

## BAB III METODE PENELITIAN

### 3.1 Objek Penelitian

Objek dalam penelitian Tugas Akhir ini adalah *dataset* yang berasal dari Badan Pusat Statistik (BPS), (BPS, 2021). Subjek dalam penelitian ini meliputi data dari bulan Maret 2020 sampai dengan Desember 2021 dimana sedang terjadinya peristiwa besar pandemi *Covid-19*, *dataset* untuk penelitian ini diambil pada jam 14:12 WIB 10 Mei Tahun 2022. Berikut jadwal penelitian yang direncanakan untuk mengerjakan Tugas Akhir (TA):

Tabel 3.1 Jadwal Penelitian

| No. | Kegiatan                             | Bulan       |              |              |                 |
|-----|--------------------------------------|-------------|--------------|--------------|-----------------|
|     |                                      | Mei<br>2022 | Juni<br>2022 | Juli<br>2022 | Agustus<br>2022 |
| 1   | Pengumpulan Data                     |             |              |              |                 |
| 2   | Seleksi Data                         |             |              |              |                 |
| 3   | Implementasi<br>Algoritma<br>K-Means |             |              |              |                 |
| 4   | Implementasi<br>Algoritma<br>DBSCAN  |             |              |              |                 |
| 5   | Evaluasi Data                        |             |              |              |                 |
| 6   | Hasil                                |             |              |              |                 |

#### 3.1.1 Peralatan Penelitian

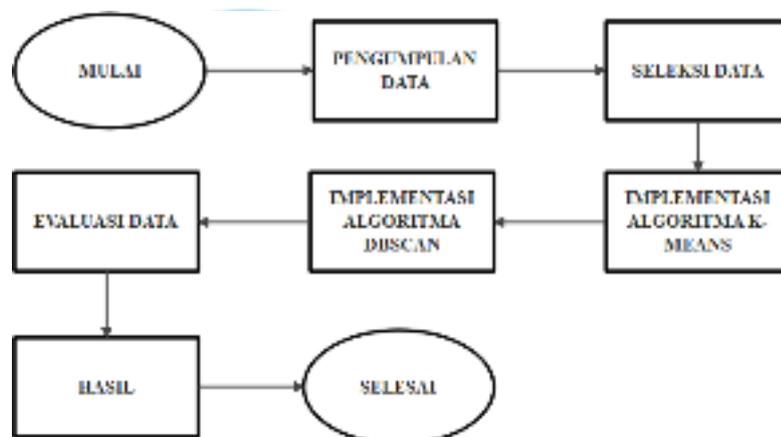
*Hardware* dan *software* yang digunakan untuk penelitian ini menggunakan laptop Asus model X505ZA dengan spesifikasi sebagai berikut:

- a) *Processor* : AMD Ryzen 5 2500U, 2.0 GHz
- b) *RAM* : 8GB Onboard Memory
- c) *HDD* : 1TB 5400RPM

- d) *Graphic : AMD Radeon Vega 8*
- e) *Screen : 15.6" (16:9) LED backlit HD (1366x768) 60Hz*
- f) *connectivity: Wifi 802.11 & Bluetooth*
- g) *Windows 10 Home 64bit*
- h) *Microsoft Word*
- i) *Microsoft Excel*
- j) *Google Chrome*
- k) *Python*

### 3.2 Prosedur Penelitian

Tahapan prosedur penelitian Tugas Akhir yang akan dilakukan dapat di gambar kan sebagai berikut:



Gambar 3.1 Alur Penelitian

#### 3.2.1 Teknik Pengumpulan Data

Pada penelitian ini membutuhkan dataset yang sumbernya valid, relevan, dan terupdate, maka untuk mendapatkan data tersebut diputuskan untuk mengambil dataset dari situs resmi pemerintah yaitu Badan Pusat Statistik (BPS). Dataset yang akan digunakan dalam penelitian ini adalah data kunjungan wisatawan mancanegara yang berkunjung ke Indonesia pada Maret 2020 sampai dengan Desember 2021, *dataset* untuk penelitian ini diambil pada jam 2:12 PM 10 Mei Tahun 2022.

#### 3.2.2 Seleksi Data

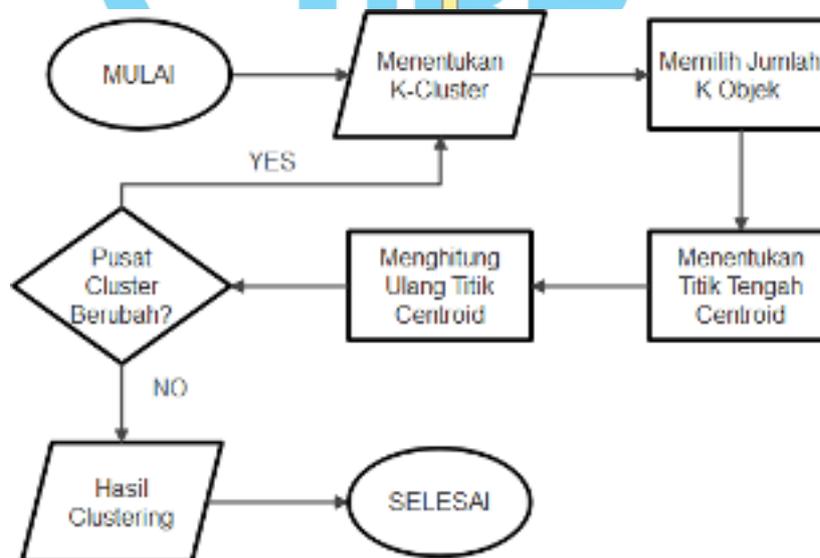
Pada tahap ini akan melakukan pemilihan data yang ada pada dataset karena tidak semuanya dipakai untuk penelitian yang akan dilakukan, hanya data yang sesuai yang akan dipilih untuk di analisa. Dalam penelitian ini hanya

menggunakan data Maret 2020 sampai dengan Desember 2021, karena menyesuaikan dengan tema yang diambil berkaitan dengan peristiwa terjadinya pandemi *Covid-19*, kemudian juga melakukan seleksi data dengan menggunakan *Microsoft Excel* dan *Python*.

### 3.2.3 Implementasi Algoritma K-Means

Pada tahap ini dataset akan diolah dengan menggunakan Algoritma K-Means. Dengan urutan sebagai berikut:

1. Menentukan K *cluster*
2. Memilih sejumlah K objek secara acak untuk dijadikan titik *centroid*
3. Menentukan K *centroid* (titik tengah)
4. Mengelompokkan objek ke *centroid cluster* yang paling dekat sesuai jarak objek ke *centroid* (*Euclidean distance*)
5. Menghitung kembali semua nilai titik *centroid*
6. Mengulang langkah 3 sampai 5 hingga titik *centroid* nilainya tidak berubah lagi.



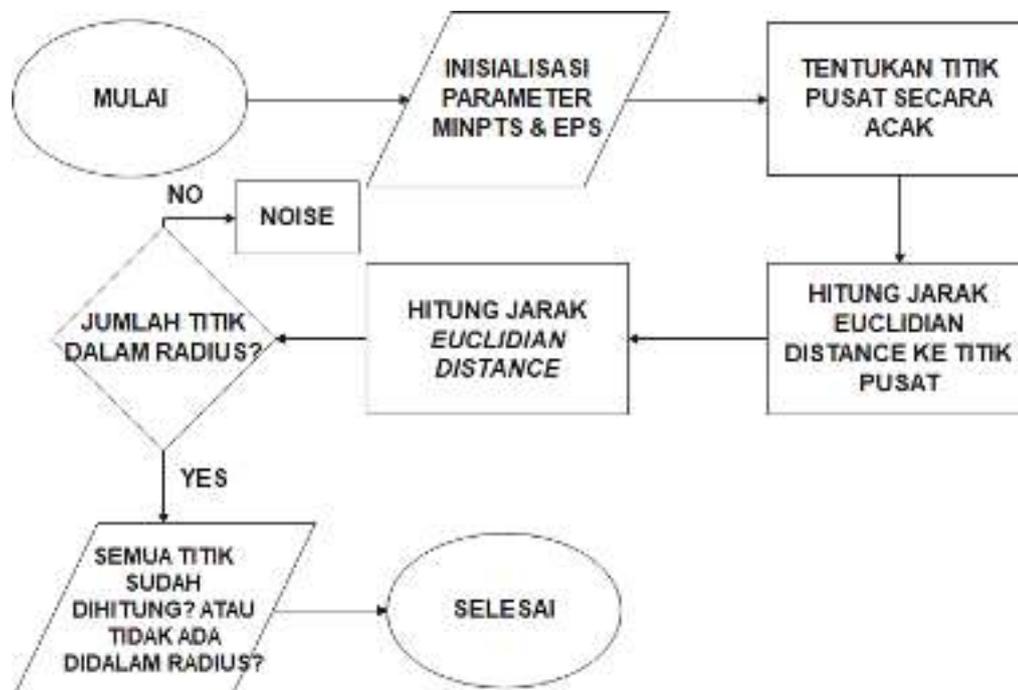
Gambar 3.2. Tahapan Algoritma K-Means

Pada penelitian ini mengimplimentasikan Algoritma K-Means dengan menggunakan 2 cara perhitungan yaitu perhitungan menggunakan *Microsoft Excel* dan juga menggunakan *Python*.

### 3.2.4 Implementasi Algoritma DBSCAN

Pada tahap ini dataset akan diolah dengan menggunakan Algoritma *DBSCAN*. Dengan urutan sebagai berikut:

1. Tentukan nilai *minPts* dan epsilon (*eps*) yang akan digunakan.
2. Pilih data awal “p” secara acak
3. Hitung jarak antara data “p” terhadap semua data menggunakan rumus *Euclidian distance*.
4. Ambil semua point yang density-reachable dengan “p”.
5. Jika memenuhi nilai *epsilon* lebih dari jumlah minimal point maka titik “p” dikategorikan sebagai *core point*.
6. Lanjutkan proses sampai selesai semua *point* terhitung atau tidak ada yang memenuhi syarat lagi titik pusat *point* untuk perhitungan selanjutnya.



Gambar 3.3. Tahapan Algoritma *DBSCAN*

### 3.2.5 Evaluasi Data

Untuk tahap ini pengelompokan data wisatawan asing yang berkunjung ke Indonesia akan dianalisis dan dievaluasi dengan menggunakan *Silhouette Coefficient* dengan bahasa pemrograman *python*.

Evaluasi dalam *clustering* dapat dihitung tingkat akurasi menggunakan *Silhouette Coefficient*. Metode evaluasi ini memiliki nilai dari -1 hingga 1 yang dimana jika hasil lebih mendekati nilai 1 maka

metode *Clustering* lebih bagus dan sebaliknya jika hasil mendekati nilai -1 maka metode *Clustering* sangat buruk (Paembunan, Abduh, 2021). Metode *Silhouette Coefficient* ini dapat di terapkan dengan persamaan 5 berikut:

$$sil(c) = \frac{sil(k) - 1}{(|k|)} \left[ \sum_{i=1}^k sil(C_i) \right]$$

Keterangan dari persamaan 5 di atas yaitu  $sil(k)$  adalah nilai *silhouette* semua *cluster*,  $|k|$  adalah banyaknya *cluster*, dan  $sil(c_i)$  adalah rata-rata nilai *silhouette*.

### 3.2.6 Hasil

Setelah melakukan pengumpulan data, seleksi data, dan evaluasi data maka akan diketahui hasil dari pengelompokan data kunjungan wisatawan tersebut dengan menggunakan Algoritma K-Means dan *DBSCAN*, akan terlihat nilai akurasi dari hasil evaluasi data dan bisa dibandingkan mana nilai akurasi yang lebih baik antara Algoritma K-Means dengan *DBSCAN*.

