

## BAB IV HASIL DAN PEMBAHASAN

### 1.1 Collecting atau Extract.

Pada tahapan *Collecting* Data yang dipakai adalah data *Mahasiswa* tahun 2015-2021 berdasarkan Nilai TA 2015/2021 sebanyak 262489 yang di ambil berdasarkan kurikulum tahun 2015/2021 semua prodi, Dari data mahasiswa tersebut di pilih mata kuliah peminatan dengan atribut yaitu : Nama, Npm, Jurusan, KDTA, Semester, KODEMK, Nama Mata Kuliah. Tabel 4.1 adalah data yang diperoleh dari BAAK DARMAJAYA sebagai berikut:

Tabel 4.3 Data set Mahasiswa tahun 2015-2021

kdta	npm	nama	ta	semester	kodemk	namamk	sks	waktu	kelas	ruang	quis1	quis2	uts	utsprakti	tugas	tugaspra	uas	uasprakti	presensi	total	huruf
20201	13110101	Edy Julia	2020/2021	15	IBI20612	Skripsi	6	13:00-14:00	77I-P1	F.3.4   F.3.4											
20201	14110100	Rizal Mu	2020/2021	13	IBI20612	Skripsi	6	18:30-20:00	177I-S1	F.2.7   F.2.7											
20201	14110101	Doly Nay	2020/2021	13	TIF20209	Mobile C	2	16:30-18:00	137I-S1	F.3.1			80	0	60	0	70	0	72	70,2	B+
20201	14110101	Doly Nay	2020/2021	13	TIF20213	Multime	2	18:30-20:00	137I-S1	Lab.Das B (G.2)			78	0	50	0	80	0	88	71,2	B+
20201	14110101	Doly Nay	2020/2021	13	IBI20216	Paten da	2	07:10-08:00	77I-P3	F.4.6			70	0	80	0	75	0	75	75	A-
20201	14110101	Doly Nay	2020/2021	13	TIF20404	Matemat	4	08:50-10:00	17I-P4	G.1.3   G.1.3			60	0	60	0	60	0	70	61	C
20201	14110101	Doly Nay	2020/2021	13	IBI20204	Pengemt	2	18:30-20:00	137I-S1	F.3.1			70	0	70	0	70	0	100	73	B+
20201	14110500	Wahyu A	2020/2021	13	SIF20224	Teknik V	2	08:50-10:00	75I-P2	F.2.2			0	0	0	0	0	0	0	0	E
20201	14110500	Wahyu A	2020/2021	13	IBI20612	Skripsi	6	13:00-14:00	75I-P1	F.3.2   F.3.2											
20201	14110500	Wahyu A	2020/2021	13	TIF20421	Rekayasa	4	13:00-14:00	75I-P1	G.1.6   G.1.6			0	0	0	0	0	0	0	0	E
20201	14110500	Wahyu A	2020/2021	13	TIF20212	Kecerdas	2	07:10-08:00	37I-P2	F.2.2			0	0	0	0	0	0	0	0	E
20201	14110500	Wahyu A	2020/2021	13	SIF20225	Sistem Ir	2	10:30-12:00	75I-P1	Lab.Das E (G.2)			0	75	0	85	0	78	85	79,9	A
20201	14110500	Heru Ard	2020/2021	13	SIF20225	Sistem Ir	2	10:30-12:00	75I-P1	Lab.Das.E (G.2)			0	80	0	80	0	85	75	81	A
20201	14110500	Heru Ard	2020/2021	13	TIF20212	Kecerdas	2	16:30-18:00	137I-S1	F.2.2			0	0	0	0	0	0	43,75	4,38	E
20201	14110500	Heru Ard	2020/2021	13	SIF20420	Strategi	4	16:30-18:00	155I-S1	F.2.5   F.2.5			60	0	65	0	70	0	75	66	B
20201	14110500	Heru Ard	2020/2021	13	SIF20403	Pengant	4	07:10-08:00	380G-P1	D.3.3   D.3.3			70	0	70	0	85	0	100	77,5	A-
20201	14110500	Heru Ard	2020/2021	13	SIF20226	Audit Sis	2	14:40-16:00	75I-P1	F.3.5			60	0	60	0	80	0	97	69,7	B+
20201	14110500	Heru Ard	2020/2021	13	SIF20222	E-Busine	2	13:00-14:00	35I-P1	Lab.Lanjut D (G.4)			0	70	0	70	0	70	75	70,5	B+
20201	14110500	Heru Ard	2020/2021	13	SIF20419	Data Wa	4	16:30-18:00	155I-S1	F.3.3   F.3.3			0	73	0	78	0	78	100	78,7	A-
20201	14110500	Heru Ard	2020/2021	13	IBI20207	Pengemt	2	18:30-20:00	75I-S1	E.3.1			0	82	0	70	0	82	80	78,2	A-
20201	14110500	Rio Darm	2020/2021	13	SIF20419	Data Wa	4	13:00-14:00	155I-P1	D.3.6   D.3.6			0	74	0	78	0	78	100	79	A-
20201	14110500	Rio Darm	2020/2021	13	TIF20421	Rekayasa	4	16:30-18:00	175I-S1	G.1.7   G.1.7			70	0	70	0	75	0	85	73	B+
20201	14110500	Rio Darm	2020/2021	13	IBI20612	Skripsi	6	18:30-20:00	175I-S1	F.2.6   F.2.6											
20201	14110500	Rio Darm	2020/2021	13	SIF20420	Strategi	4	16:30-18:00	155I-S1	F.2.5   F.2.5			80	0	65	0	85	0	80	77	A-
20201	14110500	Rio Darm	2020/2021	13	SIF20421	Data Mir	4	13:00-14:00	155I-P1	F.2.2   F.2.2			78	0	80	0	80	0	90	80,4	A
20201	14110500	Pramudy	2020/2021	13	IBI20612	Skripsi	6	18:30-20:00	175I-S1	F.2.6   F.2.6											
20201	14110501	Rendi De	2020/2021	13	SIF20419	Data Wa	4	10:30-12:00	155I-P2	F.4.4   F.4.4											
20201	14110501	Rendi De	2020/2021	13	SIF20423	Pemrogr	4	07:10-08:00	155I-P1	E.3.1   Lab.Lanjut D (G.4)											

Data tersebut akan dilakukan pengabungan data dari tahun 2015 – 2021 Sebelum melanjutkan ke langkah selanjut nya yaitu processing data dengan memilih atribut yang akan di gunakan untuk perhitungan dalam *K-Means*.

## 1.2. Selection atau Transform.

Pada *selection* dilakukan Penggabungan data mahasiswa dari tahun 2015 - 2021. untuk mengambil atribut yang akan di gunakan dalam perhitungan dengan versi Ms.Exel Untuk memilih berdasarkan Mata kuliah peminatan dan Prodi.

Seperti keterangan sebelumnya bahwa jurusan yang memiliki mata kuliah peminatan terdiri dari tiga prodi saja maka data di pilih berdasarkan data pe jurusan sebanyak 22.439.

Tabel 4.6 Data set Mahasiswa per tiga prodi .

NAMA	NPM	JURUSAN	KDTA	SEMEST	KODEM	NAMAMK	TOTAL	HURUF	Mata Kuliah Peminatan
Unike Reza Windriyani	1,61E+09	Akuntansi	20191	7	AKT19206	Akuntansi Manajemen Desa	83,09	A	Akuntansi Publik
Eva Christiana	1,61E+09	Akuntansi	20191	7	AKT19206	Akuntansi Manajemen Desa	80,87	A	Akuntansi Publik
Rhegita Indah Purnama	1,61E+09	Akuntansi	20191	7	AKT19206	Akuntansi Manajemen Desa	92,5	A	Akuntansi Publik
Fauzia Nurjana	1,61E+09	Akuntansi	20191	7	AKT19206	Akuntansi Manajemen Desa	87,4	A	Akuntansi Publik
Monica Rizki Wulandari	1,61E+09	Akuntansi	20191	7	AKT19206	Akuntansi Manajemen Desa	80,57	A	Akuntansi Publik
Sugesti Nabela	1,61E+09	Akuntansi	20191	7	AKT19206	Akuntansi Manajemen Desa	84,47	A	Akuntansi Publik
Tiger Tri Sakti	1,61E+09	Akuntansi	20191	7	AKT19206	Akuntansi Manajemen Desa	85,9	A	Akuntansi Publik
Feby Syaras Sambudi	1,61E+09	Akuntansi	20191	7	AKT19206	Akuntansi Manajemen Desa	89,46	A	Akuntansi Publik
TRISKA MARSTELLA MA	191212801	Akuntansi	20201	3	AKT20206	Akuntansi Manajemen Desa	78,38	A-	Akuntansi Publik
FADILA AYU PUTRI	191212802	Akuntansi	20201	3	AKT20206	Akuntansi Manajemen Desa	78,68	A-	Akuntansi Publik
Virelly Aidhil Sachputra	191212802	Akuntansi	20201	3	AKT20206	Akuntansi Manajemen Desa	77,5	A-	Akuntansi Publik
Elisabeth Venny Hartan	191212802	Akuntansi	20201	3	AKT20206	Akuntansi Manajemen Desa	76,6	A-	Akuntansi Publik
Niluh Okta Raswita	191212803	Akuntansi	20201	3	AKT20206	Akuntansi Manajemen Desa	78,1	A-	Akuntansi Publik
Vidya Trianita	191212803	Akuntansi	20201	3	AKT20206	Akuntansi Manajemen Desa	79	A-	Akuntansi Publik
Agus Rahman	1,61E+09	Akuntansi	20191	7	AKT19206	Akuntansi Manajemen Desa	77,87	B	Akuntansi Publik
Gusti Ayu Putu Messa L	1,61E+09	Akuntansi	20191	7	AKT19206	Akuntansi Manajemen Desa	78,89	B	Akuntansi Publik
Nabila Pebri Dwi Putri K	1,61E+09	Akuntansi	20191	7	AKT19206	Akuntansi Manajemen Desa	78,66	B	Akuntansi Publik
Defi Rentama	1,61E+09	Akuntansi	20191	7	AKT19206	Akuntansi Manajemen Desa	78,7	B	Akuntansi Publik
Meisy Ramadani	1,61E+09	Akuntansi	20191	7	AKT19206	Akuntansi Manajemen Desa	75,17	B	Akuntansi Publik
Bresly Harianja	1,61E+09	Akuntansi	20191	7	AKT19206	Akuntansi Manajemen Desa	74,16	B	Akuntansi Publik
LARAS HATI	181212011	Akuntansi	20201	5	AKT20206	Akuntansi Manajemen Desa	70,58	B+	Akuntansi Publik
ALIA NUR AMINI	181212011	Akuntansi	20201	5	AKT20206	Akuntansi Manajemen Desa	69,35	B+	Akuntansi Publik

### 4.2.1 Sample Data.

Setelah melakukan proses pengambilan data perprodi Diperoleh data sebanyak 22.439 Record.

Pada tahap *clustering* yang memakai metode *K-Means* diperlukan penentuan *sampling data*. Dalam menentukan sampling data dipilih secara *Proportionate Stratified Random* dari objek data yang dipakai. Hasil pemilihan pengambilan sampling data dengan menggunakan *metode Slovin* dengan perhitungan sebagai berikut:

$$n = \frac{N}{1 + Ne^2}$$

$$n = \frac{22.439}{1 + 22.439(0.001)^2}$$

$$n = \frac{22.439}{1 + (22439 \times 0.0001)}$$

$$n = \frac{22.439}{1 + 2.2439}$$

$$n = \frac{22.439}{3.2439} = 6.917 \text{ record}$$

Setelah di dapatkan hasil dari perhitungan sample dari tiga prodi per mata kuliah konsentrasi sebagai berikut :

#### Peminatan Konsentrasi Jurusan SI

- BI (611 / 22.439) x 6917 = 188 sampel
- IS (2046 / 22.439) x 6917 = 630 sampel

#### Peminatan Konsentrasi Jurusan MA

- Manajemen Hospitaly (68 / 22.439) x 6917 = 9 sampel
- Manajemen Keuangan (7003/22.439) x 6917 = 2156 sampel
- Manajemen Operasi (78/22.439) x 6917 = 24 sampel
- Manajemen Pemasaran (4296/22.439) x 6917 = 1.324 sampel
- Manajemen SDM (5304 / 22.439) x 6917 = 1.635 sampel

#### Peminatan Konsentrasi Jurusan AK

- Akuntansi Keuangan (608 / 22.439) x 6917 = 187 sampel
- Akuntansi Manajemen (2139 / 22.439) x 6917 = 660 sampel
- Akuntansi Publik (168 / 22.439) x 6917 = 52 sampel

- Akuntansi Syariah  $(42 / 22.439) \times 6917 = 13$  sampel
- Sistem Informasi Akuntansi  $(116 / 22.439) \times 6917 = 36$  sampel

Total populasi = 22.439

Sample = 6917

Jumlah sample mata kuliah konsentrasi yang memiliki data terbanyak adalah (Informasi Sistem, Manajemen Pemasaran, Manajemen SDM, Manajemen Keuangan, Akuntansi Manajemen) sehingga data lima mata kuliah konsentrasi tersebut saja yang mewakili data untuk di kurangi, dengan perhitungan sebagai berikut:

$$\begin{aligned}
 \text{Rumus} &= \text{Sample} - 5000 \\
 &= 6917 - 5000 \\
 &= 1917 / 5 \text{ (Sample MK data terbanyak)} \\
 &= 384
 \end{aligned}$$

Table hasil perhitungan semua mata kuliah konsentrasi.

No	Nama Mata Kuliah	Total Sample	Total yang Di Olah
1	Informasi Sistem	630	247
2	Manajemen Pemasaran	1324	940
3	Manajemen SDM	1635	1251
4	Akuntansi Manajemen	660	277
5	Manajemen keuangan	2156	1776
6	Busines Intelegence	188	188
7	Manajemen Hospitality	9	9
8	Manajemen Operasi	24	24
9	Akuntansi Keuangan	187	187
10	Akuntansi Publik	52	52
11	Akuntansi Syariah	13	13
12	Sistem Informasi Akuntansi	36	36

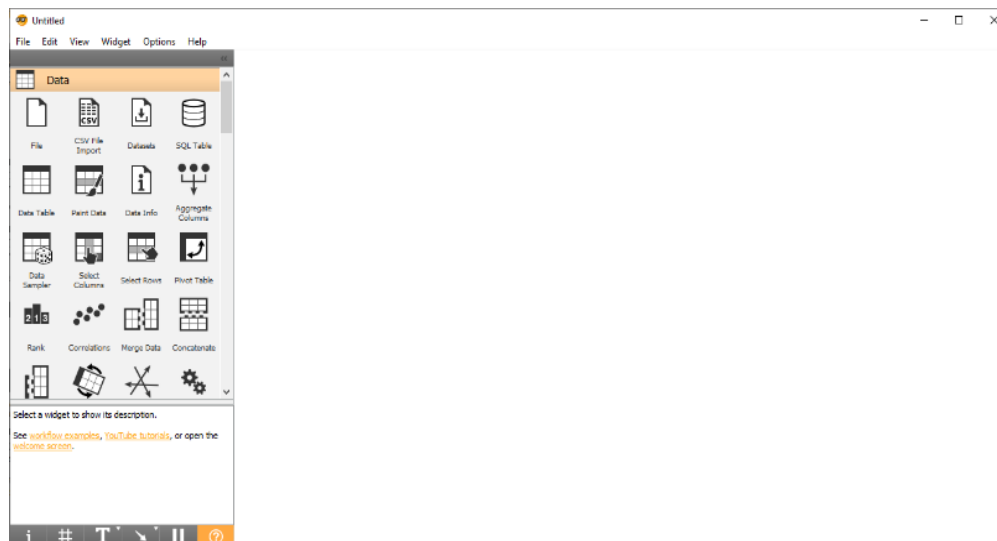
		<b>Total</b>	<b>5000</b>
--	--	--------------	-------------

NAMA	NPM	JURUSA	KOTA	SEMEST	KODEM	NAMAMK	TOTAL	HURUF	Mata Kuliah Peminatan
CIKITA HA	181211033	Manajema	20201	5	MAN2041	Manajemen Fasilitas	80,6	A	Manajemen Hospitality
MALVIN S	181211052	Manajema	20201	5	MAN2041	Manajemen Fasilitas	79,7	A	Manajemen Hospitality
ELISA NUR	1,81E+09	Manajema	20192	4	MAN1944	Teknologi Manajem	80,2	A	Manajemen Hospitality
HANDY Cl	1,81E+09	Manajema	20192	4	MAN1944	Teknologi Manajem	79,6	A	Manajemen Hospitality
ADINDA C	181211018	Manajema	20201	5	MAN2045	Perilaku Organisasi	75,1	A-	Manajemen Hospitality
ELISA NUR	181211025	Manajema	20201	5	MAN2045	Perilaku Organisasi	76	A-	Manajemen Hospitality
MICHAEL	181211006	Manajema	20201	5	MAN2045	Perilaku Organisasi	71,7	B+	Manajemen Hospitality
KURNIA A	181211031	Manajema	20201	5	MAN2045	Perilaku Organisasi	74,1	B+	Manajemen Hospitality
MICHAEL	181211006	Manajema	20201	5	MAN2041	Manajemen Fasilitas	56,6	C	Manajemen Hospitality
Ani Erlina	181212800	Manajema	20182	2	AKT15202	Akuntansi Internasi	80,5	A	Manajemen Keuangan
Siska Putr	1,61E+09	Manajema	20182	6	AKT15202	Akuntansi Internasi	80,5	A	Manajemen Keuangan
Aprilia Kri	1,61E+09	Manajema	20182	6	AKT15202	Akuntansi Internasi	80,5	A	Manajemen Keuangan
Indri Dwi	1,61E+09	Manajema	20182	6	AKT15202	Akuntansi Internasi	85	A	Manajemen Keuangan
Robita Raf	1,61E+09	Manajema	20182	6	AKT15202	Akuntansi Internasi	82	A	Manajemen Keuangan
Rudi Lesta	1,61E+09	Manajema	20182	6	AKT15202	Akuntansi Internasi	85	A	Manajemen Keuangan
Ardika Tio	1,31E+09	Manajema	20162	8	AKT15202	Akuntansi Internasi	71,5	B	Manajemen Keuangan
Muhamm	1,31E+09	Manajema	20162	8	AKT15202	Akuntansi Internasi	71,5	B	Manajemen Keuangan
Ricardo	1,31E+09	Manajema	20162	8	AKT15202	Akuntansi Internasi	70,88	B	Manajemen Keuangan
Putri Men	1,41E+09	Manajema	20162	6	AKT15202	Akuntansi Internasi	77,5	B	Manajemen Keuangan
Lica Septik	1,51E+09	Manajema	20162	4	AKT15202	Akuntansi Internasi	70,88	B	Manajemen Keuangan
Lorensia L	1,41E+09	Manajema	20162	6	AKT15202	Akuntansi Internasi	74,5	B	Manajemen Keuangan
Dian Yura	1,41E+09	Manajema	20162	6	AKT15202	Akuntansi Internasi	75,38	B	Manajemen Keuangan

Tabel 4.5. Data setelah melakukan proses slovin.

### 4.3 K-Means

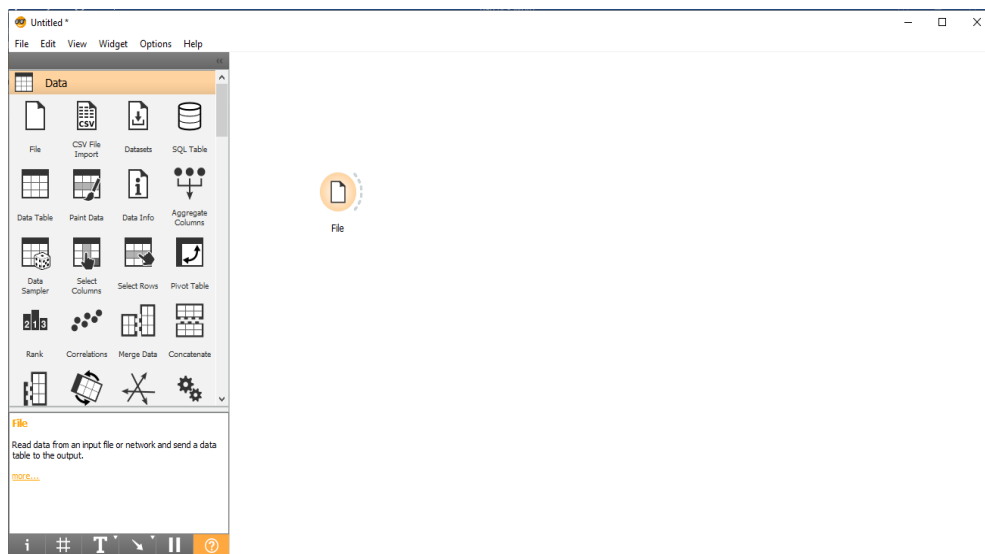
Setelah melakukan processing data kemudian dilakukan proses *K-Means* dengan menggunakan *Aplikasi Orange*. Pada halaman utama *Aplikasi Orange* terdapat beberapa



widget dan lembar kerja di dalam nya, yang akan membantu dalam mengolah data mahasiswa TA.2015-2021 Pada Gambar 4.2.1 berikut ini.

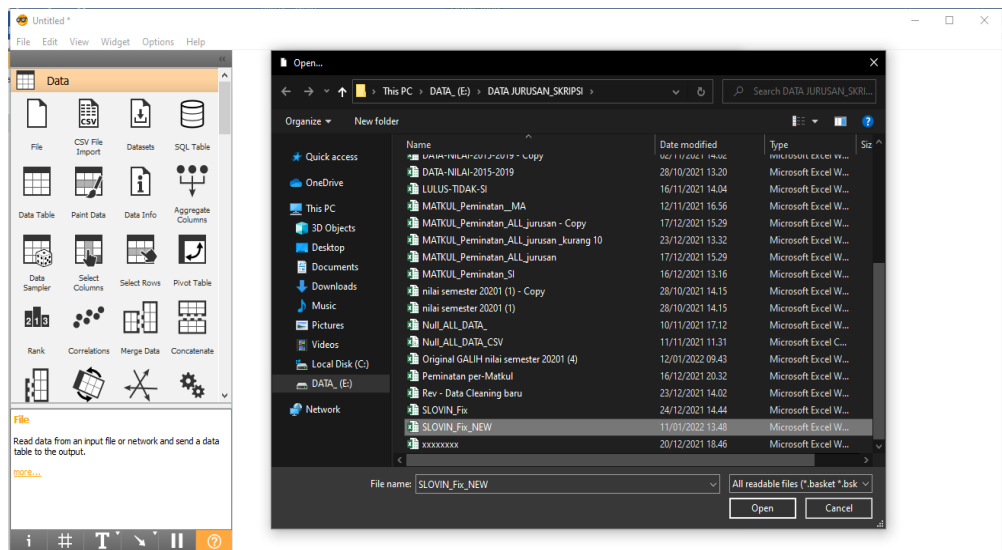
Gambar 4.2.1. Lembar halaman Aplikasi Orange

Untuk memasuk kan data digunakan widget file seperti pada gambar berikut:



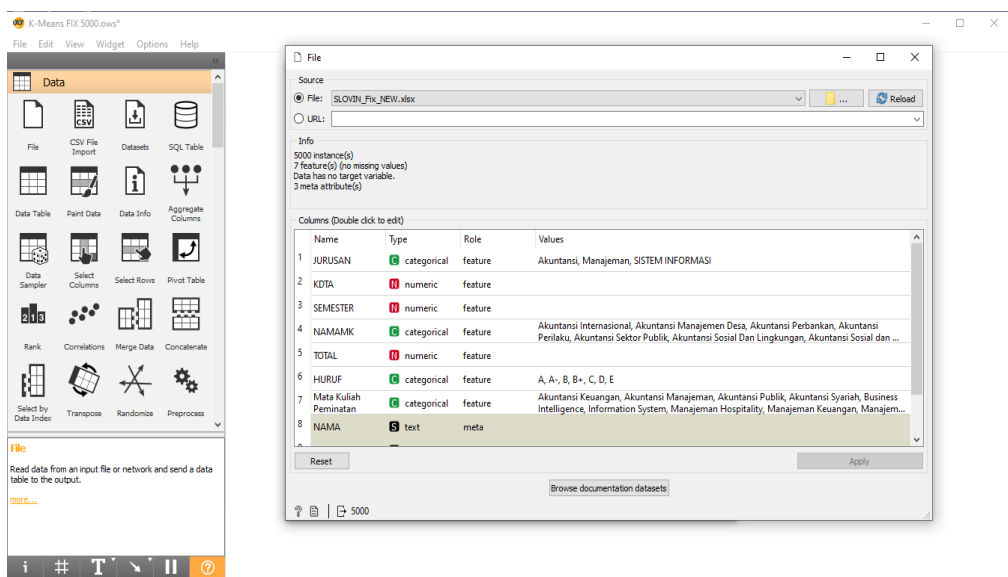
Gambar 4.2.2. Widget file Aplikasi Orange

Widget File untuk membaca file data mahasiswa TA.2015-2021 dan mengirimkan kumpulan data ke saluran outputnya. Widget ini juga menyertakan direktori dengan kumpulan data sampel mahasiswa yang sudah diinstal sebelumnya dengan Orange. Widget membaca data mahasiswa 2015-2021 pada komputer dengan versi file excel seperti pada gambar 4.2.3 sebagai berikut:

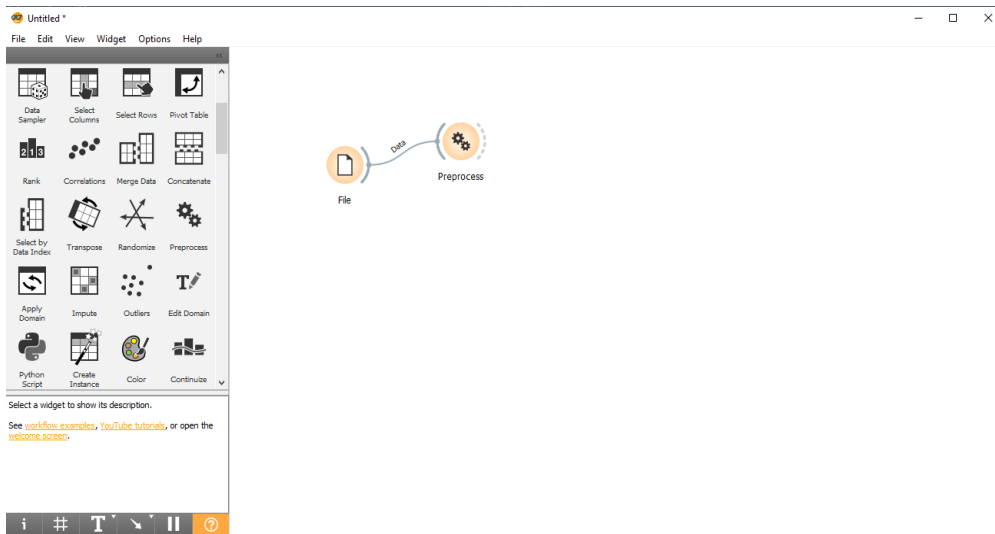


Gambar.4.2.3 Lembar Halaman masukan File.

Data yang di gunakan yaitu data mahasiswa dengan Pilih kolom file data mahasiswa yang sudah di kelompokkan kemudian tarik file ke dalam lembar kerja, atau bisa juga dengan cara lain yaitu klik kiri pada mouse kemudian pilih file. Kemudian Klik Data Table lalu hubungkan dengan file dengan tampilan sebagai berikut:



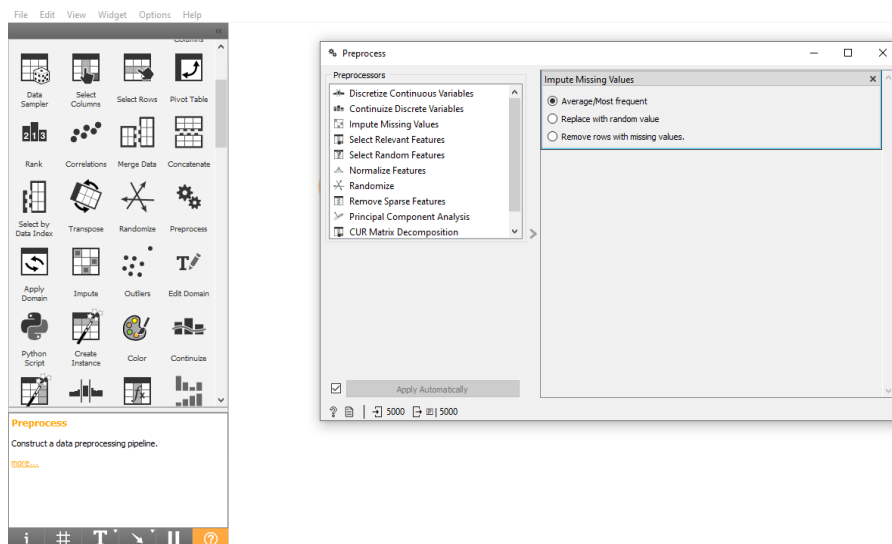
Gambar.4.2.4 Aplikasi Orange Telusuri file data.



Gambar.4.2.4

Lembar Halaman metode prapemrosesan.

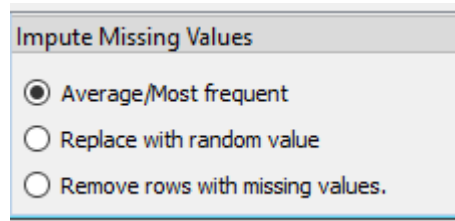
Data mahasiswa akan diproses sebelumnya yaitu pra-pemrosesan untuk mencapai hasil analisis yang lebih baik.



Gambar.4.2.5 Widget Preprocess



Dapat di lakukan Dengan klik widget Preprocess lalu pilih Menghilangkan Value yang hilang, Berdasarkan data yang ada dilakukan proses untuk Impute missing values kemudian akan muncul tiga pemberitahuan .

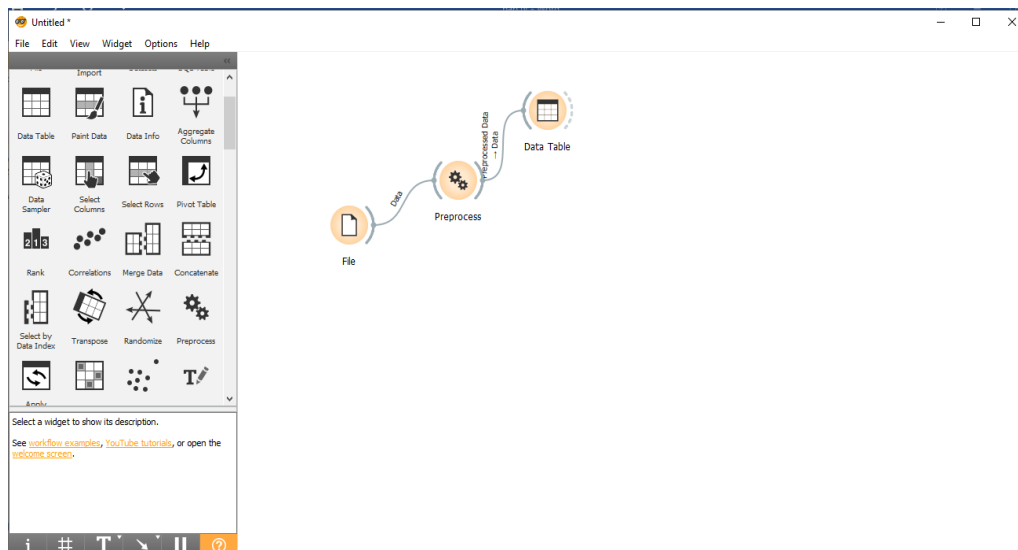


Gambar Impute Missing Values

Keterangan :

1. Avarage/Most frequent : Rata-Rata / Paling sering
2. Replace with random value : Mengganti nilai value yg lain
3. Remove rows with missing values : Menghapus baris dengan menghilangkan nilai value

Kemudian memilih avarage / most frequent (Rata-Rata/Paling sering), dan kemudian data akan di proses oleh *orange*.



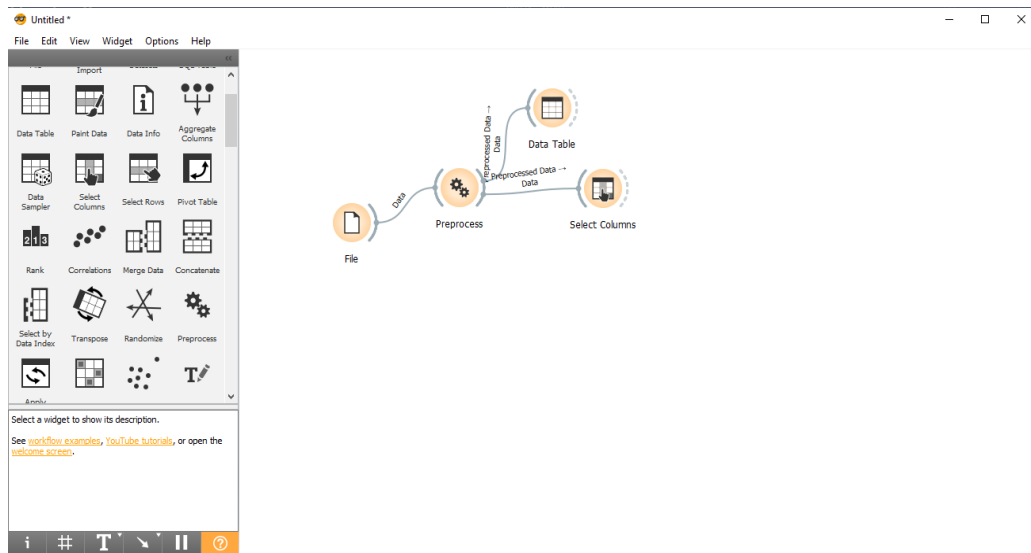
Gambar.4.2.6 Widget data table.

Untuk memastikan data yang di masukkan sudah tidak ada data value, dengan menampilkan data table kemudian bisa doble klik, maka akan terlihat apakah data value mash ada atau tidak dalam data table, Setelah melakukan proses tersebut maka akan menghasilkan data sebagai berikut:

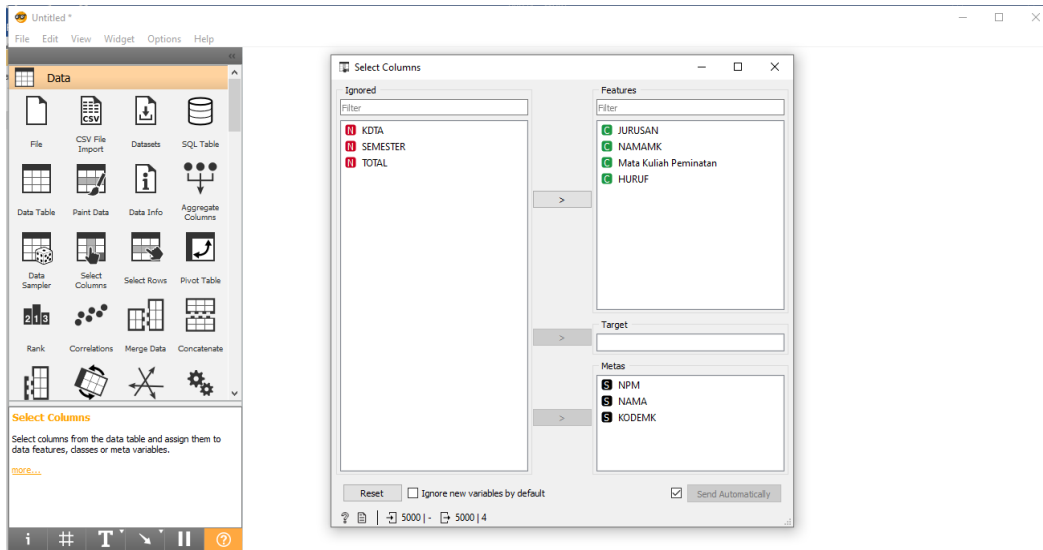
	NAMA	NPM	KODEMK	JURISAN	KDTA	SEMESTER	NAMAMK	TOTAL	HURUF	ata Kuliah Peminata
1	CIKITA HANDA...	1812110331	MAN20417	Manajemen	20201	5	Manajemen Fas...	80.6	A	Manajemen Ho...
2	MAUVIN SUHA...	1812110529	MAN20417	Manajemen	20201	5	Manajemen Fas...	79.7	A	Manajemen Ho...
3	ELISA NURIYANA	1812110299	MAN19449	Manajemen	20192	4	Teknologi Man...	80.2	A	Manajemen Ho...
4	HANDY CIAYADI	1812110011	MAN19449	Manajemen	20192	4	Teknologi Man...	79.6	A	Manajemen Ho...
5	ADINDA OKTAVIA	1812110189	MAN20454	Manajemen	20201	5	Penilaku Organi...	75.1	A-	Manajemen Ho...
6	ELISA NURIYANA	1812110299	MAN20454	Manajemen	20201	5	Penilaku Organi...	76	A-	Manajemen Ho...
7	MICHAEL WEN...	1812110066	MAN20454	Manajemen	20201	5	Penilaku Organi...	71.7	B+	Manajemen Ho...
8	KURNIA APRELL...	1812110313	MAN20454	Manajemen	20201	5	Penilaku Organi...	74.1	B+	Manajemen Ho...
9	MICHAEL WEN...	1812110066	MAN20417	Manajemen	20201	5	Manajemen Fas...	56.6	C	Manajemen Ho...
10	Ani Erlina Siregar	1812128002p	AKT15202	Manajemen	20182	2	Akuntansi Inter...	80.5	A	Manajemen Ke...
11	Siska Putri Meir...	1612120153	AKT15202	Manajemen	20182	6	Akuntansi Inter...	80.5	A	Manajemen Ke...
12	Aprilia Kristina	1612120116	AKT15202	Manajemen	20182	6	Akuntansi Inter...	80.5	A	Manajemen Ke...
13	Indri Dwi Putri	1612120101	AKT15202	Manajemen	20182	6	Akuntansi Inter...	85	A	Manajemen Ke...
14	Robita Rahman	1612120060	AKT15202	Manajemen	20182	6	Akuntansi Inter...	82	A	Manajemen Ke...
15	Rudi Lestari	1612120217	AKT15202	Manajemen	20182	6	Akuntansi Inter...	85	A	Manajemen Ke...
16	Ardika Tio Septi...	1312120198	AKT15202	Manajemen	20162	8	Akuntansi Inter...	71.5	B	Manajemen Ke...
17	Muhammad Bu...	1312120133	AKT15202	Manajemen	20162	8	Akuntansi Inter...	71.5	B	Manajemen Ke...
18	Ricardo	1312120124	AKT15202	Manajemen	20162	8	Akuntansi Inter...	70.88	B	Manajemen Ke...
19	Putri Meydary...	1412120045	AKT15202	Manajemen	20162	6	Akuntansi Inter...	77.5	B	Manajemen Ke...
20	Lica Septiary R...	1512120046	AKT15202	Manajemen	20162	4	Akuntansi Inter...	70.88	B	Manajemen Ke...
21	Lorensia Lilik P...	1412120124	AKT15202	Manajemen	20162	6	Akuntansi Inter...	74.5	B	Manajemen Ke...
22	Dian Yura Triana	1412120176	AKT15202	Manajemen	20162	6	Akuntansi Inter...	75.38	B	Manajemen Ke...
23	Misyra Ananda	1512120229	AKT15202	Manajemen	20162	4	Akuntansi Inter...	74.5	B	Manajemen Ke...
24	Hafidz Ar - Razi	1512129007	AKT15202	Manajemen	20162	4	Akuntansi Inter...	76	B	Manajemen Ke...
25	Danny Ferdians...	1512128001	AKT15202	Manajemen	20162	4	Akuntansi Inter...	70.88	B	Manajemen Ke...
26	Monalisa Gultorn	1612129010P	AKT15202	Manajemen	20162	2	Akuntansi Inter...	77.5	B	Manajemen Ke...
27	Liya Pitriyanti	1512120048	AKT15202	Manajemen	20172	6	Akuntansi Inter...	70.88	B	Manajemen Ke...
28	Susanti	1512120041	AKT15202	Manajemen	20172	6	Akuntansi Inter...	70	B	Manajemen Ke...
29	Devi Lili Marcell...	1512120124	AKT15202	Manajemen	20172	6	Akuntansi Inter...	70	B	Manajemen Ke...

Gambar.4.2.7 Lembar Halaman Data Table

Kemudian di perlukan proses pemilihan atribut data dan komposisi domain data, oleh karena itu dapat di gunakan widget select collums.

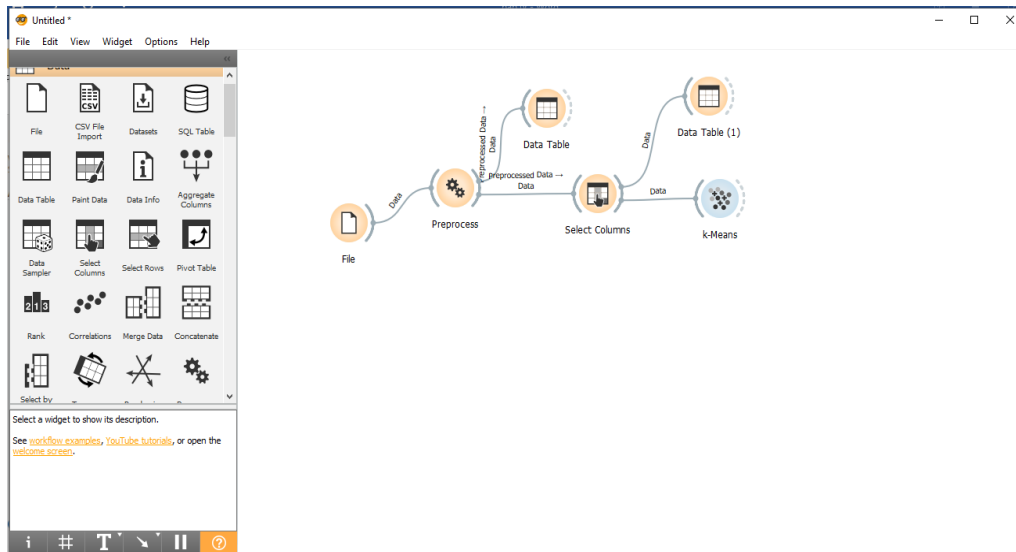


Gambar.4.2.8 Lembar Halaman Select Columns



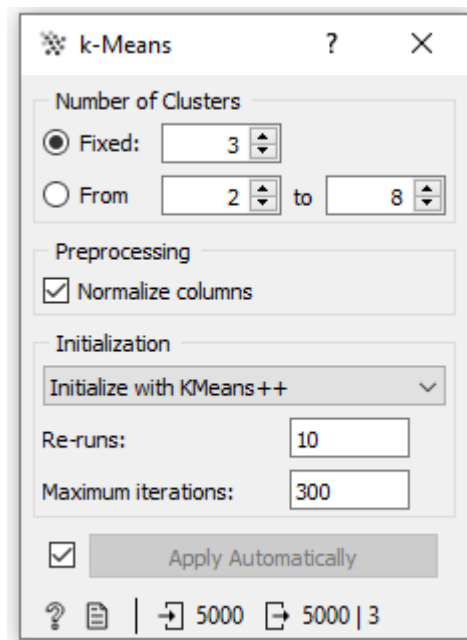
Gambar.4.2.9 Isi Lembar Halaman Select Collums.

Di orange akan memisahkan sendiri atribut yang tidak berpengaruh terhadap analisis, Oranye membedakan antara *Ignored* yaitu data yang tidak di ambil atau di gunakan di antaranya: KDTA, Semester, Total, *Fitures* menampilkan field yang ada pada data mahasiswa di antaranya: jurusan, namamk, mata kuliah peminatan, huruf dan *Metas* Atribut meta dari file data baru. Atribut-atribut ini termasuk dalam kumpulan data tetapi, untuk sebagian besar metode, tidak dipertimbangkan dalam analisis diantara nya : Nama, NPM, Kodemk. dapat menggunakannya sebagai label instan dan akan ada pada komponen cluster.



Gambar.4.2.7 Lembar Halaman *K-Means*

Pertama, kita memuat Widget k-Means dan kemudian double klik pada widget *K-Means* maka akan tampil sebagai berikut :



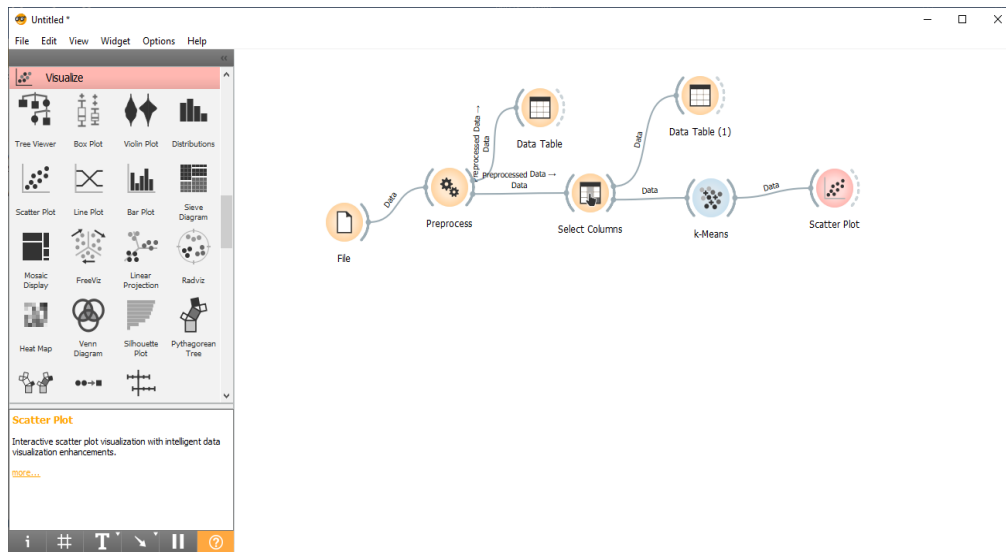
Penjelasan dari tiap fitur pada gambar sebagai berikut:

*Number of clusters*: jumlah cluster, *Fixed*: mengelompokkan data ke sejumlah kluster tertentu, *Prapemrosesan*: Jika opsi dipilih, kolom dinormalisasi, *Metode inisialisasi*: cara algoritme memulai pengelompokan, *k-Means++* : pusat pertama dipilih secara acak, selanjutnya dipilih

dari titik-titik yang tersisa dengan probabilitas sebanding dengan kuadrat jarak dari pusat terdekat, *Re-runs*: berapa kali algoritme dijalankan dari posisi awal acak; hasil dengan jumlah kuadrat terkecil dalam cluster. *Iterasi maksimum*: jumlah maksimum iterasi dalam setiap algoritma yang dijalankan.

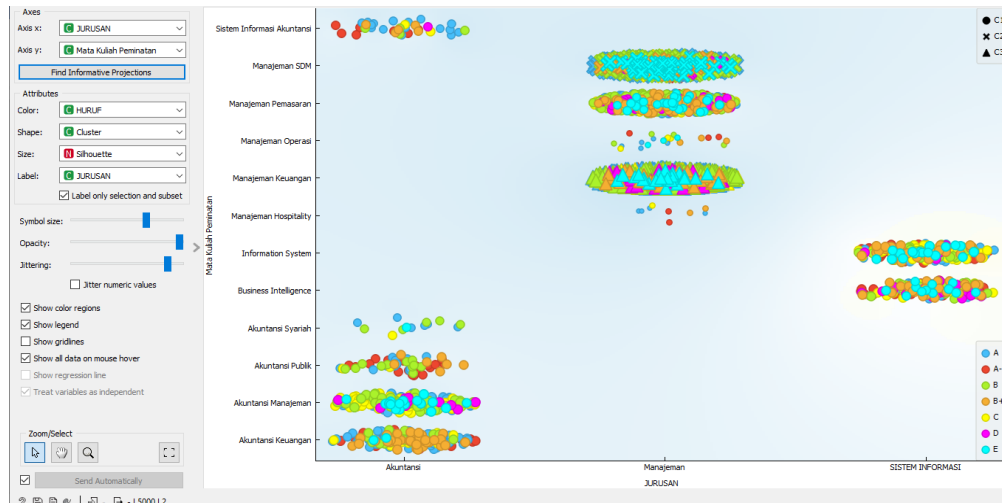
#### 4.5 Load

Setelah di setup pada *widget K-Means*, dapat dilakukan proses untuk menampilkan visualisasi data menggunakan *widget scatter plot*.



Gambar.4.2.10 Lembar Halaman Scatter plot

Keterangan gambar Scater plot dari gambar di bawah ini, dapat merekomendasikan ke prodi banyak mahasiswa yang mengambil mata kuliah konsentrasi / peminatan dari total populasi sample data setiap prodi. Dengan cara lakukan doble klik pada widget scatter plot akan muncul tampilan sebagai berikut :



Gambar.4.2.11 Lembar Halaman Scatter plot Clustering.

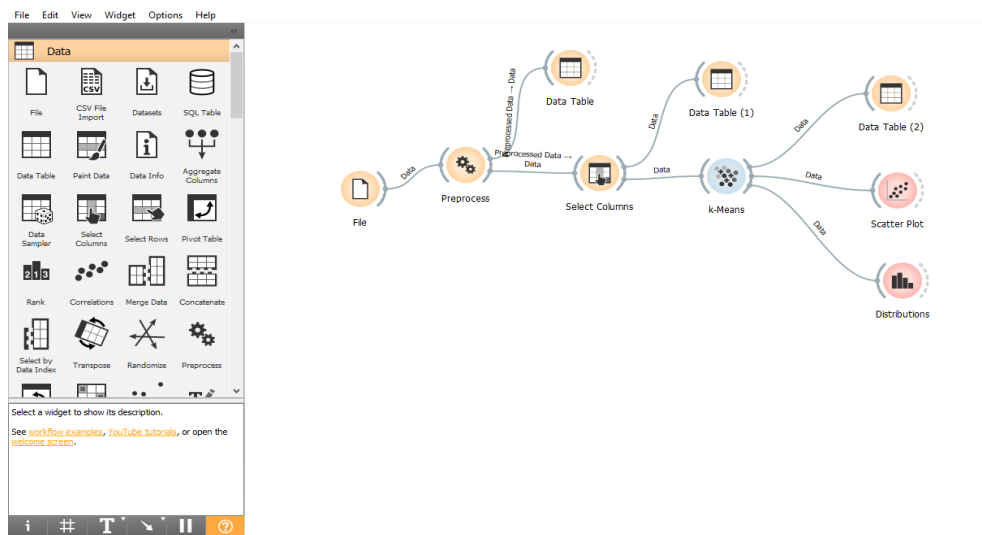
Langkah selanjutnya dengan Visualisasi Scatter plot akan ada bentuk visualisasi yang di tampilkan di antaranya terdiri dari nilai mahasiswa sumbu  $x$  Prodi dan sumbu  $y$  mata kuliah peminatan dengan analisis eksplorasi dan peningkatan visualisasi data cerdas. Inputs (masukan Data) masukan kumpulan data, Subset Data: subset dari instance, Fitur: daftar atribut Keluaran. Data yang Dipilih dari Scatter plot, Data dengan kolom tambahan yang menunjukkan apakah suatu titik dipilih.

Widget Scatter Plot menyediakan visualisasi scatter plot 2 dimensi. Data ditampilkan sebagai kumpulan titik yang masing-masing memiliki nilai atribut sumbu  $x$  yang menentukan posisi pada sumbu horizontal dimana sumbu  $x$  sebagai prodi dan nilai atribut sumbu  $y$  yang menentukan posisi pada sumbu vertikal sebagai sumbu mata kuliah peminatan. Berbagai properti grafik, seperti warna, ukuran dan bentuk titik, judul sumbu, ukuran titik maksimum, dan jittering dapat disesuaikan di sisi kiri widget. berbagai macam warna sebagai tampilan nya antara lain :

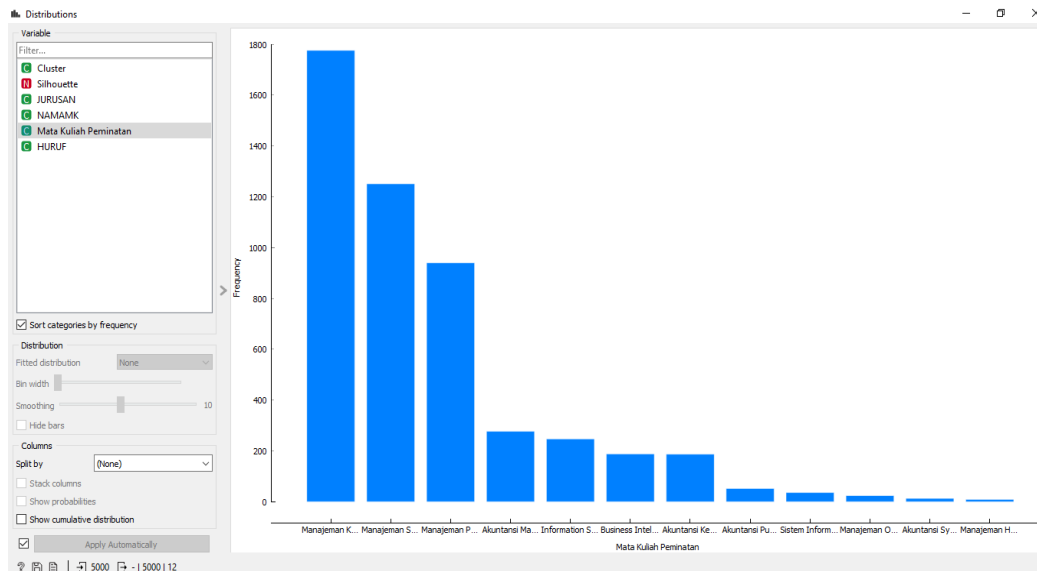
1. Biru dengan Nilai A,
2. Merah dengan Nilai A-,
3. Hijau dengan nilai B,
4. Orange dengan nilai B+,
5. Kuning dengan nilai C,
6. Ungu dengan nilai D,
7. dan Biru muda dengan nilai E.

Dan berikut lambang-lambang pada cluster menunjukkan C1, C2, C3 diantaranya :

1. C1 menunjukkan mahasiswa yang paling sedikit mengambil mata kuliah peminatan, tanda bulat pada cluster.
2. C2 menunjukkan mahasiswa dengan jumlah sedang pada mata kuliah peminatan, dan lambang silang pada cluster.
3. C3 menunjukkan mahasiswa yang paling banyak mengambil mata kuliah peminatan dengan lambang segitiga pada cluster.



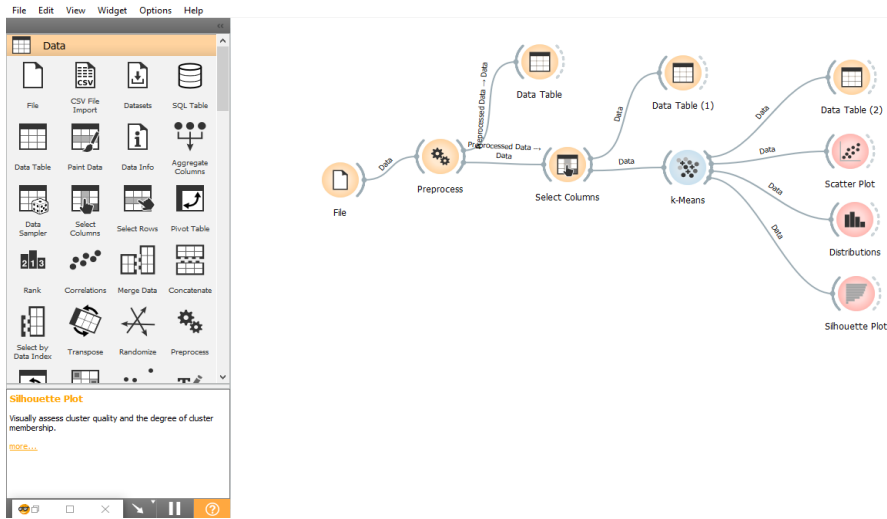
Gambar.4.2.16 Lembar Halaman Distribution



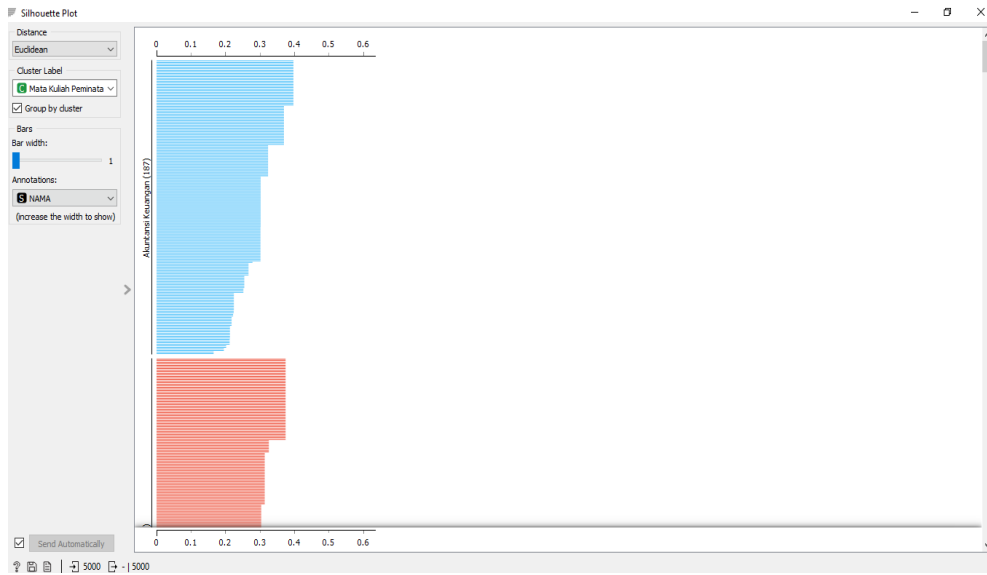
Gambar.4.2.17 Lembar Halaman Distribution

Langkah berikutnya dengan Visualisasi Distribution akan ada bentuk Grafik yang di tampilkan di antaranya terdiri dari daftar mata kuliah peminatan konsentrasi dimana grafik tertinggi sampai terendah mata kuliah Manajemen keuangan, Manajemen SDM, Manajemen pemasaran, Akuntansi Manajemen, Informasi Manajemen, Busines Intellegence, Akuntansi Keuangan, Akuntansi Publik, Sistem Informasi Akuntansi, Manajemen Operasi, Akuntansi Syariah, Manajemen Hospitality.





Gambar.4.2.19 Lembar Halaman Silhouette Plot



Gambar.4.2.20 Lembar Halaman Silhouette Plot

Widget Silhouette Plot menawarkan representasi grafis dari konsistensi dalam kluster data dan memberi pengguna sarana untuk menilai kualitas kluster secara visual. Skor siluet adalah ukuran

seberapa mirip suatu objek dengan clusternya sendiri dibandingkan dengan cluster lain dan sangat penting dalam pembuatan plot siluet. Skor siluet mendekati 1 menunjukkan bahwa instance data dekat dengan pusat cluster dan instance yang memiliki skor siluet mendekati 0 berada di perbatasan antara dua cluster. Dimana Akuntansi Keuangan 0,4 Akuntansi Manajemen 0,39 Akuntansi Publik 0,45 Akuntansi Syariah 0,62 Busines Intelegence 0,3 Informasi Sistem 0,4 Manajemen Hospitality 0,32 Manajeman Keuangan 0,3 Manajeman Operasi 0,4 Manajeman Pemasaran 0,3 Manajeman SDM 0,31 Sistem Informasi Akuntansi 0,3.