

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Sumber utama data kependudukan adalah sensus penduduk yang dilaksanakan setiap sepuluh tahun sekali. Sensus penduduk telah dilaksanakan sebanyak tujuh kali sejak Indonesia merdeka, yaitu tahun 1961, 1971, 1980, 1990, 2000, 2010, dan 2020. Di dalam sensus penduduk, pencacahan dilakukan terhadap seluruh penduduk yang berdomsili di wilayah teritorial Indonesia termasuk warga negara asing kecuali anggota korps diplomatik negara sahabat beserta keluarganya. Metode pengumpulan data dalam sensus dilakukan dengan wawancara antara petugas sensus dengan responden dan juga melalui e-census. Pencatatan penduduk menggunakan konsep usual residence, yaitu konsep di mana penduduk biasa bertempat tinggal.

Bagi penduduk yang bertempat tinggal tetap dicacah di mana mereka biasa tinggal, sedangkan untuk penduduk yang tidak bertempat tinggal tetap dicacah di tempat di mana mereka ditemukan petugas sensus pada malam 'Hari Sensus'. Termasuk penduduk yang tidak bertempat tinggal tetap adalah tuna wisma, awak kapal berbendera Indonesia, penghuni perahu/rumah apung, masyarakat terpencil/terasing, dan pengungsi. Bagi mereka yang mempunyai tempat tinggal tetap dan sedang berpergian ke luar wilayah lebih dari enam bulan, tidak dicacah di tempat tinggalnya, tetapi dicacah di tempat tujuannya.

Untuk tahun yang tidak dilaksanakan sensus penduduk, data kependudukan diperoleh dari hasil proyeksi penduduk. Proyeksi penduduk merupakan suatu perhitungan ilmiah yang didasarkan pada asumsi dari komponen-komponen perubahan penduduk, yaitu kelahiran, kematian, dan migrasi. Proyeksi penduduk Indonesia 2010-2035 menggunakan data dasar penduduk hasil SP2020.

Jumlah penduduk Provinsi Jawa Timur pada tahun 2020 berdasarkan hasil sensus penduduk berjumlah sekitar 40,666 juta jiwa. Kepadatan penduduk per km<sup>2</sup> sebesar 851 jiwa/km<sup>2</sup>. Sedangkan rasio jenis kelamin adalah 99,60. Dari total penduduk Provinsi Jawa Timur pada tahun 2020, kelompok umur yang berjumlah paling banyak adalah kelompok umur 30-34 tahun dengan jumlah sekitar 3,150 jiwa.

Jumlah penduduk tahun 2020 yang paling banyak terdapat di Kota Surabaya dengan jumlah penduduk sekitar 2,874 juta jiwa. Sedangkan jumlah penduduk paling sedikit ada di Kota Mojokerto dengan jumlah penduduk sekitar 132 ribu jiwa. Laju pertumbuhan penduduk per tahun antara tahun 2010-2020 yang paling tinggi ada di Kabupaten Bangkalan dengan laju pertumbuhan 1,53 persen. Sedangkan laju pertumbuhan terendah ada di Kota Malang dengan laju pertumbuhan sebesar 0,27 persen.

Jumlah penduduk usia 15 tahun keatas yang menganggur di Provinsi Jawa Timur pada tahun 2020 adalah sebesar 1,301,145 orang, dari total Angkatan kerja sejumlah 22,264 juta orang. Jumlah penganggur sebagian besar berada di Kabupaten/Kota besar yang ada di Provinsi Jawa Timur seperti Kota Surabaya dengan jumlah pengangguran sebesar 154,896 orang, diikuti Kabupaten Sidoarjo sebesar 134,444 orang, Kabupaten Gresik sebesar 56,261 orang, Kabupaten Pasuruan sebesar 54,181 orang, Kabupaten Banyuwangi sebesar 49,252 orang, Kabupaten Kediri sebesar 46,061 orang dan Kota Malang sebesar 45,242 orang.

Berdasarkan pendidikannya, jumlah penduduk di atas 15 tahun yang menganggur paling tinggi adalah lulusan Sekolah Menengah Atas (SMA) yaitu 663.486 orang. Jumlah penduduk diatas 15 tahun yang bekerja paling banyak mempunyai status Buruh/Karyawan/Pegawai dengan jumlah 6,877 juta orang, sedangkan yang paling sedikit adalah berusaha dibantu buruh tetap/buruh dibayar dengan jumlah 744.415 orang.

Angkatan kerja adalah penduduk usia kerja (15 tahun ke atas) yang bekerja, punya pekerjaan. Bekerja adalah kegiatan ekonomi yang dilakukan oleh seseorang dengan maksud memperoleh atau membantu memperoleh pendapatan atau keuntungan dan lamanya bekerja paling sedikit 1 jam secara terus menerus dalam seminggu yang lalu (termasuk pekerja keluarga tanpa upah yang membantu dalam suatu usaha/kegiatan ekonomi).

Di Kabupaten/Kota besar yang ada di Provinsi Jawa Timur terlihat bahwa rata-rata pendapatan bersih sebulan pekerja informal lebih besar dibandingkan di Kabupaten/Kota lainnya. Rata-rata pendapatan bersih pekerja informal per bulan di Kota Surabaya, Kabupaten Sidoarjo, dan Kabupaten Gresik berkisar di atas 2 Juta rupiah per bulan. Bahkan di Kabupaten Sidoarjo merupakan yang paling tinggi dengan 2,281 juta rupiah per bulan (Provinsi Jawa Timur Dalam Angka, 2021).

Variabel yang digunakan dalam penelitian ini adalah Penduduk Berumur 15 Tahun ke Atas Menurut Kabupaten/Kota dan Jenis Kegiatan Selama Seminggu yang Lalu di Provinsi Jawa Timur 2020 dengan kategori 1 yaitu mencari pekerjaan/*looking for work*, kategori 2 yaitu mempersiapkan usaha/*establishing a new business/firm*, kategori 3 yaitu merasa tidak mungkin mendapat pekerjaan/*hopeless of job*, kategori 4 yaitu sudah punya pekerjaan tetapi belum mulai bekerja/*have a job in future start*.

Rata-rata Upah/Gaji Bersih Sebulan Pekerja Formal Menurut Kabupaten/Kota dan Lapangan Pekerjaan Utama (rupiah) 2020 dengan kategori 2 yaitu Pertambangan dan Penggalian; Industri Pengolahan; Pengadaan Listrik dan Gas; Pengadaan Air; Pengelolaan Sampah, Limbah, dan Daur Ulang; Konstruksi. Kategori 3 yaitu Perdagangan Besar dan Eceran; Reparasi Mobil dan Sepeda Motor; Transportasi dan Pergudangan; Penyediaan Akomodasi dan Makan Minum; Informasi dan Komunikasi; Jasa Keuangan dan Asuransi; Real Estat; Jasa Perusahaan; Administrasi Pemerintahan, Pertahanan, dan Jaminan Sosial Wajib; Jasa Pendidikan; Jasa Kesehatan dan Kegiatan Sosial; Jasa Lainnya.

Rata-rata Pendapatan Bersih Sebulan Pekerja Informal Menurut Kabupaten/Kota dan Lapangan Pekerjaan Utama (rupiah) 2020 dengan kategori 1 yaitu Pertanian, Kehutanan, Perikanan/*Agriculture, Forestry, and Fishing*. Kategori 2 yaitu Pertambangan dan Penggalian; Industri Pengolahan; Pengadaan Listrik dan Gas; Pengadaan Air; Pengelolaan Sampah, Limbah, dan Daur Ulang; Konstruksi. Kategori 3 yaitu Perdagangan Besar dan Eceran; Reparasi Mobil dan Sepeda Motor; Transportasi dan Pergudangan; Penyediaan Akomodasi dan Makan Minum; Informasi dan Komunikasi; Jasa Keuangan dan Asuransi; Real Estat; Jasa Perusahaan; Administrasi Pemerintahan, Pertahanan, dan Jaminan Sosial Wajib; Jasa Pendidikan; Jasa Kesehatan dan Kegiatan Sosial; Jasa Lainnya.

Rata-rata Pendapatan Bersih Sebulan Pekerja Informal Menurut Kabupaten/Kota dan Pendidikan Tertinggi yang Damatkan (rupiah) 2020 dengan kategori Tidak Pernah Sekolah/Belum Tamat SD, SD/MI, SMP/MTs, SMA ke Atas. Pekerja Informal/*Informal Employee*: Pekerja yang berstatus berusaha sendiri dan pekerja bebas di sektor pertanian dan nonpertanian/*Employment status are self employed, casual agricultural worker, and casual non-agricultural worker*.

Penelitian-penelitian yang berkaitan dengan jumlah penduduk yang bekerja dan rata-rata pendapatan sebelumnya sudah pernah dilakukan. Penelitian oleh Profil Perempuan Indonesia 2021 yaitu persentase perempuan berumur 15 tahun ke atas yang bekerja seminggu yang lalu sebagian besar berstatus kawin. Secara total, persentase pekerja perempuan yang berstatus kawin sebesar 69,52 persen, sedangkan yang berstatus belum kawin sebesar 16,34 persen, cerai hidup sebesar 4,01 persen, dan cerai mati sebesar 10,13 persen.

Rata-rata upah/gaji bersih pekerja di perkotaan lebih tinggi dibandingkan di pedesaan. Rata-rata upah/gaji pekerja pada tahun 2020 sebesar Rp. 2.756.345,- sedangkan di pedesaan sebesar Rp. 2.131.717,- lebih rendah dibandingkan pedesaan sebesar Rp. 3.026.853,-. Pola ini juga mengikuti bila dilihat menurut jenis kelamin, dimana baik laki-laki dan perempuan rata-rata upah/gaji bersih pekerjaanya lebih tinggi di perkotaan dibandingkan di pedesaan.

Penelitian oleh Deasy Dwi Ramiayu yaitu Rata-rata lama sekolah berpengaruh positif dan signifikan terhadap tingkat pengangguran terbuka. Di Jawa Timur, meningkatnya pendidikan masyarakat tidak menjamin tenaga kerja untuk memperoleh pekerjaan. Data menunjukkan tingkat pengangguran terbuka didominasi oleh penduduk lulusan SMA, SMK, diploma, dan perguruan tinggi. Meningkatnya pendidikan masyarakat menyebabkan mereka lebih selektif dalam memilih pekerjaan. Masyarakat dengan pendidikan tinggi cenderung memilih menjadi pengangguran daripada bekerja yang kualifikasinya tidak sesuai dengan pendidikan mereka. Sehingga walaupun rata-rata lama sekolah meningkat setiap tahun, tidak bisa mengurangi tingkat pengangguran.

Upah minimum berpengaruh negatif dan tidak signifikan terhadap tingkat pengangguran terbuka. Upah minimum tidak berpengaruh secara nyata terhadap pengangguran. Dikarenakan upah yang bersifat kaku, upah minimum yang diberlakukan tidak langsung berpengaruh pada tahun tersebut, tetapi berpengaruh pada jangka panjang. Sehingga ketika upah minimum diterapkan, upah minimum tidak langsung mendorong perusahaan untuk mengurangi tenaga kerja. Akan tetapi, dalam jangka panjang upah minimum akan mendorong perusahaan untuk melakukan PHK dan meningkatkan pengangguran.

Penelitian ini mengenai perbandingan antara dua metode *Clustering* yaitu *K-Means* dan *K-Medoids*. *K-Means* adalah salah satu prosedur non hierarki yang populer. *K-Means* digunakan untuk

mendeskripsikan suatu algoritma yang membagi setiap item ke suatu *cluster* tertentu yang memiliki pusat *cluster optimum* (*mean*) terdekat (Johnson & Wichern, 2007). *K-Medoids* atau dikenal juga dengan algoritma PAM (*Partitioning Around Medoids*) adalah algoritma pengelompokan yang berkaitan dengan algoritma *K-Means* dan algoritma *Medoids*. Algoritma ini mirip dengan *K-Means* karena kedua algoritma ini *partitional* yang memecah *dataset* menjadi kelompok-kelompok. Perbedaan dari algoritma *K-Means* dengan algoritma *K-Medoids* terletak pada penentuan pusat *cluster*. Algoritma *K-Means* menggunakan nilai *mean* atau rata-rata dari setiap *cluster* sebagai pusat *cluster*, sedangkan algoritma *K-Medoids* menggunakan objek representative (*medoids*) sebagai pusat *cluster* (Kaur, Kaur & Singh, 2014). Pusat *cluster K-Medoids* ditentukan secara random hingga jarak Euclidean lebih besar dari *k-cluster* sebelumnya, maka *k-cluster* yang memiliki jarak Euclidean terbesar adalah kelompok optimum.

Penelitian tentang *K-Means* dan *K-Medoids* sebelumnya pernah dilakukan oleh Kaur, dkk (2014), dengan berkesimpulan bahwa metode *K-Medoids* memiliki *performance* yang lebih baik daripada metode *K-Means* karena metode *K-Medoids* tidak sensitive terhadap *outlier*, lebih *robust*, dan waktu eksekusi per *Cluster* lebih sedikit. Penelitian lainnya dilakukan oleh Sangga (2018) yaitu mengenai perbandingan metode *K-Means* dan *K-Medoids* dalam pengelompokan kabupaten/kota berdasarkan variabel jenis ternak di Provinsi Jawa Tengah. Hasil yang didapatkan yaitu metode *K-Medoids* terindikasi lebih baik karena memiliki nilai variansi *Cluster* yang lebih kecil daripada metode *K-Means*. Ogbuabor dan Ugwoke (2018) meneliti *healthcare dataset* menggunakan *K-Means* dan DBSCAN kemudian membandingkan hasil keduanya menggunakan nilai skor *Silhouette*, dengan kesimpulan bahwa *K-Means* memiliki *performance* lebih baik karena memiliki skor *Silhouette* lebih tinggi dan waktu eksekusi lebih sebentar daripada DBSCAN.

Penelitian ini dilakukan bertujuan untuk mengelompokkan kabupaten/kota di Jawa Timur menjadi sejumlah *cluster*. Data yang digunakan dalam penelitian adalah data Jumlah Penduduk Berumur 15 Tahun ke Atas yang Bekerja dan Rata-Rata Pendapatan di Jawa Timur tahun 2020 dengan variabel berjumlah 12. Analisis statistika deskriptif menggunakan diagram profil dan peta tematik. Metode *clustering* yang digunakan adalah *K-Means* dan *K-Medoids clustering*. Penentuan jumlah *cluster optimum* dan evaluasi hasil *clustering* menggunakan *Silhouette*

*Coefficient (SC)*. Analisis lanjutan terhadap pengelompokan terbaik dilakukan secara deskriptif baik numerik (uji *Kolmogorov Smirnov*, *mean*, *median*) maupun visualisasi dalam bentuk peta tematik.

## 1.2 Rumusan Masalah

Rumusan masalah yang dirumuskan berdasarkan latar belakang tersebut adalah :

1. Bagaimana deskriptif Jumlah Penduduk Berumur 15 Tahun ke Atas yang Bekerja dan Rata-Rata Pendapatan di Jawa Timur tahun 2020 ?
2. Bagaimana hasil pengelompokan Jumlah Penduduk Berumur 15 Tahun ke Atas yang Bekerja dan Rata-Rata Pendapatan di Jawa Timur tahun 2020 berdasarkan hasil *Clustering* terbaik antara metode *K Means* dan *K-Medoids Clustering* ?

## 1.3 Tujuan Penelitian

Tujuan yang ingin dicapai berdasarkan rumusan masalah yang telah dipaparkan pada poin 1.2 antara lain adalah sebagai berikut :

1. Memperoleh deskriptif Jumlah Penduduk Berumur 15 Tahun ke Atas yang Bekerja dan Rata-Rata Pendapatan di Jawa Timur tahun 2020.
2. Memperoleh pengelompokan Jumlah Penduduk Berumur 15 Tahun ke Atas yang Bekerja dan Rata-Rata Pendapatan di Jawa Timur tahun 2020 berdasarkan hasil *Clustering* terbaik antara metode *K Means* dan *K-Medoids Clustering*.

## 1.4 Manfaat Penelitian

Manfaat yang ingin diberikan dari hasil penelitian yang dilakukan antara lain adalah sebagai berikut :

1. Bagi Lembaga Terkait

Manfaat yang ingin diberikan bagi lembaga terkait adalah Pendapatan merupakan indikator utama dalam Laporan Pemerintahan. Hal ini menunjukkan betapa pentingnya data Jumlah Penduduk dan Rata-Rata Pendapatan dalam menyusun laporan sebagai bahan evaluasi kebijakan pemerintah sekaligus menjadi acuan dalam perencanaan pembangunan jangka menengah dan jangka panjang.

Bagi Indonesia, Pendapatan merupakan data strategis karena sebagai ukuran kinerja Pemerintah dan Masyarakat. Pendapatan dapat menentukan peringkat atau level pembangunan suatu wilayah atau negara. Peringkat tersebut bukan sekedar angka semata, namun benar-benar menggambarkan capaian suatu wilayah atau negara di bidang ekonomi.

## **2. Bagi Peneliti**

Manfaat yang diperoleh bagi peneliti adalah agar dapat memahami kegunaan dan mengimplementasikan ilmu statistika khususnya metode *K-Means* dan *K-Medoids Clustering* terhadap permasalahan di kehidupan masyarakat yang sesungguhnya, yaitu bidang Jumlah Penduduk Berumur 15 Tahun ke Atas yang Bekerja dan Rata-Rata Pendapatan di Jawa Timur.

### **1.5 Batasan Masalah**

Batasan masalah dalam penelitian ini adalah data Jumlah Penduduk Berumur 15 Tahun ke Atas yang Bekerja dan Rata-Rata Pendapatan di Jawa Timur tahun 2020 yang digunakan yaitu variabel Penduduk Berumur 15 Tahun ke Atas yang Bekerja Selama Seminggu yang Lalu Menurut Kabupaten/Kota dan Lapangan Pekerjaan Utama yaitu kategori 1, 2, 3; Rata-rata Upah/Gaji Bersih Sebulan Pekerja Formal Menurut Kabupaten/Kota dan Lapangan Pekerjaan Utama (rupiah) yaitu kategori 2, 3; Rata-rata Pendapatan Bersih Sebulan Pekerja Informal Menurut Kabupaten/Kota dan Lapangan Pekerjaan Utama (rupiah) yaitu kategori 1, 2, 3; Rata-rata Pendapatan Bersih Sebulan Pekerja Informal Menurut Kabupaten/Kota dan Pendidikan Tertinggi yang Ditamatkan yaitu kategori Belum Tamat SD, Tamat SD/MI, Tamat SMP/MTs, Tamat SMA Ke Atas sebanyak 12 variabel. Unit penelitian adalah Kabupaten/Kota di Jawa Timur.