

## **BAB III**

### **METODOLOGI PENELITIAN**

Bab ini menjelaskan tentang prosedur yang digunakan untuk mengumpulkan, mengelola dan menganalisa data guna menunjang penelitian. Tahapan yang akan dilakukan dalam penelitian ini sebagai berikut :

#### **3.1. Analisis Pengumpulan Data**

##### **3.1.1. Waktu dan Tempat**

Penelitian ini dilaksanakan di Kantor Kecamatan Tingkir yang beralamat di Jalan Marditomo No. 37 Salatiga. Penelitian dilaksanakan mulai bulan September sampai dengan bulan Desember 2019.

##### **3.1.2. Objek Penelitian**

Penelitian dilakukan dengan mengambil objek penelitian data presensi Pegawai Negeri Sipil (PNS) Kecamatan Tingkir tahun 2018.

##### **3.1.3. Metode Pengumpulan Data**

Tahapan ini bertujuan untuk mendapatkan informasi dari sebuah data untuk keperluan penelitian. Peneliti menggunakan teknik untuk mengumpulkan data antara lain *survey*, *study literature*, dan *research and site visit*.

###### **3.1.3.1 Survey**

Pengumpulan data dilakukan dengan wawancara dengan Ibu Nunuk Dartini, S.Pd., M.Si selaku Camat Tingkir dan bapak Wahono selaku staf Sub Bagian Umum dan Kepegawaian yang merupakan pihak yang bertanggung jawab terhadap masalah yang sedang diteliti. Bagian yang akan dilakukan *survey* yaitu tentang kedisiplinan PNS Kecamatan Tingkir. Hasil yang sudah didapatkan dari *survey* tersebut nantinya akan diolah untuk *clustering* tingkat kedisiplinan pegawai.

###### **3.1.3.2 Study Literature**

Pengumpulan data dilakukan guna menunjang pembuatan laporan skripsi dengan cara mempelajari teori dari buku-buku yang berkaitan dengan masalah yang dibahas. Penelitian dilakukan dengan cara mempelajari karangan ilmiah dan juga buku yang relevan dengan pokok bahasan. Peneliti menggunakan beberapa jurnal

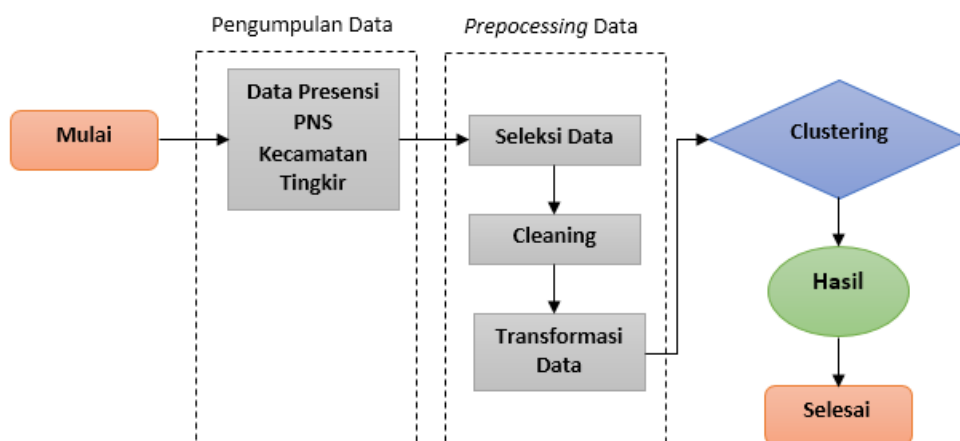
mengenai kasus *clustering* dan *algoritma k-means* guna menunjang pembuatan laporan skripsi.

### 3.1.3.3 Research And Site Visit

Pengumpulan data untuk mendapatkan informasi yang berhubungan dengan penelitian dengan cara mengunjungi situs-situs yang berada di internet. Kunjungan pada situs tersebut merupakan suatu bentuk penelitian yang khusus. Peneliti dapat memperoleh banyak sumber referensi dan informasi yang berkaitan dengan penelitian dengan memanfaatkan teknologi internet ini.

## 3.2. Teknik Analisis Data

Setelah data terkumpul maka langkah selanjutnya melakukan analisis terhadap data tersebut. Tahapan penelitian dapat dilihat pada Gambar 3.1.



Gambar 3.1. Tahapan Penelitian

### 3.2.1. Pengumpulan Data

Tahapan pertama dalam penelitian adalah pengumpulan data presensi masing-masing pegawai Kecamatan Tingkir secara lengkap setiap bulannya. Selain itu juga dilakukan pencarian kriteria-kriteria yang dibutuhkan dalam penelitian masing-masing pegawai seperti jenis kelamin, status perkawinan, jabatan, umur dan masa kerja. Data diperoleh dari admin yang bertugas mengelola mesin presensi yaitu staf Sub Bagian Umum dan Kepegawaian. Data diperoleh dengan cara mengunduh dari aplikasi presensi *online* dan Sistem Informasi Kepegawaian

(SIMPEG) sehingga akhirnya didapatkan data seluruh pegawai yang nanti akan menjadi *dataset* dengan jumlah 912 data.

Pengumpulan data awal bisa dilihat pada Tabel 3.1. Data pegawai berisi rincian lengkap Finger ID, NIP, Nama Staff, Telat [Menit], Telat [kali], Pulang Cepat [menit], Pulang Cepat [kali], Aturan Jam Kerja, Total Jam Kerja, Total Kurang Jam, Total Hari Kerja, Total Hari Alpha, Total Hari Lupa, Total Istirahat, Total Ijin/Cuti/DL/Diklat, Total Hari Libur, Rincian Cuti/DL/Diklat yang berisi Dinas Luar/Penugasan, Cuti Tahunan, Cuti Sakit, Cuti Besar, Cuti Melahirkan, Cuti Alasan Penting, Cuti DTN, Tugas Belajar, Bintek/Diklat, TKL, Bulan, Jenis Kelamin, Status Perkawinan, Jabatan, Umur dan Masa Kerja.

### **3.2.2. Preprocessing Data**

Pada bagian ini data akan dipersiapkan agar lebih mudah dalam proses *mining*. Proses *preprocessing* akan mencakup 3 hal mendasar yaitu seleksi data, *cleaning* dan transformasi data.

#### **3.2.2.1 Seleksi Data**

Seleksi data merupakan suatu pemilihan himpunan data yang nantinya digunakan dalam penulisan, yaitu data Nama, Telat, Pulang\_Cepat, Total\_Hari\_Lupa, Total\_Hari\_Alpha, Total\_Kurang\_Jam\_Menit, Bulan, Jenis\_Kelamin, Status\_Perkawinan, Jabatan, Usia dan Masa\_Kerja.

Tidak semua data yang diperoleh nantinya akan diproses. Beberapa data ada yang dihilangkan tanpa mengurangi nilai dari tingkat kedisiplinan pegawai. Hal ini bertujuan agar hasil yang diperoleh menjadi lebih terfokus kepada tingkat kedisiplinan pegawai dengan tingkat presensi kehadiran baik dan pegawai dengan tingkat presensi kehadiran kurang baik. Sehingga dari data pegawai yang diperoleh beberapa kolom saja yang akan dipakai sebagai atribut dalam *clustering*.

Tabel 3.1. Rekap Presensi Pegawai

		FORM REKAPITULASI PRESENSI																											
		UNIT KERJA : KECAMATAN TINGKIR																											
		PERIODE : 01/01/2018 s.d 31/12/2018																											
Finger ID	NIP	Nama Staff	Telat [Menit]	Telat [leif]	Pulang Cepat [Menit]	Pulang Cepat [leif]	Aturan Jam Kerja	Total Jam Kerja	Total Kurang Jam Kerja	Total Hari Kerja	Total Hari Alpha	Total Hari Lupa	Total Hari Istirahat	Total Cuti/DL/Diklat	Dinas Luar/Penggunaan	Cuti Tahunan	Cuti Sakit	Cuti Melahirkan	Cuti Alasan Penting	Cuti DTN Belajar	Tugas Belajar	Blntsk/Diklat	TKL	BULAN	JENIS KELAMIN	STATUS PERKAWINAN	JABATAN	USIA	MASA KERJA
1038	197403081993111001	SULISTYONO	00:00	0	00:00	0	169:00	148:00	00:00	22	0	0	0	1										JANUARI	L	KAWIN	Struktural	44	25
1038	197403081993111001	SULISTYONO	00:00	0	00:00	0	148:00	148:00	00:00	19	0	0	0	1										FEBRUARI	L	KAWIN	Struktural	44	25
1038	197403081993111001	SULISTYONO	00:00	0	00:00	0	160:30	152:00	08:30	20	0	1	9	0	1									MARET	L	KAWIN	Struktural	44	25
1038	197403081993111001	SULISTYONO	00:00	0	00:00	0	160:30	160:30	00:00	21	0	0	9	0	0									APRIL	L	KAWIN	Struktural	44	25
1038	197403081993111001	SULISTYONO	00:00	0	00:00	0	141:00	141:00	00:00	20	0	0	8	0	3									MEI	L	KAWIN	Struktural	44	25
1038	197403081993111001	SULISTYONO	00:00	0	00:00	0	74:30	74:30	00:00	11	0	0	9	0	10									JUNI	L	KAWIN	Struktural	44	25
1038	197403081993111001	SULISTYONO	00:00	0	00:00	0	148:00	148:00	00:00	19	0	0	9	3	0	3								JULI	L	KAWIN	Struktural	44	25
1038	197403081993111001	SULISTYONO	00:00	0	00:00	0	139:30	139:30	00:00	18	0	0	8	3	2	3								AGUSTUS	L	KAWIN	Struktural	44	25
1038	197403081993111001	SULISTYONO	00:00	0	00:00	0	143:30	143:30	00:00	19	0	0	10	0	1									SEPTEMBER	L	KAWIN	Struktural	44	25
1038	197403081993111001	SULISTYONO	00:00	0	00:00	0	177:30	177:30	00:00	23	0	0	8	0	0									OKTOBER	L	KAWIN	Struktural	44	25
1038	197403081993111001	SULISTYONO	00:00	0	00:00	0	147:30	147:30	00:00	20	0	0	8	1	1									NOVEMBER	L	KAWIN	Struktural	44	25
1038	197403081993111001	SULISTYONO	00:00	0	00:00	0	114:00	114:00	00:00	15	0	0	10	4	2	4								DESEMBER	L	KAWIN	Struktural	44	25
1336	196006271991031003	TRI PRADEKSO	00:04	3	00:00	0	169:00	164:36	04:04	21	0	1	8	0	1									JANUARI	L	KAWIN	Struktural	44	25
1336	196006271991031003	TRI PRADEKSO	00:21	4	00:00	0	148:00	147:39	00:21	19	0	0	8	0	1									FEBRUARI	L	KAWIN	Struktural	44	25
1336	196006271991031003	TRI PRADEKSO	00:57	6	00:00	0	160:30	159:33	00:57	21	0	0	9	0	1									MARET	L	KAWIN	Struktural	44	25

Kolom Telat [Menit] dan Pulang Cepat [Menit] tidak dimasukkan karena data-data tersebut nanti sudah terakumulasi dalam kolom Total Jam Kerja. Kolom Total Jam Kerja merupakan hasil pengurangan dari kolom Aturan Jam Kerja dengan kolom Telat [Menit], Pulang Cepat [Menit], Total Hari Alpha dan Total Hari Lupa. Kolom Aturan Jam Kerja, Total Hari Kerja dan Total Jam Kerja nanti juga tidak akan dimasukkan dalam *dataset* karena data tersebut sudah terakumulasi dalam kolom Total Kurang Jam. Perlu diketahui pada kolom Total Hari Alpha dan Total Hari Lupa jika Alpha/Lupa satu hari pada hari Senin sampai dengan Kamis maka kolom Total Kurang Jam akan ditambahkan 8:30 jam, sedangkan jika Alpha/Lupa satu hari pada hari Jumat maka kolom Total Kurang Jam akan ditambahkan 4 jam sesuai dengan jam kerjanya.

Kolom Cuti (Cuti Tahunan, Cuti Sakit, Cuti Besar, Cuti Melahirkan, Cuti DTN) tidak dimasukkan karena pegawai yang tidak berangkat dengan alasan tersebut memiliki hak untuk tidak berangkat sebagaimana diatur dalam Peraturan Badan Kepegawaian Negara Republik Indonesia Nomor 24 Tahun 2017 Tentang Tata Cara Pemberian Cuti Pegawai Negeri Sipil. Sama halnya dengan kolom Dinas Luar/Penugasan, Tugas Belajar, Bintek/Diklat tidak dimasukkan karena pegawai yang tidak berangkat ke kantor dengan alasan tersebut telah mendapatkan Surat Perintah dari pimpinan, sehingga dianggap hadir dengan alasan yang sah. Hasil kolom yang dipakai menjadi *dataset* bisa dilihat pada Tabel 3.2.

Tabel 3.2. Seleksi Data

	Nama	Telat	Pulang_Cepat	Total_Hari_Lupa	Total_Hari_Alpha	Total_Kurang_Jam_Menit	Bulan	Jenis_Kelamin	Status_Perkawinan	Jabatan	Usia	Masa_Kerja
5												
6	SULISTIYONO	0	0	0	0	00:00	JANUARI	L	KAWIN	Struktural	44	25
7	SULISTIYONO	0	0	0	0	00:00	FEBRUARI	L	KAWIN	Struktural	44	25
8	SULISTIYONO	0	0	1	0	08:30	MARET	L	KAWIN	Struktural	44	25
9	SULISTIYONO	0	0	0	0	00:00	APRIL	L	KAWIN	Struktural	44	25
10	SULISTIYONO	0	0	0	0	00:00	MEI	L	KAWIN	Struktural	44	25
11	SULISTIYONO	0	0	0	0	00:00	JUNI	L	KAWIN	Struktural	44	25
12	SULISTIYONO	0	0	0	0	00:00	JULI	L	KAWIN	Struktural	44	25
13	SULISTIYONO	0	0	0	0	00:00	AGUSTUS	L	KAWIN	Struktural	44	25
14	SULISTIYONO	0	0	0	0	00:00	SEPTEMBER	L	KAWIN	Struktural	44	25
15	SULISTIYONO	0	0	0	0	00:00	OKTOBER	L	KAWIN	Struktural	44	25
16	SULISTIYONO	0	0	0	0	00:00	NOVEMBER	L	KAWIN	Struktural	44	25
17	SULISTIYONO	0	0	0	0	00:00	DESEMBER	L	KAWIN	Struktural	44	25

### 3.2.2.2 Cleaning

Pada proses ini, data akan dibersihkan meliputi kelengkapan data, penghapusan data *redundant* (data yang sama dalam *database*) dan penghilangan noise.

Pembersihan data yang pertama dilakukan pada kolom *Total\_Kurang\_Jam*, karena pada kolom tersebut waktu masih menunjukkan gabungan antara jam dan menit. Pada kolom tersebut data jam akan dikonversikan kedalam bentuk menit sehingga nanti data akan lebih mudah ditransformasi. Mengkonversi kolom *Total\_Kurang\_Jam* dari jam ke menit penulis memanfaatkan rumus yang terdapat pada *Microsoft Excel* sehingga konversi data bisa lebih akurat. Contoh penggunaan rumus yaitu  $=F6*1440$ . Keterangan formula 1440 adalah jumlah menit dalam sehari, yaitu 60 menit x 24. Kolom *Total\_Kurang\_Jam* yang sudah di konversi terlihat pada Tabel 3.3.

Tabel 3.3. Konversi Jam ke Menit

		=F6*1440												
	A	B	C	D	E	F	G	H	I	J	K	L	M	
	Nama	Telat	Pulang_Cepat	Total_Hari_Lupa	Total_Hari_Alpha	Total_Kurang_Jam_Menit	Total_Kurang_Jam_Menit	Bulan	Jenis_Kelamin	Status_Perkawinan	Jabatan	Usia	Masa_Kerja	
5														
6	SULISTIYONO	0	0	0	0	00:00	0	JANUARI	L	KAWIN	Struktural	44	25	
7	SULISTIYONO	0	0	0	0	00:00	0	FEBRUARI	L	KAWIN	Struktural	44	25	
8	SULISTIYONO	0	0	1	0	08:30	510	MARET	L	KAWIN	Struktural	44	25	
9	SULISTIYONO	0	0	0	0	00:00	0	APRIL	L	KAWIN	Struktural	44	25	
10	SULISTIYONO	0	0	0	0	00:00	0	MEI	L	KAWIN	Struktural	44	25	
11	SULISTIYONO	0	0	0	0	00:00	0	JUNI	L	KAWIN	Struktural	44	25	
12	SULISTIYONO	0	0	0	0	00:00	0	JULI	L	KAWIN	Struktural	44	25	
13	SULISTIYONO	0	0	0	0	00:00	0	AGUSTUS	L	KAWIN	Struktural	44	25	
14	SULISTIYONO	0	0	0	0	00:00	0	SEPTEMBER	L	KAWIN	Struktural	44	25	
15	SULISTIYONO	0	0	0	0	00:00	0	OKTOBER	L	KAWIN	Struktural	44	25	
16	SULISTIYONO	0	0	0	0	00:00	0	NOVEMBER	L	KAWIN	Struktural	44	25	
17	SULISTIYONO	0	0	0	0	00:00	0	DESEMBER	L	KAWIN	Struktural	44	25	
18	TRI PRADEKSO	3	0	1	0	04:04	244	JANUARI	L	KAWIN	Struktural	BUP		

Pembersihan data kedua yaitu menghilangkan kolom *Bulan* karena nanti kolom tersebut tidak dipakai sebagai bahan *clustering*. Sebagai gantinya untuk menunjukkan bulan, pada kolom *Nama* ditambah dengan angka sesuai bulan yang berkenaan. Perlu diketahui dalam data tidak semua pegawai memiliki data genap 12 bulan. Karena pada tahun 2018 bisa terjadi mutasi pegawai baik pindah datang maupun pindah pergi, meninggal dan pegawai yang mengalami pensiun. Kolom *Nama* yang sudah di gabung dengan data bulan terlihat pada Tabel 3.4.

Tabel 3.4. Konversi Nama

	Nama	Telat	Pulang_Cepat	Total_Hari_Lupa	Total_Hari_Alpha	Total_Kurang_Jam_Menit	Bulan	Jenis_Kelamin	Status_Perkawinan	Jabatan	Usia	Masa_Kerja
5												
6	SULISTIYONO_1	0	0	0	0	0	JANUARI	L	KAWIN	Struktural	44	25
7	SULISTIYONO_2	0	0	0	0	0	FEBRUARI	L	KAWIN	Struktural	44	25
8	SULISTIYONO_3	0	0	1	0	510	MARET	L	KAWIN	Struktural	44	25
9	SULISTIYONO_4	0	0	0	0	0	APRIL	L	KAWIN	Struktural	44	25
10	SULISTIYONO_5	0	0	0	0	0	MEI	L	KAWIN	Struktural	44	25
11	SULISTIYONO_6	0	0	0	0	0	JUNI	L	KAWIN	Struktural	44	25
12	SULISTIYONO_7	0	0	0	0	0	JULI	L	KAWIN	Struktural	44	25
13	SULISTIYONO_8	0	0	0	0	0	AGUSTUS	L	KAWIN	Struktural	44	25
14	SULISTIYONO_9	0	0	0	0	0	SEPTEMBER	L	KAWIN	Struktural	44	25
15	SULISTIYONO_10	0	0	0	0	0	OKTOBER	L	KAWIN	Struktural	44	25
16	SULISTIYONO_11	0	0	0	0	0	NOVEMBER	L	KAWIN	Struktural	44	25
17	SULISTIYONO_12	0	0	0	0	0	DESEMBER	L	KAWIN	Struktural	44	25
18	TRI PRADEKSO_1	3	0	1	0	244	JANUARI	L	KAWIN	Struktural	BUP	
19	TRI PRADEKSO_2	4	0	0	0	21	FEBRUARI	L	KAWIN	Struktural		
20	TRI PRADEKSO_3	6	0	0	0	57	MARET	L	KAWIN	Struktural		
21	TRI PRADEKSO_4	5	0	0	0	43	APRIL	L	KAWIN	Struktural		
22	TRI PRADEKSO_5	8	0	0	0	72	MEI	L	KAWIN	Struktural		
23	TRI PRADEKSO_6	0	0	0	0	0	JUNI	L	KAWIN	Struktural		
24	Putut Andi Setiyawan_1	3	0	0	0	40	JANUARI	L	KAWIN	Struktural	46	13

Pembersihan data selanjutnya yaitu melengkapi data yang masih kosong. Pada Tabel 3.4 terlihat kolom Usia dan Masa\_Kerja sebagian pegawai yang memasuki Batas Usia Pensiun (BUP) masih kosong sehingga perlu dilengkapi datanya. Batas usia pensiun diberhentikan dengan hormat sebagai PNS berdasarkan Peraturan Pemerintah Nomor 11 Tahun 2017 adalah 58 (lima puluh delapan) tahun bagi pejabat administrasi, pejabat fungsional ahli muda, pejabat fungsional ahli pertama, dan pejabat fungsional keterampilan; 60 (enam puluh) tahun bagi pejabat pimpinan tinggi dan pejabat fungsional madya; dan 65 (enam puluh lima) tahun bagi PNS yang memangku pejabat fungsional ahli utama. Kolom Usia dan Masa\_Kerja yang sudah dilengkapi datanya bisa dilihat pada Tabel 3.5.

Tabel 3.5. *Cleaning* Data Usia dan Masa\_Kerja

	Nama	Telat	Pulang_Cepat	Total_Hari_Lupa	Total_Hari_Alpha	Total_Kurang_Jam_Menit	Jenis_Kelamin	Status_Perkawinan	Jabatan	Usia	Masa_Kerja
5											
6	SULISTİYONO_1	0	0	0	0	0	L	KAWIN	Struktural	44	25
7	SULISTİYONO_2	0	0	0	0	0	L	KAWIN	Struktural	44	25
8	SULISTİYONO_3	0	0	1	0	510	L	KAWIN	Struktural	44	25
9	SULISTİYONO_4	0	0	0	0	0	L	KAWIN	Struktural	44	25
10	SULISTİYONO_5	0	0	0	0	0	L	KAWIN	Struktural	44	25
11	SULISTİYONO_6	0	0	0	0	0	L	KAWIN	Struktural	44	25
12	SULISTİYONO_7	0	0	0	0	0	L	KAWIN	Struktural	44	25
13	SULISTİYONO_8	0	0	0	0	0	L	KAWIN	Struktural	44	25
14	SULISTİYONO_9	0	0	0	0	0	L	KAWIN	Struktural	44	25
15	SULISTİYONO_10	0	0	0	0	0	L	KAWIN	Struktural	44	25
16	SULISTİYONO_11	0	0	0	0	0	L	KAWIN	Struktural	44	25
17	SULISTİYONO_12	0	0	0	0	0	L	KAWIN	Struktural	44	25
18	TRI PRADEKSO_1	3	0	1	0	244	L	KAWIN	Struktural	58	25
19	TRI PRADEKSO_2	4	0	0	0	21	L	KAWIN	Struktural	58	25
20	TRI PRADEKSO_3	6	0	0	0	57	L	KAWIN	Struktural	58	25
21	TRI PRADEKSO_4	5	0	0	0	43	L	KAWIN	Struktural	58	25
22	TRI PRADEKSO_5	8	0	0	0	72	L	KAWIN	Struktural	58	25
23	TRI PRADEKSO_6	0	0	0	0	0	L	KAWIN	Struktural	58	25
24	Putut Andi Setiyawan_1	3	0	0	0	40	L	KAWIN	Struktural	45	12

### 3.2.2.3 Transformasi Data

Tahapan ini data akan di transformasi agar nantinya data siap di *clustering*. Dilakukan beberapa normalisasi terhadap kolom-kolom yang memiliki data bervariasi ke bentuk angka atau numerik. Skala pada masing-masing kolom menjadi bervariasi sesuai dengan setiap isi kolom tersebut. Transformasi data yang akan dilakukan terlihat seperti Tabel 3.6.

Tabel 3.6. Transformasi Data

Atribut	Uraian	Keterangan
Jenis_Kelamin	Laki-Laki	0
	Perempuan	1
Usia	20 - 30 Tahun	0
	31 - 40 Tahun	1
	41 - 50 Tahun	2
	51 - 60 Tahun	3
Status_Perkawinan	Lajang/Tidak Kawin	0
	Kawin	1
	Duda/Janda/Cerai	2
Masa_Kerja	0 - 10 Tahun	0
	11 - 20 Tahun	1
	21 - 30 Tahun	2
	31 Tahun Keatas	3



Lanjutan Tabel 3.6

Atribut	Uraian	Keterangan
Jabatan	Pelaksana	0
	Fungsional	1
	Struktural	2
Telat	= 0 Kali	0
	<= 3 Kali	1
	< 5 Kali	2
	< 10 Kali	3
	>= 10 Kali	4
Pulang_Cepat	= 0 Kali	0
	<= 3 Kali	1
	< 5 Kali	2
	< 10 Kali	3
	>= 10 Kali	4
Total_Hari_Lupa	= 0 Kali	0
	<= 3 Kali	1
	< 5 Kali	2
	< 10 Kali	3
	>= 10 Kali	4
Total_Hari_Alpha	= 0 Kali	0
	<= 3 Kali	1
	< 5 Kali	2
	< 10 Kali	3
	>= 10 Kali	4
Total_Kurang_Jam_Menit	= 0 Menit	0
	<= 120 Menit	1
	< 240 Menit	2
	< 360 Menit	3
	< 540 Menit	4
	>= 540 Menit	5

Normalisasi dibuat secara manual dengan rumus *if* pada masing-masing kolom *file excel* setelah semua data pegawai terkumpul. Rumus yang digunakan untuk normalisasi kolom Telat yaitu =IF(B7=0,"0",IF(B7<=3,"1", IF(B7<5,"2", IF(B7<10,"3",IF(B7>=10,"4")))). Kolom Pulang\_Cepat yaitu =IF(D7=0,"0", IF(D7<=3,"1",IF(D7<5,"2",IF(D7<10,"3",IF(D7>=10,"4")))). Kolom Total\_Hari\_Lupa =IF(F7=0,"0",IF(F7<=3,"1",IF(F7<5,"2",IF(F7<10,"3",IF(F7>=10,"4"

))))). Kolom Total\_Hari\_Alpha =IF(H7=0,"0",IF(H7<=3,"1", IF(H7<5,"2", IF(H7<10,"3",IF(H7>=10,"4")))). Kolom Total\_Kurang\_Jam\_Menit =IF(J7=0,"0",IF(J7<=120,"1",IF(J7<240,"2",IF(J7<360,"3",IF(J7<540,"4",IF(J7>=540,"5")))). Kolom Jenis\_Kelamin =IF(L7="L","0",IF(L7="P","1")). Kolom Status\_Perkawinan =IF(N7="TIDAK KAWIN","0",IF(N7="KAWIN","1", IF(N7="JANDA","2",IF(N7="DUDA","2")))). Kolom Jabatan =IF(P7="Pelaksana","0",IF(P7="Fungsional","1",IF(P7="Struktural","2"))). Kolom Usia =IF(AND(R7>=20,R7<=30),"0",IF(AND(R7>=31,R7<=40),"1",IF(AND(R7>=41 ,R7<=50),"2",IF(AND(R7>=51,R7<=60),"3")))). Kolom Masa\_Kerja =IF(AND (T7>=0,T7<=10),"0",IF(AND(T7>=11,T7<=20),"1",IF(AND(T7>=21,T7<=30)," 2",IF(AND(T7>30),"3")))). *Dataset* yang merupakan hasil dari transformasi data dapat dilihat pada Tabel 3.7.

Tabel 3.7. *Dataset* Penelitian

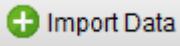
Nama	Telat	Pulang_Cepat	Total_Hari_Lupa	Total_Hari_Alpha	Total_Kurang_Jam_Menit	Jenis_Kelamin	Status_Perkawinan	Jabatan	Usia	Masa_Kerja
SULISTIYONO_1	0	0	0	0	0	0	1	2	2	2
SULISTIYONO_2	0	0	0	0	0	0	1	2	2	2
SULISTIYONO_3	0	0	1	0	4	0	1	2	2	2
SULISTIYONO_4	0	0	0	0	0	0	1	2	2	2
SULISTIYONO_5	0	0	0	0	0	0	1	2	2	2
SULISTIYONO_6	0	0	0	0	0	0	1	2	2	2
SULISTIYONO_7	0	0	0	0	0	0	1	2	2	2
SULISTIYONO_8	0	0	0	0	0	0	1	2	2	2
SULISTIYONO_9	0	0	0	0	0	0	1	2	2	2
SULISTIYONO_10	0	0	0	0	0	0	1	2	2	2
SULISTIYONO_11	0	0	0	0	0	0	1	2	2	2
SULISTIYONO_12	0	0	0	0	0	0	1	2	2	2
TRI PRADEKSO_1	1	0	1	0	3	0	1	2	3	2
TRI PRADEKSO_2	2	0	0	0	1	0	1	2	3	2
TRI PRADEKSO_3	3	0	0	0	1	0	1	2	3	2
TRI PRADEKSO_4	3	0	0	0	1	0	1	2	3	2
TRI PRADEKSO_5	3	0	0	0	1	0	1	2	3	2
TRI PRADEKSO_6	0	0	0	0	0	0	1	2	3	2

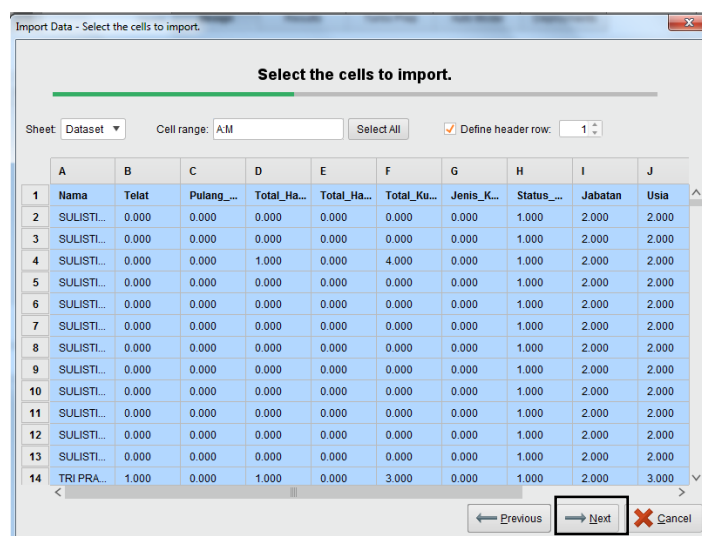
*Dataset* penelitian seperti yang terlihat pada Tabel 3.7 merupakan data yang nanti akan digunakan untuk tahap berikutnya yaitu *clustering* dengan menggunakan aplikasi Rapidminer Studio. *Clustering* data akan lebih difokuskan kepada kolom Total\_Kurang\_Jam untuk mengetahui pegawai yang memiliki tingkat kedisiplinan yang sangat baik dan pegawai yang memiliki tingkat kedisiplinan buruk.

### 3.2.3. Clustering Data

*Clustering* data menggunakan algoritma *k-means* terdapat pada aplikasi Rapidminer Studio. Berikut merupakan langkah-langkah dalam penyelesaian pengklasteran menggunakan algoritma *k-means*:

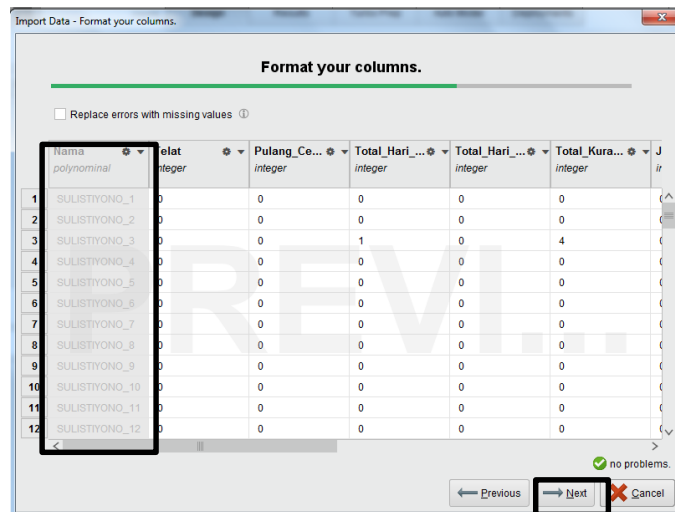
- 1) Import *dataset* ke aplikasi RapidMiner Studio;
- 2) Perancangan model *clustering k-means*;
- 3) Pengujian dalam pembagian kelompok;
- 4) Analisa hasil *clustering*.

Langkah pertama yang dilakukan adalah *import* data Dataset\_Presensi\_2018\_Kecamatan\_Tingkir pada aplikasi RapidMiner Studio dengan memilih  kemudian pilih *file excel dataset* yang sebelumnya sudah di transformasi. Pada kotak *Import Data – Select the cells to import* pastikan semua *cell* sudah terpilih lalu pilih *Next* seperti terlihat pada Gambar 3.2.



Gambar 3.2. *Import Data - Select the cells to import*

Lakukan *Exclude* pada kotak *Import Data – Format your columns* terhadap kolom Nama karena tidak berpengaruh terhadap hasil pengelompokan nantinya. Selanjutnya ubah kolom dengan tipe data menjadi *integer* lalu pilih *Next* seperti Gambar 3.3.



Gambar 3.3. *Import Data – Format your columns*

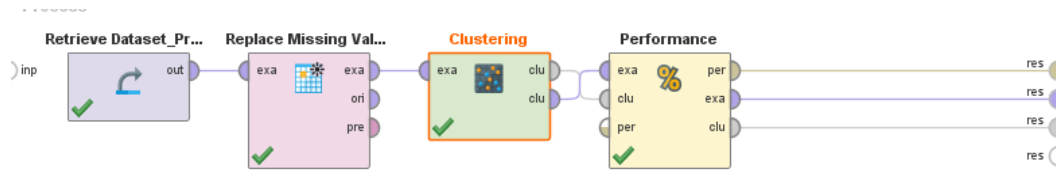
Gambar 3.4 menampilkan hasil *import dataset* yang akan dianalisa. Kolom yang telah ditiadakan tidak akan muncul. Kolom yang digunakan yaitu Telat, Pulang\_Cepat, Total\_Hari\_Lupa, Total\_Kurang\_Jam\_Menit, Jenis\_Kelamin, Status\_Perkawinan, Jabatan, Usia dan Masa\_Kerja. Kolom atau atribut ini yang nanti akan digunakan untuk mengetahui tingkat kedisiplinan PNS Kecamatan Tingkir berdasarkan presensinya.

Row No.	Telat	Pulang_Cepat	Total_Hari_Lupa	Total_Hari_Kurang	Total_Kurang_Jam_Menit	Jenis_Kelamin	Status_Perkawinan	Jabatan	Usia
1	0	0	0	0	0	0	1	2	2
2	0	0	0	0	0	0	1	2	2
3	0	0	1	0	4	0	1	2	2
4	0	0	0	0	0	0	1	2	2
5	0	0	0	0	0	0	1	2	2
6	0	0	0	0	0	0	1	2	2
7	0	0	0	0	0	0	1	2	2
8	0	0	0	0	0	0	1	2	2
9	0	0	0	0	0	0	1	2	2

Gambar 3.4. *Dataset Analisis*

Setelah *import dataset* selesai, langkah selanjutnya adalah dengan melakukan perancangan model *clustering*. Rancangan terdiri dari *dataset presensi\_2018\_Kecamatan\_Tingkir*, *Operators tool replace missing value*, *tool k-mean*, dan *tool performance*. Perancangan *clustering* harus disilangkan ke *tool*

*performance (Cluster Distance Performance)* untuk memperoleh hasil yang sesuai. Proses perancangan *clustering* apabila tidak terdapat masalah maka ketika di klik tombol *Run* masing-masing *Operators* akan menunjukkan tanda centang seperti yang terlihat pada Gambar 3.5.



Gambar 3.5. Rancangan Proses *Clustering*

Langkah selanjutnya dilakukan pengujian dalam pembagian kelompoknya untuk memperoleh hasil terbaik dengan mengganti nilai  $k$  dari 2 sampai 10 seperti Gambar 3.6 sampai Gambar 3.14. Kemudian pilih kelompok yang paling bisa menggambarkan tingkat kedisiplinan PNS Kecamatan Tingkir berdasarkan tingkat presensinya dengan membandingkan *Attribute Total\_Kurang\_Jam\_Menit* pada masing-masing kelompok. Dipilih *Attribute Total\_Kurang\_Jam\_Menit* karena *Attribute* tersebut sudah mempresentasikan *Attribute* yang lain (Telat, Pulang\_Cepat, Total\_Hari\_Lupa, Total\_Hari\_Alpha) yang diakumulasikan dalam jumlah Menit.

Attribute	cluster_0	cluster_1
Telat	0.333	0.512
Pulang_Cepat	0.017	0.040
Total_Hari_Lupa	0.002	0.939
Total_Hari_Alpha	0	0.155
Total_Kurang_Jam_Menit	0.263	4.328
Jenis_Kelamin	0.419	0.451
Status_Perkawinan	1.024	1.027
Jabatan	1.114	1.301
Usia	2.115	2.165
Masa_Kerja	1.391	1.451

Gambar 3.6. Pengujian Parameter  $K=2$

Attribute	cluster_0	cluster_1	cluster_2
Telat	0.159	0.511	0.570
Pulang_Cepat	0.013	0.040	0.022
Total_Hari_Lupa	0.003	0.939	0.004
Total_Hari_Alpha	0	0.155	0
<b>Total_Kurang_Jam_Menit</b>	<b>0.143</b>	<b>4.332</b>	<b>0.435</b>
Jenis_Kelamin	0.367	0.449	0.491
Status_Perkawinan	1.029	1.027	1.017
Jabatan	1.390	1.305	0.739
Usia	2.701	2.171	1.322
Masa_Kerja	2.084	1.455	0.457

Gambar 3.7. Pengujian Parameter K=3

Attribute	cluster_0	cluster_1	cluster_2	cluster_3
Telat	0.597	0.159	0.419	0.568
Pulang_Cepat	0.041	0.013	0.039	0.022
Total_Hari_Lupa	0.934	0.003	0.944	0
Total_Hari_Alpha	0.214	0	0.089	0
<b>Total_Kurang_Jam_Menit</b>	<b>4.372</b>	<b>0.143</b>	<b>4.279</b>	<b>0.424</b>
Jenis_Kelamin	0.526	0.367	0.369	0.489
Status_Perkawinan	1.071	1.029	0.978	1.017
Jabatan	0.929	1.390	1.709	0.742
Usia	1.582	2.701	2.804	1.328
Masa_Kerja	0.561	2.084	2.425	0.459

Gambar 3.8. Pengujian Parameter K=4

Attribute	cluster_0	cluster_1	cluster_2	cluster_3	cluster_4
Telat	0.202	0.421	0.442	0.500	0.599
Pulang_Cepat	0.016	0.039	0.017	0	0.043
Total_Hari_Lupa	0.004	0.944	0	0.500	0.957
Total_Hari_Alpha	0	0.067	0	3.300	0.070
<b>Total_Kurang_Jam_Menit</b>	<b>0.181</b>	<b>4.275</b>	<b>0.330</b>	<b>5</b>	<b>4.342</b>
Jenis_Kelamin	0.342	0.371	0.483	0.500	0.524
Status_Perkawinan	0.967	0.978	1.071	1.300	1.059
Jabatan	1.547	1.708	0.755	0.700	0.947
Usia	2.728	2.803	1.609	1.700	1.583
Masa_Kerja	2.374	2.427	0.578	0.800	0.556

Gambar 3.9. Pengujian Parameter K=5

Attribute	cluster_0	cluster_1	cluster_2	cluster_3	cluster_4	cluster_5
Telat	0.421	0	0.581	1.542	0.500	0.147
Pulang_Cepat	0.039	0.015	0.043	0.021	0	0.017
Total_Hari_Lupa	0.944	0	0.957	0.010	0.500	0.004
Total_Hari_Alpha	0.067	0	0.070	0	3.300	0
Total_Kurang_Jam_Menit	4.275	0.015	4.349	1.062	5	0.164
Jenis_Kelamin	0.371	0.451	0.522	0.531	0.500	0.349
Status_Perkawinan	0.978	1.118	1.059	0.969	1.300	0.966
Jabatan	1.708	0.676	0.946	0.948	0.700	1.555
Usia	2.803	1.618	1.586	1.656	1.700	2.723
Masa_Kerja	2.427	0.637	0.559	0.521	0.800	2.382

Gambar 3.10. Pengujian Parameter K=6

Attribute	cluster_0	cluster_1	cluster_2	cluster_3	cluster_4	cluster_5	cluster_6
Telat	0	0.638	0.268	0.617	0.500	1.558	0.126
Pulang_Cepat	0.024	0.054	0.039	0.028	0	0.011	0.014
Total_Hari_Lupa	0	0.908	0.937	1.019	0.500	0.011	0.004
Total_Hari_Alpha	0	0.085	0.031	0.093	3.300	0	0
Total_Kurang_Jam_Menit	0.030	4.354	4.118	4.495	5	1.053	0.141
Jenis_Kelamin	0.458	0.554	0.378	0.402	0.500	0.526	0.361
Status_Perkawinan	1.042	1.085	0.976	0.991	1.300	0.968	1.032
Jabatan	0.831	0.492	2	1.514	0.700	0.958	1.336
Usia	1.283	1.492	2.197	3	1.700	1.674	2.762
Masa_Kerja	0.548	0.408	1.520	2.710	0.800	0.526	2.188

Gambar 3.11. Pengujian Parameter K=7

Attribute	cluster_0	cluster_1	cluster_2	cluster_3	cluster_4	cluster_5	cluster_6	cluster_7
Telat	0.126	0.720	0.527	0.391	1.964	0.250	0.137	0.196
Pulang_Cepat	0.019	0.070	0	0.042	0	0	0.017	0.019
Total_Hari_Lupa	0	0.950	0.878	0.974	0.018	0.500	0	0.006
Total_Hari_Alpha	0	0.110	0.108	0.052	0	3.625	0	0
Total_Kurang_Jam_Menit	0.145	4.360	4.338	4.286	1.071	5	0.171	0.215
Jenis_Kelamin	0.401	0.630	0.297	0.417	0.357	0.375	0.615	0.323
Status_Perkawinan	1.116	0.940	1.216	0.979	1	1.500	0.974	0.949
Jabatan	1.034	0.970	0	2	0.768	0.750	0.838	1.544
Usia	2.309	0.960	2.676	2.609	1.643	2	0.821	2.981
Masa_Kerja	1.367	0.160	1.635	2.078	0.446	1	0.274	2.576

Gambar 3.12. Pengujian Parameter K=8

Attribute	cluster_0	cluster_1	cluster_2	cluster_3	cluster_4	cluster_5	cluster_6	cluster_7	cluster_8
Telat	0.787	0.151	0.453	2.767	0.245	0.222	0.333	0.500	0.366
Pulang_Cepat	0.093	0.013	0.022	0	0.014	0.028	0	0	0.054
Total_Hari_Lupa	0.933	0	1.007	0.033	0	0.007	0.444	0.850	0.946
Total_Hari_Alpha	0.120	0	0.051	0	0	0	3.444	0.100	0.054
<b>Total_Kurang_Jam_Menit</b>	<b>4.400</b>	<b>0.164</b>	<b>4.314</b>	<b>1.067</b>	<b>0.281</b>	<b>0.250</b>	<b>5</b>	<b>4.300</b>	<b>4.258</b>
Jenis_Kelamin	0.720	0.400	0.409	0.533	0.547	0.306	0.333	0.333	0.376
Status_Perkawinan	0.920	0.964	0.971	0.900	0.978	1.188	1.444	1.267	1
Jabatan	0.813	2	1.693	0.867	0.885	0	0.889	0	2
Usia	0.800	2.529	3	1.733	0.914	2.701	2.111	2.433	1.882
Masa_Kerja	0.027	2.067	2.555	0.533	0.230	1.625	1.111	1.017	1.301

Gambar 3.13. Pengujian Parameter K=9

Attribute	cluster_0	cluster_1	cluster_2	cluster_3	cluster_4	cluster_5	cluster_6	cluster_7	cluster_8	cluster_9
Telat	0.281	0.333	0.098	0	0.649	0.429	0.299	2.643	1.263	2.692
Pulang_Cepat	0.012	0	0.016	0.024	0.072	0	0.043	0	0.026	0
Total_Hari_Lupa	0	0.444	0	0	0.948	0.900	0.940	1.357	0.026	0.038
Total_Hari_Alpha	0	3.444	0	0	0.103	0.086	0.038	0.286	0	0
<b>Total_Kurang_Jam_Menit</b>	<b>0.305</b>	<b>5</b>	<b>0.120</b>	<b>0.024</b>	<b>4.351</b>	<b>4.329</b>	<b>4.261</b>	<b>4.714</b>	<b>1.026</b>	<b>1.115</b>
Jenis_Kelamin	0.575	0.333	0.375	0.293	0.619	0.314	0.435	0.214	0.237	0.615
Status_Perkawinan	0.982	1.444	0.957	1.195	0.948	1.214	0.978	1	1.079	0.885
Jabatan	1.216	0.889	2	0	0.969	0	2	1.214	0.211	0.769
Usia	1.102	2.111	2.641	2.618	0.938	2.686	2.598	2.500	2.816	1.462
Masa_Kerja	0.431	1.111	2.299	1.407	0.155	1.629	2.054	1.929	1.895	0.269

Gambar 3.14. Pengujian Parameter K=10

### 3.2.4. Analisis Hasil Pengujian

Analisis yang dilakukan adalah hasil pengujian *clustering* menggunakan parameter k untuk memperoleh hasil yang terbaik dengan adanya variasi hasil pada kolom Total\_Kurang\_Jam\_Menit. Setelah dilakukan beberapa pengujian dengan menggunakan parameter k dari k=2 sampai k=10 diperoleh hasil yang paling baik yaitu pengujian dengan parameter k=6. Pemilihan dilakukan dengan melihat adanya variasi hasil pada kolom Total\_Kurang\_Jam\_Menit, terutama untuk mengetahui pegawai yang memiliki nilai Total\_Kurang\_Jam\_Menit tersedikit (tingkat kedisiplinan paling baik) dan terbanyak (tingkat kedisiplinan paling buruk).

Pengujian k=2 sampai dengan k=4 diperoleh hasil *cluster* yang kurang variatif dan juga kurang spesifik. Pengujian k=5, k=7, k=8, k=9 dan k=10 hasilnya cukup bagus namun masih kurang menggambarkan kolom dengan nilai



Total\_Kurang\_Jam\_Menit tersedikit. Selanjutnya akan dilakukan analisis lebih dalam lagi pada karakteristik dari masing-masing *cluster* berdasarkan jumlah frekuensi tiap atributnya.