

# BAB III

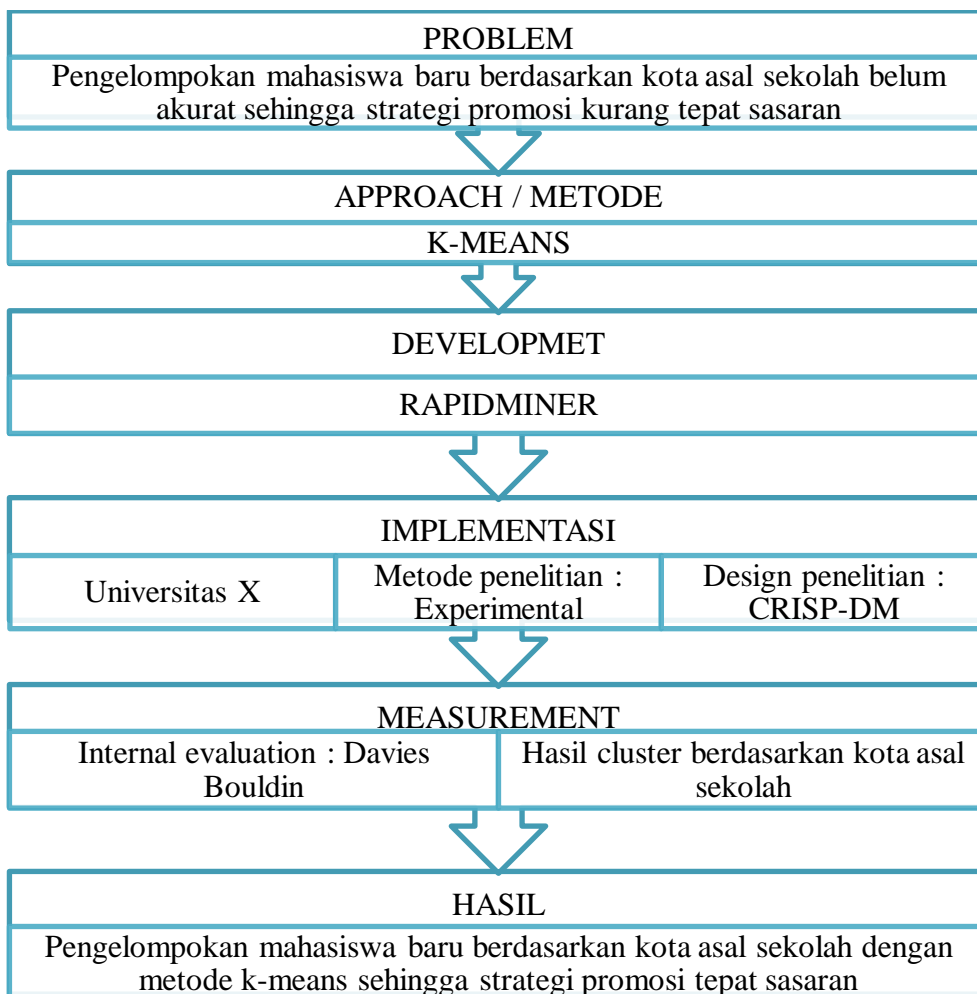
## METODOLOGI PENELITIAN

### 3.1 Objek Penelitian

Penelitian ini mengambil objek yaitu pada data penerimaan mahasiswa baru (PMB) angkatan 2016 sampai dengan angkatan 2019 di Universitas X. Metode yang digunakan dalam penulisan tugas akhir ini adalah metode data mining K – Means Clustering.

### 3.2 Kerangka Pemikiran

Penulis membuat kerangka penelitian untuk mengidentifikasi permasalahan



*Gambar 3.1 Kerangka Pemikiran*

### **3.3 Analisa Sistem Berjalan**

Pada penulisan skripsi ini menggunakan metode CRISP – DM yang terdiri dari *business understanding*, *data understanding*, *data preparation*, *modeling*, dan *evaluation*. Pada sub bab ini akan diuraikan tentang *business understanding* dan *data understanding*.

#### **3.3.1 Business Understanding**

Permasalahan proses penerimaan mahasiswa baru menghasilkan data yang berlimpah melingkupi data - data mahasiswa tersebut, hal ini akan terus berlangsung setiap tahun sehingga data yang tersimpan akan semakin banyak dan jika data yang ada tidak dimanfaatkan dengan baik sebagai hal positif bagi universitas. Sama halnya yang terjadi di lingkungan Universitas X setiap tahun menerima hampir 2.000 lebih mahasiswa baru dari berbagai kota.

Data yang digunakan merupakan data mahasiswa selama 4 tahun yaitu 2016, 2017, 2018, dan 2019. Data tersebut akan diolah dengan metode data mining k – means sehingga nantinya dapat digunakan untuk strategi promosi sebaran wilayah berdasarkan kota asal sekolah.

#### **3.3.2 Data Understanding**

File yang didapatkan berupa data mentah yang belum diolah. file yang didapatkan berjumlah satu file yang terdapat empat sheet excel terdiri dari sheet mahasiswa angkatan 2016, sheet mahasiswa angkatan 2017, sheet mahasiswa angkatan 2018, dan sheet mahasiswa angkatan 2019. Pada atribut berikut yang memiliki bercabangan yaitu atribut keterangan yang memiliki bercabangan (umum, keluarga polisi, polisi/polri), atribut jenis kelamin yang memiliki bercabangan (P dan L), atribut kelas yang memiliki bercabangan (pagi, sore, p2k), atribut jenis pendaftar yang memiliki bercabangan (mahasiswa baru, konversi, readmisi), atribut gelombang yang memiliki bercabangan (1,2,3,4,5), dan atribut status yang memiliki bercabangan (daftar ulang, lulus tes, registrasi).

- A. Pada tahun 2016 berjumlah 2340 data yang terdapat 29 atribut. Atributnya terdiri dari No, ID Registrasi, Nama, Keterangan, NIK, Tanggal Lahir, Jenis Kelamin, Jenis Pendaftaran, Kelas, Telepon, Telepon wali, Alamat, Pekerjaan, Nama Ibu, Agama, Kota Asal Sekolah, Pendidikan Ayah, Pendidikan Ibu, Pekerjaan Ayah, Pekerjaan Ibu, Asal Sekolah, Kategori Sekolah, Jenis Sekolah, Penghasilan, Prodi, Kampus, Gelombang, Status, NPM, dan Tahun Lulus.

*Table 3.1 Table Data 2016*

| No   | ID Registrasi | Nama                  | Keterangan | NIK              | Tanggal Lahir | Jenis Kelamin | Jenis Pendaftaran | Kelas | :   | Tahun Lulus |
|------|---------------|-----------------------|------------|------------------|---------------|---------------|-------------------|-------|-----|-------------|
| 1    | 16.S1.1.0001  | Daryus                | Umum       | 3175060712750000 | 07/12/1975    | L             | Mahasiswa Baru    | Sore  | :   | 1994        |
| 2    | 16.S1.1.0005  | Muhammad Iqbal        | Umum       | 3275032507980020 | 25/07/1998    | L             | Mahasiswa Baru    | Pagi  | :   | 2013        |
| 3    | 16.S1.1.0003  | Ida Faridah           | Umum       | 3216024804940000 | 08/04/1994    | P             | Mahasiswa Baru    | Sore  | :   | 2016        |
| 4    | 16.S1.1.0004  | Dian Fahruly Nur      | Umum       | 3216020303980000 | 03/03/1998    | L             | Mahasiswa Baru    | Pagi  | :   | 2016        |
| ...  | ...           | ...                   | ...        | ...              | ...           | ...           | ...               | ...   | ... | ...         |
| 2340 | 16.S1.5.2299  | Adityo Harits Rinanto | Umum       | 3275122705960000 | 27/05/1996    | L             | Mahasiswa Baru    | Pagi  | :   | 2015        |

- B. Pada tahun 2017 berjumlah 2347 data yang terdapat 29 atribut. Atributnya terdiri dari No, ID Registrasi, Nama, Keterangan, NIK, Tanggal Lahir, Jenis Kelamin, Jenis Pendaftaran, Kelas, Telepon, Telepon wali, Alamat, Pekerjaan, Nama Ibu, Agama, Kota Asal Sekolah, Pendidikan Ayah, Pendidikan Ibu, Pekerjaan Ayah, Pekerjaan Ibu, Asal Sekolah, Kategori Sekolah, Jenis Sekolah, Penghasilan, Prodi, Kampus, Gelombang, Status, NPM, dan Tahun Lulus.

*Table 3.2 Table Data 2017*

| No   | ID Registrasi | Nama                      | Keterangan         | NIK              | Tanggal Lahir | Jenis Kelamin | Jenis Pendaftaran | Kelas | :   | Tahun Lulus |
|------|---------------|---------------------------|--------------------|------------------|---------------|---------------|-------------------|-------|-----|-------------|
| 1    | 17.S1.1.0001  | Nurjanah                  | Umum               | 3275056907940000 | 29/07/1994    | P             | Mahasiswa Baru    | Pagi  | :   | 2012        |
| 2    | 17.S1.1.0002  | Husnul Khotimah           | Umum               | 3275094210970010 | 02/10/1997    | P             | Mahasiswa Baru    | Sore  | :   | 2015        |
| 3    | 17.S1.1.0003  | Sinta Purnamasari         | Umum               | 3275026207970010 | 22/07/1997    | P             | Mahasiswa Baru    | P2k   | :   | 2015        |
| 4    | 17.S1.1.0004  | Nancy Gabrela Noviani     | Umum               | 1256194811970000 | 08/11/1997    | P             | Mahasiswa Baru    | Pagi  | :   | 2015        |
| ...  | ...           | ...                       | ...                | ...              | ...           | ...           | ...               | ...   | ... | ...         |
| 2347 | 17.S1.5.2313  | Febri Prasetyo Ari Wibowo | POLISI / PNS POLRI | 3174040302970000 | 03/02/1997    | L             | Mahasiswa Baru    | Pagi  | :   | 2014        |

- C. Pada tahun 2018 berjumlah 2688 data yang terdapat 30 atribut. Atributnya terdiri dari No, ID Registrasi, Nama, Keterangan, NIK, Tanggal Lahir, Jenis Kelamin, Jenis Pendaftaran, Kelas, Telepon, Telepon wali, Alamat, Pekerjaan, Nama Ibu, Agama, Kota Asal Sekolah, Pendidikan Ayah, Pendidikan Ibu, Pekerjaan Ayah, Pekerjaan Ibu, Asal Sekolah, Kategori Sekolah, Jenis Sekolah, Penghasilan, Prodi, Kampus, Gelombang, Status, NPM, Tahun Lulus, dan Email.

*Table 3.3 Table Data 2018*

| No | ID Registrasi | Nama                       | Keterangan         | NIK              | Tanggal Lahir | Jenis Kelamin | Jenis Pendaftaran | Kelas | : | Email   |
|----|---------------|----------------------------|--------------------|------------------|---------------|---------------|-------------------|-------|---|---|
| 1  | 18.S1.1.0001  | Wachyudi Tresna Septiannie | POLISI / PNS POLRI | 3212081609950000 | 16/09/1995    | L             | Mahasiswa Baru    | P2k   | : | wachyuditr<br>esnaseptiya<br>nnie@gmai<br>l.com |
| 2  | 18.S1.1.0002  | Wahyu Putri Fitrianiingsih | Umum               | 3216055801890000 | 18/01/1999    | P             | Mahasiswa Baru    | Pagi  | : | wahyuputri<br>f36@gmail<br>.com                 |
| 3  | 18.S1.1.0003  | Kevin Nathaniel            | Umum               | 3275032009970030 | 20/09/1997    | L             | Mahasiswa Baru    | Pagi  | : | kevinthnl<br>@gmail.co<br>m                     |

| No   | ID Registrasi | Nama                 | Keterangan | NIK              | Tanggal Lahir | Jenis Kelamin | Jenis Pendaftaran | Kelas | :   | Email                        |
|------|---------------|----------------------|------------|------------------|---------------|---------------|-------------------|-------|-----|------------------------------|
| 4    | 18.S1.1.0004  | Ahmad Rizky Ramadhan | Umum       | 3174102302940000 | 23/02/1994    | L             | Mahasiswa Baru    | Pagi  | :   | ahmadrizky023@gmail.com      |
| ...  | ...           | ...                  | ...        | ...              | ...           | ...           | ...               | ...   | ... | ...                          |
| 2688 | 18.S1.2.2661  | Novia Permata Sari   | Umum       | 3216064811980010 | 08/11/1998    | P             | Mahasiswa Baru    | P2k   | :   | noviapermatasari69@gmail.com |

- D. Pada tahun 2019 berjumlah 3237 data yang terdapat 30 atribut. Atributnya terdiri dari No, ID Registrasi, Nama, Keterangan, NIK, Tanggal Lahir, Jenis Kelamin, Jenis Pendaftaran, Kelas, Telepon, Telepon wali, Alamat, Pekerjaan, Nama Ibu, Agama, Kota Asal Sekolah, Pendidikan Ayah, Pendidikan Ibu, Pekerjaan Ayah, Pekerjaan Ibu, Asal Sekolah, Kategori Sekolah, Jenis Sekolah, Penghasilan, Prodi, Kampus, Gelombang, Status, NPM, Tahun Lulus, dan Email.

*Table 3.4 Table Data 2019*

| No  | ID Registrasi | Nama                            | Keterangan | NIK              | Tanggal Lahir | Jenis Kelamin | Jenis Pendaftaran | Kelas | :   | Email                   |
|-----|---------------|---------------------------------|------------|------------------|---------------|---------------|-------------------|-------|-----|-------------------------|
| 1   | 19.S1.1.0009  | Eka Putri Christiani Hutagalung | Umum       | 3275114307990000 | 03/07/1999    | P             | Mahasiswa Baru    | Sore  | :   | ekaputriak3@gmail.com   |
| 2   | 19.S1.1.0001  | Bella Nusa Bangsa               | Umum       | 3275055604980010 | 16/04/1998    | P             | Mahasiswa Baru    | Pagi  | :   | bellanusa3@gmail.com    |
| 3   | 19.S1.1.0002  | Riza Fauziah                    | Umum       | 3275036808950020 | 28/08/1995    | P             | Mahasiswa Baru    | Sore  | :   | rizafauziah08@gmail.com |
| 4   | 19.S1.1.0003  | Roy Fernando                    | Umum       | 1301100908910000 | 06/08/1990    | L             | Mahasiswa Baru    | P2k   | :   | royfernando21@gmail.com |
| ... | ...           | ...                             | ...        | ...              | ...           | ...           | ...               | ...   | ... | ...                     |

| No   | ID Registrasi | Nama                | Keterangan | NIK              | Tanggal Lahir | Jenis Kelamin | Jenis Pendaftaran | Kelas | : | Email                   |
|------|---------------|---------------------|------------|------------------|---------------|---------------|-------------------|-------|---|-------------------------|
| 3237 | 19.S1.5.4996  | Muhammad Alfiansyah | Umum       | 3171041010000000 | 10/10/2000    | L             | Mahasiswa Baru    | Pagi  | : | anisanisah.aa@gmail.com |



### **3.4 Permasalahan**

Data yang ada pada proses penerimaan mahasiswa baru menghasilkan data yang berlimpah melingkupi data - data mahasiswa tersebut, hampir setiap tahun Universitas X menerima mahasiswa baru sehingga data yang tersimpan akan semakin banyak dan jika data yang ada tidak dimanfaatkan dengan baik sebagai hal positif bagi universitas. Sama halnya yang terjadi di lingkungan Universitas X setiap tahun menerima hampir 2.000 lebih mahasiswa baru dari berbagai kota. Untuk itu perlu diolah dengan data mining clustering k – means sehingga nantinya dapat digunakan untuk strategi promosi sebaran wilayah berdasarkan kota asal sekolah.

### **3.5 Analisis Usulan Sistem**

Pada sub bab ini akan diuraikan tentang tahapan yang dilakukan untuk mengolah data mentah menjadi data yang baik untuk dilanjutkan ke proses perhitungan data mining clustering k - means. Pada tahapan data ini memiliki beberapa proses :

#### **3.5.1 Data Preparation**

##### **1. Data Integration**

Data integration atau integrasi data adalah suatu proses untuk menggabungkan data dari beberapa file. Pada pengambilan data ini, penulis mendapatkan 1 file excel dan memiliki 4 sheet excel selama empat tahun dari 2016 sampai dengan 2019. Data tersebut harus digabungkan sesuai dengan kebutuhan penulisan menjadi satu file dan satu sheet excel agar mudah diolah ke dalam metode k – means clustering. Setelah di gabungkan data tersebut berjumlah 10612 data.

Table 3.5 Table Data Integration

| No    | ID Registrasi | Nama                 | Keterangan | NIK              | Tanggal Lahir | Jenis Kelamin | Jenis Pendaftaran | Kelas | :   | Tahun Lulus |
|-------|---------------|----------------------|------------|------------------|---------------|---------------|-------------------|-------|-----|-------------|
| 1     | 16.S1.1.0001  | Daryus               | Umum       | 3175060712750000 | 07/12/1975    | L             | Mahasiswa Baru    | Sore  | :   | 1994        |
| 2     | 16.S1.1.0003  | Ida Faridah          | Umum       | 3216024804940000 | 08/04/1994    | P             | Mahasiswa Baru    | Sore  | :   | 2013        |
| 3     | 16.S1.1.0004  | Dian Fahruly Nur     | Umum       | 3216020303980000 | 03/03/1998    | L             | Mahasiswa Baru    | Pagi  | :   | 2016        |
| 4     | 16.S1.1.0005  | Muhammad Iqbal       | Umum       | 3275032507980020 | 25/07/1998    | L             | Mahasiswa Baru    | Pagi  | :   | 2016        |
| 5     | 17.S1.1.0007  | Imam Miftahudin      | Umum       | 3208160307950000 | 02/07/1995    | L             | Mahasiswa Baru    | Sore  | :   | 2014        |
| ...   | ...           | ...                  | ...        | ...              | ...           | ...           | ...               | ...   | ... | ...         |
| 10612 | 19.S1.4.0005  | Nabil Rifqi Ramdhani | Umum       | 3216011408010000 | 14/08/2001    | L             | Mahasiswa Baru    | Pagi  | :   | 2019        |

## 2. Data Reduction

Data reduction adalah proses untuk mereduksi atau mengurangi dimensi, atribut, ataupun sejumlah data yang tidak digunakan untuk perhitungan clustering  $k - means$ . Data reduction sangat berguna untuk mendapatkan atribut dan sejumlah data yang akan digunakan dalam penulisan ini. Dalam penulisan ini atribut yang digunakan adalah id registrasi, kota asal sekolah, kategori sekolah, dan jenis sekolah.

## 3. Data Cleaning

Data cleaning adalah proses untuk mengisi *missing value* atau isian yang hilang atau kurang pas. peneliti melakukan pembersihan data terhadap data mahasiswa baru dengan cara memfilter setiap kolom dan mencari data yang kosong atau *missing*. Pada data mahasiswa yang sudah melewati proses pembersihan data terdapat 213 data yang dihapus karena kosong atau *missing*, dan terdapat 10399 data mahasiswa yang digunakan untuk penulisan ini selama 4 tahun dari angkatan 2016 sampai dengan angkatan 2019. Dapat dilihat data yang sudah di cleaning sebagai berikut:

Table 3.6 Table Cleaning Data

| No | ID Registrasi   | Kota Asal Sekolah | Kategori Sekolah | Jenis Sekolah |
|----|-----------------|-------------------|------------------|---------------|
| 1  | 16.KS1.1.0001   |                   | Swasta           | LAIN          |
| 2  | 16.KS1.1.0002   |                   | Swasta           | LAIN          |
| 3  | 16.RA.S1.1.0001 |                   | Swasta           | LAIN          |
| 4  | 16.RA.S1.1.0002 |                   | Swasta           | LAIN          |
| 5  | 16.KS1.2.0003   |                   | Swasta           | LAIN          |
| 6  | 16.KS1.2.0004   |                   | Swasta           | LAIN          |
| 7  | 16.KS1.4.0005   |                   | Swasta           | LAIN          |
| 8  | 16.RA.S1.4.0003 |                   | Swasta           | LAIN          |
| 9  | 16.KS1.4.0006   |                   | Swasta           | LAIN          |
| 10 | 19.S1.5.0004    |                   | Swasta           |               |

|       |              |       |        |       |
|-------|--------------|-------|--------|-------|
| ..... | .....        | ..... | .....  | ..... |
| 213   | 19.S1.5.0005 |       | Swasta |       |

#### 4. Sampel dan Populasi

Dalam proses selanjutnya peneliti menggunakan sampel dan populasi. Dengan cara menentukan jumlah sampel dan populasi dari 10399 data mahasiswa, dan diambil dengan taraf kesalahan (*significance level*) 5% yaitu 336 data yang digunakan untuk clustering k – means.

*Table 3.7 Table Sampel dan Populasi*

| N   | S   |    |     | N   | S   |     |     | N      | S   |     |     |
|-----|-----|----|-----|-----|-----|-----|-----|--------|-----|-----|-----|
|     | 1%  | 5% | 10% |     | 1%  | 5%  | 10% |        | 1%  | 5%  | 10% |
| 10  | 10  | 10 | 10  | 280 | 197 | 155 | 138 | 2800   | 537 | 310 | 247 |
| 15  | 15  | 14 | 14  | 290 | 202 | 158 | 140 | 3000   | 543 | 312 | 248 |
| 20  | 19  | 19 | 19  | 300 | 207 | 161 | 143 | 3500   | 558 | 317 | 251 |
| 25  | 24  | 25 | 23  | 320 | 216 | 167 | 147 | 4000   | 569 | 320 | 254 |
| 30  | 29  | 28 | 27  | 340 | 225 | 172 | 151 | 4500   | 578 | 323 | 255 |
| 35  | 33  | 32 | 31  | 360 | 234 | 177 | 155 | 5000   | 586 | 326 | 257 |
| 40  | 38  | 36 | 35  | 380 | 242 | 182 | 158 | 6000   | 598 | 329 | 259 |
| 45  | 42  | 40 | 39  | 400 | 250 | 186 | 162 | 7000   | 606 | 332 | 261 |
| 50  | 47  | 44 | 42  | 420 | 257 | 191 | 165 | 8000   | 613 | 334 | 263 |
| 55  | 51  | 48 | 46  | 440 | 265 | 195 | 168 | 9000   | 618 | 335 | 263 |
| 60  | 55  | 51 | 49  | 460 | 272 | 198 | 171 | 10000  | 622 | 336 | 263 |
| 65  | 59  | 55 | 53  | 480 | 279 | 202 | 173 | 15000  | 635 | 340 | 266 |
| 70  | 63  | 58 | 56  | 500 | 285 | 205 | 176 | 20000  | 642 | 342 | 267 |
| 75  | 67  | 62 | 59  | 550 | 301 | 213 | 182 | 30000  | 649 | 344 | 268 |
| 80  | 71  | 65 | 62  | 600 | 315 | 221 | 187 | 40000  | 563 | 345 | 269 |
| 85  | 75  | 68 | 65  | 650 | 329 | 227 | 191 | 50000  | 655 | 346 | 269 |
| 90  | 79  | 72 | 68  | 700 | 341 | 233 | 195 | 75000  | 658 | 346 | 270 |
| 95  | 83  | 75 | 71  | 750 | 352 | 238 | 199 | 100000 | 659 | 347 | 270 |
| 100 | 87  | 78 | 73  | 800 | 363 | 243 | 202 | 150000 | 661 | 347 | 270 |
| 110 | 94  | 84 | 78  | 850 | 373 | 247 | 205 | 200000 | 661 | 347 | 270 |
| 120 | 102 | 89 | 83  | 900 | 382 | 251 | 208 | 250000 | 662 | 348 | 270 |
| 130 | 109 | 95 | 88  | 950 | 391 | 255 | 211 | 300000 | 662 | 348 | 270 |

|     |     |     |     |      |     |     |     |         |     |     |     |
|-----|-----|-----|-----|------|-----|-----|-----|---------|-----|-----|-----|
| 140 | 116 | 100 | 92  | 1000 | 399 | 258 | 213 | 350000  | 662 | 348 | 270 |
| 150 | 122 | 105 | 97  | 1100 | 414 | 265 | 217 | 400000  | 662 | 348 | 270 |
| 160 | 129 | 110 | 101 | 1200 | 427 | 270 | 221 | 450000  | 663 | 348 | 270 |
| 170 | 135 | 114 | 105 | 1300 | 440 | 275 | 224 | 500000  | 663 | 348 | 270 |
| 180 | 142 | 119 | 108 | 1400 | 450 | 279 | 227 | 550000  | 663 | 348 | 270 |
| 190 | 148 | 123 | 112 | 1500 | 460 | 283 | 229 | 600000  | 663 | 348 | 270 |
| 200 | 154 | 127 | 115 | 1600 | 469 | 286 | 232 | 650000  | 663 | 348 | 270 |
| 210 | 160 | 131 | 118 | 1700 | 477 | 289 | 234 | 700000  | 663 | 348 | 270 |
| 220 | 165 | 135 | 122 | 1800 | 485 | 292 | 235 | 750000  | 663 | 348 | 270 |
| 230 | 171 | 139 | 125 | 1900 | 492 | 294 | 237 | 800000  | 663 | 348 | 271 |
| 240 | 176 | 142 | 127 | 2000 | 498 | 297 | 238 | 850000  | 663 | 348 | 271 |
| 250 | 182 | 146 | 130 | 2200 | 510 | 301 | 241 | 900000  | 663 | 348 | 271 |
| 260 | 187 | 149 | 133 | 2400 | 520 | 304 | 243 | 950000  | 663 | 348 | 271 |
| 270 | 192 | 152 | 135 | 2600 | 529 | 307 | 245 | 1000000 | 663 | 348 | 271 |
|     |     |     |     |      |     |     |     | X       | 664 | 349 | 272 |

Refrensi : prof.dr.sugiyono

## 5. Transformasi Data

Dalam proses metode clustering k – means data yang bisa diolah hanya data yang berupa angka sedangkan data yang ada pada atribut kota asal sekolah, kategori sekolah, dan jenis sekolah adalah data yang berupa text sehingga peneliti harus mengubah data tersebut ke dalam angka agar bisa dilanjutkan ke dalam metode clustering k – means. Pada atribut kota asal sekolah peneliti mendapatkan 38 data kota asal sekolah dengan inisial sebagai berikut :

*Table 3.8 Table Transformasi Kota Asal Sekolah*

| Kota Asal Sekolah | Inisial |
|-------------------|---------|
| Bekasi Utara      | 1       |
| Bekasi Barat      | 2       |
| Bekasi Timur      | 3       |
| Bekasi Selatan    | 4       |
| Babelan           | 5       |

| Kota Asal Sekolah | Inisial |
|-------------------|---------|
| Bantargebang      | 6       |
| Tambun Selatan    | 7       |
| Medan Satria      | 8       |
| Rawalumbu         | 9       |
| Tambun Utara      | 10      |
| Mustika Jaya      | 11      |
| Tarumajaya        | 12      |
| Jatiasih          | 13      |
| Cikarang Barat    | 14      |
| Cikarang Selatan  | 15      |
| Margahayu         | 16      |
| Pondokgede        | 17      |
| Cibitung          | 18      |
| Jakarta Utara     | 19      |
| Pondok Melati     | 20      |
| Tambelang         | 21      |
| Jakarta Pusat     | 22      |
| Jakarta Timur     | 23      |
| Tegal             | 24      |
| Setu              | 25      |
| Garut             | 26      |
| Baamang           | 27      |
| Binjai Utara      | 28      |
| Ponorogo          | 29      |
| Babakan           | 30      |
| Indramayu         | 31      |
| Kebumen           | 32      |

| Kota Asal Sekolah | Inisial |
|-------------------|---------|
| Kuningan          | 33      |
| Manokwari         | 34      |
| Pamekasan         | 35      |
| Cilacap           | 36      |
| Temanggung        | 37      |
| Torgamba          | 38      |

Pada atribut kategori sekolah peneliti mendapatkan 2 data ketegori sekolah dengan inisial sebagai berikut :

*Table 3.9 Table Transformasi Kategori Sekolah*

| Kategori Sekolah | Inisial |
|------------------|---------|
| Negeri           | 1       |
| Swasta           | 2       |

Pada atribut jenis sekolah peneliti mendapatkan 4 data jenis sekolah dengan inisial sebagai berikut :

*Table 3.10 Table Transformasi Jenis Sekolah*

| Jenis Sekolah | Inisial |
|---------------|---------|
| SMK           | 1       |
| SMA           | 2       |
| MA            | 3       |
| LAIN          | 4       |

## 6. Perhitungan Clustering K - Means

Data yang akan diproses untuk data mining clustering k – means terdapat 336 data dengan 4 atribut yaitu id registrasi, kota asal sekolah, kategori sekolah, dan jenis sekolah. Dari 336 data sampel dibagi menjadi data training dan data testing, Data *training* dan data *testing* adalah 90% (302 data) untuk data *training* dan 10% (34 data) untuk data *testing*.

*Table 3.11 Table Training Clustering K - Means*

| No    | ID Registrasi | Kota Asal Sekolah | Kategori Sekolah | Jenis Sekolah |
|-------|---------------|-------------------|------------------|---------------|
| 1     | 16.S1.1.0001  | 19                | 2                | 4             |
| 2     | 16.S1.1.0002  | 7                 | 1                | 2             |
| 3     | 16.S1.1.0003  | 3                 | 2                | 1             |
| 4     | 16.S1.1.0004  | 8                 | 2                | 1             |
| 5     | 16.S1.1.0005  | 1                 | 2                | 1             |
| 6     | 16.S1.1.0006  | 8                 | 2                | 2             |
| 7     | 16.S1.1.0007  | 4                 | 2                | 1             |
| 8     | 16.S1.1.0012  | 9                 | 2                | 1             |
| 9     | 16.S1.1.0013  | 1                 | 1                | 1             |
| 10    | 16.S1.1.0014  | 11                | 1                | 1             |
| 11    | 16.S1.1.0015  | 27                | 2                | 2             |
| 12    | 16.S1.1.0016  | 1                 | 2                | 1             |
| 13    | 16.S1.1.0017  | 12                | 1                | 1             |
| 14    | 16.S1.1.0018  | 1                 | 2                | 1             |
| 15    | 16.S1.1.0019  | 28                | 2                | 1             |
| 16    | 16.S1.1.0020  | 7                 | 1                | 2             |
| 17    | 16.S1.1.0021  | 1                 | 2                | 1             |
| 18    | 16.S1.1.0022  | 3                 | 2                | 1             |
| 19    | 16.S1.1.0061  | 7                 | 2                | 4             |
| 20    | 16.S1.1.0062  | 23                | 2                | 1             |
| ..... | .....         | .....             | .....            | .....         |
| 302   | 19.S1.2.0472  | 23                | 2                | 1             |



a. Menentukan Nilai K (Cluster)

Data – data yang ada akan dikelompokkan menjadi 3 cluster (K = 3) terdiri dari cluster 1, cluster 2, dan cluster 3.

b. Menentukan Titik K Centroid (Titik Pusat Cluster) Awal

Dalam penulisan ini titik pusat awal cluster yang sudah ditentukan secara *random* (acak) dan didapat titik pusat dari setiap cluster dapat dilihat sebagai berikut :

*Table 3.12 Table Titik Pusat Awal Cluster*

| Titik Pusat Awal Cluster | ID Registrasi | Kota Asal Sekolah | Kategori Sekolah | Jenis Sekolah |
|--------------------------|---------------|-------------------|------------------|---------------|
| Cluster 1                | 16.S1.1.0002  | 7                 | 1                | 2             |
| Cluster 2                | 16.S1.1.0015  | 27                | 2                | 2             |
| Cluster 3                | 16.S1.1.0021  | 1                 | 2                | 1             |

c. Menghitung Jarak Data Ke Centroid

Pada proses ini, setelah diketahui nilai k (cluster) dan titik pusat cluster awal selanjutnya mengukur jarak antara pusat cluster menggunakan euclidian distance.

Rumus Euclidian Distance :

$$d(x,y) = |x-y| = \sqrt{\sum_{i=1}^n (Xi - Yi)^2}$$

d. Perhitungan setiap data ke setiap cluster

1) Hasil perhitungan data ke setiap cluster literasi 1

*Table 3.13 Table literasi 1*

| Data ke i | c1    | c2    | c3    | Jarak | Cluster |
|-----------|-------|-------|-------|-------|---------|
| 1         | 12,21 | 8,25  | 18,25 | 8,25  | 2       |
| 2         | 0,00  | 20,02 | 6,16  | 0,00  | 1       |
| 3         | 4,24  | 24,02 | 2,00  | 2,00  | 3       |

|       |       |       |       |       |       |
|-------|-------|-------|-------|-------|-------|
| 4     | 1,73  | 19,03 | 7,00  | 1,73  | 1     |
| 5     | 6,16  | 26,02 | 0,00  | 0,00  | 3     |
| 6     | 1,41  | 19,00 | 7,07  | 1,41  | 1     |
| 7     | 3,32  | 23,02 | 3,00  | 3,00  | 3     |
| 8     | 2,45  | 18,03 | 8,00  | 2,45  | 1     |
| 9     | 6,08  | 26,04 | 1,00  | 1,00  | 3     |
| 10    | 4,12  | 16,06 | 10,05 | 4,12  | 1     |
| 11    | 20,02 | 0,00  | 26,02 | 0,00  | 2     |
| 12    | 6,16  | 26,02 | 0,00  | 0,00  | 3     |
| 13    | 5,10  | 15,07 | 11,05 | 5,10  | 1     |
| 14    | 6,16  | 26,02 | 0,00  | 0,00  | 3     |
| 15    | 21,05 | 1,41  | 27,00 | 1,41  | 2     |
| 16    | 0,00  | 20,02 | 6,16  | 0,00  | 1     |
| 17    | 6,16  | 26,02 | 0,00  | 0,00  | 3     |
| 18    | 4,24  | 24,02 | 2,00  | 2,00  | 3     |
| 19    | 2,24  | 20,10 | 6,71  | 2,24  | 1     |
| 20    | 16,06 | 4,12  | 22,00 | 4,12  | 2     |
| ..... | ..... | ..... | ..... | ..... | ..... |
| 302   | 16,06 | 4,12  | 22,00 | 4,12  | 2     |

*Table 3.14 Cluster Baru Literasi 1*

| Cluster 1 |      |      | Cluster 2 |      |      | Cluster 3 |      |      |
|-----------|------|------|-----------|------|------|-----------|------|------|
| 8,77      | 1,61 | 1,39 | 23,23     | 1,75 | 1,64 | 1,83      | 1,80 | 1,34 |

Pada tahap ini centroid baru yang dibangkitkan ternyata belum konvergen, sehingga literasi harus dilanjutkan.

2) Hasil Perhitungan data ke setiap cluster literasi 2

*Table 3.15 Table Literasi 2*

| Data ke i | c1    | c2    | c3    | Jarak | Cluster |
|-----------|-------|-------|-------|-------|---------|
| 1         | 10,57 | 4,85  | 17,38 | 4,85  | 2       |
| 2         | 1,97  | 16,26 | 5,63  | 1,97  | 1       |
| 3         | 5,80  | 20,34 | 2,52  | 2,52  | 3       |
| 4         | 0,95  | 15,37 | 6,56  | 0,95  | 1       |
| 5         | 7,79  | 22,33 | 2,38  | 2,38  | 3       |
| 6         | 1,06  | 15,36 | 6,79  | 1,06  | 1       |
| 7         | 4,80  | 19,34 | 3,12  | 3,12  | 3       |
| 8         | 0,60  | 14,38 | 7,51  | 0,60  | 1       |
| 9         | 7,80  | 22,26 | 1,64  | 1,64  | 3       |
| 10        | 2,34  | 12,28 | 9,28  | 2,34  | 1       |
| 11        | 18,24 | 4,29  | 25,33 | 4,29  | 2       |
| 12        | 7,79  | 22,33 | 2,38  | 2,38  | 3       |
| 13        | 3,31  | 11,29 | 10,27 | 3,31  | 1       |
| 14        | 7,79  | 22,33 | 2,38  | 2,38  | 3       |
| 15        | 19,24 | 5,21  | 26,27 | 5,21  | 2       |
| 16        | 1,97  | 16,26 | 5,63  | 1,97  | 1       |
| 17        | 7,79  | 22,33 | 2,38  | 2,38  | 3       |
| 18        | 5,80  | 20,34 | 2,52  | 2,52  | 3       |
| 19        | 3,18  | 16,52 | 6,84  | 3,18  | 1       |
| 20        | 14,24 | 2,11  | 21,29 | 2,11  | 2       |
| .....     | ..... | ..... | ..... | ..... | .....   |
| 302       | 14,24 | 2,11  | 21,29 | 2,11  | 2       |

*Table 3.16 Cluster Baru Literasi 2*

| Cluster 1 |      |      | Cluster 2 |      |      | Cluster 3 |      |      |
|-----------|------|------|-----------|------|------|-----------|------|------|
| 8,88      | 1,62 | 1,36 | 22,71     | 1,71 | 1,60 | 1,91      | 1,76 | 1,35 |

Pada tahap ini centroid baru yang dibangkitkan ternyata masih belum konvergen, sehingga literasi harus dilanjutkan, ke literasi 3.

3) Hasil perhitungan data ke setiap cluster literasi 3

*Table 3.17 Table Literasi 3*

| Data ke i | c1    | c2    | c3    | Jarak | Cluster |
|-----------|-------|-------|-------|-------|---------|
| 1         | 10,47 | 4,43  | 17,30 | 4,43  | 2       |
| 2         | 2,08  | 15,73 | 5,19  | 2,08  | 1       |
| 3         | 5,90  | 19,72 | 1,17  | 1,17  | 3       |
| 4         | 1,02  | 14,73 | 6,11  | 1,02  | 1       |
| 5         | 7,90  | 21,72 | 1,00  | 1,00  | 3       |
| 6         | 1,15  | 14,72 | 6,13  | 1,15  | 1       |
| 7         | 4,91  | 18,72 | 2,14  | 2,14  | 3       |
| 8         | 0,54  | 13,73 | 7,11  | 0,54  | 1       |
| 9         | 7,91  | 21,73 | 1,23  | 1,23  | 3       |
| 10        | 2,24  | 11,75 | 9,13  | 2,24  | 1       |
| 11        | 18,14 | 4,32  | 25,10 | 4,32  | 2       |
| 12        | 7,90  | 21,72 | 1,00  | 1,00  | 3       |
| 13        | 3,20  | 10,75 | 10,13 | 3,20  | 1       |
| 14        | 7,90  | 21,72 | 1,00  | 1,00  | 3       |
| 15        | 19,13 | 5,33  | 26,10 | 5,33  | 2       |
| 16        | 2,08  | 15,73 | 5,19  | 2,08  | 1       |
| 17        | 7,90  | 21,72 | 1,00  | 1,00  | 3       |
| 18        | 5,90  | 19,72 | 1,17  | 1,17  | 3       |
| 19        | 3,26  | 15,90 | 5,75  | 3,26  | 1       |
| 20        | 14,13 | 0,73  | 21,10 | 0,73  | 2       |
| .....     | ..... | ..... | ..... | ..... | .....   |
| 302       | 14,13 | 0,73  | 21,10 | 0,73  | 2       |

*Table 3.18 Cluster Baru Literasi 3*

| Cluster 1 |      |      | Cluster 2 |      |      | Cluster 3 |      |      |
|-----------|------|------|-----------|------|------|-----------|------|------|
| 9,37      | 1,66 | 1,31 | 23,07     | 1,76 | 1,69 | 2,23      | 1,73 | 1,39 |

Pada tahap ini centroid baru yang dibangkitkan ternyata masih belum konvergen, sehingga literasi harus dilanjutkan, ke literasi 4.

4) Hasil perhitungan data ke setiap cluster literasi 4

*Table 3.19 Table Literasi 4*

| Data ke i | c1    | c2    | c3    | Jarak | Cluster |
|-----------|-------|-------|-------|-------|---------|
| 1         | 10,01 | 4,68  | 16,98 | 4,68  | 2       |
| 2         | 2,55  | 16,09 | 4,87  | 2,55  | 1       |
| 3         | 6,38  | 20,08 | 0,91  | 0,91  | 3       |
| 4         | 1,44  | 15,08 | 5,79  | 1,44  | 1       |
| 5         | 8,38  | 22,08 | 1,31  | 1,31  | 3       |
| 6         | 1,57  | 15,07 | 5,81  | 1,57  | 1       |
| 7         | 5,39  | 19,08 | 1,84  | 1,84  | 3       |
| 8         | 0,59  | 14,09 | 6,79  | 0,59  | 1       |
| 9         | 8,40  | 22,09 | 1,48  | 1,48  | 3       |
| 10        | 1,79  | 12,11 | 8,81  | 1,79  | 1       |
| 11        | 17,65 | 3,95  | 24,78 | 3,95  | 2       |
| 12        | 8,38  | 22,08 | 1,31  | 1,31  | 3       |
| 13        | 2,73  | 11,11 | 9,81  | 2,73  | 1       |
| 14        | 8,38  | 22,08 | 1,31  | 1,31  | 3       |
| 15        | 18,64 | 4,99  | 25,78 | 4,99  | 2       |
| 16        | 2,55  | 16,09 | 4,87  | 2,55  | 1       |
| 17        | 8,38  | 22,08 | 1,31  | 1,31  | 3       |
| 18        | 6,38  | 20,08 | 0,91  | 0,91  | 3       |
| 19        | 3,60  | 16,23 | 5,45  | 3,60  | 1       |
| 20        | 13,64 | 0,73  | 20,78 | 0,73  | 2       |
| .....     | ..... | ..... | ..... | ..... | .....   |
| 302       | 13,64 | 0,73  | 20,78 | 0,73  | 2       |

*Table 3. 20 Cluster Baru Literasi 4*

| Cluster 1 |      |      | Cluster 2 |      |      | Cluster 3 |      |      |
|-----------|------|------|-----------|------|------|-----------|------|------|
| 9,42      | 1,66 | 1,34 | 23,23     | 1,75 | 1,64 | 2,23      | 1,73 | 1,39 |

Pada tahap ini centroid baru yang dibangkitkan ternyata masih belum konvergen, sehingga literasi harus dilanjutkan, ke literasi 5.

5) Hasil perhitungan data ke setiap cluster literasi 5

*Table 3.21 Table Literasi 5*

| Data ke i | c1    | c2    | c3    | Jarak | Cluster |
|-----------|-------|-------|-------|-------|---------|
| 1         | 9,95  | 4,85  | 16,98 | 4,85  | 2       |
| 2         | 2,60  | 16,25 | 4,87  | 2,60  | 1       |
| 3         | 6,44  | 20,24 | 0,91  | 0,91  | 3       |
| 4         | 1,50  | 15,24 | 5,79  | 1,50  | 1       |
| 5         | 8,44  | 22,24 | 1,31  | 1,31  | 3       |
| 6         | 1,61  | 15,23 | 5,81  | 1,61  | 1       |
| 7         | 5,44  | 19,24 | 1,84  | 1,84  | 3       |
| 8         | 0,64  | 14,24 | 6,79  | 0,64  | 1       |
| 9         | 8,46  | 22,25 | 1,48  | 1,48  | 3       |
| 10        | 1,74  | 12,27 | 8,81  | 1,74  | 1       |
| 11        | 17,59 | 3,80  | 24,78 | 3,80  | 2       |
| 12        | 8,44  | 22,24 | 1,31  | 1,31  | 3       |
| 13        | 2,68  | 11,27 | 9,81  | 2,68  | 1       |
| 14        | 8,44  | 22,24 | 1,31  | 1,31  | 3       |
| 15        | 18,58 | 4,82  | 25,78 | 4,82  | 2       |
| 16        | 2,60  | 16,25 | 4,87  | 2,60  | 1       |
| 17        | 8,44  | 22,24 | 1,31  | 1,31  | 3       |
| 18        | 6,44  | 20,24 | 0,91  | 0,91  | 3       |
| 19        | 3,62  | 16,40 | 5,45  | 3,62  | 1       |
| 20        | 13,59 | 0,72  | 20,78 | 0,72  | 2       |
| .....     | ..... | ..... | ..... | ..... | .....   |
| 302       | 13,59 | 0,72  | 20,78 | 0,72  | 2       |

*Table 3.22 Cluster Baru Literasi 5*

| Cluster 1 |      |      | Cluster 2 |      |      | Cluster 3 |      |      |
|-----------|------|------|-----------|------|------|-----------|------|------|
| 9,42      | 1,66 | 1,34 | 23,23     | 1,75 | 1,64 | 2,23      | 1,73 | 1,39 |

Pada perhitungan ini literasi berhenti pada literasi ke-5 centroid baru yang sudah dihitung sudah tidak berubah (stabil) dan sudah konvergen.

Langkah selanjutnya melakukan perhitungan untuk menunjukkan bahwa data tersebut berada dalam satu kelompok dengan pusat cluster.

*Table 3.23 Pengelompokan Pusat Cluster*

| Data ke i | c1    | c2    | c3    | Jarak | Cluster 1 | Cluster 2 | Cluster 3 |
|-----------|-------|-------|-------|-------|-----------|-----------|-----------|
| 1         | 9,95  | 4,85  | 16,98 | 4,85  |           | ok        |           |
| 2         | 2,60  | 16,25 | 4,87  | 2,60  | ok        |           |           |
| 3         | 6,44  | 20,24 | 0,91  | 0,91  |           |           | ok        |
| 4         | 1,50  | 15,24 | 5,79  | 1,50  | ok        |           |           |
| 5         | 8,44  | 22,24 | 1,31  | 1,31  |           |           | ok        |
| 6         | 1,61  | 15,23 | 5,81  | 1,61  | ok        |           |           |
| 7         | 5,44  | 19,24 | 1,84  | 1,84  |           |           | ok        |
| 8         | 0,64  | 14,24 | 6,79  | 0,64  | ok        |           |           |
| 9         | 8,46  | 22,25 | 1,48  | 1,48  |           |           | ok        |
| 10        | 1,74  | 12,27 | 8,81  | 1,74  | ok        |           |           |
| 11        | 17,59 | 3,80  | 24,78 | 3,80  |           | ok        |           |
| 12        | 8,44  | 22,24 | 1,31  | 1,31  |           |           | ok        |
| 13        | 2,68  | 11,27 | 9,81  | 2,68  | ok        |           |           |
| 14        | 8,44  | 22,24 | 1,31  | 1,31  |           |           | ok        |
| 15        | 18,58 | 4,82  | 25,78 | 4,82  |           | ok        |           |
| 16        | 2,60  | 16,25 | 4,87  | 2,60  | ok        |           |           |
| 17        | 8,44  | 22,24 | 1,31  | 1,31  |           |           | ok        |
| 18        | 6,44  | 20,24 | 0,91  | 0,91  |           |           | ok        |
| 19        | 3,62  | 16,40 | 5,45  | 3,62  | ok        |           |           |
| 20        | 13,59 | 0,72  | 20,78 | 0,72  |           | ok        |           |
| .....     | ..... | ..... | ..... | ..... | .....     | .....     | .....     |
| 302       | 13,59 | 0,72  | 20,78 | 0,72  |           | ok        |           |

Dalam proses selanjutnya peneliti mencari nilai rata – rata berdasarkan data yang sudah dikelompokkan ke setiap cluster masing – masing dan mendapatkan hasil sebagai berikut :

- a) Pada cluster 1 dengan jumlah data 116 data

*Table 3.24 Table Titik Pusat Cluster Baru (Cluster 1)*

| Titik Pusat Cluster Baru |      |                     |      |                  |      |
|--------------------------|------|---------------------|------|------------------|------|
| Kota<br>Asal<br>Sekolah  | 9,42 | Kategori<br>Sekolah | 1,66 | Jenis<br>Sekolah | 1,34 |

- b) Pada cluster 2 dengan jumlah data 44 data

*Table 3.25 Table Titik Pusat Cluster Baru (Cluster 2)*

| Titik Pusat Cluster Baru |       |                     |      |                  |      |
|--------------------------|-------|---------------------|------|------------------|------|
| Kota<br>Asal<br>Sekolah  | 23,23 | Kategori<br>Sekolah | 1,75 | Jenis<br>Sekolah | 1,64 |

- c) Pada cluster 3 dengan jumlah data 142 data

*Table 3.26 Table Titik Pusat Cluster Baru (Cluster 3)*

| Titik Pusat Cluster Baru |      |                     |      |                  |      |
|--------------------------|------|---------------------|------|------------------|------|
| Kota<br>Asal<br>Sekolah  | 2,23 | Kategori<br>Sekolah | 1,73 | Jenis<br>Sekolah | 1,39 |



e. Hasil Perhitungan Clustering K – Means

Setelah melewati tahapan proses clustering k – means menggunakan microsoft excel maka proses berikutnya adalah menampilkan hasil dari perhitungan tersebut. Untuk mendapatkan hasil yang diolah menggunakan clustering k – means penulis melewati perulangan proses literasi sebanyak 5 kali proses literasi. Hasil dari penerapan ini adalah sebagai berikut :

*Table 3.27 Table Hasil Cluster 1*

| Hasil Cluster 1  |     |                  |     |               |     |
|--|-----|------------------|-----|---------------|-----|
| CLUSTER 1 TERDIRI DARI 116 MAHASISWA YANG BERASAL DARI |     |                  |     |               |     |
| Kota Asal Sekolah                                      |     | KATEGORI SEKOLAH |     | Jenis Sekolah |     |
| Bantargebang   | 3   | Negeri           | 39  | SMK           | 81  |
| Tambun Selatan   | 30  | Swasta           | 77  | SMA           | 33  |
| Medan Satria   | 20  |                  |     | Lain          | 2   |
| Rawalumbu  | 17  |                  |     |               |     |
| Tambun Utara   | 10  |                  |     |               |     |
| Mustika Jaya   | 8   |                  |     |               |     |
| Tarumajaya   | 14  |                  |     |               |     |
| Jatiasih   | 4   |                  |     |               |     |
| Cikarang Barat   | 7   |                  |     |               |     |
| Cikarang Selatan                                       | 2   |                  |     |               |     |
| Margahayu  | 1   |                  |     |               |     |
| Total  | 116 |                  | 116 |               | 116 |

Berdasarkan hasil cluster dapat diketahui terdapat 116 data mahasiswa yang termasuk kedalam kelompok cluster satu. Dalam table tersebut terdapat informasi tentang persebaran wilayah kota asal sekolah, kategori sekolah, dan jenis sekolah beserta jumlah mahasiswa. Jumlah data mahasiswa pada cluster satu adalah jumlah data yang kurang dari cluster tiga tetapi lebih dari cluster dua. Sehingga dapat mempertimbangkan strategi promosi ditahun berikutnya.

Table 3.28 Table Hasil Cluster 2

| Hasil Cluster 2                                       |    |                  |    |               |    |
|---|----|------------------|----|---------------|----|
| CLUSTER 2 TERDIRI DARI 44 MAHASISWA YANG BERASAL DARI |    |                  |    |               |    |
| Kota Asal Sekolah                                     |    | Kategori Sekolah |    | Jenis Sekolah |    |
| Pondokgede  | 2  | Negeri           | 11 | SMK           | 23 |
| Cibitung  | 1  | Swasta           | 33 | SMA           | 16 |
| Jakarta Utara   | 7  |                  |    | MA            | 3  |
| Pondok Melati   | 1  |                  |    | Lain          | 2  |
| Tambelang   | 1  |                  |    |               |    |
| Jakarta Pusat   | 1  |                  |    |               |    |
| Jakarta Timur   | 20 |                  |    |               |    |
| Tegal   | 1  |                  |    |               |    |
| Setu  | 3  |                  |    |               |    |
| Baamang   | 1  |                  |    |               |    |
| Binjai Utara  | 1  |                  |    |               |    |
| Ponorogo  | 1  |                  |    |               |    |
| Indramayu   | 1  |                  |    |               |    |
| Kebumen   | 1  |                  |    |               |    |
| Kuningan  | 1  |                  |    |               |    |
| Pamekasan   | 1  |                  |    |               |    |
| Total   | 44 |                  | 44 |               | 44 |

Berdasarkan hasil cluster dua dapat diketahui bahawa terdapat 44 data mahasiswa yang termasuk ke dalam kelompok cluster dua. Dalam table tersebut terdapat informasi tentang persebaran wilayah kota asal sekolah, kategori sekolah, dan jenis sekolah beserta jumlah mahasiswa. Jumlah data mahasiswa pada cluster dua adalah jumlah data yang paling sedikit dibandingkan dengan cluster satu dan cluster tiga. Sehingga dapat mempertimbangkan strategi promosi ditahun berikutnya.

Table 3.29 Table Hasil Cluster 3

| Hasil Cluster 3  |     |                  |     |               |     |
|--|-----|------------------|-----|---------------|-----|
| CLUSTER 3 TERDIRI DARI 142 MAHASISWA YANG BERASAL DARI |     |                  |     |               |     |
| Kota Asal Sekolah                                      |     | Kategori Sekolah |     | Jenis Sekolah |     |
| Bekasi Utara   | 70  | Negeri           | 38  | SMK           | 98  |
| Bekasi Barat   | 8   | Swasta           | 104 | SMA           | 35  |
| Bekasi Timur   | 41  |                  |     | MA            | 7   |
| Bekasi Selatan   | 8   |                  |     | Lain          | 2   |
| Babelan  | 15  |                  |     |               |     |
| Total  | 142 |                  | 142 |               | 142 |

Berdasarkan hasil cluster tiga dapat diketahui bahwa terdapat 142 data mahasiswa yang termasuk ke dalam kelompok cluster tiga. Dalam table tersebut terdapat informasi tentang persebaran wilayah kota asal sekolah, kategori sekolah, dan jenis sekolah beserta jumlah mahasiswa. Jumlah data mahasiswa pada cluster tiga adalah jumlah data yang paling banyak di antara cluster yang lain sehingga dapat mempertimbangkan strategi promosi ditahun berikutnya.

f. Tabel data testing

Pada Tabel ini adalah table data *testing* dari sampel data 336 yang dengan 10% (34 data) untuk data *testing*.

Table 3.30 Table Data Testing

| No | ID Registrasi | Kota Asal Sekolah | Kategori Sekolah | Jenis Sekolah |
|----|---------------|-------------------|------------------|---------------|
| 1  | 17.S1.1.0186  | 38                | 2                | 2             |
| 2  | 18.S1.1.0234  | 34                | 1                | 2             |
| 3  | 16.S1.1.0142  | 30                | 1                | 2             |
| 4  | 16.S1.1.0008  | 26                | 2                | 1             |
| 5  | 17.S1.1.0043  | 22                | 1                | 1             |
| 6  | 17.S1.1.0048  | 22                | 2                | 1             |
| 7  | 19.S1.1.0265  | 18                | 2                | 1             |
| 8  | 19.S1.1.0280  | 18                | 1                | 2             |
| 9  | 16.S1.1.0145  | 14                | 1                | 2             |

| No    | ID Registrasi | Kota Asal Sekolah | Kategori Sekolah | Jenis Sekolah |
|-------|---------------|-------------------|------------------|---------------|
| 10    | 19.S1.2.0483  | 10                | 1                | 2             |
| 11    | 19.S1.2.0473  | 1                 | 2                | 1             |
| 12    | 19.S1.2.0474  | 1                 | 2                | 1             |
| 13    | 19.S1.2.0475  | 11                | 2                | 1             |
| 14    | 19.S1.2.0476  | 3                 | 2                | 1             |
| 15    | 19.S1.2.0477  | 3                 | 2                | 1             |
| 16    | 19.S1.2.0478  | 36                | 2                | 1             |
| 17    | 19.S1.2.0479  | 23                | 2                | 1             |
| 18    | 19.S1.2.0480  | 8                 | 1                | 2             |
| 19    | 19.S1.2.0481  | 1                 | 2                | 1             |
| 20    | 19.S1.2.0482  | 2                 | 1                | 1             |
| 21    | 19.S1.2.0484  | 1                 | 1                | 1             |
| 22    | 19.S1.2.0485  | 11                | 2                | 2             |
| 23    | 19.S1.2.0486  | 11                | 2                | 2             |
| 24    | 19.S1.2.0487  | 1                 | 2                | 1             |
| 25    | 19.S1.2.0488  | 2                 | 1                | 1             |
| 26    | 19.S1.2.0489  | 37                | 1                | 1             |
| 27    | 19.S1.2.0490  | 11                | 2                | 2             |
| 28    | 19.S1.2.0491  | 1                 | 2                | 2             |
| ..... | .....         | .....             | .....            | .....         |
| 34    | 19.S1.2.0497  | 1                 | 2                | 1             |

### 3.6 Analisis Kebutuhan Sistem

Dalam melakukan perhitungan data mining clustering k – means tidak akan berhasil jika tidak didukung dengan sarana pendukung yang baik pula. lainnya. Maka dari itu untuk mendukung sistem tersebut membutuhkan suatu perangkat keras (*hardware*) dan perangkat lunak (*software*) berupa komputer dan aplikasinya untuk menginput pekerjaan yang sedang dilakukan.

### 3.6.1 Kebutuhan Perangkat Keras

Komponen perangkat keras (*hardware*) merupakan salah satu sarana pendukung yang wajib untuk sebuah sistem berbasis komputer, dengan dukungan perangkat keras yang baik maka sistem akan mempunyai kinerja yang baik pula. Berikut merupakan komponen perangkat keras yang digunakan untuk melakukan penulisan data mining clustering k – means :

1. Processor : AMD E1 - 1200
2. RAM : 4GB
3. HDD : 298GB
4. Display : AMD Radeon

### 3.6.2 Kebutuhan Perangkat Lunak

Berikut merupakan perangkat lunak (*software*) yang penulis gunakan untuk melakukan perhitungan data mining clustering k – means :

1. Operating System Windows 7 Ultimate 64-bit

Operasi sistem merupakan sistem utama dalam komputer yang menjalankan operasi secara keseluruhan

2. MS. Excel 2010

Untuk melakukan perhitungan, menganalisa, dan mengolah data dengan data mining clustering k - means

3. Rapidminer 5.3

Untuk melakukan pengelolaan data mining