

BAB III

METODE PENELITIAN

3.1 Desain Penelitian

Penelitian ini merupakan pendekatan kuantitatif dimana pengambilan data dari penelitian ini dilakukan secara kuantitatif. Kuantitatif merupakan sebuah metode penelitian untuk menguji teori-teori tertentu dengan cara meneliti hubungan antar variabel. Variabel-variabel ini diukur sehingga data yang terdiri dari angka-angka dapat dianalisis berdasarkan prosedur-prosedur statistik (Creswell, 2012).

3.2 Lokasi Penelitian

Penelitian ini mengambil lokasi di *WU Backroom Bar* Bandung. Tempat hiburan malam ini berada di Jl. Sulanjana No.7, Tamansari, Kec. Bandung Wetan, Kota Bandung, Jawa Barat 40116. *WU Backroom Bar* Bandung terletak di daerah yang memang menjadi tempat hiburan malam lainnya di Kota Bandung. Aksesibilitas menuju lokasi juga terbilang sangat mudah. Terdapat beberapa angkutan umum yang melewati lokasi tersebut. Selain itu juga, jalanan kota yang sudah bagus memberikan kenyamanan bagi pengunjung yang akan berkunjung ke *WU Backroom Bar* Bandung. Peta lokasi penelitian yang diambil dapat dilihat pada Gambar 3.1.



Gambar 3. 1 Lokasi WU Backroom Bar

Sumber : GoogleMaps (2022)

Rian Turmudzi, 2025

TIPOLOGI PENGUNJUNG HIBURAN MALAM BERDASARKAN MOTIVASI DAN KONTAK SOSIAL DI BANDUNG (STUDI KASUS CLUB WU BACKROOM BANDUNG)

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

3.3 Populasi dan Sampel

Menurut Sugiyono (2011) populasi adalah wilayah generalisasi yang terdiri atas objek atau subjek yang mempunyai kualitas dan karakteristik tertentu yang ditetapkan oleh peneliti untuk dipelajari dan kemudian ditarik kesimpulannya.

Subjek penelitian ini adalah individu atau responden yang pernah mengunjungi klub malam *WU Backroom Bar Bandung*. Sedangkan objek dalam penelitian ini adalah tipologi pengunjung berdasarkan motivasi dan kontak sosial.

Seseorang yang pernah mengunjungi klub malam *WU Backroom Bar Bandung* menjadi populasi dalam penelitian ini. Artinya tidak terbatas jumlah populasinya selama seseorang tersebut pernah berkunjung ke lokasi tersebut.

Sampel adalah sebagian yang diambil dari keseluruhan objek yang diteliti dan dianggap mewakili keseluruhan populasi. Agar sampel tersebut mewakili populasi maka diperlukan metode atau teknik yang tepat pada proses pengambilan sampel (Notoatmodjo, 2018). teknik sampel yang digunakan yaitu *quota sampling*. *Quota sampling* artinya teknik untuk menentukan sampel dari populasi yang mempunyai ciri-ciri tertentu sampai jumlah (kuota) yang diinginkan (Sugiyono, 2016). Adapun caranya adalah dengan memberikan kuesioner online kepada seseorang yang pernah berkunjung ke klub malam *WU Backroom Bandung*. Penentuan jumlah sampel yang representatif adalah bahwa jumlah sampel penelitian yang tidak diketahui jumlah populasi pastinya, dapat dihitung dengan cara mengalikan satu angka diantara 5 sampai 10 dikali jumlah indikator. Rumusnya adalah $(5-10) \times \text{jumlah indikator}$.

Maka akan didapatkan jumlah sampel sebanyak dibawah ini:

$$\begin{aligned}\text{Jumlah Sampel} &= (5-10) \times \text{jumlah indikator} \\ &= 5 \times 23 \\ &= 115\end{aligned}$$

Dari rumus tersebut, didapatkan jumlah sampel sebanyak 115 sampel. Maka sampel yang akan diteliti adalah sebanyak 115 responden.

3.4 Jenis Sumber Data

Untuk mengumpulkan data yang benar, maka mengecek kebenaran data agar validitasnya dapat terbukti. Jenis data terbagi atas data primer dan data sekunder (Wardiyanta, 2006). Maka dalam penelitian ini peneliti memakai teknik pengumpulan data sebagai berikut:

1. Data Primer

Data primer merupakan data yang diperoleh langsung di lapangan oleh peneliti sebagai obyek penulisan. Dalam penelitian ini, data primer yang digunakan diantaranya Kuesioner. Kuesioner adalah daftar pertanyaan yang akan diajukan pada responden untuk diisi sendiri oleh responden. Yang menjadi responden dalam penelitian ini adalah orang yang pernah mengunjungi klub malam *WU Backroom* Bar Bandung. Kuesionernya berisi beberapa pertanyaan seputar motivasi dan kontak sosial di *WU Backroom* Bar Bandung.

2. Data Sekunder

Data sekunder adalah data yang tidak langsung memberikan data kepada peneliti, yaitu penelitian harus melalui orang lain atau mencari melalui dokumen. Jadi data tersebut bukanlah data yang diperoleh secara langsung oleh peneliti, melainkan data yang sudah diperoleh oleh peneliti lain atau dari berbagai sumber lain seperti studi literatur yang dilakukan terhadap banyak buku dan diperoleh berdasarkan catatan-catatan yang berhubungan dengan penelitian, selain itu peneliti mempergunakan data yang diperoleh dari internet. (Sugiyono, 2005).

3.5. Instrumen Penelitian

Menurut Sugiyono (2012), instrument penelitian adalah suatu alat yang digunakan untuk mengukur fenomena alam maupun sosial yang diamati. Pengumpulan data dalam penelitian ini didapatkan dari studi literatur dan website. Alat yang digunakan peneliti dalam melakukan penelitian adalah kuesioner. Menurut Kusumah (2011), kuesioner adalah daftar pertanyaan tertulis yang diberikan kepada subjek yang diteliti untuk mengumpulkan informasi yang dibutuhkan peneliti. Dalam penelitian ini, kuesioner yang digunakan adalah

kuesioner tertutup, dimana pertanyaan tersebut sudah dipersiapkan jawabannya, sehingga responden hanya memilih dari alternatif jawaban yang sesuai dengan pendapat atau pilihannya. Cara pengumpulan data dalam penelitian ini dilakukan dengan prosedur yaitu responden diberi kuesioner, kemudian setelah responden mengisi kuesioner, jawaban tersebut diketahui, diolah, dianalisa, dan dikumpulkan. Penyebaran kuesioner penelitian ini menggunakan google form.

Kuesioner penelitian ini berupa bentuk pendapat atas pernyataan yang diberi nilai dengan skala likert untuk jawabannya. Skala likert digunakan untuk mengukur sikap, pendapat, dan persepsi seseorang atau sekelompok tentang fenomena sosial Sarwono (2006,). Fenomena ini telah ditetapkan secara spesifik oleh peneliti sebagai variabel penelitian. Dengan skala likert, variabel yang akan diukur dijabarkan menjadi indikator variabel. Kemudian indikator tersebut dijadikan titik tolak untuk menyusun item-item instrumen yang dapat berupa pertanyaan atau pernyataan. Jawaban setiap item instrument yang menggunakan skala likert mempunyai gradasi dari sangat positif sampai sangat negatif, yang contohnya dapat berupa seperti kata-kata dibawah ini:

1. Sangat Rendah
2. Rendah
3. Netral
4. Tinggi
5. Sangat Tinggi

3.6 Operasional Variabel

Variabel adalah suatu atribut atau sifat nilai dari orang, objek atau kegiatan yang mempunyai variasi tertentu yang diterapkan oleh peneliti untuk dipelajari dan kemudian ditarik kesimpulannya (Sugiyono, 2014). Dalam penelitian ini variabel yang digunakan adalah variabel tunggal yaitu tipologi pengunjung berdasarkan motivasi perjalanan dan kontak sosial: studi kasus klub *WU Backroom* Bandung. Operasional variabel merupakan penjelasan secara rinci mengenai variabel, sub variabel, indikator variabel, dan skala pengukuran dengan tujuan untuk

memperoleh nilai variabel penelitian. Untuk lebih jelasnya dapat dilihat pada tabel di bawah ini:

Tabel 3. 1 Operasional Variabel

Variabel	Sub variable	Indikator	Skala	No item
Motivasi Perjalanan	Motivasi Intrinsik (Buhalis,2019)	Mencari ilmu	Ordinal	1.
		Memahami orang lain	Ordinal	2.
		Bersantai	Ordinal	3.
		Melepaskan diri dari rutinitas sehari-hari	Ordinal	4.
		Pemulihan	Ordinal	5.
		Menjalin kedekatan pasangan/kerabat	Ordinal	6.
		merefleksikan diri	Ordinal	7.
		Membebaskan diri	Ordinal	8.
		Menikmati Hidup	Ordinal	9.
	Motivasi Ekstrinsik (Jackson & Ryan, 1988 dan Zeng, 2015)	Menikmati Makanan & Minuman	Ordinal	10
		Merasakan budaya lokal	Ordinal	11
		Mencari Prestis/Status	Ordinal	12
		Menghadiri acara spesial	Ordinal	13
Kontak Sosial	Kontak sosial primer (<i>face to face contact</i>) (Buhalis,2019)	Bertemu Pengunjung lain	Ordinal	14
		Bertemu penyedia layanan	Ordinal	15
		Bertemu Teman dan Kerabat lokal	Ordinal	16
		Bertemu Masyarakat lokal	Ordinal	17
		Bercengkrama dengan Teman perjalanan	Ordinal	18
		Bercengkrama dengan Pengunjung yang dikenal selama perjalanan	Ordinal	19
	Kontak sosial Sekunder (<i>Online Contact</i>) (Buhalis,2019)	Melakukan kontak dengan dengan Teman	Ordinal	20
		Melakukan kontak dengan Keluarga	Ordinal	21

		Melakukan kontak dengan Kolega/klien	Ordinal	22
		Berbagi foto/video (momen) di sosial media	Ordinal	23

Sumber: Hasil Olahan Peneliti (2022)

3.7 Uji Validitas dan Uji Realibilitas

Sanusi (2011) mengatakan bahwa agar data yang diperoleh mempunyai tingkat akurasi yang tinggi dan konsistensi yang tinggi, instrumen penelitian yang digunakan harus valid dan reliabel. Berikut ini adalah uji validitas dan uji realibilitas dalam penelitian Tipologi Pengunjung Hiburan Malam Berdasarkan Pola Perilaku Kontak Sosial di Bandung (studi kasus *WU Backroom*).

1. Uji Validitas

Validitas adalah suatu ukuran yang menunjukkan tingkat-tingkat kevalidan dan kesahihan suatu instrumen (Arikunto, 2010). Rumus untuk menghitung kevalidan dari suatu instrumen adalah rumus korelasi product moment, yaitu sebagai berikut

Kriteria Uji jika $r_{hitung} > r_{tabel}$ maka data dinyatakan valid. Dimana skor r tabel yang dianggap menjadi syarat minimum adalah jika $r = 0,361$. Setelah instrumen dinyatakan valid, maka dapat dilihat kriteria penafsiran mengenai indeks korelasinya (r) pada tabel sebagai berikut:

Tabel 3. 2 Koefisien Korelasi nilai r

Interval Koefisien	Tingkat Hubungan
0,800 – 1,000	Sangat Kuat
0,600 – 0,799	Kuat
0,400 – 0,599	Cukup Kuat
0,200 – 0,399	Rendah
0,000 – 0,199	Sangat Rendah

Sumber : Sugiyono (2009)

Pengujian validitas responden dilakukan dengan menggunakan taraf signifikan sebagai berikut:

- Item pertanyaan yang diteliti dikatakan valid jika $r_{hitung} > r_{tabel}$
- Item pertanyaan yang diteliti dikatakan tidak valid jika $r_{hitung} < r_{tabel}$

Berikut ini adalah hasil uji validitas dari variabel motivasi pengunjung dan kontak sosial pengunjung pada klub malam WU Backroom:

Tabel 3. 3 Rekapitulasi Hasil Uji Validitas Variabel Penelitian

Variabel	Item	r_{hitung}	r_{kritis}	Kesimpulan
Motivasi Perjalanan	1	0.395	0,361	Valid
	2	0.626	0,361	Valid
	3	0.748	0,361	Valid
	4	0.730	0,361	Valid
	5	0.660	0,361	Valid
	6	0.622	0,361	Valid
	7	0.679	0,361	Valid
	8	0.620	0,361	Valid
	9	0.713	0,361	Valid
	10	0.802	0,361	Valid
	11	0.634	0,361	Valid
	12	0.468	0,361	Valid
	13	0.416	0,361	Valid
Kontak sosial	14	0.366	0,361	Valid
	15	0.603	0,361	Valid
	16	0.673	0,361	Valid
	17	0.617	0,361	Valid
	18	0.669	0,361	Valid
	19	0.373	0,361	Valid
	20	0.393	0,361	Valid
	21	0.363	0,361	Valid
	22	0.675	0,361	Valid
	23	0.631	0,361	Valid

Sumber: Olahan Peneliti (2023)

Berdasarkan pada tabel diatas, didapatkan bahwa semua indikator dalam masing-masing variabel memiliki nilai r hitung lebih tinggi dari r tabelnya. Maka dari itu setiap indikator dalam penelitian ini tetap digunakan dalam pengolahan data analisis selanjutnya.

2. Uji Reliabilitas

Arikunto (2009), mengungkapkan bahwa, reliabilitas menunjukan suatu pengertian, yaitu instrumen cukup dapat dipercaya untuk digunakan sebagai alat pengumpulan data karna instrumen tersebut sudah baik. Sedangkan menurut

Sugiyono (2010), reliabilitas berkenaan dengan derajat konsistensi dan stabilitas data atau temuan. Suatu data dinyatakan reliabel apabila dua atau lebih peneliti dalam objek yang sama menghasilkan data yang sama atau peneliti sama dalam waktu berbeda menghasilkan data yang sama, atau sekelompok data bila dipecahkan menjadi dua menunjukkan data yang tidak berbeda.

Berdasarkan penjelasan diatas maka dapat disimpulkan bahwa instrument cukup dapat dipercaya untuk digunakan sebagai alat pengumpulan data dan apabila instrument sudah dapat dipercaya maka akan dapat menghasilkan data yang dapat dipercaya. Pada penelitian kali ini untuk menguji reliabilitas menggunakan rumus alpha atau cronbach's alpha (α).

Keputusan uji reliabilitas ditentukan dengan ketentuan sebagai berikut:

- a. Jika koefisien internal seluruh item rhitung $>$ rtabel dengan tingkat signifikan 5 % maka item pertanyaan dinyatakan reliabel.
- b. Jika koefisien internal seluruh item rhitung $<$ rtabel dengan tingkat signifikan 5 % maka item pertanyaan dinyatakan tidak reliable.

Kaplan (1993) mengemukakan bahwa kelompok item dalam suatu dimensi dinyatakan reliabel jika koefisien reliabilitasnya tidak lebih rendah dari 0,70. Bila koefisien reliabilitas telah dihitung, maka menentukan keeratan hubungan bisa digunakan dengan kriteria Guilford (1956), yaitu:

0	: Tidak ada hubungan
0.01 - 0,19	: Hubungan sangat kecil dan bisa diabaikan
0,20 - 0,39	: Hubungan yang kecil (tidak erat)
0,40 - 0,69	: Hubungan yang cukup erat
0,70 - 0,89	: Hubungan yang erat (reliabel)
0,90 - 0,99	: Hubungan yang sangat erat
1,00	: Hubungan yang sempurna

Berikut ini merupakan hasil uji realibilitas dari variabel motivasi perjalanan dan kontak sosial dalam penelitian ini menggunakan SPSS:

Tabel 3. 4 Hasil Uji Realibilitas

<i>Reliability Statistics</i>

Variabel	<i>Cronbach's Alpha</i>	<i>N of Items</i>
Motivasi Perjalanan	0,877	13
Kontak Sosial	0,755	10

Sumber : Olahan Peneliti (2023)

Berdasarkan tabel diatas dapat dilihat bahwa nilai *Cronbach's Alpha* variabel memiliki nilai diatas 0,6 untuk variabel motivasi perjalanan dan juga untuk kontak sosial. Maka dari itu kedua variabel dalam penelitian ini memiliki nilai hubungan yang erat atau reliabel dari nilai item nya, yang artinya kedua variabel memenuhi uji realibilitas dan dapat dilakukan analisis selanjutnya.

3.8 Teknik Analisis Data

Teknik analisis data yang digunakan dalam penelitian ini yaitu analisis deskriptif dan analisis verifikatif. Berikut ini adalah penjelasan mengenai kedua analisis tersebut:

3.8.1. Analisis Deskriptif

Analisis deskriptif digunakan untuk memaparkan karakteristik variabel yang diamati demi mendukung pemecahan masalah dan memperoleh saran secara operasional. Objek yang dideskripsikan akan digambarkan dalam bentuk tabel, grafik atau diagram beserta analisis statistik sederhana.

Peneliti melakukan interpretasi variabel yang diteliti, disajikan dalam bentuk rekapitulasi skor rata-rata jawaban unit analisis untuk setiap indikator, dimensi dan variabel. Pembahasan pada analisis deskriptif disusun berdasar skala Likert. Skala Likert dianggap sebagai simetris dan jarak yang sama, maka akan berperilaku seperti skala interval (Hair et al, 2017:9).

Metode deskriptif menggambarkan keadaan yang sebenarnya secara sistematis. Metode yang digunakan berupa hasil pengoperasian variabel yang disusun dalam bentuk pernyataan (kuesioner/ angket). Dimana masing-masing variabel penelitian, setiap item dari kuesioner tersebut memiliki lima jawaban dengan bobot atau nilai yang berbeda.

Analisis deskriptif dilakukan dengan perhitungan skor konstruk dengan langkah-langkah sebagai berikut:

- 1) Menghitung frekuensi dari setiap pilihan jawaban berdasarkan bobot skor.

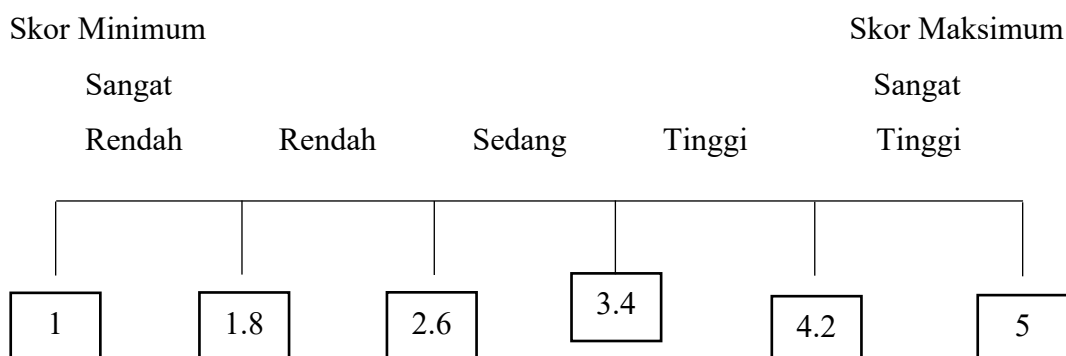
- 2) Menghitung skor aktual dengan mengalikan bobot skor dengan frekuensi tiap bobot skor.
- 3) Menjumlahkan skor aktual yang ada pada setiap pilihan jawaban.
- 4) Menghitung jumlah skor frekuensi yang ada pada setiap pilihan jawaban.
- 5) Jumlah total skor aktual (hasil langkah 3) dibagi dengan jumlah total skor (hasil langkah 4).
- 6) Hasil skor berdasarkan langkah 5 dimasukkan ke dalam kriteria pengategorian.

Tabel 3.5 Kriteria Kategori Jawaban Responden

Interval	Kategori Kriteria
1 – 1,8	Sangat Rendah
1,8 – 2,6	Rendah
2,6 – 3,4	Sedang
3,4 – 4,2	Tinggi
4,2 – 5,0	Sangat Tinggi

Sumber: Olahan Penulis (2022)

Selanjutnya dibuat garis kontinum sebagai berikut:



Gambar 3.2 Garis Kontinum

Sumber: Olahan Penulis (2022)

Analisis deskriptif yang disusun adalah menggambarkan dimensi dan indikator dari variabel yang diteliti untuk seluruh sampel yang digunakan.

3.8.2. Analisis Verifikatif

Analisis penelitian ini menggunakan analisis cluster. Analisis cluster menurut Ediyanto (2013) merupakan teknik multivariat yang mempunyai tujuan

utama untuk mengelompokkan objek-objek berdasarkan karakteristik yang dimilikinya. Analisis cluster mengklasifikasi objek sehingga setiap objek yang paling dekat kesamaannya dengan objek lain berada dalam cluster yang sama. Cluster-cluster yang terbentuk memiliki homogenitas internal yang tinggi dan heterogenitas eksternal yang tinggi. Berbeda dengan teknik multivariat lainnya, analisis ini tidak mengestimasi set variabel secara empiris sebaliknya menggunakan set variabel yang ditentukan oleh peneliti itu sendiri. Fokus dari analisis cluster adalah membandingkan objek berdasarkan set variabel, hal inilah yang menyebabkan para ahli mendefinisikan set variabel sebagai tahap kritis dalam analisis cluster. Set variabel cluster adalah suatu set variabel yang merpresentasikan karakteristik yang dipakai objek-objek. Solusi analisis cluster bersifat tidak unik, anggota cluster untuk tiap penyelesaian/solusi tergantung pada beberapa elemen prosedur dan beberapa solusi yang berbeda dapat diperoleh dengan mengubah satu elemen atau lebih. Solusi *cluster* secara keseluruhan bergantung pada variabel-variabel yang digunakan sebagai dasar untuk menilai kesamaan. Penambahan atau pengurangan variabel-variabel yang relevan dapat mempengaruhi substansi hasil analisis cluster.

Menurut Agusta (2007) ada dua jenis data clustering yang sering dipergunakan dalam proses pengelompokan data yaitu *hierarchical* (hirarki) data clustering dan *non-hierarchical* (non hirarki) data clustering. K-Means merupakan salah satu metode *data clustering* non hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih cluster atau kelompok. Metode ini mempartisi data ke dalam cluster atau kelompok sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu cluster yang sama dan data yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam kelompok yang lain. Adapun tujuan dari data clustering ini adalah untuk meminimalisasikan *objective function* yang diset dalam proses clustering, yang pada umumnya berusaha meminimalisasikan variasi di dalam suatu cluster dan memaksimalkan variasi antar cluster.

Data clustering menggunakan metode K-Means ini secara umum dilakukan dengan algoritma dasar sebagai berikut (Agusta, 2007):

1. Tentukan jumlah cluster
2. Alokasikan data ke dalam cluster secara random
3. Hitung centroid/ rata-rata dari data yang ada di masing-masing cluster
4. Alokasikan masing-masing data ke centroid/ rata-rata terdekat
5. Kembali ke langkah 3, apabila masih ada data yang berpindah cluster atau apabila perubahan nilai centroid, ada yang di atas nilai threshold yang ditentukan atau apabila perubahan nilai pada objective function yang digunakan di atas nilai threshold yang ditentukan.

Karakteristik *K-means*:

1. K-means sangat cepat dalam proses clustering.
2. K-means sangat sensitif pada pembangkitan centroid awal secara random.
3. Memungkinkan suatu cluster tidak mempunyai anggota.
4. Hasil clustering dengan K-means bersifat unik (selalu berubah-ubah, terkadang baik, terkadang jelek).

Analisis cluster dalam penelitian ini menggunakan metode non-hirarkis atau K-Means Cluster dengan menentukan jumlah cluster terlebih dahulu. Alasan digunakannya Metode Non-Hirarkis adalah untuk menghindari kelemahan pada metode hirarkis yaitu seringnya terdapat kesalahan pada data outlier, perbedaan ukuran yang jarak digunakan, dan variabel yang tidak relevan. Sedangkan keuntungan metode non-hierarki dapat melakukan analisis sampel dalam ukuran yang lebih besar dengan lebih efisien. Selain itu, non-hirarkis hanya memiliki sedikit kelemahan pada data outlier, ukuran jarak yang digunakan, dan variabel tak relevan atau variabel yang tidak tepat.

Metode Non-Hirarkis pengelompokan dimulai dengan menentukan terlebih dahulu jumlah cluster yang diinginkan (dua, tiga, empat, atau yang lain). Setelah jumlah cluster ditentukan kemudian dilakukan proses analisis cluster dengan tanpa mengikuti proses hirarki. Jika data yang digunakan memiliki variabilitas satuan maka harus dilakukan proses standarisasi/transformasi data terlebih dahulu dan sebaliknya, jika tidak mempunyai variabilitas satuan maka proses analisis cluster dapat langsung dilakukan tanpa terlebih dahulu melakukan proses

standarisasi/transformasi data (Rahayu, 2005). Metode analisis cluster dilakukan dengan program SPSS.

Langkah-langkah dalam analisis cluster menurut Malhotra (2006) terdiri dari :

1. Memformulasikan masalah.
2. Memilih ukuran jarak atau ukuran kesamaan.
3. Memilih prosedur pengelompokan.
4. Menetapkan jumlah kelompok.
5. Menafsirkan dan membuat profil kelompok.
6. Melakukan kajian validitas pengelompokan.

Selanjutnya proses analisis cluster menurut Talakua et al (2017) dijelaskan sebagai berikut :

1. Menetapkan ukuran jarak antar data. Mengukur kesamaan antara objek. Sesuai prinsip dasar cluster yang mengeelompokan objek yang mempunyai kemiripan, maka proses pertama adalah mengukur seberapa jauh ada kesamaan objek. Pengukuran jarak yang populer adalah metode *Euclidean distance*. Pada dasarnya, cara ini akan memasukan sebuah data ke dalam cluster tertentu dengan mengukur jarak data tersebut ke pusat cluster.
2. Melakukan proses standarisasi data jika diperlukan.
3. Melakukan pengklasteran.

Proses inti clustering adalah pengelompokan data, yang biasa dilakukan dengan dua metode yaitu:

a) Metode Hierarki

Metode ini memulai pengelompokan dengan dua atau lebih objek yang mempunyai kesamaan paling dekat. Kemudian proses diteruskan ke objek lain yang mempunyai kedekatan kedua. Demikian seterusnya sehingga cluster akan membentuk semacam pohon dimana ada hierarki (tingkatan yang jelas) antara objek. Dendogram biasanya digunakan untuk membantu memperjelas proses hierarki tersebut.

b) Metode Non hierarki

Berbeda dengan metode hierarki, metode ini justru di mulai dengan menentukan terlebih dahulu jumlah cluster yang diinginkan. Setelah jumlah cluster diketahui, baru proses cluster dilakukan

tanpa mengikuti proses hierarki. Metode ini juga disebut K-Means cluster.

4. Melakukan penanaman cluster-cluster yang terbentuk.

5. Melakukan validasi dan profiling cluster.

Adapun ciri-ciri cluster adalah homogenitas (kesamaan) yang tinggi antara anggota dalam satu cluster (*within cluster*) dan heterogenitas (perbedaan) yang tinggi antara cluster yang satu dengan cluster lain (*between-cluster*).

Analisis cluster memiliki beberapa istilah penting, antara lain:

- 1) *Distances between cluster centers*, yaitu jarak yang menunjukkan bagaimana terpisahnya pasangan individu cluster.
- 2) Keanggotaan cluster (*cluster membership*), ialah keanggotaan yang menunjukkan cluster untuk setiap objek yang menjadi anngotanya
- 3) Pusat cluster (*cluster centers*), ialah titik awal di mulai pengelompokan dalam cluster nonhierarki.
- 4) Rata-rata lama cluster (*cluster centroid*), ialah nilai rata-rata variabel dari semua objek atau observasi dalam cluster tertentu.
- 5) Skedul aglomerasi (*agglomeration schedule*), ialah jadwal yang memberikan informasi tentang objek atau kasus yang dikelompokan pada setiap tahap pada suatu proses analisis cluster yang hierarki.