

BAB III

METEDOLOGI PENELITIAN

3.1 Jenis dan Sumber Data

Data yang digunakan dalam penelitian ini adalah data yang diperoleh dari Badan Pusat Statistik (BPS) yang berjudul “Perkembangan Beberapa Indikator Utama Ekonomi dan Sosial-Ekonomi Indonesia” pada Tahun 2020 dan Tahun 2021. Variabel yang digunakan variabel produksi pertanian padi dalam ton, produk domestik regional bruto (PDRB) perkapita dalam ribu rupiah, rata-rata upah perbulan dalam rupiah, persentase penduduk yang bekerja di sektor formal dalam persen. Data pada penelitian ini menggunakan 34 Provinsi yang ada di Indonesia yang dapat dilihat pada Tabel 3.1.

Tabel 3. 1 Daftar Nama 34 Provini yang digunakan dalam Penelitian

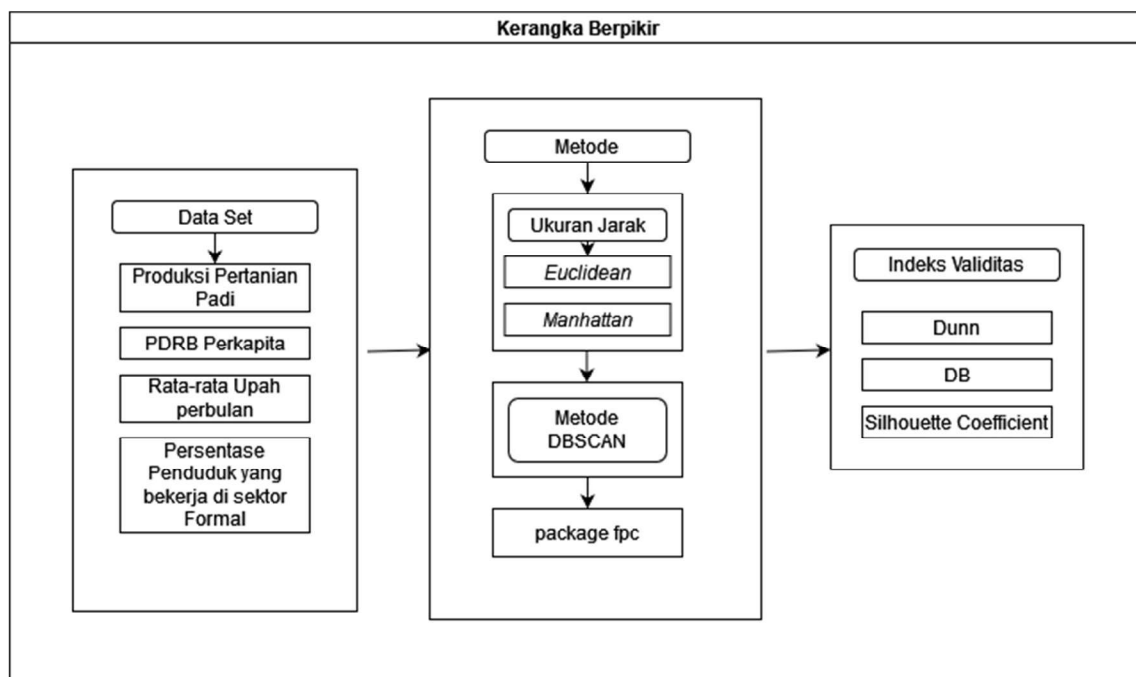
No	Provinsi	No	Provinsi
1	ACEH	21	KALIMANTAN TENGAH
2	SUMATERA UTARA	22	KALIMANTAN SELATAN
3	SUMATERA BARAT	23	KALIMANTAN TIMUR
4	RIAU	24	KALIMANTAN UTARA
5	JAMBI	25	SULAWESI UTARA
6	SUMATERA SELATAN	26	SULAWESI TENGAH
7	BENGKULU	27	SULAWESI SELATAN
8	LAMPUNG	28	SULAWESI TENGGARA
9	KEP. BANGKA BELITUNG	29	GORONTALO
10	KEP. RIAU	30	SULAWESI BARAT
11	DKI JAKARTA	31	MALUKU
12	JAWA BARAT	32	MALUKU UTARA
13	JAWA TENGAH	33	PAPUA BARAT
14	DI YOGYAKARTA	34	PAPUA
15	JAWA TIMUR		
16	BANTEN		
17	BALI		
18	NUSA TENGGARA BARAT		
19	NUSA TENGGARA TIMUR		
20	KALIMANTAN BARAT		

3.2 Pendekatan dan Variabel Penelitian

Pada penelitian ini variable yang digunakan meliputi nilai *epsilon* (ϵ) dan *minimum points* (*Minpts*) dan variabel produksi pertanian padi dalam ton, produk domestik regional bruto (PDRB) perkapita dalam ribu rupiah, rata-rata upah perbulan dalam rupiah, persentase penduduk yang bekerja di sektor formal dalam persen.

3.3 Kerangka Berpikir

Gambar di bawah ini menggambarkan kerangka berpikir pada penelitian ini

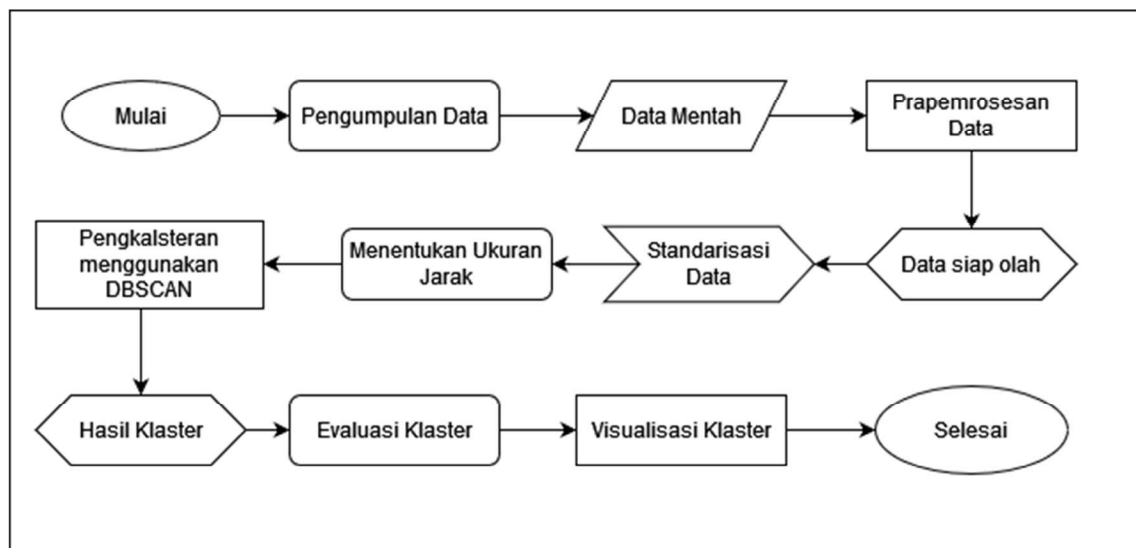


Gambar 3. 1 Kerangka Berpikir

Penelitian ini menggunakan data pada 34 provinsi di Indonesia Tahun 2020/2021. Selanjutnya, dilakukan *clustering* menggunakan metode DBSCAN yang terdapat pada *package* ‘fpc’ di *software* R. Dataset akan digunakan untuk menghitung jarak yang nantinya dipakai untuk *clustering*, adapun ukuran jarak yang digunakan adalah jarak *Euclidean* dan jarak *manhattan*. Klaster yang dihasilkan akan dievaluasi menggunakan Indeks Dunn dan Indeks DB dan Indeks *Silhouette Coefficient*.

3.4 Metode Analisis Data

Penelitian ini menggunakan metode DBSCAN untuk mengelompokkan provinsi di Indonesia periode 2020/2021 ke dalam kluster. Alur penelitian ini digambarkan pada gambar 3.2.



Gambar 3. 2 Alur Penelitian

3.5 Tahapan Penelitian

Penelitian ini akan menggunakan DBSCAN untuk mengelompokkan provinsi berdasarkan kondisi perekonomian provinsi tersebut. Adapun tahapan penelitian akan diuraikan berikut ini.

1. Penyusunan Deskripsi Data

Penelitian diawali dengan penyusunan setiap variabel yang digunakan pada penelitian ini digunakan untuk mengetahui gambaran suatu yang akan diteliti.

2. Standarisasi Data

Data perekonomian setiap provinsi di Indonesia ini perlu dilakukan standarisasi terlebih dahulu menggunakan rumus *z-score* pada persamaan 2.1, karena jarak antar datanya berbeda jauh.

3. Menghitung matriks jarak

Setelah data distandarisasi kemudian menghitung jarak antar provinsi berdasarkan variabel Produksi Pertanian Padi, Produk Domestik Regional Bruto (PDRB) Perkapita, Rata-Rata Upah Perbulan, Persentase Penduduk yang bekerja di sektor Formal. Perhitungan jarak menggunakan ukuran jarak *Euclidean* dan Jarak *Manhattan*. Jarak antar provinsi selanjutnya disajikan dalam bentuk matriks 2 dimensi berukuran 34×34 di mana elemen (a,b) (baris ke-a kolom ke-b) akan bernilai sama dengan elemen (b,a).

4. Menentukan parameter DBSCAN

Parameter DBSCAN terdiri atas epsilon dan minPts, penentuan parameter tersebut memanfaatkan matriks jarak yang dihasilkan sebelumnya. Matriks jarak digunakan untuk membuat grafik jarak ke-k. Grafik ini menggambarkan jarak suatu amatan (provinsi) dengan tetangga ke-k. Kemungkinan nilai epsilon dan minPts berdasarkan grafik tersebut. Nilai epsilon didapat berdasarkan besar ukuran jarak pada posisi menyiku di grafik jarak ke-k dan nilai minPts didapatkan dari nilai k.

5. Melakukan *clustering*

Tahapan selanjutnya adalah melakukan *clustering* terhadap 34 provinsi menggunakan berbagai kemungkinan parameter yang telah terpilih pada tahapan sebelumnya. *Clustering* dilakukan menggunakan metode DBSCAN pada *software* R. Metode tersebut terdapat pada *package* 'fpc'. Klaster yang dihasilkan selanjutnya dianalisis untuk melihat jumlah klaster yang dihasilkan, anggota kelompok dalam suatu klaster, dan profil dari setiap klaster.

6. Mengevaluasi hasil klaster

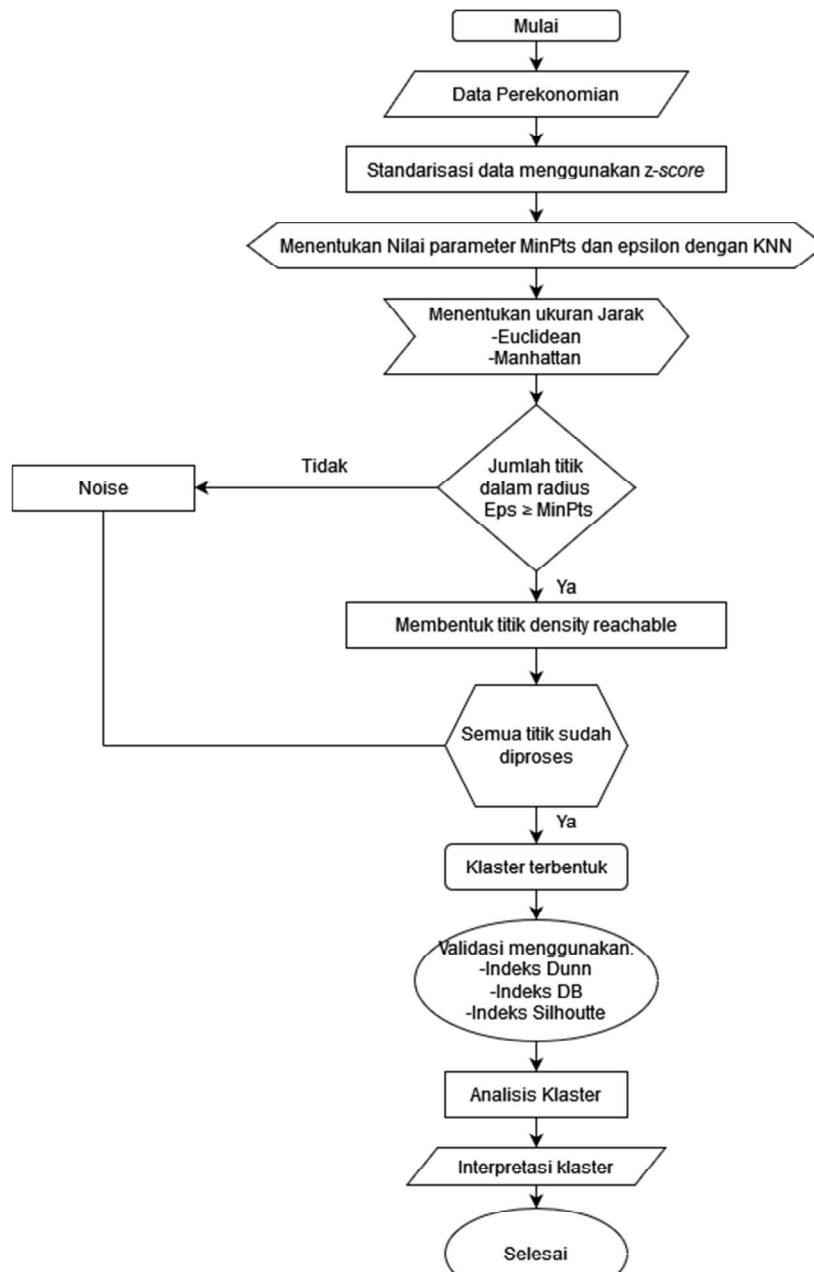
Klaster yang dihasilkan pada tahap sebelumnya dievaluasi menggunakan indeks validitas. Indeks validitas pada setiap klaster yang dihasilkan akan dibandingkan untuk menentukan klaster terbaik.

7. Memetakan kelompok Provinsi berdasarkan profil perekonomiannya.

Klaster terbaik akan dipetakan agar lebih mudah untuk dipahami oleh pembaca. Selain itu, setiap kelompok provinsi akan dilihat agregat kondisi

perekonomiannya meliputi variabel produksi pertanian padi, produk domestik regional bruto (PDRB) perkapita, rata-rata upah perbulan, persentase Penduduk yang bekerja di sektor formal. Kelompok provinsi dengan kondisi variabel produksi pertanian padi, produk domestik regional bruto (PDRB) perkapita, Rata-rata upah perbulan, persentase penduduk yang bekerja di sektor formal yang rendah akan menjadi prioritas pemerintah dalam rangka pemulihan kondisi perekonomian paska pandemi.

Berikut flowchart metode DBSCAN:



Gambar 3. 3 Flowchart DBSCAN