

BAB III METODE PENELITIAN

3.1. Objek Penelitian

Objek penelitian ini yaitu dataset kunjungan wisatawan asing berdasarkan jalur masuk udara, darat, laut ke Indonesia perbulannya pada tahun 2017 sampai 2021, diperoleh dari website Badan Pusat Statistik dengan jumlah data sebanyak 1.572. Data diambil pada hari sabtu 05 Februari 2022, kemudian data tersebut akan diproses menggunakan perhitungan excel dan *Python* dengan melalui beberapa tahapan data mining.

Badan Pusat Statistik (BPS) merupakan lembaga non-kementerian yang bertanggung jawab langsung kepada Presiden. BPS merupakan Badan Pusat Statistik yang dibuat berdasarkan Undang-Undang Sensus Nomor 6 Tahun 1960 dan Undang-Undang Statistik Nomor 7 Tahun 1960.

3.2. Prosedur Penelitian

Prosedur penelitian ini dilakukan, agar penelitian tersusun rapih untuk memperoleh hasil sesuai dengan tujuan penelitian. Berikut merupakan prosedur penelitian:



Gambar 3.1 Prosedur Penelitian

3.2.1 Pengumpulan Data

Dalam proses ini maka dibutuhkan data yang sumbernya valid dan relevan. Maka untuk memperoleh data tersebut, telah dilakukan pencarian data melalui situs resmi pemerintah yaitu Badan Pusat Statistik (BPS) di www.bps.go.id. Data yang

diperoleh merupakan data kunjungan wisatawan asing berdasarkan jalur masuk udara, darat, laut ke Indonesia perbulannya pada tahun 2017 sampai 2021.

Tabel 3.1 Sample Data Kunjungan Wisatawan Berdasarkan Jalur Masuk

Pintu Masuk	Januari	Februari	.	.	November	Desember
Ngurah Rai	451708	436266	.	.	492904	544726
Soekarno-Hatta	174963	196183	.	.	183759	186723
Juanda	13792	17389	.	.	20780	20546
Kualanamu	16253	21815	.	.	20798	22431
Husein Sastranegara	12524	15127	.	.	14616	14951

3.2.2 Pembersihan Data

Pada tahap pembersihan data, dilakukan secara manual dengan memisahkan antara data yang digunakan dan data yang tidak digunakan. Kemudian data yang tidak digunakan akan dibersihkan agar data menjadi efektif saat melakukan proses perhitungan. Berikut sample data sebelum dan sesudah dibersihkan:

Tabel 3.2 Sample Data Sebelum Dibersihkan

Kunjungan Wisatawan Asing perbulan ke Indonesia berdasarkan jalur masuk (2019)					
Pintu Masuk	Januari	.	.	Desember	Tahunan
A. Pintu Udara	706704	.	.	838978	-
Ngurah Rai	451708	.	.	544726	-
B. Pintu Laut	304479	.	.	388495	-
Batam	134415	.	.	190232	-
C. Pintu Darat	190552	.	.	149594	-
Jayapura	4418	.	.	6155	-
Jumlah (A+B+C)	1201735	.	.	1377067	-

Sebelum dibersihkan, terdapat beberapa variabel dan atribut yang tidak diperlukan seperti data kosong, tahunan, jumlah total, dan keterangan yang tidak diperlukan.

Tabel 4.1 Sample Data Setelah Dibersihkan

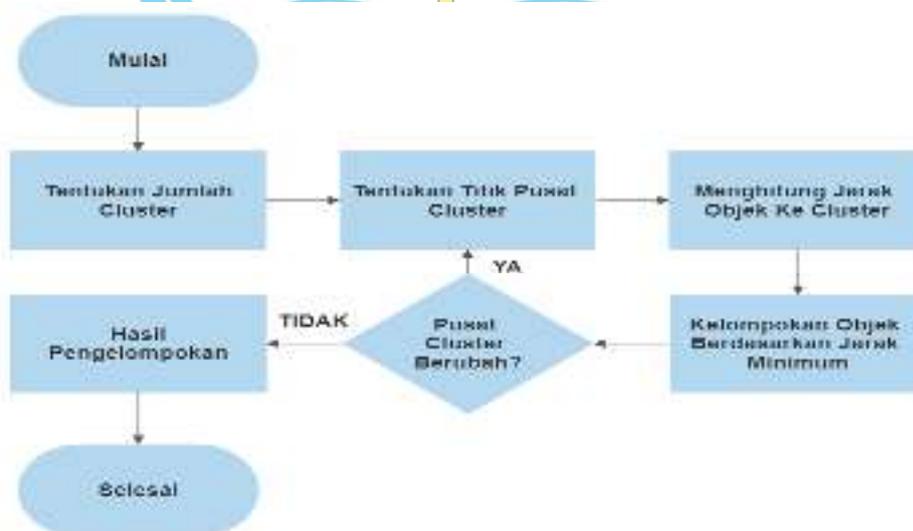
Pintu Masuk	Januari	Februari	.	.	November	Desember
Ngurah Rai	451708	436266	.	.	492904	544726
Pintu Udara Lainnya	1972	2742	.	.	4006	3593
Batam	134415	159248	.	.	167288	190232
Pintu Laut Lainnya	106115	104039	.	.	109792	98769
Jayapura	4418	3889	.	.	5006	6155
Pintu Darat Lainnya	174745	156591	.	.	126789	124408

Setelah dibersihkan, hanya terdapat atribut jalur masuk, kunjungan perbulan dan variabel data yang diperlukan.

3.2.3 Penerapan Algoritma K-Means

Berikut tahapan Algoritma K-Means dengan rumus euclidean distance:

1. Menentukan jumlah K cluster dan menentukan titik pusat cluster.
2. Menghitung jarak objek ke titik pusat cluster (*centroid*).
3. Mengelompokkan objek ke pusat cluster terdekat.
4. Hitung ulang semua nilai di Cluster sampai hasil iterasi tidak berubah.

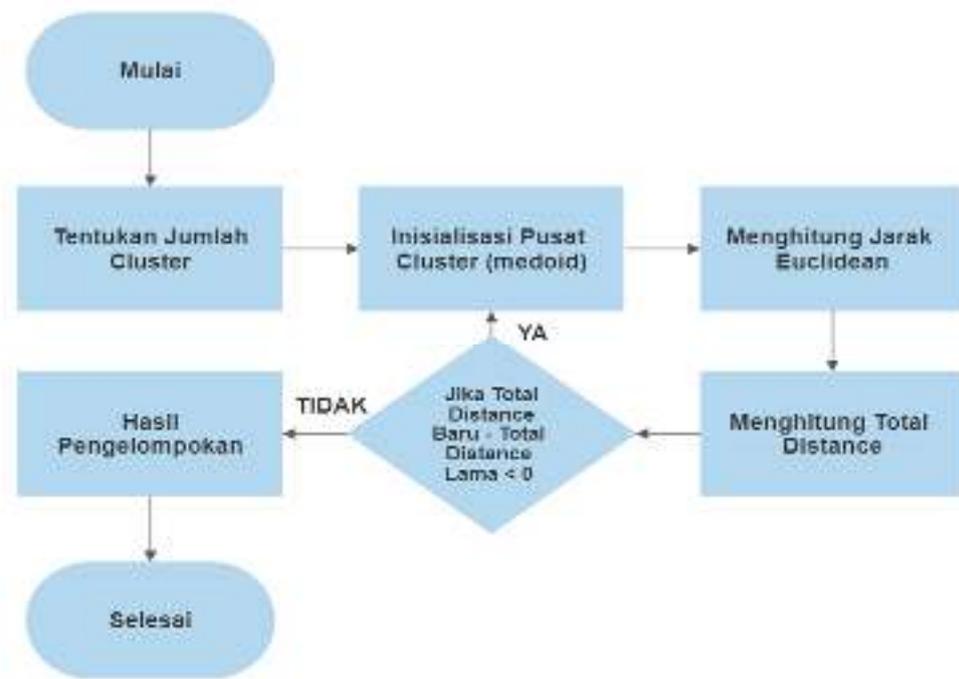


Gambar 3.2 Alur Algoritma K-Means

Sumber: (Rahman et al., 2017)

3.2.4 Penerapan Algoritma K-Medoid

Pada Tahap ini dataset akan diproses dengan menerapkan algoritma k-medoid menggunakan perhitungan manual dan bahasa pemrograman *python* menggunakan rumus *euclidean distance* dengan tahapan sebagai berikut:



Gambar 3.3 Alur Algoritma K-Medoid

Sumber: (Rahman et al., 2017)

3.2.5 Evaluasi Data

Pada tahapan evaluasi, dapat dihitung tingkat akurasi menggunakan *Silhouette Coefficient* dengan nilai dari -1 hingga 1, dimana jika hasil lebih mendekati nilai 1 maka metode pengelompokan lebih bagus dan jika hasil mendekati nilai -1 maka metode pengelompokan sangat buruk (Paembonan & Abduh, 2021).

3.2.6 Hasil

Setelah melakukan beberapa proses tahapan seperti pengumpulan data, pembersihan data, penerapan algoritma k-means, penerapan algoritma k-medoid, evaluasi data dan hasil. Maka akan keluar hasil dari pengelompokan data kunjungan wisatawan asing berdasarkan jalur masuk udara, darat, laut ke Indonesia dibagi menjadi 3 pengelompokan yaitu C1 tingkat kunjungan tinggi, C2 tingkat kunjungan sedang dan C3 tingkat kunjungan rendah.