rahma, 2020), banyak pulau di Indonesia yang memiliki keindahan alam yang dapat menarik perhatian wisatawan dalam negeri ataupun wisatawan mancanegara. Destinasi wisata yang ada di pulau Sumatera antara lain Pesisir Kalinda di Lampung, Danau Ranau di Sumatera Selatan, Lembah Harau, Payakumbuh di Sumatera Barat, Kawasan Danau Toba dan Pulau Samosir di Sumatera Utara, sedangkan di pulau Jawa antara lain Borobudur, dan Kepulauan Karimunjawa yang berada di Jawa Tengah, Taman Nasional Bromo Tengger Semeru yang berada di Jawa Timur.

Negara Indonesia termasuk dalam Negara yang memiliki potensi wisata yang beragam, keberagaman bahasa, suku bangsa, serta kuliner yang didukung dengan keindahan alam, budaya leluhur serta kultur yang ada di Indonesia menjadi daya

tarik yang menjanjikan untuk menjadikan Indonesia Negara pantas dikunjungi untuk berwisata, namun banyak daerah di Indonesia yang belum mampu mengelola serta memanfaatkan potensi yang dimiliki, dan sumber daya manusia masih menjadi peranan terpenting yang diandalkan dalam pengembangan sektor pariwisata (Melinda, 2019), dalam sebuah keterangan tertulis pada Selasa 28 November 2023 Wakil ketua MPR RI, Lestari Moerdijat memaparkan bahwa “Peningkatan keterampilan dan kualitas SDM adalah salah satu kunci dalam pengembangan sektor pariwisata nasional”.

Bagi sebuah Negara sektor pariwisata menjadi gagasan yang dikembangkan agar pendapatan yang cukup besar (Kiwang & Arif, 2020), kedatangan wisatawan mancanegara ke Indonesia dapat menambah devisa Negara serta membuka banyak lapangan kerja bagi masyarakat sehingga memiliki pengaruh besar disektor ekonomi dengan terbukanya kesempatan berusaha yang lebih luas, pendapatan yang lebih merata, serta meningkatnya pajak dan penerimaan daerah melalui retribusi yang berdampak meningkatnya penerimaan Negara. Kementerian Pariwisata dan Ekonomi Kreatif mencatat pada tahun 2022 pendapatan devisa berasal dari sektor pariwisata di Indonesia bahkan dapat mencapai nilai US\$6,72 miliar, dan meningkat 55,65% pada tahun 2023 dengan tercatatnya pendapatan devisa mencapai US\$10,46 miliar terhitung hingga September 2023, meskipun sempat mengalami penurunan cukup besar ditahun 2020 sebesar 80,97% dengan pendapatan devisa mencapai US\$3,38 miliar karena adanya pandemi *Covid-19* saat ini pendapatan devisa dari sektor pariwisata telah kembali meningkat.

Persaingan bisnis saat ini sangat memerlukan proses pengolahan data yang menuntut untuk mendapatkan informasi dengan cepat, di mana informasi tersebut terkait data yang digunakan dalam proses bisnis, penentuan strategi . Pengelolaan data yang besar memerlukan *data mining* dengan tujuan untuk mempermudah aktivitas *recording* pada transaksi dan dalam proses data *warehousing* sehingga informasi yang diberikan kepada *user* menjadi akurat (Riskha & Farokhah, 2023), *data mining* sendiri merupakan proses untuk memperoleh informasi dan data tertentu dalam jumlah besar yang telah dikumpulkan. Di dalam prosesnya seringkali memanfaatkan metode-metode antara lain, statistika, matematika, serta pemanfaatan terhadap berkembangnya teknologi *artificial intelligence (AI)*. *Data*

*mining* secara umum juga disebut dengan *Knowledge Discovery in Databases (KKD)*, *Data Analysis*. *Data mining* secara umum memiliki fungsi yaitu deskriptif dan prediktif, namun *data mining* juga memiliki beberapa fungsi lainnya antara lain, *clustering*, klasifikasi, asosiasi, *forecasting*, serta *sequencing*. Pada penelitian ini akan memfokuskan fungsi *data mining* dalam proses *clustering*. Penelitian serupa sebelumnya pernah dilakukan pada tahun 2019 dengan penelitian berjudul “Penerapan *Clustering* dalam Mengelompokkan Jumlah Kunjungan Wisatawan dengan Metode *K-Means*” (Satria et al., 2019), namun pada penelitian ini menggunakan data Data Jumlah Kunjungan Wisatawan Mancanegara Tahun 2017/2018.

*Clustering* sering disebut dengan *segmentation*, metode ini berguna dalam melakukan identifikasi kelompok alami dalam suatu kasus yang menjadikan suatu kelompok atribut pengelompokan data yang mempunyai kemiripan antar atribut. *Clustering* merupakan sebuah metode *data mining unsupervised*, dalam memandu proses pembelajaran tidak ada satu atribut pun yang digunakan, sehingga keseluruhan atribut input adalah sama. Secara umum *clustering* adalah kumpulan teknik yang digunakan dalam proses pembagian data menjadi kelompok (Ni Luh Wiwik Sri Rahayu Ginantra, Fatimah Nur Arifah Anggi Hadi Wijaya, Ri Sabti Septarini, Nazaruddin Ahmad Dewa Putu Yudhi Ardiana, Faried Effendy, Akbar Iskandar Hazriani, Ika Yusnita Sari & Cahyo Prianto, Dudih Gustian, 2021). *Clustering* dalam penerapannya digunakan dalam memprediksi dan menganalisa permasalahan bisnis, seperti segmentasi pasar, pemetaan zonasi wilayah, *marketing*, identifikasi objek seperti *computer vision* dan *image processing*. Terdapat beberapa tipe data yang digunakan dalam *clustering* antara lain, *Variable* berskala *interval*, *Variabel biber*, *Variable nominal*, *ordinal*, dan *rasio*, Variabel dengan tipe lainnya.

Algoritma *clustering* memiliki syarat yang menjadi tantangan yang harus dipenuhi menurut Han dan Kamber pada 2012 antara lain, skalabilitas di mana *clustering* diharuskan dapat menangani data dengan kapasitas yang besar, serta kemampuan dalam menganalisa bermacam-macam bentuk data di mana *clustering* diharuskan dapat diimplementasikan terhadap semua bentuk data, bahkan ketika ditemukan *cluster* dengan bentuk yang tidak dapat diduga, kemampuan dalam

proses penanganan dalam menangani *noise* karena terdapat kemungkinan sebuah data dalam keadaan yang rusak, tidak bisa dipahami, ataupun hilang, *sensitive* terhadap perubahan input karena ketika menggunakan algoritma *clustering* dengan tingkat *sensitive* rendah akan menimbulkan perubahan yang mencolok ketika terdapat perubahan atau penambahan data pada *input*, maka memiliki kemampuan melakukan *clustering* terhadap data dimensi tinggi, *interpretasi* dan kegunaan karena *clustering* harus dapat menginterpretasikan hasilnya serta memiliki hasil yang berguna. Umumnya *clustering* mempunyai dua metode populer yaitu, *hierarchical clustering*, dan *partitional clustering*. Namun terdapat pula metode lain yang digunakan dalam implementasi *clustering* yaitu *Density-Based*, dan *Grid-Based*.

*K-Means* merupakan algoritma yang proses pengelompokannya dilakukan secara berulang. Pada prosesnya suatu nilai ditetapkan ke suatu *cluster*, proses ini dilakukan secara acak atau secara tidak teratur, dengan sebuah nilai yang dijadikan nilai tengah disebut *centroid*, atau *mean*. Algoritma yang dianggap sebagai algoritma yang cukup mudah untuk difahami adalah *K-Means Clustering*, hal tersebut disebabkan karena *K-Means* memiliki kemampuan mengolah data dalam jumlah besar dengan cepat, efisien, dan efektif. *K-Means* juga memiliki kekurangan yaitu saat penentuan *cluster* awal sangat bergantung pada data masukan yang diberikan (Hellyana & Fadlilah, 2023). Pada dasarnya *K-Means* juga mempunyai beberapa titik kelemahan yaitu tidak mudah untuk menentukan secara mutlak jumlah *cluster* yang tepat, serta terjadinya kegagalan untuk menemukan titik pusat.

*K-Means* merupakan sebuah metode yang dapat dimanfaatkan dalam proses pengelompokan data yang diolah menjadi beberapa bagian, dalam hal ini metode *K-Means* yang menjadikan jarak untuk membagi *cluster-cluster* yang hanya dapat digunakan ketika atribut tersebut berupa *numeric*, *K-Means* juga merupakan sebuah metode yang *non-hierarki* di mana dalam tahapannya sebuah pusat *cluster* didapatkan dengan memilih secara *random* yang berasal dari kumpulan data komponen pada populasi data sehingga komponen tersebut ditetapkan sebagai salah satu pusat *cluster* yang nantinya didefinisikan sesuai dengan jarak minimum antara komponen dengan setiap *clusternya* (Febriyanto et al., 2021). *K-Means* memiliki beberapa karakteristik antara lain, cepat dalam proses *clustering*, sensitif terhadap

nilai *centroid*, memiliki hasil yang selalu berubah-ubah, sulitnya mendapat *global optimum* (di mana sebuah algoritma mendapat nilai tertinggi atau nilai terendah untuk seluruh kandidat solusi dalam suatu masalah tertentu. Metode *K-Means Clustering* ini telah digunakan dalam beberapa penelitian antara lain, “Penerapan Metode *K-Means* untuk *Clustering* Pengunjung Perpustakaan Itn Malang”(Febriyanto et al., 2021), “Implementasi Algoritma *K-Means* dalam Pengelompokan Jumlah Wisatawan Akomodasi di Jawa Barat”(Lusianah et al., 2023), “Implementasi *K-Means Clustering* Dalam Mengelompokkan Minat Membaca Penduduk Menurut Wilayah”(Rizki et al., 2021), “Segmentasi Pengunjung Pusat Perbelanjaan Menggunakan Metode *K-Means Clustering*”(Maulana et al., 2022), “Implementasi *Data Mining* Untuk Pengelompokan Buku Menggunakan Algoritma *K-Means Clustering* (Studi Kasus : Perpustakaan Politeknik LPP Yogyakarta)”(Hasanah & Purnomo, 2022), “Penerapan *K-Means Clustering* Pada Pariwisata Kabupaten Bojonegoro Untuk Mendukung Keputusan Strategi Pemasaran”(Al-Fahmi et al., 2023).

Merujuk kembali pada penelitian yang pernah dilakukan sebelumnya maka di dalam proses penelitian ini akan dilakukan pengklusteran pada data jumlah kunjungan wisatawan mancanegara ke Indonesia berdasarkan pintu gerbang kedatangannya pada Januari 2017 sampai Oktober 2023 dengan menerapkan algoritma *K-Means clustering*.

## 1.2 Rumusan Masalah

Penelitian ini memiliki rumusan masalah antara lain:

1. Bagaimana cara mengklusterkan wisatawan mancanegara ke Indonesia yang datang melalui 15 pintu masuk dari jalur udara dengan metode *K-Means Clustering*?
2. Bagaimana proses pembuatan *website* yang di dalamnya menerapkan metode *K-Means* untuk membantu pengklusteran data wisatawan mancanegara ke Indonesia melalui 15 pintu masuk dari jalur udara?

### 1.3 Tujuan penelitian

Penelitian ini memiliki tujuan penelitian antara lain:

1. Untuk mengetahui hasil *clustering* wisatawan mancanegara ke Indonesia yang datang melalui 15 pintu masuk dari jalur udara dengan metode *K-Menas Clustering*.
2. Untuk mengetahui proses pembuatan *website* yang didalamnya menerapkan metode *K-Means* untuk membantu pengklusteran data wisatawan mancanegara ke Indonesia melalui 15 pintu masuk dari jalur udara.

### 1.4 Batasan Masalah

Penelitian ini memiliki batasan masalah antara lain:

1. Penelitian ini nantinya akan mengfokuskan penelitiannya terhadap jumlah wisatawan mancanegara yang datang ke Indonesia melewati 15 pintu masuk dari jalur udara mulai Januari 2017 – oktober 2023.
2. Penelitian hanya berfokus pada kategori pintu masuk jalur udara

### 1.5 Manfaat Penelitian

Penelitian ini memiliki manfaat penelitian antara lain.

1. Dapat dijadikan sebagai bahan evaluasi dalam meningkatkan jumlah wisatawan mancanegara ke Indonesia.
2. Pihak-pihak yang bersangkutan dapat melakukan sebuah perencanaan dan pengambilan keputusan mengenai strategi yang dapat meningkatkan jumlah wisatawan yang datang ke Indonesia yang memberikan dampak positif.
3. Dapat membantu industri pariwisata di Indonesia untuk setiap daerah pada proses pengembangan strategi yang dapat meningkatkan keuntungan seperti memperluas dan menemukan ide-ide kreatif untuk mengembangkan pariwisata yang ada di setiap daerah yang dapat dijadikan daya tarik wisatawan mancanegara untuk datang ke Indonesia pada kesempatan yang akan datang.

4. Dapat dijadikan informasi yang digunakan bahan acuan untuk perbaikan infrastruktur dan sarana penunjang kenyamanan wisatawan di area dengan *cluster* jumlah kunjungan tinggi, dan pengembangan pariwisata pada area dengan *cluster* jumlah kunjungan rendah.



# UNUGIRI