

# Modul ANALISIS DAN INTERPRETASI DATA

Pelatihan Jabatan Fungsional  
Analisis Data Ilmiah



Fungsi Layanan Pengembangan Kompetensi Kedinasan

**Penanggung Jawab:**

1. Edy Giri Racman Putra, Ph.D.
2. Nining Setyowati Dwi Andayani, S.E., M.M.
3. Raden Arthur Ario Lelono, Ph.D.
4. Alpha Fadila Juliana Rahman, S.Pd., M.Pd.

**Tim Penyusun Modul:**

1. Rizka Rahmaida, M.Si.
2. Shiddiq Sugiono S.Si., M.Si
3. Anggih Tangkas Wibowo S.T., MMSi
4. Naily Kamaliah, M.Si.

**Diterbitkan oleh:**

Direktorat Pengembangan Kompetensi - BRIN  
Gedung B.J. Habibie, Jalan M.H. Thamrin Nomor 8,  
Jakarta Pusat 10340

**Diterbitkan pertama kali tahun 2022**

# KATA PENGANTAR

Pengembangan kompetensi Aparatur Sipil Negara (ASN) khususnya Pegawai Negeri Sipil (PNS) dalam mengembangkan karier jabatan fungsionalnya menjadi suatu tuntutan sehingga mampu menjalankan tugas dan fungsinya dengan sebaik – baiknya sehingga mampu memberikan kontribusi yang nyata terhadap bangsa dan Negara Kesatuan Republik Indonesia (NKRI).

Badan Riset dan Inovasi Nasional (BRIN) sebagai Instansi pembina 11 (sebelas) jabatan fungsional yang meliputi peneliti, perekayasa, pengembangan teknologi nuklir, analis perkebunrayaan, analis pemanfaatan iptek, analis data ilmiah, kurator koleksi hayati, penata penerbitan ilmiah, teknisi perkebunrayaan, teknisi penelitian dan perekayasaan, dan pranata nuklir. BRIN melalui kedeputian Sumber Daya Manusia Ilmu Pengetahuan dan Teknologi (SDMI) bertanggung jawab dalam penyelenggaraan pengembangan kompetensi 11 (sebelas) jabatan fungsional tersebut.

Kedeputian SDMI – BRIN melalui Direktorat Pengembangan Kompetensi sebagai penyelenggara pengembangan kompetensi jabatan fungsional bertanggung jawab dalam menyiapkan kebutuhan tersebut baik berupa pengelolaan pembelajaran, fasilitator, modul, bahan ajar dan sebagainya, yang merujuk kepada regulasi peraturan BRIN nomor 28 tahun 2022 tentang pedoman pelatihan pembentukan jabatan fungsional peneliti, peraturan BRIN nomor 29 tahun 2022 tentang pedoman pelatihan jabatan fungsional kurator koleksi hayati, peraturan BRIN nomor 30 tahun 2022 tentang pedoman pelatihan jabatan fungsional analis pemanfaatan iptek, peraturan BRIN nomor 31 tahun 2022 tentang pedoman pelatihan jabatan fungsional analis data ilmiah, peraturan BRIN nomor 32 tahun 2022 tentang pedoman pelatihan jabatan fungsional analis perkebunrayaan, peraturan BRIN nomor 33 tahun 2022 tentang pedoman pelatihan jabatan fungsional teknisi perkebunrayaan, dan peraturan BRIN nomor 34 tahun 2022 tentang pedoman pelatihan jabatan fungsional penata penerbitan ilmiah.

Kami mengucapkan syukur ke hadirat Tuhan Yang Maha Esa, atas berkat rahmat-Nya, modul pelatihan jabatan fungsional analis data ilmiah yang berjudul “ANALISIS

DAN INTERPRETASI DATA” dapat diselesaikan tepat waktu. Modul ini digunakan dalam pelatihan jabatan fungsional yang dibina oleh BRIN yang diselenggarakan oleh Kedeputian SDMI - BRIN melalui Direktorat Pengembangan Kompetensi. Kami berharap modul ini dapat memberikan manfaat dan kontribusi dalam meningkatkan dan mengembangkan kompetensi jabatan fungsional yang dibina BRIN.

Jakarta, Desember 2022

Plt. Deputi Sumber Daya Manusia  
Ilmu Pengetahuan dan Teknologi  
Badan Riset dan Inovasi Nasional

*(Tanda tangan)*

Edy Giri Rachman Putra, Ph.D.

# DAFTAR ISI

|                                                       |     |
|-------------------------------------------------------|-----|
| <b>KATA PENGANTAR</b> .....                           | iii |
| <b>DAFTAR ISI</b> .....                               | v   |
| <b>DAFTAR GAMBAR</b> .....                            | vii |
| <b>DAFTAR TABEL</b> .....                             | ix  |
| <br><b>PENDAHULUAN</b> .....                          | 1   |
| A. Deskripsi Singkat .....                            | 1   |
| B. Alokasi Waktu .....                                | 1   |
| C. Tujuan Pembelajaran.....                           | 1   |
| D. Materi Pokok.....                                  | 2   |
| <br><b>MATERI POKOK 1: PROSES ANALISIS DATA</b> ..... | 3   |
| A. Pengantar Analisis Data .....                      | 3   |
| B. Analisis Data Sebagai Proses.....                  | 8   |
| C. Analisis Data Dalam Organisasi .....               | 12  |
| D. Studi Kasus Analisis Data.....                     | 13  |
| E. Rangkuman .....                                    | 18  |
| F. Evaluasi .....                                     | 18  |
| <br><b>MATERI POKOK 2: ANALISIS DESKRIPTIF</b> .....  | 19  |
| A. Pengantar Analisis Deskriptif .....                | 19  |
| B. Jenis Data dan Skala Pengukuran.....               | 20  |
| C. Deskripsi Data Univariat: Grafik.....              | 24  |
| D. Deskripsi Data Univariat: Numerik .....            | 31  |
| E. Analisis Deskriptif untuk Dua Variabel .....       | 38  |
| F. Grafik yang Efektif .....                          | 41  |
| G. Rangkuman .....                                    | 42  |
| H. Evaluasi .....                                     | 42  |
| <br><b>MATERI POKOK 3 VISUALISASI DATA</b> .....      | 44  |
| A. Pengantar Visualisasi Data.....                    | 44  |
| B. Jenis Visualisasi Data .....                       | 45  |
| C. Memilih visualisasi data yang efektif.....         | 50  |

|                                                                          |           |
|--------------------------------------------------------------------------|-----------|
| D. Rangkuman .....                                                       | 56        |
| E. Evaluasi .....                                                        | 56        |
| <b>MATERI POKOK 4 : INTERPRETASI HASIL ANALISIS .....</b>                | <b>57</b> |
| A. Pengantar Interpretasi Hasil Analisis .....                           | 57        |
| B. Tahap Interpretasi Hasil Analisis .....                               | 61        |
| C. Hirarki Pengetahuan .....                                             | 63        |
| D. Studi Kasus Interpretasi Hasil Analisis Data .....                    | 65        |
| E. Rangkuman .....                                                       | 68        |
| F. Evaluasi .....                                                        | 68        |
| <b>MATERI POKOK 5: PRAKTIK VISUALISASI DATA .....</b>                    | <b>69</b> |
| A. Pengenalan Aplikasi Tableau untuk Analisis dan Visualisasi Data ..... | 69        |
| B. Analisis Deskriptif Univariat .....                                   | 75        |
| C. Analisis Deskriptif Lebih dari Satu Variabel .....                    | 85        |
| D. Rangkuman .....                                                       | 90        |
| E. Evaluasi .....                                                        | 90        |
| <b>DAFTAR PUSTAKA .....</b>                                              | <b>91</b> |

## DAFTAR GAMBAR

|                                                                                                             |    |
|-------------------------------------------------------------------------------------------------------------|----|
| <b>Gambar 1.</b> Karakteristik Analisis Data antara Nilai yang ditawarkan dengan kecanggihan komputasi..... | 5  |
| <b>Gambar 2.</b> Proses Linier Analisis Data .....                                                          | 9  |
| <b>Gambar 3.</b> Proses Sirkuler Analisis Data.....                                                         | 11 |
| <b>Gambar 4.</b> Hasil Contoh Analisis Data Deskriptif.....                                                 | 15 |
| <b>Gambar 5.</b> Contoh Penyajian Hasil Analisis Data.....                                                  | 16 |
| <b>Gambar 6.</b> Jenis data berdasarkan skala pengukurannya.....                                            | 24 |
| <b>Gambar 7.</b> Contoh diagram kue.....                                                                    | 26 |
| <b>Gambar 8.</b> Contoh diagram batang .....                                                                | 27 |
| <b>Gambar 9.</b> Perbandingan dua histogram .....                                                           | 29 |
| <b>Gambar 10.</b> Distribusi unimodal dan bimodal .....                                                     | 30 |
| <b>Gambar 11.</b> Histogram dengan data pencilan .....                                                      | 31 |
| <b>Gambar 12.</b> Sebaran data dan titik tumpu .....                                                        | 33 |
| <b>Gambar 13.</b> Hubungan mean dan median, pada bentuk Distribusi .....                                    | 33 |
| <b>Gambar 14.</b> Posisi kuartil pada kurva distribusi data .....                                           | 37 |
| <b>Gambar 15.</b> Contoh Boxplot .....                                                                      | 38 |
| <b>Gambar 16.</b> Contoh diagram batang tidak berkelompok (kiri) dan berkelompok (kanan) .....              | 39 |
| <b>Gambar 17.</b> Boxplot dua kelompok.....                                                                 | 40 |
| <b>Gambar 18.</b> Histogram dua kelompok .....                                                              | 40 |
| <b>Gambar 19.</b> Contoh diagram pecar (Agresti, dkk, 2017) .....                                           | 41 |
| <b>Gambar 20</b> Visualisasi temporal .....                                                                 | 46 |
| <b>Gambar 21.</b> Visualisasi Hierarki.....                                                                 | 47 |
| <b>Gambar 22.</b> Visualisasi data jaringan .....                                                           | 48 |
| <b>Gambar 23.</b> Visualisasi Data Multidimensi .....                                                       | 49 |
| <b>Gambar 24.</b> Visualisasi Data Geospasial .....                                                         | 50 |
| <b>Gambar 25.</b> Grafik tiga dimensi (3D) dan Grafik dua dimensi (2D).....                                 | 51 |

|                                                                                          |    |
|------------------------------------------------------------------------------------------|----|
| <b>Gambar 26.</b> <i>Value encoding attribute</i> dalam visualisasi data .....           | 52 |
| <b>Gambar 27.</b> Grafik line chart dan Grafik heatmap .....                             | 53 |
| <b>Gambar 28.</b> Perbedaan Grafik dengan rentang sumbu titik nol .....                  | 58 |
| <b>Gambar 29.</b> Teknik visualisasi data dengan skala linier dan skala logaritmik ..... | 58 |
| <b>Gambar 30.</b> Proses Interpretasi Hasil Analisis.....                                | 58 |
| <b>Gambar 31.</b> Piramida Data-Information-Knowledge-Wisdom (DIKW) .....                | 64 |
| <b>Gambar 32 .</b> Piramida Data Information-Knowledge-Wisdom (DIKW).....                | 66 |
| <b>Gambar 33.</b> Hasil Analisis Deskriptif berupa grafik batang .....                   | 77 |
| <b>Gambar 34.</b> Hasil Analisis Deskriptif berupa stacked bars .....                    | 78 |
| <b>Gambar 35.</b> Hasil Analisis Deskriptif berupa grafik grouped bar/column chart ....  | 79 |
| <b>Gambar 36.</b> Hasil Analisis Deskriptif berupa grafik garis .....                    | 80 |
| <b>Gambar 37.</b> Hasil Analisis Deskriptif berupa grafik boxplot .....                  | 82 |
| <b>Gambar 38.</b> Hasil Analisis Deskriptif berupa grafik boxplot .....                  | 83 |
| <b>Gambar 39.</b> Hasil Analisis Deskriptif berupa diagram heatmap.....                  | 84 |
| <b>Gambar 40.</b> Hasil Analisis Deskriptif berupa grafik dual combination.....          | 85 |
| <b>Gambar 41.</b> Hasil Analisis Deskriptif berupa diagram pencar/scatter plot .....     | 86 |
| <b>Gambar 42.</b> Hasil Analisis Deskriptif berupa matrix scatter plot .....             | 87 |
| <b>Gambar 43.</b> Hasil Analisis Deskriptif berupa bubble chart.....                     | 88 |

## DAFTAR TABEL

|                                                                          |    |
|--------------------------------------------------------------------------|----|
| <b>Tabel 1.</b> Pengertian Analisis Data.....                            | 4  |
| <b>Tabel 2.</b> Perbedaan Analisis Data Kuantitatif dan Kualitatif ..... | 7  |
| <b>Tabel 3.</b> Contoh Hasil Pengoleksian Data.....                      | 14 |
| <b>Tabel 4.</b> Tingkat Pendapatan .....                                 | 25 |
| <b>Tabel 5.</b> Pengkategorian Tingkat Pendapatan (2).....               | 25 |
| <b>Tabel 6.</b> Interpretasi Nilai Statistik Data Kuantitatif .....      | 59 |
| <b>Tabel 7.</b> Interpretasi Pada Jenis Data Kuantitatif .....           | 60 |
| <b>Tabel 8.</b> Perbedaan Tableau Public dan Tableau Desktop .....       | 70 |

# PENDAHULUAN

## A. Deskripsi Singkat

Mata diklat ini menjelaskan tentang Proses Analisis Data, Analisis Deskriptif, Visualisasi Data, Interpretasi Hasil Analisis, dan Praktik Visualisasi Data.

## B. Alokasi Waktu

Pelatihan Jabatan Fungsional Analis Data Ilmiah dapat diselenggarakan dengan tiga metode. Setiap metode memiliki alokasi waktu yang berbeda. Berikut adalah alokasi waktu pembelajaran untuk mata pelatihan Analisis dan Interpretasi Data.

| Metode     | On Kampus | Asinkronus | Sinkronus | Total |
|------------|-----------|------------|-----------|-------|
| Klasikal   | 8 JP      | -          | -         | 8 JP  |
| Bauran     | 3 JP      | 3 JP       | 3 JP      | 9 JP  |
| Jarak Jauh | -         | 3 JP       | 6 JP      | 9 JP  |

## C. Tujuan Pembelajaran

### 1. Hasil Belajar

Setelah mengikuti pembelajaran ini, peserta mampu melakukan analisis dan interpretasi data dengan tepat.

### 2. Indikator Hasil Belajar

Setelah selesai pembelajaran diharapkan peserta mampu:

- Merancang proses analisis data dengan benar;
- Menerapkan berbagai metode analisis deskriptif, dengan tepat
- Memilih dan memaksimalkan penyajian data dengan visualisasi data yang sesuai
- Melakukan interpretasi dari hasil analisis data dengan benar
- Membuat visualisasi data dengan aplikasi tableau dengan benar

#### **D. Materi Pokok**

Mata pelatihan ini terdiri dari 5 (Lima) Materi Pokok, yaitu:

1. Proses Analisis Data
  - a. Pengantar Analisis Data
  - b. Analisis Data sebagai Proses
  - c. Analisis Data dalam Organisasi
  - d. Studi Kasus Analisis Data
2. Analisis Deskriptif
  - a. Pengantar Analisis Deskriptif
  - b. Jenis Data dan Skala Pengukuran
  - c. Deskripsi Data Univariat: Grafik
  - d. Deskripsi Data Univariat: Numerik
  - e. Analisis Deskriptif untuk Dua Variabel
  - f. Grafik yang Efektif
3. Visualisasi Data
  - a. Pengantar Visualisasi Data
  - b. Jenis Visualisasi Data
  - c. Memilih visualisasi data yang efektif
4. Interpretasi Data
  - a. Pengantar Interpretasi Hasil Analisis
  - b. Tahap Interpretasi Hasil Analisis
  - c. Hirarki Pengetahuan
  - d. Studi Kasus Interpretasi Hasil Analisis Data
5. Praktik Visualisasi Data
  - a. Pengenalan Aplikasi Tableau untuk Analisis dan Visualisasi Data
  - b. Analisis Deskriptif Univariate
  - c. Analisis Deskriptif Lebih dari Satu Variabel

## MATERI POKOK 1:

# PROSES ANALISIS DATA

Indikator Hasil Belajar:

Peserta mampu **merancang proses analisis data dengan benar**, meliputi:

1. Memahami konsep analisis data dengan benar;
2. Menjelaskan proses analisis data dengan benar;
3. Menyebutkan pentingnya analisis data dalam organisasi;
4. Merancang proses pemilihan analisis data sesuai dengan kebutuhan.

### A. Pengantar Analisis Data

Data merupakan sebuah sumber daya yang harus dimanfaatkan oleh setiap pihak karena mampu memberikan informasi yang berguna sebagai dasar pengambilan keputusan atau kesimpulan. Analisis data dalam hal ini memiliki peran yang penting sebagai alat yang memungkinkan seseorang atau organisasi untuk mengambil informasi yang berguna dari suatu kumpulan data. Pemahaman yang baik mengenai analisis data dapat memengaruhi kualitas kesimpulan yang ditarik dari suatu kumpulan data.

Saat ini cukup banyak ditemukan istilah-istilah yang menyerupai analisis data, sebagai contoh interpretasi data, data analytics, visualisasi data dan lain-lain. Pendefinisian yang tepat akan memberikan pemahaman yang baik mengenai fungsi dari masing-masing istilah tersebut. Tabel 1 menyajikan berbagai definisi istilah analisis data dari berbagai sumber sebagai referensi untuk membentuk definisi analisis data yang komprehensif:

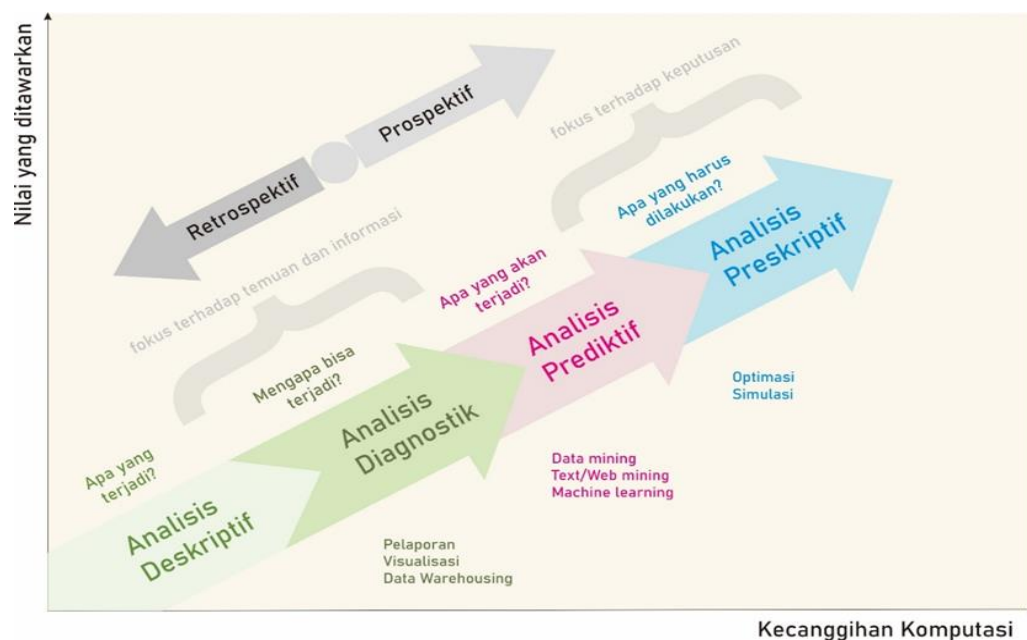
**Tabel 1.** Pengertian Analisis Data

| No | Definisi                                                                                                                                                                                         | Sumber                  | Kata Kunci                                                               |
|----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------|--------------------------------------------------------------------------|
| 1  | Penggunaan metode dan alat analisis yang tepat untuk memproses dan menganalisis data, mengambil informasi yang bernilai, dan membentuk kesimpulan yang efektif berdasarkan tujuan dari analisis. | Li et al. (2020)        | Memproses data, mengambil informasi, dan membentuk kesimpulan            |
| 2  | Proses dari riset yang terperinci dan proses meringkas data untuk mengambil informasi yang berguna dalam membentuk kesimpulan.                                                                   | Bai et al. (2019)       | Proses riset, meringkas data, mengambil informasi, membentuk kesimpulan  |
| 3  | Sebuah proses dalam mengembangkan jawaban atas pertanyaan melalui pemeriksaan dan interpretasi data.                                                                                             | Sharma (2018)           | Menjawab Pertanyaan, Pemeriksaan Data, Interpretasi Data                 |
| 4  | Proses memeriksa, mentransformasi dan menyusun kumpulan data melalui cara tertentu untuk mengkaji bagian individual dan mengambil informasi yang berguna.                                        | Vali Rajasekaran (2020) | & Memeriksa data, mentransformasi data, mengambil informasi yang berguna |

Berdasarkan empat definisi analisis data pada tabel 1, secara umum analisis data dapat didefinisikan sebagai sebuah proses untuk menarik kesimpulan dari sebuah pertanyaan melalui interpretasi informasi yang diambil dari suatu data yang dikumpulkan dan dianalisis dengan metode tertentu. Adapun secara singkatnya analisis data adalah rangkaian proses untuk mengubah data menjadi informasi. Dengan adanya fenomena ledakan data yang terjadi saat ini maka analisis data menjadi suatu proses yang perlu dikuasai oleh setiap orang maupun organisasi sehingga semakin banyak informasi berharga yang dihasilkan.

## 1. Analisis Data Berdasarkan Tujuan

Proses analisis data pada dasarnya memiliki berbagai jenis dan pendekatan sesuai dengan tujuan yang ingin dicapai. Masing-masing jenis analisis data memiliki manfaat dan metode analisis yang melekat sehingga penting untuk mengetahui kapan baiknya suatu jenis analisis data digunakan. Secara umum terdapat empat jenis analisis data yang dapat digunakan yakni analisis deskriptif, analisis diagnostik, analisis prediktif dan analisis preskriptif (Banerjee et al., 2013; Delen & Ram, 2018).



Sumber: Delen & Ram (2018)

**Gambar 1.** Karakteristik Analisis Data antara Nilai yang ditawarkan dengan kecanggihan komputasi

Gambar 1 menjelaskan empat jenis analisis data berdasarkan tujuan yang ingin dicapai dalam analisis data. Secara umum setiap analisis memiliki fungsi untuk menjawab jenis pertanyaan tertentu. Berikut adalah penjelasan dari masing-masing jenis analisis data (Banerjee et al., 2013; Delen & Ram, 2018):

### a. Analisis Data Deskriptif

Analisis deskriptif merupakan tingkatan awal dari taksonomi analisis data. Analisis ini bertujuan untuk menggambarkan suatu fenomena melalui berbagai pengukuran yang dapat ditangkap melalui dimensi-dimensi tertentu. Secara sederhana jenis analisis ini mencoba untuk membuka apa yang terjadi atau

mewaspadaai sesuatu yang akan terjadi. Dalam ranah bisnis, spektrum analisis deskriptif antara lain tangkapan dari transaksi bisnis dalam waktu tertentu, penyampaian indikator performa kinerja secara dinamis/*dashboard*, dan lain sebagainya. Berdasarkan gambar 1, analisis deskriptif merupakan jenis analisis data yang membutuhkan kecanggihan komputasi paling rendah dan juga memiliki nilai hasil analisis yang paling rendah.

**b. Analisis Data Diagnostik**

Analisis ini bertujuan untuk mengevaluasi mengapa sesuatu dapat terjadi. Adapun analisis ini membutuhkan metode analisis data eksploratori dan teknik visualisasi untuk menemukan akar permasalahan dari suatu masalah. Analisis terhadap hubungan sebab akibat merupakan salah satu metode yang secara umum dapat digunakan dalam jenis analisis ini. Baik analisis deskriptif maupun diagnostik, keduanya berjenis retrospektif dimana hasil yang didapat adalah gambaran mengenai fenomena yang telah terjadi.

**c. Analisis Data Prediktif**

Jenis analisis ini bertujuan untuk memprediksi potensi dari dampak di masa depan dan mencari faktor yang dapat memengaruhi suatu fenomena melalui teknik statistika dan data mining berdasarkan data historis. Analisis prediktif dapat dilakukan untuk variabel yang berifat kategorikal (klasifikasi), kontinu (regresi) maupun bergantung pada waktu/time dependent (time-series) (Delen & Ram, 2018). Adapun Sheng et al. (2021) menyampaikan bahwa terdapat tiga jenis analisis prediktif yakni: statistika inferensia, mesin pembelajar/machine learning, metode untuk menganalisis data tidak terstruktur. Singkatnya analisis prediktif digunakan untuk mengestimasi suatu fenomena di masa depan.

**d. Analisis Data Preskriptif**

Jenis analisis ini merupakan analisis yang berada setelah tiga analisis sebelumnya dimana bertujuan untuk menyarankan tindakan apa yang perlu diambil untuk mencapai suatu tujuan. Dengan kata lain analisis ini berikaitan

dengan alternatif keputusan dari hasil analisis prediktif. Analisis ini dapat menggunakan metode optimasi maupun simulasi. Dalam diagram pada gambar 1, analisis ini merupakan analisis yang membutuhkan kecanggihan komputer paling tinggi, namun begitu hasil analisis ini akan mendapatkan informasi yang paling bernilai. Analisis prediktif maupun preskriptif merupakan jenis analisis yang bersifat prospective dimana hasil analisis berfokus pada fenomena yang akan terjadi di masa depan.

## 2. Analisis Data Berdasarkan Jenis Data

Secara umum jenis data dapat dibedakan menjadi data kuantitatif dan data kualitatif. Data kuantitatif adalah data numerik hasil pengukuran yang dapat digunakan untuk mengkuantifikasi jawaban dari suatu pertanyaan. Adapun data kualitatif merujuk pada informasi non-numerik seperti teks, video dan suata yang dapat digunakan untuk mengambil temuan tentang sebuah topik atau membuat suatu ide baru untuk penelitian. Perbedaan jenis data tersebut turut menjadikan proses analisis keduanya berbeda. Berikut penjelasan secara umum mengenai analisis data kuantitatif dan analisis data kualitatif:

**Tabel 2.** Perbedaan Analisis Data Kuantitatif dan Kualitatif

|                      | Analisis Data Kuantitatif                                                                                                                                                                                                                                  | Analisis Data Kualitatif                                                                                                                                                                                            |
|----------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Kerangka kerja utama | <ul style="list-style-type: none"> <li>• Mengonfirmasi hipotesis tentang suatu fenomena</li> <li>• Gaya instrumen yang digunakan lebih ketat dalam mengkategorisasi jawaban pertanyaan</li> <li>• Menggunakan instrumen yang sangat terstruktur</li> </ul> | <ul style="list-style-type: none"> <li>• Mengeksplorasi fenomena</li> <li>• Gaya instrumen lebih fleksibel dan kategorisasi jawaban yang iteratif</li> <li>• Menggunakan instrumen yang semi terstruktur</li> </ul> |
| Tujuan Analitis      | <ul style="list-style-type: none"> <li>• Mengkuantifikasi variasi</li> <li>• Memprediksi hubungan sebab akibat</li> <li>• Menggambarkan karakteristik populasi</li> </ul>                                                                                  | <ul style="list-style-type: none"> <li>• Menggambarkan variasi</li> <li>• Menggambarkan hubungan sebab akibat</li> <li>• Menggambarkan pengalaman pribadi</li> <li>• Menggambarkan norma kelompok</li> </ul>        |
| Format pertanyaan    | <i>Close-ended</i>                                                                                                                                                                                                                                         | <i>Open-ended</i>                                                                                                                                                                                                   |

| Format Data                | Numerik                                                                                                                                                                                                                                                   | Tekstual                                                                                                                                                                                                                                                                  |
|----------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Fleksibilitas dalam desain | <ul style="list-style-type: none"> <li>• Desain studi stabil dari awal hingga akhir</li> <li>• Respon responden tidak mengubah apa yang ditanyakan selanjutnya dalam kuesioner</li> <li>• Desain studi berjalan sesuai dengan asumsi statistik</li> </ul> | <ul style="list-style-type: none"> <li>• Beberapa aspek dari studi bersifat fleksibel</li> <li>• Respon informan sangat berpengaruh terhadap apa yang ditanyakan selanjutnya</li> <li>• Desain studi bersifat iteratif tergantung informasi apa yang diperoleh</li> </ul> |

Sumber: Woodsong (2005)

#### a. Analisis Data Kuantitatif

Analisis data kuantitatif merujuk pada proses mengubah data menjadi informasi melalui penyajian angka dari hasil observasi untuk mendeskripsikan dan menjelaskan suatu fakta (Bilgin, 2017). Jenis analisis ini digunakan metode statistika untuk mengolah data numerik dan menentukan sebuah kesimpulan. Analisis data kuantitatif lebih bersifat kaku dimana desain studi akan terus stabil dari awal hingga akhir. Secara umum terdapat dua metode yang digunakan dalam analisis data kuantitatif yakni analisis deskriptif dan analisis inferensia (humansofdata, 2018).

#### b. Analisis Data Kualitatif

Analisis data kuantitatif merujuk pada proses mengubah data menjadi informasi melalui penyajian angka dari hasil observasi untuk mendeskripsikan dan menjelaskan suatu fakta (Bilgin, 2017). Jenis analisis ini digunakan metode statistika untuk mengolah data numerik dan menentukan sebuah kesimpulan. Analisis data kuantitatif lebih bersifat kaku dimana desain studi akan terus stabil dari awal hingga akhir. Secara umum terdapat dua metode yang digunakan dalam analisis data kuantitatif yakni analisis deskriptif dan analisis inferensia (humansofdata, 2018).

### B. Analisis Data Sebagai Proses

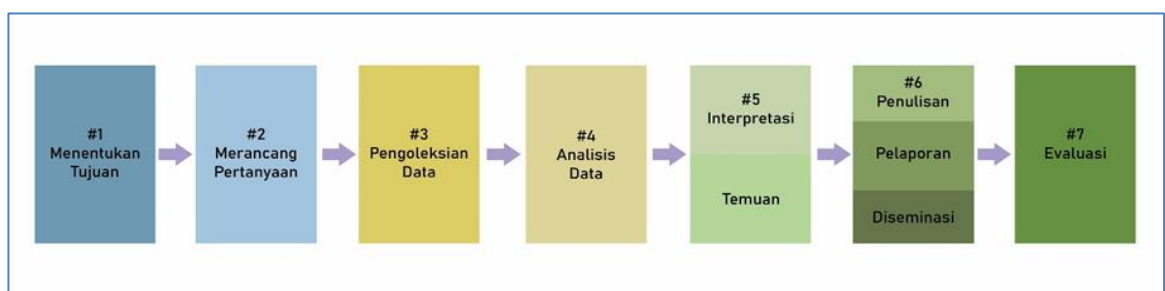
Pemahaman mengenai definisi analisis data harus dilihat melalui

perspektif proses/mekanisme. Hal ini didukung dari definisi analisis data pada tabel 1 yang sebagian besar menjelaskan bahwa kegiatan analisis data merupakan sebuah rangkaian proses untuk mendapatkan informasi yang berguna. Penggambaran proses maupun mekanisme dalam suatu analisis data menjadi penting karena kegiatan tersebut pada dasarnya terdiri atas beberapa komponen yang saling berhubungan satu dengan lainnya. Richmond (2006) secara umum membagi proses analisis data menjadi 2 jenis yakni proses secara linier dan proses secara sirkuler. Berikut penjelasan dari masing-masing proses tersebut:

### 1. Proses Linier Analisis Data

Pendekatan proses analisis data secara linier adalah suatu pekerjaan analisis data dengan melewati komponen-komponen yang tersusun secara sekuensial dari awal hingga akhir. Manfaat yang diperoleh dari pendekatan ini adalah tersedianya proses yang terstruktur dan terorganisir dimana setiap langkah diatur dalam susunan yang sudah pasti. Adapun langkah ini dinilai cukup mudah untuk dipelajari sehingga dapat menjadi referensi bagi analis yang ingin mendalami analisis data. Namun begitu langkah yang dijalankan satu per satu memiliki kelemahan karena membatasi kekuatan untuk analisis dengan kata lain mengurangi fleksibilitas dan efektivitas proses.

Richmond (2006) menjelaskan bahwa sebagai sebuah proses, kegiatan analisis data perlu memiliki tujuh komponen sebagai berikut: menentukan tujuan, menyusun pertanyaan, pengoleksian data, prosedur metode analisis data, interpretasi, pelaporan dan evaluasi. Proses ini bersifat sekuensial dimana seseorang tidak boleh kembali kepada langkah sebelumnya jika sudah menyelesaikan suatu langkah.



Sumber: Richmond (2006)

**Gambar 2.** Proses Linier Analisis Data

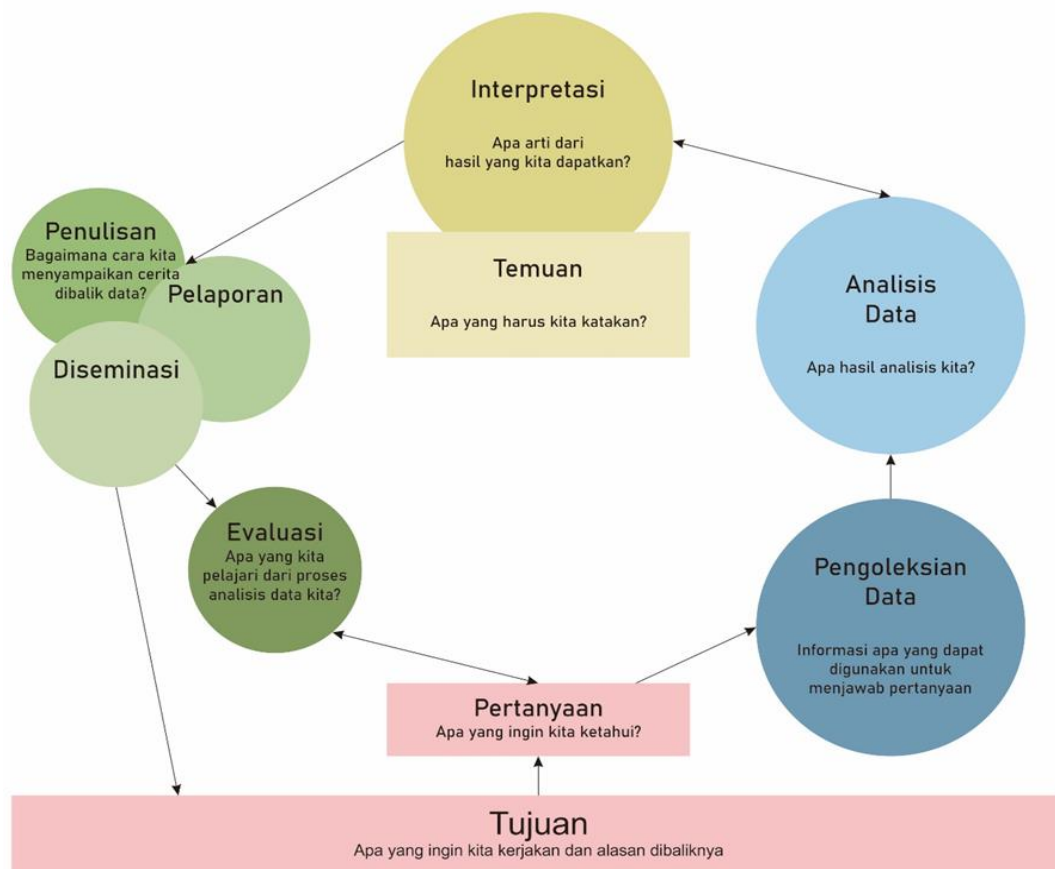
Berikut penjelasan dari setiap langkah pada gambar 2:

- a. Penentuan tujuan merupakan hal utama dalam analisis data karena kegiatan tersebut harus menjadi suatu proses yang tidak hanya bersifat strukturat tetapi fungsional dimana memberikan informasi dan nilai tambah bagi individu maupun organisasi. Pada fase ini kita harus mampu menentukan urgensi dari analisis data yang kita kerjakan sehingga mampu memunculkan implikasi diakhir proses.
- b. Merancang pertanyaan merupakan sebuah basis yang harus diformulasikan untuk mendapatkan hasil analisis data yang fungsional. Melalui pertanyaan yang baik maka dapat ditentukan data apa saja yang akan di kumpulkan dan jenis analisis yang diterapkan pada data tersebut.
- c. Pengoleksian data merupakan suatu langkah untuk mengumpulkan berbagai macam data yang diperlukan untuk menjawab pertanyaan melalui instrumen pengumpul data yang tepat.
- d. Langkah analisis data dalam proses ini merupakan suatu kumpulan prosedur atau metode yang dapat diaplikasikan kepada data yang telah dikumpulkan dalam rangka mendapatkan satu atau lebih hasil. Analisis data yang dilakukan dapat dilakukan dalam prosedur tertentu sesuai dengan pertanyaan yang akan dijawab.
- e. Terdapat dua langkah yang masuk dalam fase interpretasi data: pertama menyampaikan hasil analisis dan kedua adalah menyampaikan informasi tambahan yang memungkinkan orang lain memahami arti dari hasil analisis. Poses analisis data dapat menghasilkan banyak keluaran, tetapi temuan/*findings* yang disampaikan hanya sebagian kecil dari hasil analisis tersebut.
- f. Fase penulisan, pelaporan dan diseminasi merupakan proses dimana fase analisis,interpretasi data dan penarikan temuan sudah dilakukan. Penulisan laporan harus memiliki alur yang logis dan menyesuaikan dengan khalayak yang akan membaca. Fase ini merupakan kesempatan bagi analis untuk menyampaikan cerita dibalik data yang dianalisisnya.
- g. Evaluasi merupakan fase akhir dalam suatu proses analisis data dimana fase ini merupakan tahap untuk menguji kembali kesesuaian dan efektivitas

setiap fase yang diambil pada proses analisis data sehingga dapat menjadi masukan bagi pihak lain. Fase ini turut ditujukan untuk mengukur efektivitas pelaporan yang disampaikan kepada khalayak.

## 2. Proses Sirkuler Analisis Data

Pendekatan sirkuler dari proses analisis data memberikan fleksibilitas yang lebih dalam proses pembuatan keputusan dan juga memberikan lebih banyak jenis keputusan yang dapat dibuat. Pada pendekatan ini, suatu langkah yang berbeda dapat dikerjakan dalam waktu yang berbeda dan dalam urutan yang berbeda. Kelebihan proses ini memungkinkan seorang analis tidak terpaku dalam suatu langkah sehingga dapat memaksimalkan hasil dari langkah yang lainnya. Adapun proses ini mendukung pola belajar *learning by doing* serta meningkatkan kualitas dari suatu proses sebelum masuk ke proses selanjutnya.



Sumber: Richmond (2006)

**Gambar 3.** Proses Sirkuler Analisis Data

Secara umum proses sirkuler dari analisis data mengandung komponen-komponen yang sama seperti pada proses linier, mulai dari menentukan tujuan, menentukan pertanyaan hingga melakukan penulisan laporan/fase diseminasi. Hal yang berbeda pada proses ini adalah munculnya umpan balik/*feedback loop* pada 2 proses yakni antara analisis dengan interpretasi data dan evaluasi dengan penentuan pertanyaan. Fase interpretasi dan analisis data menjadi suatu proses yang berulang hingga analisis mendapatkan hasil yang sesuai dengan pertanyaan yang ditentukan. Begitu pula dengan fase antara evaluasi dengan penentuan pertanyaan dimana hasil dari proses analisis data bisa saja menimbulkan pertanyaan baru yang lebih mendalam sehingga perlu dilakukan analisis dari awal kembali.

### **C. Analisis Data Dalam Organisasi**

Pemanfaatan dan implementasi analisis data telah dilakukan oleh berbagai sektor organisasi dalam menentukan suatu keputusan, tak terkecuali sektor pemerintahan. Bagi suatu organisasi, data merupakan sumber daya yang perlu dimaksimalkan dalam menciptakan berbagai informasi yang dapat diimplementasikan atau menjadi sebuah instruksi.

Institusi Pemerintahan merupakan salah satu sektor organisasi yang mulai mengaplikasikan analisis data dalam menentukan sebuah keputusan. Terlebih dengan pemanfaatan teknologi digital, volume data yang tersimpan di sektor pemerintahan terus bertambah setiap waktunya. Hal ini menjadi kesempatan emas bagi pemerintah untuk mengambil informasi yang bermanfaat bagi pembentukan kebijakan maupun pekerjaan teknis lainnya. Berikut 2 contoh pemanfaatan analisis data yang sudah dilakukan oleh sektor pemerintah (Sirait, 2016).

#### **1. Pemerintah Kota Bandung**

Pemerintah Kota Bandung telah menerapkan *digital command center* di tahun 2015 dimana terobosan tersebut mengintegrasikan akses data dari 34 Satuan Kerja Perangkat Daerah/SKPD dan 30 Kecamatan sehingga dapat dievaluasi proses pengadaan barang dan jasa yang dilakukan oleh masing-masing instansi. Selain itu *digital command center* turut melakukan analisis

terhadap data-data di media sosial sehingga dapat diketahui topik-topik yang sedang dibicarakan oleh netizen mengenai Kota Bandung. Hal tersebut dapat memberikan evaluasi bagi Pemerintah Kota Bandung akan respon dari program atau kebijakan yang diterapkan kepada masyarakat di media digital.

## **2. Direktorat Jenderal Pajak/DJP Kementerian Keuangan**

Tujuan dari DJP melakukan analisis data adalah untuk meningkatkan penerimaan negara serta mengurangi kecurangan-kecurangan terkait pajak. Mereka menganalisis data dari media sosial untuk dicocokkan dengan laporan pajak dan data dari rekening tabungan. Adapun DJP melakukan analisa seperti Analisa Central Tax Analysis dan penerbitan faktur pajak. Dari hasil analisis data tersebut DJP mendapatkan berbagai macam temuan pelanggaran seperti: ditemukannya 3.563 kasus pajak senilai 32,7 Trilyun Rupiah dan ditemukannya 21.123 kasus faktur yang tidak berdasarkan transaksi sebenarnya senilai 6,2 Trilyun Rupiah.

## **D. Studi Kasus Analisis Data**

Analisis data menjadi hal yang sangat umum dilakukan oleh manusia modern dalam menentukan sebuah keputusan. Dalam contoh studi kasus ini akan dijelaskan bagaimana peran analisis data memberikan suatu informasi. Contoh kasus yang akan dijelaskan berasal dari ranah kesehatan dan menggunakan jenis analisis data deskriptif dan proses analisis yang linier. Berikut penjabaran dari contoh kasus analisis data bidang kesehatan:

Dinas Kesehatan Pemerintah Provinsi X sedang merancang program peningkatan kesehatan bagi kelompok usia anak-anak untuk tahun 2022. Dalam rangka membuat program yang tepat sasaran maka diperulkan suatu analisis data mengenai jenis penyakit apa yang paling banyak menjangkit kelompok usia anak-anak pada tahun sebelumnya (2021). Analisis data ini menjadi penting karena kelompok usia anak-anak merupakan generasi penerus yang dimana kesehatan menjadi faktor penting dalam mendukung seluruh proses belajar. Fase pada setiap analisis data perlu dirancang untuk mendapatkan hasil akhir yang sesuai dengan kebutuhan.

### 1. Menentukan Tujuan

Tujuan analisis data ini adalah untuk mengidentifikasi jenis penyakit yang paling banyak dialami anak-anak di Provinsi X pada tahun 2021. Berbagai definisi harus dipastikan dalam penentuan tujuan ini. Sebagai contoh definisi mengenai anak-anak adalah seseorang yang belum berusia 18 tahun, termasuk juga yang masih di dalam kandungan dan lain sebagainya.

### 2. Merancang Pertanyaan

Kasus penyakit apa yang paling banyak menjangkit anak-anak Provinsi X dalam rentang tahun 2021? Melalui pertanyaan ini maka dibutuhkan data yang berjenis numerik dimana satuannya adalah kasus. Pertanyaan penelitian penting untuk dirancang agar jenis data dan metode analisis dapat menjawab kebutuhan analisis data.

### 3. Pengoleksian Data

Data yang dikoleksi merupakan data sekunder yang diperoleh dari laporan profil kesehatan Provinsi X pada tahun 2021 maupun sumber-sumber kredibel lainnya. Data dapat ditampilkan dalam sebuah tabel yang menunjukkan kategori penyakit, jumlah kasusnya, dan berbagai variabel lainnya seperti unit geografis (kabupaten, kecamatan, dll).

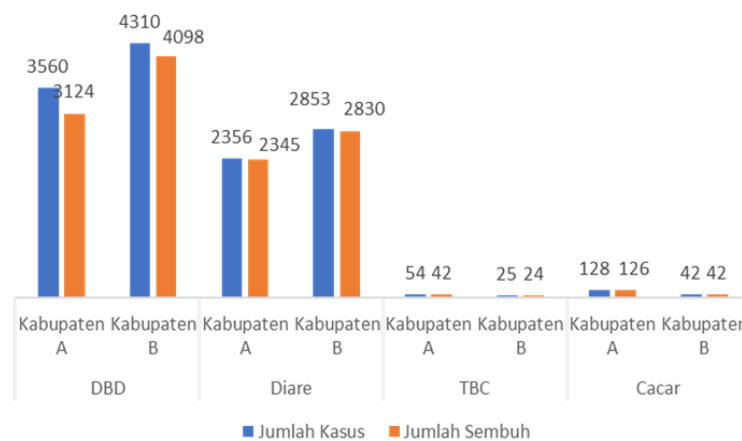
**Tabel 3.** Contoh Hasil Pengoleksian Data

| Penyakit | Jumlah Kasus | Kabupaten   | Jumlah Sembuh |
|----------|--------------|-------------|---------------|
| Diare    | 2.356        | Kabupaten A | 2.345         |
| TBC      | 54           | Kabupaten A | 42            |
| DBD      | 3.560        | Kabupaten A | 3.124         |
| Cacar    | 128          | Kabupaten A | 126           |
| Diare    | 2.853        | Kabupaten B | 2.830         |
| TBC      | 25           | Kabupaten B | 24            |
| DBD      | 4.310        | Kabupaten B | 4.098         |
| Cacar    | 42           | Kecamatan B | 42            |

Sumber: *Data dummy*

#### 4. Metode Analisis Data

Metode analisis data dilakukan secara deskriptif untuk melihat mana saja jenis penyakit yang paling banyak jumlahnya antara tahun 2021. Analisis data dapat dilakukan juga dengan memecah jumlah kasus per kabupaten. Berikut contoh visualisasi data yang digunakan untuk mendeskripsikan sebaran kasus dan jumlah pada setiap kabupaten.



Sumber: olahan data dummy

**Gambar 4.** Hasil Contoh Analisis Data Deskriptif

#### 5. Interpretasi

Berdasarkan analisis data *dummy*, kasus jenis penyakit yang paling banyak menjangkit anak-anak di Provinsi X, baik di Kabupaten A dan Kabupaten B, adalah Demam Berdarah Dengue/DBD. Temuan ini perlu dielaborasi lebih jauh dengan data pendukung lainnya sehingga memunculkan kesimpulan yang kuat bahwa pemberantasan DBD menjadi program utama di tahun 2022. Sebagai contoh, hasil analisis dapat diperkaya dengan hasil kinerja tim penanganan demam berdarah di setiap kabupaten, sudah sejauh apa efektivitas program pemberantasan demam berdarah, bagaimana tingkat kebersihan di setiap kabupaten dan lain lain. Berikut contoh interpretasi data yang dapat dibuat:

“DBD merupakan jenis penyakit yang paling banyak menjangkit anak-anak di Provinsi X pada tahun 2021, baik di Kabupaten A maupun di Kabupaten B. Hal ini didukung dengan rendahnya tingkat kebersihan lingkungan baik di Kabupaten A dan Kabupaten B pada tahun 2021. Begitupula penilaian kinerja tim tanggap DBD di Provinsi X memiliki nilai paling rendah dibanding

Provinsi yang lainnya. Hal ini harus menjadi fokus dinas kesehatan karena angka kematian karena DBD di Kabupaten A relatif tinggi jika dibandingkan dengan jenis penyakit lainnya. Pada akhirnya, analisis data ini menyimpulkan bahwa program pemberantasan DBD di Provinsi X merupakan program kesehatan utama yang harus dilaksanakan pada tahun 2022. Program ini harus didukung pula dengan peningkatan kualitas kebersihan lingkungan di setiap kabupaten serta mendisiplinkan petugas pemberantas DBD.”

## 6. Penulisan Laporan dan Diseminasi

Laporan dapat disajikan dalam suatu media cetak maupun media digital seperti buku, leaflet, bahan presentasi, maupun konten sosial media. Adapun pemilihan jenis laporan harus menyesuaikan dengan khalayak yang akan membacanya, apakah itu pemangku kepentingan dari dinas kesehatan, masyarakat umum, atau kelompok masyarakat usia muda. Sebagai contoh suatu pelaporan dapat dibuat dalam format infografis sehingga seluruh pemangku kepentingan dapat menerima informasi dengan mudah.



Sumber: indonesiabaik.id (2021)

**Gambar 5.** Contoh Penyajian Hasil Analisis Data

Gambar 5 adalah salah satu bentuk contoh penulisan laporan atau diseminasi menggunakan jenis konten infografis. Laporan analisis data tersebut tidak menampilkan data yang telah dikoleksi tetapi hasil perhitungan atau statistik-statistik tertentu seperti proporsi/persentase. Selain itu perlu adanya rekomendasi dari hasil analisis data sehingga hasil analisis data bersifat fungsional atau dapat memberikan jalan keluar bagi suatu masalah.

## **7. Evaluasi**

Evaluasi yang dilakukan adalah evaluasi terhadap setiap proses dari analisis data. Sebagai contoh, analisis data seharusnya menggunakan data seberapa banyak pasien yang dilayani oleh unit layanan kesehatan atau format pelaporan yang seharusnya lebih menarik sehingga masyarakat dapat memahami juga hasil analisis.

## **E. Rangkuman**

Analisis data merupakan suatu rangkaian proses dengan tujuan untuk mengambil informasi yang bernilai dalam suatu kumpulan data melalui metode tertentu. Secara umum terdapat empat jenis analisis data (deskriptif, diagnostik, prediktif dan preskriptif) dimana keempat jenis tersebut memiliki manfaat masing-masing. Suatu proses analisis data dapat bersifat linier maupun sirkuler. Jenis analisis sirkuler ditujukan agar diperoleh temuan maupun pertanyaan yang lebih mendalam dari suatu fenomena. Dalam suatu proses analisis data, dibutuhkan pemahaman dari tujuan analisis yang ingin dicapai dan metode yang tepat untuk mendapatkan informasi yang sesuai dengan kebutuhan atau bersifat fungsional. Adapun pemilihan jenis/format laporan dan gaya penulisan harus disesuaikan dengan khalayak yang akan membaca hasil analisis sehingga informasi dapat tersampaikan dengan baik.

## **F. Evaluasi**

1. Jelaskan yang dimaksud dengan analisis data.
2. Lakukan Evaluasi proses analisis data pada poin D.
3. Buatlah sebuah proses analisis data yang telah diimplementasikan dalam unit dimana Anda bekerja menggunakan skema analisis data linier sehingga diperoleh hasil untuk mengevaluasi kinerja.
4. Buatlah sebuah proses analisis data yang bisa diimplementasikan dalam unit dimana anda bekerja sehingga dapat memberikan nilai tambah bagi organisasi (Silakan menggunakan jenis proses analisis data linier).

## MATERI POKOK 2:

# ANALISIS DESKRIPTIF

### **Indikator Hasil Belajar:**

Peserta mampu **menerapkan berbagai metode analisis deskriptif dengan tepat**, meliputi:

1. Menjelaskan konsep, manfaat dan pentingnya analisis deskriptif dengan benar;
2. Menjelaskan jenis data kualitatif (kategorik) dan kuantitatif (numerik) serta skala pengukurannya, dengan benar;
3. Menjelaskan konsep dan jenis analisis deskriptif univariat: grafik, dengan benar;
4. Menjelaskan konsep dan jenis analisis deskriptif univariat: numerik, dengan benar;
5. Menjelaskan konsep dan jenis analisis deskriptif lebih dari satu variable, dengan benar.
6. Mengidentifikasi penggunaan grafik yang kurang tepat, dengan benar

### **A. Pengantar Analisis Deskriptif**

Statistika adalah sebuah ilmu yang mempelajari teknik-teknik pengumpulan, pengelolaan, analisis dan interpretasi data. Analisis dalam statistika dibagi menjadi dua, yaitu analisis deskriptif dan analisis inferensia. Analisis deskriptif adalah metode-metode yang berkaitan dengan pengumpulan dan penyajian suatu gugus data sehingga memberikan informasi yang berguna. Sedangkan analisis inferensia adalah metode yang berhubungan dengan analisis sebagian data untuk dapat melakukan penarikan kesimpulan terhadap keseluruhan data.

### **Ruang Lingkup**

Bagian ini akan berfokus pada pembahasan terkait analisis deskriptif. Analisis deskriptif hanya memberikan gambaran atau mendeskripsikan karakteristik atau sifat-sifat dari sekelompok data, baik data sampel maupun data

populasi. Statistika deskriptif tidak mampu melakukan generalisasi, yaitu menarik suatu kesimpulan umum terkait populasinya berdasarkan informasi data sampel dari suatu populasi.

Analisis deskriptif mencakup cara mengumpulkan data dan meringkas dan menyajikan data yang diperoleh. Selain itu, analisis deskriptif juga mencakup pengukuran pemusatan dan penyebaran data. Hal ini bermanfaat untuk memperoleh informasi terkait karakteristik data yang digunakan.

Analisis deskriptif memiliki beberapa manfaat sebagai berikut:

1. Data yang diperoleh dapat disajikan dengan ringkas dan rapi sehingga dapat memberikan informasi penting terkait data tersebut.
2. Berperan penting dalam persiapan analisis data inferensia yaitu sebagai tahap *exploratory data analysis*.

Analisis deskriptif digunakan untuk memberikan gambaran yang teratur, ringkas dan jelas, mengenai keadaan, peristiwa atau gejala tertentu sehingga dapat ditarik pengertian atau makna tertentu. Data yang telah diperoleh perlu dideskripsikan agar mudah dipahami dan agar dikenali pola sebarannya dan pola keteraturannya. Misalnya data luas provinsi atau jumlah pulau per provinsi di Indonesia. Jika data tidak disajikan secara deskriptif, akan sulit untuk mengetahui:

1. Provinsi mana yang paling luas?
2. Provinsi mana yang paling sempit?
3. Berapa rata-rata luas provinsi?
4. Berapa rata-rata jumlah pulau per provinsi?
5. Seberapa berbeda jumlah pulau dari satu provinsi ke provinsi yang lain?

## **B. Jenis Data dan Skala Pengukuran**

Sebelum melakukan proses analisis deskriptif dari suatu data, kita perlu memahami jenis data yang akan dianalisis. Hal ini penting karena jenis data menentukan metode analisis deskriptif yang tepat untuk digunakan. Dengan mengetahui jenis data, kita akan mudah menentukan metode analisis deskriptif yang tepat.

Data terbagi menjadi beberapa jenis. Berdasarkan tipe datanya, terdapat dua jenis data, yaitu data kuantitatif dan data kualitatif.

1. Kualitatif

Data kualitatif disebut juga dengan data kategorik. Penyebutan ini dikarenakan data kualitatif menggambarkan data yang sesuai dengan kategorinya.

2. Kuantitatif

Data kuantitatif disebut juga dengan data numerik. Istilah numerik digunakan karena data kuantitatif mewakili nilai numerik, misalnya seberapa banyak, seberapa sering, seberapa berat. Data kuantitatif memberikan informasi tentang jumlah, misalnya ukuran tinggi badan, berat badan, panjang jalan, dan sebagainya.

Sementara itu, berdasarkan skala pengukurannya, terdapat empat jenis data, yaitu nominal, ordinal, interval dan rasio.

1. Data Nominal

Data nominal ini merupakan hasil dari pengukuran skala nominal. Data nominal adalah data dengan beberapa kategori tertentu tanpa memperhatikan urutan. Kedudukan antar kategori dalam data nominal bersifat setara. Data nominal dapat juga diberi simbol berupa angka, tetapi angka tersebut hanya digunakan untuk membedakan setiap kategori dan tidak memiliki makna. Dengan kata lain, operasi aritmatika tidak dapat diterapkan pada data nominal. Contoh:

- a. Data jenis kelamin peserta pelatihan memiliki kategori laki-laki dan perempuan. Peserta laki-laki diberi simbol 1 sedangkan peserta perempuan diberi simbol 0.
- b. Data sektor pekerjaan peserta pelatihan terdiri dari sektor pemerintahan (ASN), perusahaan BUMN, dan perusahaan swasta. Peserta ASN diberi simbol 1, peserta dari perusahaan BUMN diberi simbol 2, dan peserta dari perusahaan swasta diberi simbol 3.

Pada contoh (a), jenis kelamin laki-laki (1) tidak berarti lebih tinggi kedudukannya dibandingkan jenis kelamin perempuan (0). Begitu pula pada contoh (b) yang berarti peserta dari perusahaan swasta (3) tidak lebih tinggi

kedudukannya dibandingkan dengan peserta ASN (1) dan peserta dari perusahaan BUMN (2).

## 2. Data Ordinal

Data ordinal merupakan data hasil pengukuran skala ordinal. Data ordinal adalah data yang cara penyusunan angkanya berdasarkan beberapa kategori, dengan memperhatikan urutan tertentu. Dengan kata lain, kategori dalam data ordinal memiliki kedudukan kategori yang tidak setara. Label, kode, atau simbol yang diberikan pada setiap kategori bermakna peringkat dari kategori tersebut. Peringkat ini menunjukkan urutan dalam penilaian. Data ordinal memiliki kesamaan dengan data nominal yaitu operasi aritmatika terhadap simbol kategori tidak berlaku.

Contoh data ordinal adalah sebagai berikut. Data tingkat pendidikan peserta pelatihan terdiri dari S1, S2, dan S3. Peserta lulusan S1 diberi simbol 1, peserta lulusan S2 diberi simbol 2, dan peserta lulusan S3 diberi simbol 3. Pada contoh tersebut, kategori tingkat pendidikan memiliki kedudukan yang tidak setara. Jika diurutkan dalam peringkat berdasarkan tingkat pendidikan yang lebih rendah, urutan kategorinya adalah S1 (1), S2 (2), dan S3 (3).

## 3. Data Interval<sup>1</sup>

Data interval merupakan data hasil pengukuran skala interval. Skala interval memiliki sifat yang dimiliki data nominal dan ordinal, tetapi perbedaan antar nilai dalam data ordinal dapat dikuantifikasi. Jenis data ini menunjukkan dua hal, yaitu urutan dan perbedaan antar nilai dalam data tersebut. Data interval dapat dijumlahkan tetapi tidak dapat dikalikan atau dibagi. Contoh data interval adalah suhu dalam satuan derajat celcius. Pada data suhu, 40 derajat Celcius tidak sama dengan dua kali 20 derajat celcius.

Data interval memiliki ciri-ciri di mana nilai nol pada data ini merupakan suatu nilai. Jika pada data ordinal, pengamatan bernilai nol berarti data tidak ada. Sedangkan pada data interval, nilai nol memiliki makna seperti nilai data yang lainnya. Pada contoh data suhu dalam ruangan, nilai nol derajat celcius

---

<sup>1</sup> <https://studyonline.unsw.edu.au/blog/types-of-data>

menunjukkan bahwa ruangan tersebut memiliki suhu yang bernilai nol derajat celcius.

Perbedaan dua nilai data interval juga dapat dikuantifikasi. Perbedaan antara 10 dan 20 derajat celcius sama dengan perbedaan antara 20 dan 30 derajat celcius.

#### 4. Data Rasio<sup>2</sup>

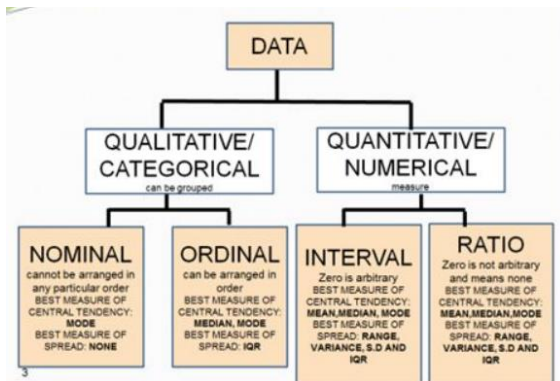
Data rasio merupakan data hasil pengukuran skala rasio. Skala rasio merupakan derajat skala data yang paling tinggi karena memiliki sifat yang paling lengkap jika dibandingkan dengan ketiga skala yang lainnya. Data rasio dapat dibedakan seperti data nominal. Data rasio juga dapat disusun berdasarkan urutan seperti data ordinal. Perbedaan antar nilai dalam data ini dapat dikuantifikasi seperti data interval. Selain dapat dijumlahkan, nilai data rasio juga dapat dikalikan. Contoh yang termasuk data rasio adalah berat, tinggi, dan jarak. Ketiga data tersebut dapat ditambahkan, dikurangkan, dan dikalikan.

Dalam hal sifat nilai nol, data rasio berbeda dengan data interval. Nilai nol pada data rasio berarti tidak ada nilai data. Misalnya, peserta dengan data berat badan nol kilogram berarti peserta tersebut tidak memiliki berat (tidak ada benda tersebut). Dengan kata lain, jika ada peserta X maka peserta tersebut tidak dapat (tidak mungkin) memiliki data berat badan nol kilogram.

Secara ringkas, pembagian data menurut jenis data dan skala pengukurannya dapat dilihat pada gambar 6.

---

<sup>2</sup> <https://studyonline.unsw.edu.au/blog/types-of-data>



Sumber: <https://medium.com/@rndayala/data-levels-of-measurement-4af33d9ab51a>  
**Gambar 6.** Jenis data berdasarkan skala pengukurannya

### C. Deskripsi Data Univariat: Grafik

Untuk menggambarkan sekumpulan data hasil pengukuran, terdapat dua metode yang umum digunakan yaitu metode grafik dan metode deskriptif numerik. Bagian ini akan membahas kedua metode tersebut.

#### Kategorisasi Data Numerik

Dalam suatu penelitian, terdapat tahap pengumpulan data. Setelah pengumpulan data selesai dilakukan, idealnya data disusun, disajikan, dan diperiksa menggunakan beberapa metode grafik. Aturan umum yang berlaku dalam deskripsi data univariat adalah data harus disusun ke dalam kategori sehingga setiap data dikategorikan ke dalam satu dan hanya satu kategori. Hal ini bertujuan untuk menghilangkan ambiguitas pada data. Misalnya, dalam pengkategorian pendapatan per bulan peserta pelatihan. Pendapatan per bulan merupakan data numerik. Selanjutnya data pendapatan setiap peserta dikategorikan ke dalam hanya satu tingkat pendapatan sehingga diperoleh data kategorik. Dengan kata lain, tidak ada peserta pelatihan yang dikategorikan ke dalam dua tingkat pendapatan.

**Tabel 4.** Tingkat Pendapatan

| Tingkat Pendapatan | Pendapatan Per Bulan            |
|--------------------|---------------------------------|
| 1                  | kurang dari Rp. 3.000.000       |
| 2                  | Rp. 3.000.000 s.d Rp. 3.999.999 |
| 3                  | Rp. 4.000.000 s.d Rp. 4.999.999 |
| 4                  | lebih dari Rp. 5.000.000        |

Berdasarkan Tabel 4 terlihat bahwa pendapatan per bulan dari setiap peserta pelatihan dikategorikan ke dalam hanya satu tingkat pendapatan. Namun, kadang kita juga menemukan tingkat pendapatan yang dikategorikan seperti pada Tabel 5.

Jika data tingkat pendapatan dikategorikan berdasarkan tabel 5, maka ada nilai yang sulit untuk dikategorikan. Sebagai contoh, peserta dengan pendapatan per bulan sebesar Rp. 4.000.000 dapat dikategorikan ke dalam tingkat pendapatan 2 atau 3.

**Tabel 5.** Pengkategorian Tingkat Pendapatan (2)

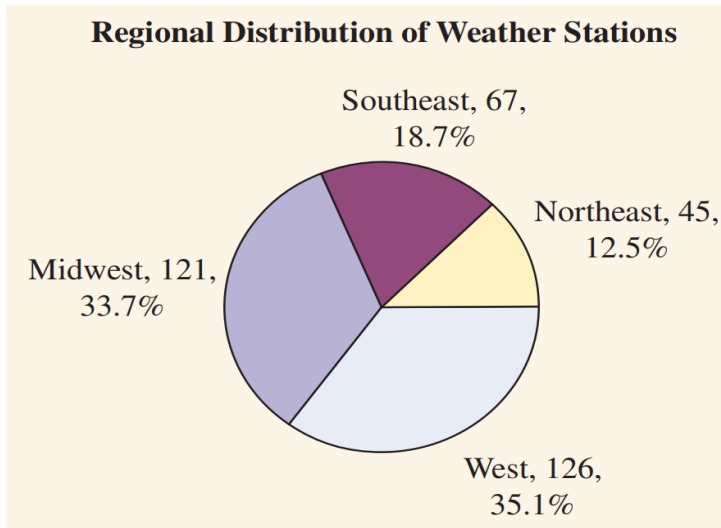
| Tingkat Pendapatan | Pendapatan Per Bulan             |
|--------------------|----------------------------------|
| 1                  | kurang dari Rp. 3.000.000        |
| 2                  | Rp. 3.000.000 s.d Rp. 4. 000.000 |
| 3                  | Rp. 4.000.000 s.d Rp. 5.000.000  |
| 4                  | lebih dari Rp. 5.000.000         |

Poin penting dalam penentuan kategori sekumpulan data adalah kategori yang didefinisikan harus menempatkan setiap data ke dalam satu kategori saja.

### **Diagram Kue (Pie Chart)**

Setelah data disusun ke dalam kategori, terdapat beberapa cara untuk menampilkan data menggunakan grafik. Grafik yang paling sederhana untuk data yang dikategorikan adalah diagram kue. Diagram ini digunakan untuk menunjukkan persentase dari jumlah data yang berada dalam setiap kategori terhadap jumlah total data. Diagram kue adalah lingkaran di mana terdapat satu

potongan kue untuk setiap kategori (Agresti, dkk., 2017). Setiap kategori digambarkan sebagai potongan kue dengan ukuran yang sesuai dengan persentase kategorinya (Gambar 7).



Sumber: Agresti, dkk (2017)

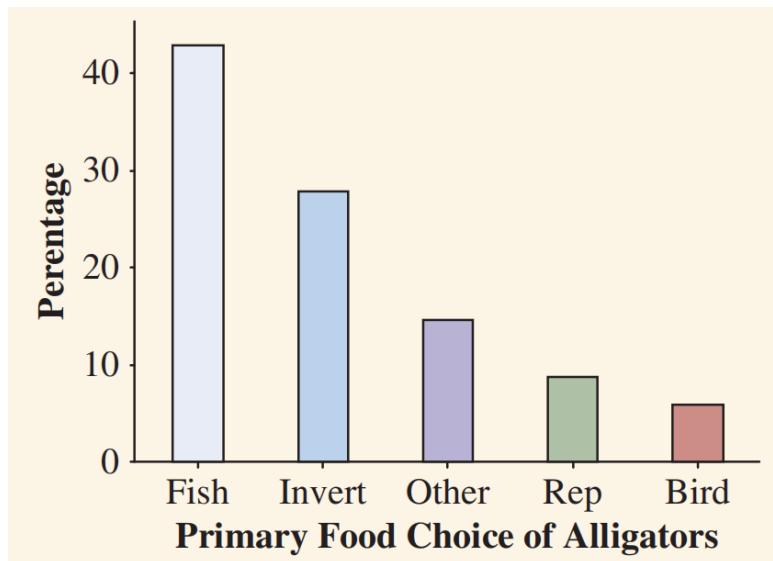
**Gambar 7.** Contoh diagram kue

Menurut Ott dan Longnecker (2010) terdapat beberapa aturan penyajian data menggunakan diagram kue yaitu sebagai berikut:

1. Kategori yang digunakan sebaiknya tidak terlalu banyak (maksimal sekitar 5 atau 6 kategori). Kategori yang terlalu banyak menyebabkan diagram kue sulit untuk diinterpretasikan.
2. Sebaiknya urutkan kategori dalam diagram kue berdasarkan urutan, baik itu menaik maupun menurun. Hal ini akan memudahkan pembaca dalam menginterpretasikan.

### **Diagram Batang (Bar Chart)**

Metode grafik yang kedua adalah diagram batang. Diagram batang menunjukkan batang vertikal untuk setiap kategori (Gambar 8). Tinggi batang menunjukkan persentase jumlah amatan dalam kategori tersebut (Agresti, dkk, 2017).



Sumber: Agresti, dkk 2017

**Gambar 8.** Contoh diagram batang

Ott dan Longnecker (2010) menyarankan langkah berikut untuk menghasilkan diagram batang yang baik:

1. Beri label yang menunjukkan frekuensi pada salah satu sumbu dan beri label yang menunjukkan kategori variabel tersebut pada sumbu lainnya.
2. Buatlah batang pada setiap kategori dengan tinggi yang sama dengan frekuensi pada kategori tersebut.
3. Beri jarak antar kategori. Hal ini berfungsi untuk menunjukkan dua kategori yang berbeda, memisahkan antar kategori, dan memperjelas tampilan diagram.

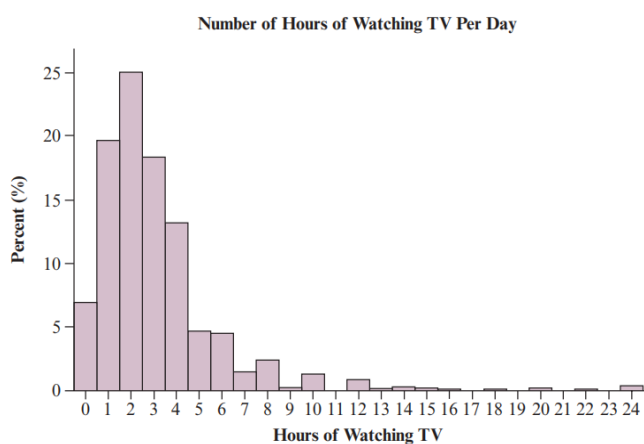
### Histogram

Histogram adalah grafik yang menggunakan batang untuk menggambarkan frekuensi atau frekuensi relatif dari setiap kemungkinan nilai untuk suatu data kuantitatif (Agresti, dkk, 2017).

Histogram secara sekilas dapat terlihat mirip seperti diagram batang. Namun perlu diingat bahwa histogram digunakan untuk menggambarkan data numerik. Sementara diagram batang digunakan untuk menggambarkan data kategorik. Pada data numerik diskrit, suatu histogram biasanya memiliki batang terpisah untuk setiap kemungkinan nilai. Pada data numerik kontinu, kita harus membagi

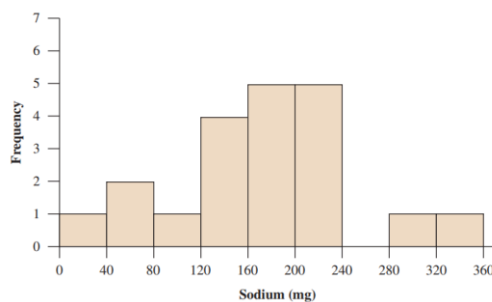
rentang nilai menjadi interval yang lebih kecil dengan jarak yang sama. Pembagian rentang nilai ke dalam interval yang lebih kecil juga dapat dilakukan pada data numerik diskrit jika memiliki banyak kemungkinan nilai. Setelah penentuan interval yang lebih kecil, kita dapat menghitung jumlah pengamatan dalam setiap interval. Tinggi batang pada histogram menggambarkan frekuensi (atau frekuensi relatif) dari pengamatan yang termasuk ke dalam masing-masing interval (Agresti, dkk, 2017).

Contoh histogram untuk data numerik diskrit:



| Hours | Count | Hours | Count |
|-------|-------|-------|-------|
| 0     | 90    | 13    | 2     |
| 1     | 255   | 14    | 4     |
| 2     | 325   | 15    | 3     |
| 3     | 238   | 16    | 1     |
| 4     | 171   | 17    | 0     |
| 5     | 61    | 18    | 1     |
| 6     | 58    | 19    | 0     |
| 7     | 19    | 20    | 2     |
| 8     | 31    | 21    | 0     |
| 9     | 3     | 22    | 1     |
| 10    | 17    | 23    | 0     |
| 11    | 0     | 24    | 5     |
| 12    | 11    |       |       |

Contoh histogram untuk data numerik kontinu:



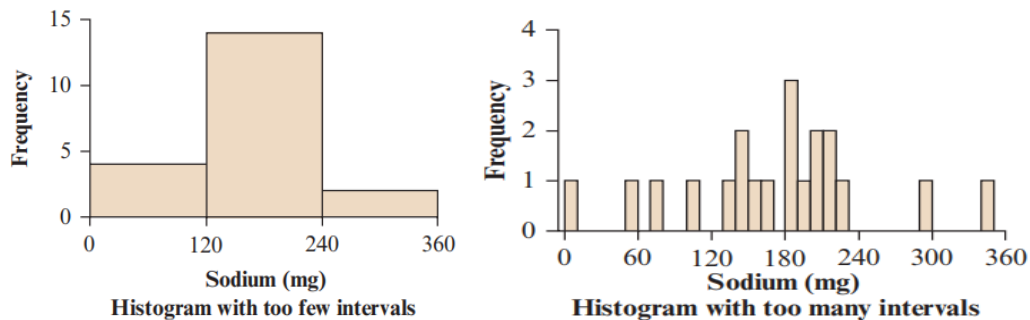
| Interval   | Frequency | Proportion | Percentage |
|------------|-----------|------------|------------|
| 0 to 39    | 1         | 0.05       | 5%         |
| 40 to 79   | 2         | 0.10       | 10%        |
| 80 to 119  | 1         | 0.05       | 5%         |
| 120 to 159 | 4         | 0.20       | 20%        |
| 160 to 199 | 5         | 0.25       | 25%        |
| 200 to 239 | 5         | 0.25       | 25%        |
| 240 to 279 | 0         | 0.00       | 0%         |
| 280 to 319 | 1         | 0.05       | 5%         |
| 320 to 359 | 1         | 0.05       | 5%         |

Langkah untuk menyusun histogram (Agresti, dkk, 2017)

1. Bagi rentang data menjadi interval dengan jarak yang sama.
2. Hitung jumlah pengamatan (frekuensi) di setiap interval sehingga membentuk tabel frekuensi.
3. Pada sumbu horizontal (mendatar), beri label untuk nilai akhir setiap interval. Gambar batang pada setiap interval dengan tinggi batang sama

dengan frekuensi (atau persentase). Nilai frekuensi (atau persentase) ditunjukkan oleh sumbu vertikal (tegak).

Bagaimana cara menentukan jumlah interval dalam penyusunan histogram? Perhatikan Gambar 9.



**Gambar 9.** Perbandingan dua histogram

Jika kita menggunakan interval yang terlalu sedikit (Gambar 9 kiri), grafik yang dihasilkan akan terlihat terlalu kasar. Hal ini terlihat dari rentang data antara 0 sampai 360 yang hanya terdiri dari 3 interval. Tiga interval tersebut membentuk batang yang cukup tinggi (lihat sumbu vertikal). Sebaliknya, jika kita menggunakan interval yang terlalu banyak (Gambar 9 kanan), grafik yang dihasilkan terlihat tidak beraturan dengan sejumlah batang yang cenderung pendek. Selain itu, terdapat jarak (*gap*) antar batang (lihat sumbu horizontal). Dengan demikian, kita akan kehilangan informasi mengenai bentuk distribusi dari data tersebut.

Jumlah interval dikatakan cukup berkisar antara 5 sampai 10 interval. Jumlah tersebut dapat ditambahkan jika data yang digunakan sangat besar. Meskipun demikian, tidak ada satu jumlah interval yang mutlak tepat untuk suatu data.

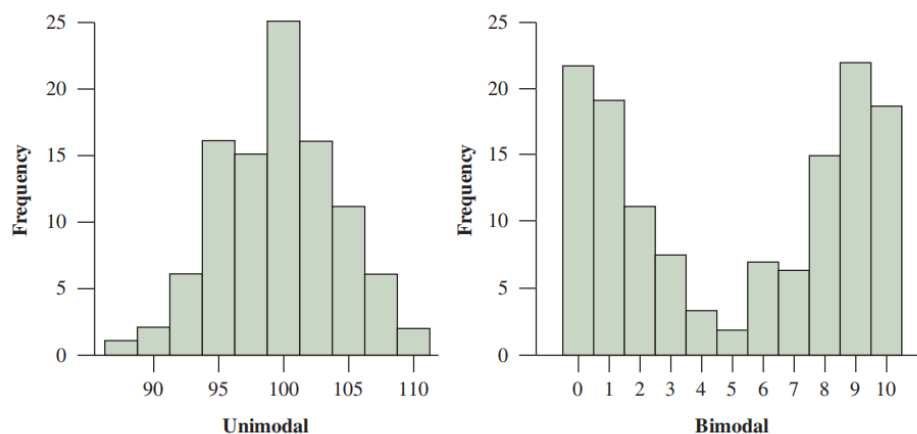
Pada praktiknya, aplikasi analisis data dapat membantu kita menentukan jumlah interval, menghitung frekuensi dan persentase, serta menghasilkan histogram. Terkait dengan kemudahan tersebut, Agresti, dkk (2017) menyarankan untuk selalu memeriksa data numerik menggunakan histogram untuk membantu memberikan gambaran tentang bentuk distribusi data. Jika jumlah data kecil (kurang dari 50), disarankan untuk menambahkan diagram batang-daun atau dot-plot untuk menunjukkan nilai numerik dari data yang

digunakan.

### Bentuk Distribusi

Seperti yang dijelaskan sebelumnya, teknologi telah membantu kita dalam membuat histogram melalui suatu aplikasi analisis data. Aplikasi dapat membantu menentukan jumlah interval serta membuat histogram. Kita tidak perlu lagi menggambar secara manual. Namun, hal yang tidak kalah penting adalah kita harus memahami bagaimana histogram tersebut digambarkan dan bagaimana menginterpretasikannya. Berikut ini beberapa hal yang perlu diperhatikan dalam pengecekan suatu histogram.

1. Apakah histogram memiliki satu puncak? Jika histogram memiliki satu puncak maka distribusi datanya disebut *unimodal*. Nilai yang paling sering muncul (memiliki frekuensi tertinggi) disebut dengan modus. Jika histogram memiliki dua puncak maka distribusi datanya disebut bimodal. Distribusi bimodal dapat terjadi pada data dengan populasi yang memiliki dua karakteristik kontras atau kepentingan yang berbeda, misalnya ketika data persepsi kualitas barang berasal dari pihak penjual dan pihak pembeli. Distribusi bimodal juga dapat terjadi jika data berasal dari dua kelompok yang berbeda. Sebagai contoh data tinggi badan dapat menunjukkan dua puncak, yang merupakan puncak data laki-laki dan puncak data perempuan.

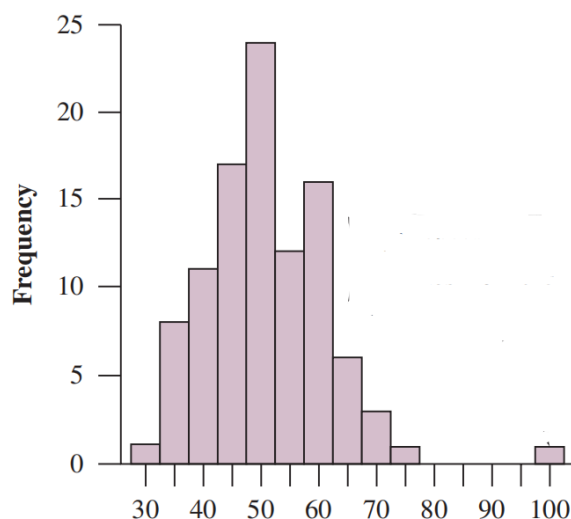


**Gambar 10.** Distribusi unimodal dan bimodal

2. Bagaimana bentuk histogram? Bentuk histogram biasanya dikategorikan

menjadi simetris atau melenceng. Histogram yang berbentuk simetris berarti distribusi datanya simetris. Data yang berdistribusi simetris berarti bagian yang lebih kecil dari pusat data berbentuk simetris terhadap bagian yang lebih besar dari pusat data. Histogram yang berbentuk melenceng berarti distribusi datanya melenceng. Data yang berdistribusi melenceng memiliki salah satu sisi yang lebih panjang dibandingkan sisi yang lainnya.

3. Apakah data mengumpul? Bentuk histogram dapat mengumpul atau berdekatan tetapi kadang ditemukan bentuk histogram yang memiliki jarak antar batangnya. Jarak pada batang ini biasanya merupakan indikasi adanya pencilan dalam data.



**Gambar 11.** Histogram dengan data pencilan

#### **D. Deskripsi Data Univariat: Numerik**

Pada bagian sebelumnya telah dibahas mengenai analisis deskriptif menggunakan diagram batang pada data univariat kategorik dan numerik. Pada data kategorik, diagram batang menunjukkan proporsi dari setiap kategori dan kategori mana yang terbanyak. Sedangkan pada data numerik, histogram menunjukkan karakteristik penting yaitu bentuk distribusi data. Untuk praktik terbaik, pada awal tahap analisis deskriptif biasakan untuk selalu memeriksa data menggunakan grafik yang tepat agar mendapatkan pemahaman tentang data yang akan diteliti.

Setelah melihat grafik yang dihasilkan, kita dapat menghitung ringkasan numerik (statistik) dari data tersebut untuk menggambarkan sifat distribusi yang belum terlihat dari grafik. Sebagai contoh: (1) berapa nilai data yang kerap muncul? (2) apakah data memiliki nilai-nilai yang mirip atau sangat berbeda?

Statistik yang menjawab pertanyaan (1) menggambarkan pusat distribusi data. Sedangkan statistik yang menjawab pertanyaan (2) menggambarkan sebaran (keragaman) dari distribusi data. Bagian ini juga akan menjelaskan bagaimana bentuk distribusi data menjadi penentu statistik yang tepat untuk menggambarkan pusat dan sebaran dari data tersebut.

### **1. Ukuran Pemusatan**

Ukuran pemusatan yang paling populer dan paling sering digunakan untuk mengukur pusat distribusi data kuantitatif adalah mean. Mean didefinisikan sebagai jumlah data pengamatan dibagi dengan banyaknya data yang diamati (Agresti, dkk, 2017). Mean diinterpretasikan sebagai titik keseimbangan dari distribusi data. Istilah mean juga merujuk pada menghitung rata-rata atau menjumlahkan semua nilai data dan membagi dengan berapa banyak data yang ada.

Ukuran pemusatan lainnya adalah median. Median adalah nilai di mana setengah dari total banyaknya data lebih kecil dan setengah sisanya lebih besar dari nilai tersebut. Median merupakan nilai tengah dari data hasil pengamatan (atau pengukuran) ketika data tersebut diurutkan mulai data yang paling kecil sampai data yang paling besar (atau sebaliknya).

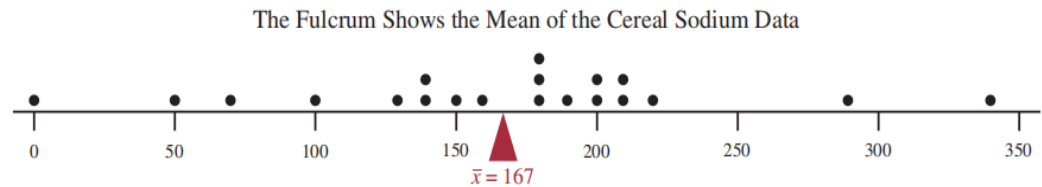
Bagaimana cara menghitung median?

- a. Susun  $n$  data berdasarkan urutan nilainya.
- b. Jika banyaknya data ( $n$ ) merupakan bilangan ganjil, median adalah data yang berada di tengah setelah diurutkan.
- c. Jika banyaknya data ( $n$ ) merupakan bilangan genap, maka terdapat dua data yang berada di tengah. Selanjutnya, median adalah rata-rata dari kedua data tersebut.

Berikut ini beberapa sifat mean:

- a. Mean adalah titik keseimbangan data. Jika kita menggunakan bandul

dengan berat yang sama untuk mewakili setiap data dan meletakkannya dalam suatu garis sesuai nilai pengukuran, maka garis tersebut akan seimbang dengan menempatkan titik tumpu pada nilai mean.



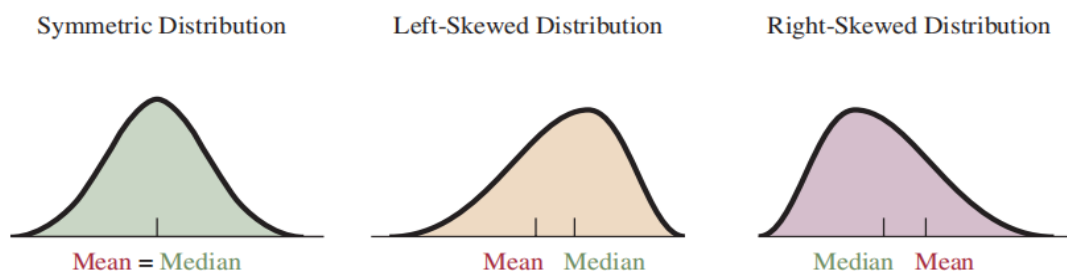
**Gambar 12.** Sebaran data dan titik tumpu

- b. Biasanya, mean tidak sama dengan nilai hasil pengamatan. Hal ini berbeda dengan median di mana nilai median akan sama dengan pengamatan (untuk data sebanyak bilangan ganjil).
- c. Untuk data dengan distribusi melenceng, posisi mean cenderung lebih dekat ke arah ekor yang panjang jika dibandingkan dengan nilai median.
- d. Mean sangat dipengaruhi oleh pencilan (*outlier*). Pencilan biasanya berupa nilai data yang sangat kecil atau sangat tinggi.

### Perbandingan Mean dan Median

Bentuk distribusi berpengaruh terhadap posisi mean dan median. Secara spesifik, bentuk distribusi menentukan manakah yang nilainya lebih besar di antara kedua ukuran tersebut. Berikut ini perbandingan mean dan median pada ketiga bentuk distribusi:

- a. Jika bentuk distribusi simetris, mean sama dengan median.
- b. Jika bentuk distribusi melenceng ke kiri, mean lebih kecil dari median.
- c. Jika bentuk distribusi melenceng ke kanan, mean lebih besar dari median.



(Sumber: Agresti, dkk, 2017)

**Gambar 13.** Hubungan mean dan median, pada bentuk Distribusi

Jika distribusi mendekati bentuk simetris, ujung (ekor) kurva distribusi akan memiliki panjang yang sama. Hal ini menyebabkan nilai median dan mean menjadi sama (atau mirip). Untuk distribusi yang melenceng, mean berada lebih dekat dengan arah ujung (ekor) yang lebih panjang dibandingkan dengan median. Hal ini disebabkan karena nilai data yang ekstrim di ujung (ekor) yang panjang tersebut mempengaruhi titik keseimbangan dari distribusi. Titik keseimbangan inilah yang disebut sebagai mean. Semakin melenceng suatu distribusi, nilai mean dan median akan semakin berbeda.

Sementara itu, median tidak dipengaruhi oleh nilai data yang ekstrim (pencilan). Mengapa demikian? Hal ini dikarenakan sejauh manapun pencilan terhadap median tidak akan mempengaruhi median. Median ditentukan dengan membagi banyaknya yang lebih besar dan lebih kecil dari median dengan jumlah yang sama. Median tidak dipengaruhi oleh pencilan sehingga disebut resisten terhadap pengaruh pencilan.

### **Manakah yang sebaiknya digunakan, mean atau median?**

Jika melihat sifat median yang resisten dibandingkan dengan mean, kita mungkin akan berpikir bahwa lebih baik selalu menggunakan median dibandingkan mean. Namun, dalam praktek analisis data, sebaiknya gunakan mean dan median untuk memeriksa ukuran pemusatan. Pada data dengan bentuk distribusi melenceng, gunakan median sebagai ukuran pemusatan. Selanjutnya, untuk informasi tambahan, hitung mean untuk seluruh data dan mean setelah pencilan dikeluarkan dari data. Hal ini akan menambah informasi dan *insight* terkait data yang digunakan.

Berikut ini panduan untuk memilih mean atau median berdasarkan bentuk distribusi (Agresti, dkk, 2017):

- a. Jika distribusi data sangat melenceng, median lebih tepat digunakan karena menggambarkan kekhasan data dengan lebih baik.
- b. Jika distribusi data simetris atau sedikit melenceng, mean lebih tepat digunakan karena menggunakan seluruh nilai data.

## **Modus**

Modus adalah nilai yang memiliki frekuensi paling tinggi. Ukuran ini menggambarkan data yang paling sering muncul (terjadi). Konsep modus paling sering digunakan untuk mendeskripsikan data kategorik. Pada data kategorik, modus berarti kategori yang memiliki frekuensi tertinggi. Sedangkan pada data numerik, modus tepat digunakan untuk data numerik diskrit yang memiliki ragam nilai yang sedikit.

## **2. Ukuran Sebaran**

Ukuran pemusatan dari suatu data belum cukup untuk menggambarkan data kuantitatif karena belum dapat memberikan informasi mengenai keragaman data. Jika diketahui mean dari nilai ujian peserta pelatihan adalah 78, apakah peserta dengan nilai 90 dikatakan tinggi, atau sangat tinggi dibandingkan rata-ratanya? Untuk menjawab pertanyaan tersebut, kita membutuhkan ukuran sebaran yang dapat menggambarkan sebaran distribusi data.

### **Rentang (Range)**

Ukuran sebaran yang paling sederhana adalah rentang. Rentang adalah selisih antara data tertinggi dan data terendah. Rentang sangat mudah untuk dihitung hanya dengan menggunakan dua nilai data ekstrim (tertinggi dan terendah). Sehingga hal ini menyebabkan rentang sangat dipengaruhi oleh pencilan. Rentang memiliki kelemahan yang sama dengan mean, yaitu tidak resisten terhadap data pencilan. Selain itu, rentang juga memiliki kelemahan yang sama dengan median, yaitu penghitungannya tidak menggunakan semua nilai data.

### **Standar Deviasi**

Deviasi menggambarkan jarak rata-rata penyimpangan data terhadap mean. Setiap data memiliki suatu penyimpangan terhadap mean. Deviasi suatu data bernilai positif jika nilai data tersebut lebih besar dari mean. Sebaliknya, deviasi bernilai negatif jika nilai data tersebut lebih kecil dari mean.

Seperti yang telah dijelaskan sebelumnya, mean memiliki makna titik keseimbangan. Dengan kata lain, deviasi positif akan menyeimbangkan deviasi negatif. Hal ini mengakibatkan jumlah dan mean dari deviasi akan selalu bernilai

nol, berapapun nilai datanya. Sehingga, deviasi tidak diukur menggunakan jumlah dan mean dari deviasi, melainkan menggunakan kuadrat deviasi atau nilai mutlak deviasi.

Rata-rata dari kuadrat deviasi disebut dengan keragaman (*variance*). Karena diperoleh dari penghitungan kuadrat, keragaman memiliki satuan kuadrat dari satuan pengukuran data. Sebagai contoh, jika data yang digunakan memiliki satuan meter, maka keragaman memiliki satuan meter<sup>2</sup>. Hal ini menjadi sulit untuk diinterpretasikan. Untuk mempermudah interpretasi ukuran keragaman, maka digunakan deviasi standar, yaitu akar kuadrat dari rata-rata dari kuadrat deviasi.

Sifat utama deviasi standar adalah sebagai berikut.

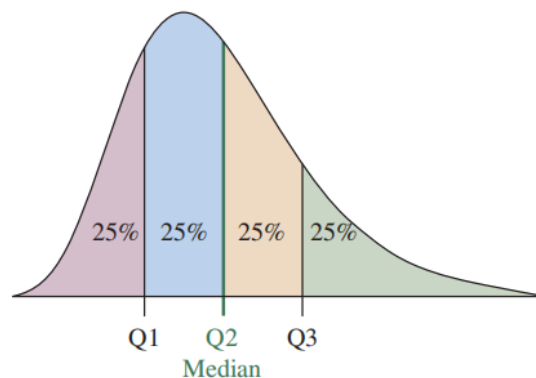
1. Semakin besar deviasi standar, semakin besar keragaman data.
2. Deviasi standar bernilai nol hanya jika semua data bernilai sama.
3. Deviasi standar dipengaruhi pencilan.

### **3. Ukuran Posisi untuk Menggambarkan Sebaran**

Mean dan median menunjukkan ukuran pemusatan dari suatu distribusi data. Sementara itu, rentang dan deviasi standar menunjukkan keragamannya. Pada bagian ini, kita akan mempelajari teknik lain untuk menggambarkan distribusi data, yaitu menggunakan ukuran posisi.

#### **Kuartil**

Seperti yang telah dipelajari sebelumnya, median menunjukkan titik di mana setengah data memiliki nilai di bawahnya dan setengah data memiliki nilai di atasnya. Sama halnya dengan kuartil yang membagi data menjadi empat bagian yang sama banyak yaitu masing-masing sebanyak 25% dari total data (Gambar 14).



**Gambar 14.** Posisi kuartil pada kurva distribusi data

Langkah menentukan kuartil:

- Susun data berurutan dari nilai terkecil ke nilai terbesar.
- Tentukan median (Q2).
- Jika data berjumlah ganjil, keluarkan median dari perhitungan (d) dan nomor (e).
- Perhatikan kelompok amatan yang bernilai kecil. Tentukan median dari kelompok data ini (Q1).
- Perhatikan kelompok amatan yang bernilai besar. Tentukan median dari kelompok data ini (Q3)

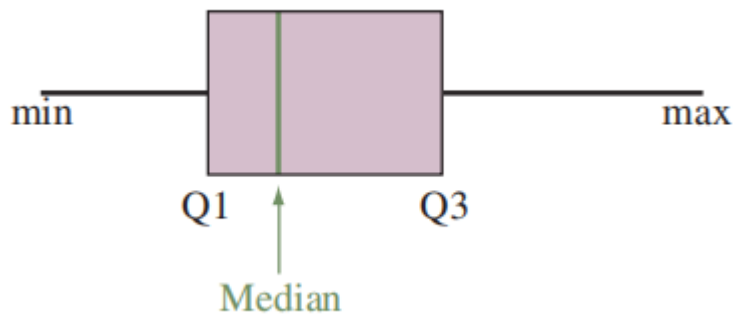
### **Rentang Antar Kuartil**

Kuartil dapat juga digunakan untuk mengukur keragaman. Kuartil bersifat lebih resisten jika dibandingkan dengan rentang dan deviasi standar. Rentang antar kuartil mengukur rentang dari setengah data yang ada di tengah-tengah distribusi data. Setengah data tersebut berada di antara Q1 dan Q3. Jarak antara Q1 dengan Q3 disebut dengan Rentang Antar Kuartil (RAK).

RAK tidak dipengaruhi oleh data yang bernilai lebih kecil dari Q1 atau data yang bernilai lebih besar dari Q3. Dengan kata lain, RAK tidak dipengaruhi oleh pencilan. Sebaliknya, RAK bergantung pada nilai minimum dan nilai maksimum (dua nilai yang paling ekstrim). Sehingga RAK berubah jika nilai ekstrim tersebut berubah.

### **Boxplot**

Boxplot adalah grafik yang disusun berdasarkan 5 ukuran, yaitu: nilai minimum, Q1, Q2, Q3 dan nilai maksimum. Gambar 15 menunjukkan contoh boxplot.



**Gambar 15.** Contoh Boxplot

Area kotak dalam boxplot mencakup setengah (50%) data yang ada di tengah-tengah distribusi data. Area kotak dimulai dari Q1 sampai dengan Q3. Garis yang digambarkan di dalam area kotak menandakan posisi median (Q2). Garis yang berada di luar kotak mewakili data selain yang ada di dalam kotak. Sedangkan untuk pencilan, biasanya digambarkan dengan tanda khusus.

Langkah menyusun Boxplot:

- a. Buat satu kotak dari Q1 ke Q3.
- b. Gambar garis di tengah kotak yang menunjukkan posisi median.
- c. Gambar garis dari Q1 ke nilai data terkecil (yang bukan merupakan pencilan).
- d. Gambar garis dari Q3 ke nilai data terbesar (yang bukan merupakan pencilan).

## **E. Analisis Deskriptif untuk Dua Variabel**

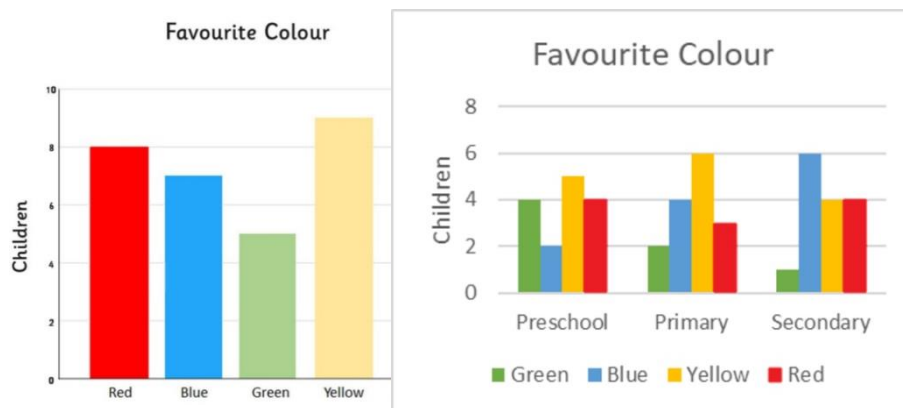
Pada praktik analisis deskriptif, seringkali kita perlu melakukan analisis deskriptif untuk dua variabel. Untuk membahas materi tersebut, akan dibagi menjadi dua bagian, yaitu analisis deskriptif untuk variabel numerik dan kategorik dan analisis deskriptif untuk dua variabel numerik.

### **1. Analisis deskriptif data numerik dan data kategorik**

Analisis deskriptif ini digunakan ketika perlu melakukan penyajian data numerik dengan mempertimbangkan data kategorik.

a. Diagram batang berkelompok

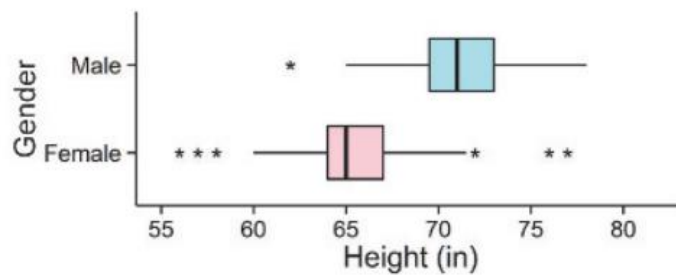
Pada Gambar 16 ditampilkan dua contoh diagram batang dari data warna favorit anak sekolah. Diagram kiri merupakan diagram batang tidak berkelompok. Diagram ini menunjukkan bahwa warna favorit anak sekolah secara umum adalah warna kuning. Untuk memperkaya informasi yang didapatkan, diagram batang dapat ditampilkan seperti diagram kanan. Diagram kanan menggambarkan warna favorit anak yang dikelompokkan berdasarkan tingkat pendidikannya. Berdasarkan diagram ini, terlihat bahwa warna kuning menjadi warna favorit pada tingkat preschool dan tingkat primary. Sedangkan pada tingkat secondary, warna favorit anak adalah warna biru.



**Gambar 16.** Contoh diagram batang tidak berkelompok (kiri) dan berkelompok (kanan)

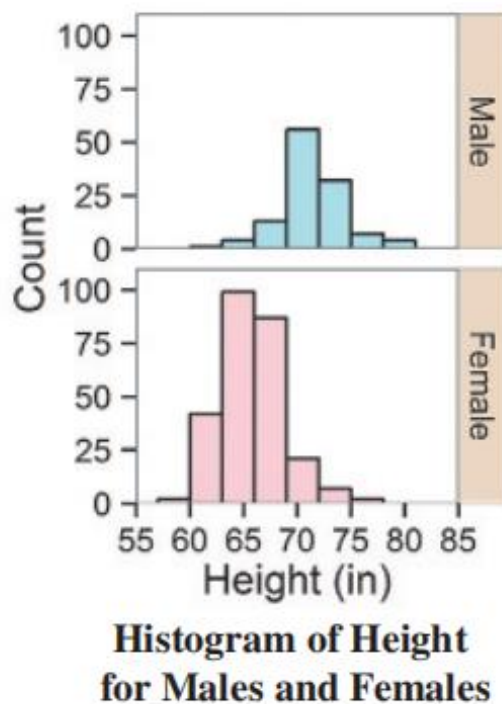
b. Boxplot dan Histogram Berkelompok

Boxplot dan histogram berkelompok digunakan untuk melakukan analisis deskriptif data numerik dengan mempertimbangkan data kategorik. Sebagai contoh, akan dianalisis data tinggi badan mahasiswa (data numerik). Untuk membandingkan tinggi badan mahasiswa pria dan wanita, kita dapat membuat dua boxplot untuk setiap jenis kelamin mahasiswa (data kategorik). Gambar 17 memperlihatkan bahwa secara umum, mahasiswa pria lebih tinggi dibandingkan dengan mahasiswa wanita. Median untuk data tinggi badan pria adalah 71 inchi sedangkan median untuk data tinggi badan wanita adalah 65 inchi.



**Gambar 17.** Boxplot dua kelompok (Sumber: Agresti, dkk, 2017)

Selain menggunakan boxplot, histogram juga dapat digunakan untuk membandingkan tinggi badan mahasiswa pria dan wanita. Berdasarkan gambar 18, terlihat bahwa sebagian besar mahasiswa pria memiliki tinggi badan 70 inchi sedangkan mahasiswa wanita 65 inchi.

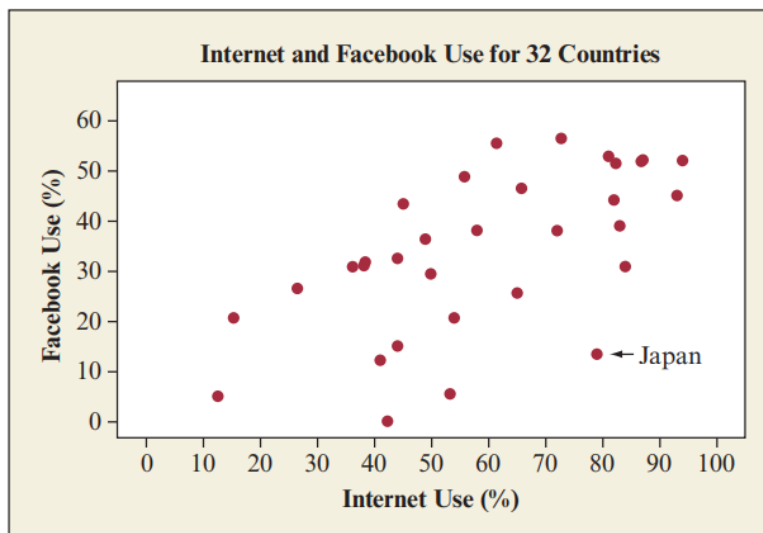


**Gambar 18.** Histogram dua kelompok (Sumber: Agresti, dkk, 2017)

## 2. Analisis deskriptif dua data numerik

Diagram pencar adalah grafik yang digunakan untuk menggambarkan dua data numerik menggunakan sumbu horizontal (x) dan sumbu vertikal (y). Nilai x dan

y untuk suatu pengamatan digambarkan menggunakan titik yang berada pada posisi sesuai dengan nilai x dan y (Gambar 19). Dengan demikian, data dengan n pengamatan mempunyai n titik pada diagram pencar. Titik-titik pada diagram pencar membentuk pola yang dapat menentukan hubungan antara nilai x dan y, apakah berpola hubungan positif atau negatif.



**Gambar 19.** Contoh diagram pencar (Agresti, dkk, 2017)

## F. Grafik yang Efektif

Setelah kita membahas mengenai analisis deskriptif, perlu kita ketahui panduan untuk menyusun grafik yang efektif, yaitu sebagai berikut (Agresti, dkk, 2017).

1. Beri label pada kedua sumbu (horizontal dan vertikal). Hal ini diperlukan untuk memperjelas data apa yang ingin digambarkan.
2. Sumbu vertikal harus dimulai dari nol. Hal ini bertujuan untuk membantu kita membandingkan ukuran secara visual dengan lebih akurat.
3. Hati-hati dalam penggunaan gambar (simbol) untuk menggantikan batang atau titik. Penggunaan simbol membuat grafik lebih menarik tetapi beresiko mengaburkan ukuran relatif dari angka yang digambarkan oleh simbol tersebut.
4. Pada beberapa kasus, jika nilai data memiliki perbedaan yang cukup besar, sulit untuk menggambarkan lebih dari 1 kelompok dalam satu grafik. Sebagai alternatif, gunakan grafik terpisah atau gunakan nilai relatif

(rasio atau persentase) untuk menggambarkan data tersebut.

## **G. Rangkuman**

Analisis deskriptif adalah metode-metode yang berkaitan dengan pengumpulan dan penyajian suatu gugus data sehingga memberikan informasi yang berguna. Sebelum melakukan proses analisis deskriptif dari suatu data, kita perlu memahami jenis data yang akan dianalisis. Hal ini penting karena jenis data menentukan metode analisis deskriptif yang tepat untuk digunakan. Dengan mengetahui jenis data, kita akan mudah menentukan metode analisis deskriptif yang tepat.

Data terbagi menjadi beberapa jenis. Berdasarkan tipe datanya, terdapat dua jenis data, yaitu data kuantitatif dan data kualitatif. Berdasarkan skala pengukurannya, terdapat empat jenis data, yaitu nominal, ordinal, interval dan rasio.

Dalam menggambarkan sekumpulan data hasil pengukuran, terdapat dua metode yang umum digunakan yaitu metode grafik dan metode deskriptif numerik. Setelah pengumpulan data selesai dilakukan, idealnya data disusun, disajikan, dan diperiksa menggunakan beberapa metode grafik. Aturan umum yang berlaku dalam deskripsi data univariat, data harus disusun ke dalam kategori sehingga setiap data dikategorikan ke dalam satu dan hanya satu kategori.

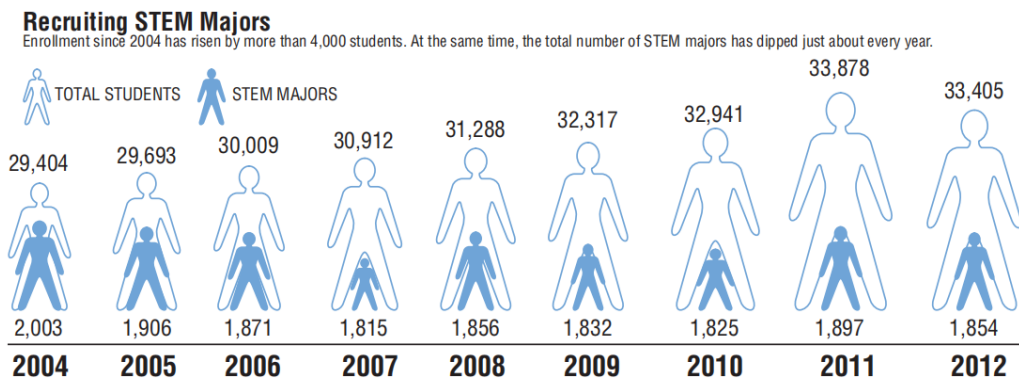
Analisis deskriptif dapat disajikan dengan menggunakan analisis univariate baik pada data univariat kategorik maupun numerik

## **H. Evaluasi**

1. Apakah yang dimaksud dengan analisis deskriptif? Apa manfaatnya?
2. Apa perbedaan antara data kategorik dan data numerik. Berikan contoh untuk masing-masing data tersebut.
3. Carilah sebuah dataset yang mencakup empat jenis data berdasarkan skala pengukurannya. Kelompokkan setiap data berdasarkan skala pengukurannya.
4. Carilah sebuah dataset yang mencakup data kategorik dan data numerik

dengan jumlah data minimal 100 amatan. (a) Lakukan analisis deskriptif (grafik) yang mencakup deskriptif univariat dan deskriptif lebih dari satu variabel. (b) Tentukan ukuran pemusatan dan ukuran penyebaran yang sesuai dengan distribusi data. Jelaskan mengapa Anda memilih ukuran tersebut.

5. Perhatikan grafik berikut ini.



(a) Berikan pendapat Anda mengenai apa yang kurang tepat dalam grafik tersebut?

(b) Apa saran Anda untuk memperbaiki grafik tersebut sehingga lebih efektif?

6. Carilah grafik dalam sebuah publikasi yang **tidak** memenuhi minimal satu ketentuan grafik yang efektif. Berikan penjelasan mengenai apa yang **kurang** tepat dalam grafik tersebut dan bagaimana cara memperbaikinya?

## MATERI POKOK 3:

# VISUALISASI DATA

### Indikator Hasil Belajar:

Peserta mampu **memilih dan memaksimalkan penyajian data dengan visualisasi data yang sesuai, melalui:**

1. Mengenal visualisasi data dengan benar;
2. Memahami visualisasi data yang efektif dengan benar;
3. Memahami dan memaksimalkan penyampaian informasi dengan efektif.

### A. Pengantar Visualisasi Data

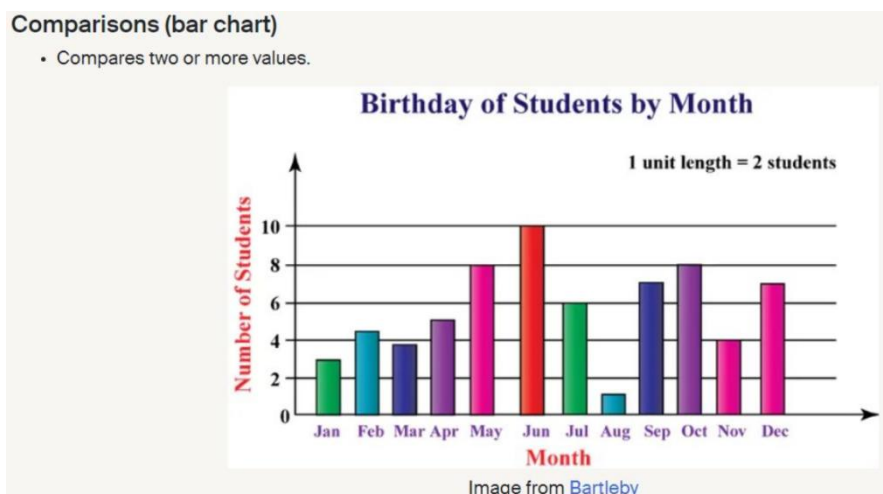
Seiring perkembangan teknologi informasi, penyajian data yang monoton kurang mendapat perhatian, dimana sebagian besar orang lebih melihat tampilan visual yang menarik. Visualisasi data (*data visualization*) dalam menyampaikan data harus tepat dan akurat. Pengertian visualisasi data adalah mempresentasikan data dan informasi dalam kumpulan grafik maupun diagram baik berupa data tabular dan juga data geospasial sehingga mudah dipahami. Estetika dalam visualisasi data harus memberikan kesan menyenangkan. Sebagai gambaran, data mentah yang kita miliki tentunya belum menjadi sebuah informasi bila belum dilakukan representasi data dengan teknik visualisasi data. Bila sebuah gambar ditampilkan dengan elemen visual yang tidak seimbang dalam warna, ukuran, dan bentuk, maka tampilan menjadi kurang diminati pengguna secara visual, baik meneliti nilai kebenarannya serta menafsirkan gambar tersebut secara benar. Dari penggunaan elemen visual tersebut, maka tren, *outliers*, dan pola data menjadi lebih mudah untuk dipahami. Elemen-elemen visual yang digunakan seperti grafik, diagram dan geospasial diperlukan untuk berkomunikasi secara visual, sehingga menjadi *insight* dalam mengambil sebuah kebijakan/keputusan.

## B. Jenis Visualisasi Data

Visualisasi data merupakan proses penyampaian hasil analisis data yang telah dilakukan. Konsep dari visualisasi data yaitu agar pemahaman objek dari visualisasi menerima pesan yang kita sampaikan, dalam hal ini audience (penerima). Visualisasi data yang menarik dan mudah dipahami, tentunya harus diketahui oleh kita, baik tipe dan jenis visualisasi data yang ada saat ini. Setiap jenis visualisasi data memiliki peran dan fungsi yang berbeda dalam menyampaikan hasil analisis. Tentunya kita perlu mempelajari jenis-jenis dari visualisasi data agar mudah menentukan jenis mana yang akan kita gunakan. Jenis dari visualisasi data diantaranya:

### 1. Visualisasi data temporal (*Temporal data visualizations*)

Visualisasi data ini sering digunakan untuk menyajikan data dalam bentuk satu dimensi atau linier. Visualisasi data temporal umumnya sering digunakan karena mudahnya untuk dipelajari. Tipe visualisasi ini umumnya berbentuk diawali garis dan diakhiri dengan titik. Garis yang dipakai bisa garis tunggal atau bisa juga garis bersinggungan dengan garis lainnya. Jenis visualisasi data temporal dapat berupa bar chart, line chart, scatter plot, timeline, time series, dan gantt chart. Misalkan dengan diagram garis dapat memperlihatkan dalam periode tertentu muncul perubahan secara kontinyu. Bisa juga pada diagram garis memperlihatkan beberapa garis yang memiliki faktor yang berbeda pada jangka waktu yang sama.



### Relationship (scatter plot)

- Shows the connection and impact between elements such as life expectancy and GDP per capita.

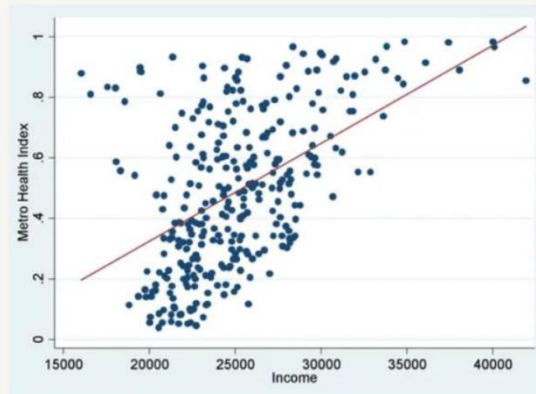


Image from [Minnesota Department of Health](#)

### Timeframe (line graph)

- Shows how something changes over time.

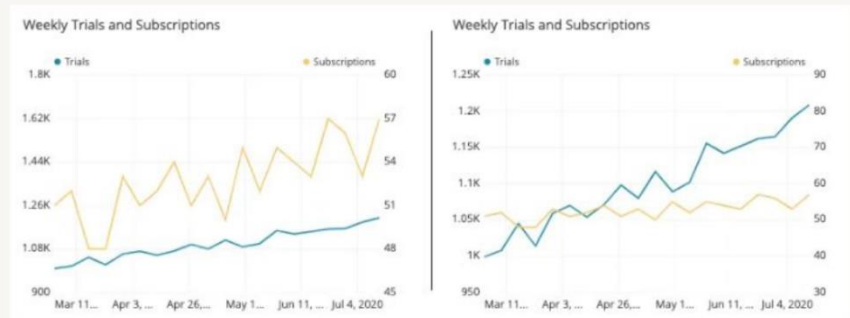
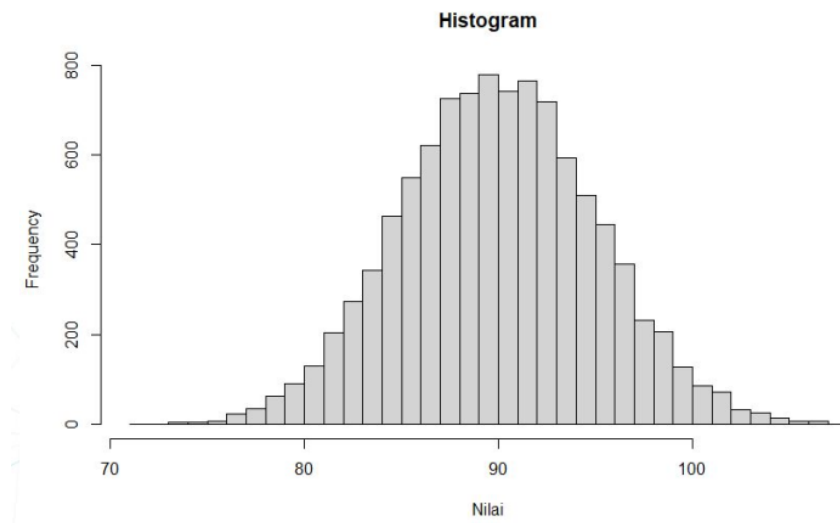


Image from [Chartio](#)

**Gambar 20.** Visualisasi temporal

## 2. Visualisasi data hierarki (*Hierarchical data visualizations*)

Visualisasi data tipe ini digunakan dalam melihat hubungan antara suatu kelompok dengan kelompok lebih besar. Visualisasi data hierarki juga memperlihatkan hubungan yang memiliki kesamaan antara item induk dengan sekelompok item dibawahnya. Tipe visualisasi ini sering dipilih karena cocok dengan munculnya data yang baru karena ada sebabnya. Jenis visualisasi ini dapat berupa grafik batang dan juga diagram pohon. Misalkan, pada diagram pohon memperlihatkan jumlah data sebuah ruangan, dimana simpul induknya (ruang kerja) kemudian simpul anak (lemari, meja, kursi).



**Gambar 21.** Visualisasi Hierarki

### 3. Visualisasi data jaringan (*network data visualizations*)

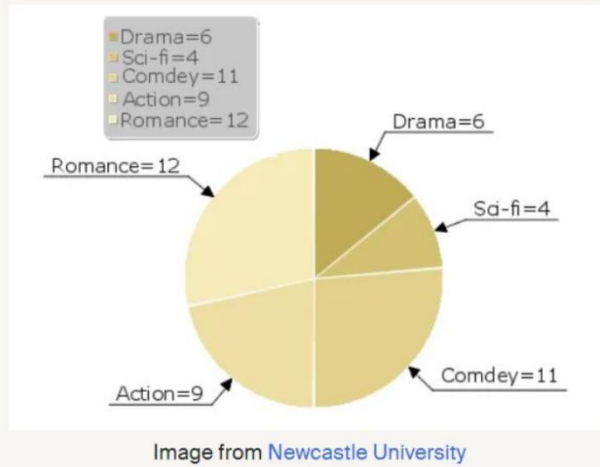
Tipe visualisasi data jaringan memperlihatkan hubungan antara bermacam jenis data secara kompleks dan terkait satu dengan lainnya. Hubungan di dalam visualisasi data jaringan sering disebut entitas dan tautan. Visualisasi data ini saling berhubungan dan berpengaruh antara data satu dengan data lainnya. Jenis visualisasi data ini seperti word cloud, node-link diagram, alluvial diagram, dan matrix chart. Implementasi visualisasi data ini misalnya:

- a. Sekumpulan data dalam bentuk titik-titik pada sebuah grafik yang menggambarkan dan merepresentasikan sebaran plot.
- b. Word cloud yang merepresentasikan jumlah dan frekuensi kata yang sering muncul ditampilkan dalam besaran ukuran yang berbeda, semakin kata tersebut muncul, maka ukurannya katanya semakin besar.



### Composition (pie chart)

- Shows the parts of a single unit.



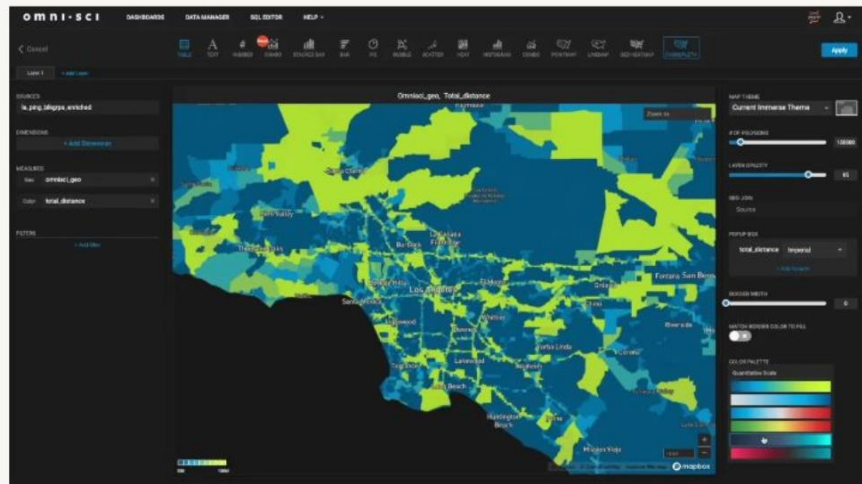
**Gambar 23.** Visualisasi Data Multidimensi

### 5. Visualisasi data geospasial (*geospatial data visualizations*)

Visualisasi data geospasial adalah wujud dari benda atau ruang yang mempunyai nilai data yang mau disajikan. Visualisasi data tipe ini biasanya ditampilkan dalam bentuk peta atau lokasi geografis tertentu. Contoh visualisasi data geospasial, diantaranya kartogram (cartogram), peta densitas (density map), dan peta panas (heat map). Visualisasi pemetaan berhubungan dengan tempat atau lokasi sebenarnya. Misalkan contoh penerapannya, visualisasi data geospasial memperlihatkan jumlah customer yang menggunakan dan berkunjung ke layanan cabang bank swasta yang berbeda. Maka akan terlihat, sebaran peta dari customer dan berkunjung di masing-masing bank swasta.

## Interactive maps

- Allows users to drill into data, explore and find more detail in a geospatial context.



**Gambar 24.** Visualisasi Data Geospasial

### C. Memilih visualisasi data yang efektif

Visualisasi data memiliki pengaruh penting dalam proses penelitian, menyampaikan presentasi, media informasi, dan komunikasi. Jumlah data yang cukup besar dapat dituangkan dalam grafik atau diagram yang efektif dan efisien. Secara umum, audience lebih mudah melihat secara visual berupa gambar atau grafik daripada angka atau kata. Beberapa teknik visualisasi data perlu diterapkan, supaya hasil yang dicapai tidak terlepas dari interpretasi data yang diinginkan.

Beberapa elemen untuk visual yang efektif diantaranya

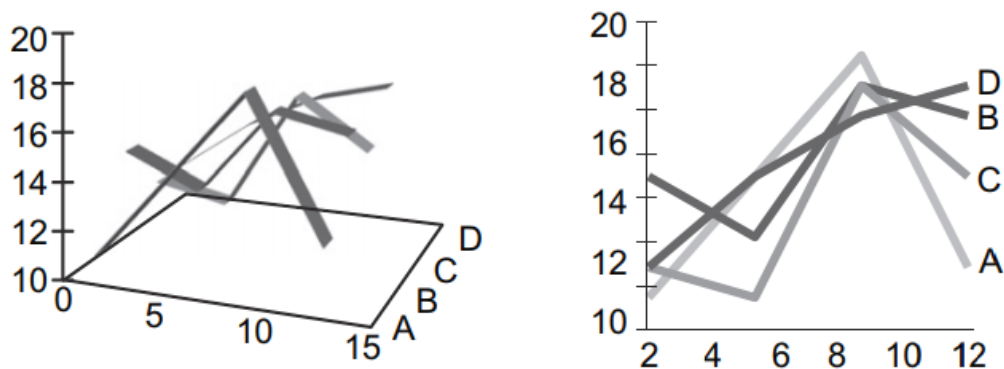
- a. **Clear meaning** (memiliki makna yang jelas), yaitu diharapkan mendapatkan *insight* dari hasil komunikasi.
- b. **Sophisticated use of contrast** (penggunaan kontras yang berbeda satu dengan lainnya), yaitu pemisahan data penting dengan data lainnya.
- c. **Refined execution** (melakukan eksekusi yang sempurna), yaitu dengan memberi perhatian lebih terhadap detail pada penggunaan elemen visual.

Teknik-teknik dalam visualisasi data supaya visualisasi yang dihasilkan efektif seperti:

**1. Pertama, membuat grafik sederhana yang berisi seluruh informasi yang akan disajikan.**

Pemilihan grafik dalam visualisasi data diperlukan dalam mempresentasikan informasi yang disajikan ke audience. Misalkan pemilihan grafik dua dimensi (2D) biasanya lebih informatif daripada grafik tiga dimensi (3D). Hal ini dikarenakan audience lebih melihat perbedaan seperti bentuk, warna, dan ketebalan garis pada grafik dan terutama tidak perlu menambahkan atau melebihkan visualisasi yang ditampilkan.

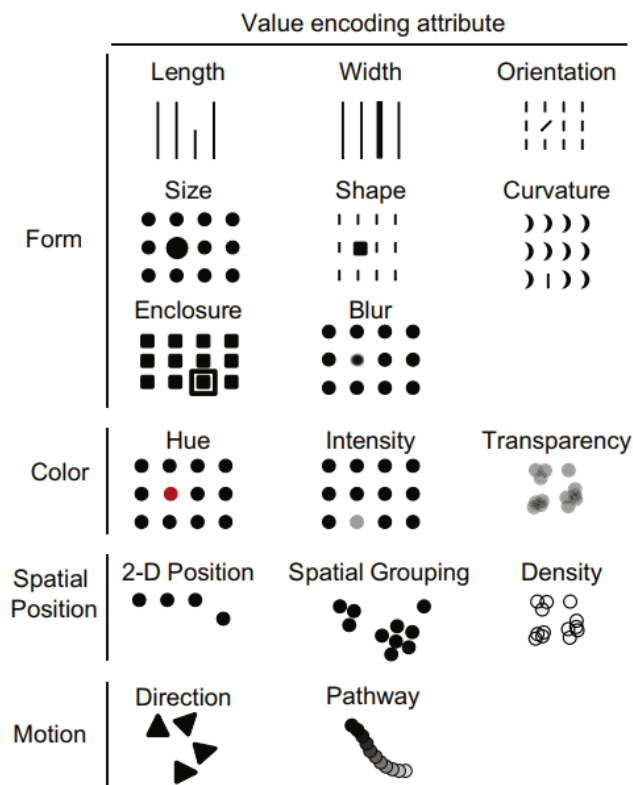
Contoh grafik sederhana dibawah ini:



**Gambar 25 .** Grafik tiga dimensi (3D) dan Grafik dua dimensi (2D)

**2. Kedua, memilih dan mempertimbangkan jenis encoding object dan value attribute yang akan digunakan dalam proses visualisasi data**

Tahap selanjutnya, dalam visualisasi data diperlukan pemilihan encoding object dan value attribute yang tepat. Encoding objects yaitu bar, line, dan point. Sedangkan value attribute berupa panjang garis, colour, point position. Pertimbangan dalam pemilihan encoding object dan value attribute melihat kebiasaan umumnya yang lebih memahami value attribute berupa panjang garis dengan grafik dua dimensi (2D), bila dibandingkan dengan lainnya. Tetapi jika kita memerlukan perbandingan atau pola umum yang terjadi, dapat menggunakan value attribute berupa ketebalan garis.

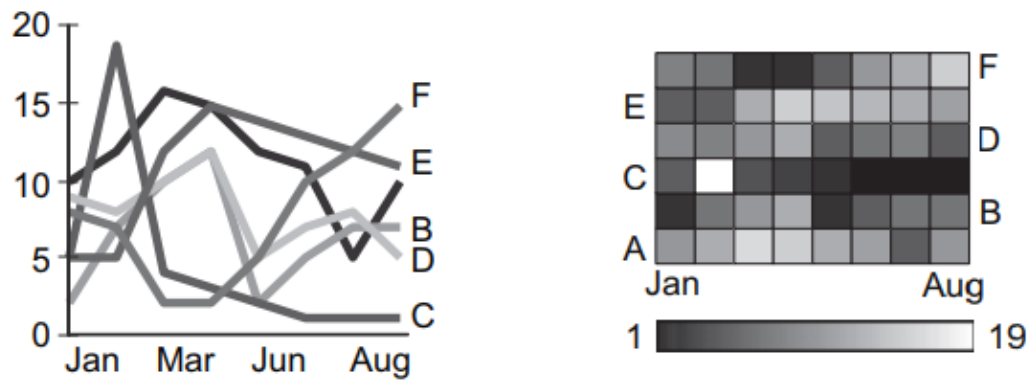


**Fig. 2.** The following attributes corresponding to different encoding objects can be used in data visualization to highlight contrasts between different parts of a graphic, to represent quantitative differences, or to categorize and group information. Adapted from [Few \(2009, p. 39\)](#).

**Gambar 26.** *Value encoding attribute* dalam visualisasi data

### 3. Ketiga, fokus pada tujuan memvisualisasi yakni pola atau detail

Dalam melakukan visualisasi, juga diperlukan memilih jenis visualisasi yang tepat. Misalkan kalau tujuan dari visualisasi itu membuat sebuah pola, maka jenis heatmap dapat memvisualisasikan dataset menjadi kotak-kotak yang memperlihatkan value/nilai dari data sedangkan magnitude data direpresentasikan dengan warna pada heatmap. Kelebihan dari heatmap adalah skala pewarnaan yang mudah dipahami sedangkan kekurangannya tidak menjelaskan secara detail perbedaan value/nilai antar data. Bila tujuan visualisasi data secara detail, misalkan ingin mengetahui perbedaan antar line, maka dengan bar atau line bisa menjadi pilihan. Gambaran line dan heatmap seperti berikut:

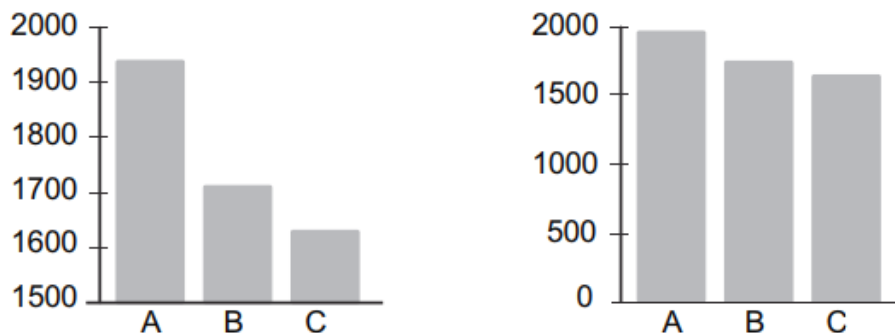


**Gambar 27.** Grafik line chart dan Grafik heatmap

#### 4. Keempat, memulai grafik dari rentang titik nol

Pada proses interpretasi data dimana informasi yang disampaikan kepada user menjadi penting dan bermanfaat, perlu melakukan teknik memulai grafik dari rentang titik nol. Teknik ini diperlukan agar user tidak salah memahami arti dari visualisasi yang telah dibuat. Jika teknik ini tidak digunakan, maka tidak menutup kemungkinan informasi yang disampaikan berbeda dengan pemahaman dari sisi user.

Dibawah ini merupakan contoh grafik yang tidak dimulai dengan rentang titik nol dan grafik yang menggunakan rentang titik nol:



**Gambar 28.** Perbedaan Grafik dengan rentang sumbu titik nol

Pada grafik diatas menggunakan dataset yang sama, tetapi divisualisasikan dengan teknik yang berbeda. Grafik sebelah kiri, pada sumbu vertikal dimulai dengan nilai 1500 sampai dengan 2000 dengan rentang yang cukup jauh antara bagan A, B, dan C.

Sedangkan pada grafik sebelah kanan, sumbu vertikal dimulai dengan rentang nilai nol (0) dengan rentang sampai dengan nilai 2000. Grafik disebelah kanan dapat terlihat perbedaan antara bagan A, B, dan C dengan rentang yang

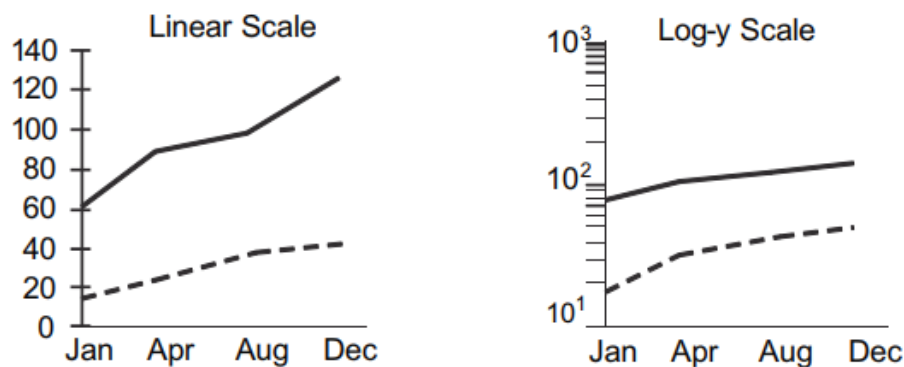
sesuai dan jarak yang tidak terlalu jauh nilainya.

Dari grafik diatas dapat disimpulkan bahwa grafik sebelah kiri terlihat melebih-lebihkan jarak dan kurang pas dalam hal memvisualisasikan data. Sedangkan grafik pada sebelah kanan terlihat sesuai dan jarak yang diambil pas, sehingga informasi yang disampaikan dapat diterima user/pengguna dengan baik.

##### 5. Kelima, teknik visualisasi data dan grafik terhadap data time series dengan bentuk skala.

Data time series merupakan data yang berubah dari waktu ke waktu yang dialami sebuah objek. Untuk melakukan visualisasi data time series memerlukan teknik visualisasi data dengan dua teknik khusus, yaitu skala linier dan skala logaritmik. Penggunaan teknik dari skala linier yakni melihat perubahan data time series dengan vertical axis. Sedangkan teknik dari skala logaritmik dilakukan apabila ingin menjadikan alternatif saat memvisualisasikan data time series yang kita miliki.

Adapun contoh penggunaan teknik skala liner dan skala logaritmik dengan melihat pada gambar dibawah ini:



**Gambar 29.** Teknik visualisasi data dengan skala linier dan skala logaritmik

Pada grafik diatas terlihat garis lurus utuh dan garis putus-putus. Grafik sebelah kiri, yakni grafik teknik skala linier memperlihatkan jarak perbedaan pada periode bulan Januari ke April dengan ditunjukkan dengan garis lurus utuh. Sedangkan grafik sebelah kanan, yakni grafik teknik skala logaritmik memperlihatkan jarak perbedaan pada periode bulan Januari ke April dengan

ditunjukkan dengan garis putus-putus.

Dari penjelasan diatas, maka dapat disimpulkan perbedaan antara grafik dengan skala logaritmik tidak memperlihatkan perubahan dari nilai asli tetapi menunjukkan persentase dari perubahan. Pada teknik visualisasi data logaritmik dengan cara menghapus skewness didalam dataset dan memperlihatkan persentase perubahan pada initial value. Maka pada grafik terdapat slope yang menunjukkan persentase perubahan.

Kesimpulan, kelima teknik dalam visualisasi data dapat digunakan agar visualisasi yang dihasilkan efektif dan informasi yang disampaikan kepada user dapat dipahami dengan baik.

#### **D. Rangkuman**

Visualisasi data adalah mempresentasikan data dan informasi dalam kumpulan grafik maupun diagram baik berupa data tabular dan juga data geospasial sehingga mudah dipahami. Estetika dalam visualisasi data harus memberikan kesan menyenangkan. Data mentah yang kita miliki tentunya belum menjadi sebuah informasi bila belum dilakukan representasi data dengan teknik visualisasi data.

Elemen-elemen visual yang digunakan seperti grafik, diagram dan geospasial diperlukan untuk berkomunikasi secara visual, sehingga menjadi *insight* dalam mengambil sebuah kebijakan/keputusan

Tentunya kita perlu mempelajari jenis-jenis dari visualisasi data agar mudah menentukan jenis mana yang akan kita gunakan. Jenis dari visualisasi data diantaranya: Visualisasi data temporal (*Temporal data visualizations*) , Visualisasi data hierarki (*Hierarchical data visualizations*), Visualisasi data jaringan (*network data visualizations*), Visualisasi data multidimensi (*multidimensional data visualizations*), Visualisasi data geospasial (*geospatial data visualizations*)

Visualisasi data memiliki pengaruh penting dalam proses penelitian, menyampaikan presentasi, sehingga perlu mengetahui Teknik-teknik dalam visualisasi data supaya visualisasi yang dihasilkan efektif

#### **E. Evaluasi**

1. Jelaskan pentingnya visualisasi data penting dalam penyajian data
2. Sebutkan dan jelaskan jenis-jenis visualisasi data
3. Jelaskan teknik-teknik dalam visualisasi data agar visualisasi yang dihasilkan efektif.

## MATERI POKOK 4:

# INTERPRETASI HASIL ANALISIS

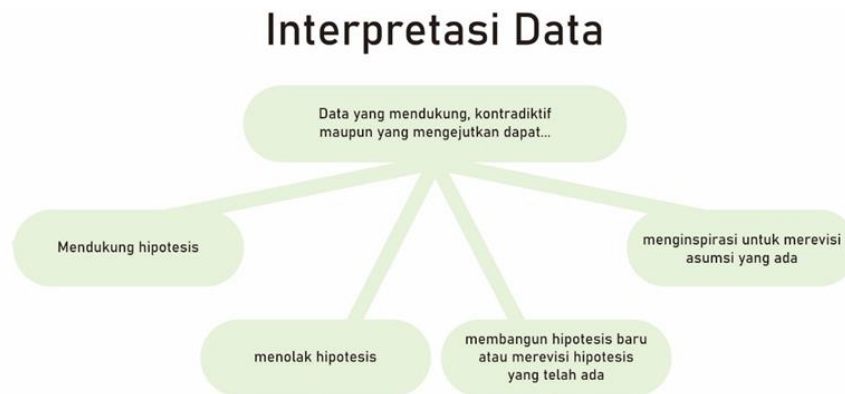
### **Indikator Hasil Belajar:**

Peserta mampu **melakukan interpretasi dari hasil analisis data dengan benar**, meliputi:

1. Memahami Konsep Interpretasi Data dengan benar;
2. Memahami Metode Interpretasi Data dengan benar;
3. Membuat Interpretasi Dari Hasil Analisis Data sesuai dengan tujuan yang ingin dicapai.

### **A. Pengantar Interpretasi Hasil Analisis**

Istilah interpretasi hasil analisis mengacu kepada suatu proses mengkritisi dan menentukan signifikansi dari informasi yang penting seperti hasil survei, temuan eksperimental, observasi ataupun laporan naratif sehingga dapat mengantarkan pada suatu kesimpulan (Islam et al., 2022). Interpretasi bertujuan untuk memberikan makna yang lebih mendalam terhadap informasi yang telah dianalisa serta menjelaskan berbagai macam implikasinya. Interpretasi dapat dikatakan sebagai langkah untuk menempatkan hasil analisis pada suatu konteks informasi yang relevan sehingga diperlukan pemahaman terhadap berbagai perspektif ilmu dalam menginterpretasi suatu hasil analisis. Richmond (2006) menyampaikan bahwa dua orang peneliti dapat memberikan interpretasi berbeda terhadap suatu analisis data yang sama tergantung perspektif yang mereka gunakan.



Sumber: Islam et al. (2022)

**Gambar 30. Proses Interpretasi Hasil Analisis**

Berdasarkan proses pada gambar 2, tahapan interpretasi data dapat mendukung suatu hipotesis, menolak suatu hipotesis, merevisi/memperbarui hipotesis maupun merevisi asumsi yang sebelumnya diyakini. Tahapan interpretasi berbeda dengan analisis yang bertujuan untuk mencari sebuah informasi atau temuan dalam kumpulan data. Interpretasi data berupa kegiatan untuk mendalami makna dari sebuah temuan sehingga bisa menjadi sebuah kesimpulan, bahkan menjadi referensi untuk membuat studi yang lebih baik kedepannya.

## **1. Interpretasi Berdasarkan Jenis Data**

### **a. Interpretasi Hasil Analisis Data Kualitatif**

Interpretasi Data Kuantitatif berarti bahwa peneliti menarik kesimpulan dari hasil analisis data untuk menjawab pertanyaan penelitian, menguji hipotesis dan makna yang lebih luas dari analisis yang dilakukan (Creswell & Creswell, 2018). Secara umum analisis data kuantitatif menghasilkan berbagai macam nilai statistik yang harus diinterpretasi sehingga memiliki makna. Tabel 5 menjelaskan cara untuk menginterpretasi beberapa contoh nilai statistik.

**Tabel 6.** Interpretasi Nilai Statistik Data Kuantitatif

| No | Nilai Statistik              | Interpretasi                                                                                                                                                                                                                                                                                                                                                                                                  |
|----|------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1  | Rata-rata                    | Menunjukkan nilai pusat pada suatu kumpulan data tertentu. Nilai ini dapat digunakan sebagai nilai ekspektasi matematis atau nilai yang mewakili sekelompok hasil pengamatan.                                                                                                                                                                                                                                 |
| 2  | Modus                        | Nilai yang paling banyak muncul pada suatu kumpulan data. Nilai ini dapat digunakan sebagai nilai pusat data nominal atau kategorikal.                                                                                                                                                                                                                                                                        |
| 3  | Deviasi Standar dan Variansi | Menunjukkan seberapa besar penyimpangan dari suatu nilai pusat. Nilai ini menggambarkan tingkat konsistensi suatu data terhadap nilai tengah/rata-ratanya. Semakin besar nilainya maka akan suatu data semakin tersebar dari nilai tengahnya.                                                                                                                                                                 |
| 4  | Jangkauan                    | Sejauh apa data tersebar antara nilai terendah dengan nilai tertinggi.                                                                                                                                                                                                                                                                                                                                        |
| 5  | Distribusi Frekuensi         | Nilai statistik ini adalah pengukuran terhadap tingkat kemunculan nilai dalam suatu kumpulan data.                                                                                                                                                                                                                                                                                                            |
| 6  | <i>P-Value</i>               | Nilai yang mengukur seberapa kuat yang kita miliki untuk menolak hipotesis nol/ <i>null hypothesis</i> . Semakin besar nilai <i>P-Value</i> maka semakin kuat bukti untuk menerima <i>null hypothesis</i> .                                                                                                                                                                                                   |
| 7  | Nilai Korelasi               | Nilai keeratan dari dua buah variabel. Terdapat tiga kemungkinan dalam penghitungan nilai korelasi:<br>1. Kedua variabel saling meningkat atau menurun maka memiliki nilai korelasi mendekati positif 1.<br>2. Jika salah satu nilai variabel meningkat dan lainnya menurun maka memiliki nilai korelasi mendekati negatif 1.<br>3. Tidak ada hubungan yang jelas antara dua variabel maka nilai mendekati 0. |

Sumber: Islam et al. (2022), Terrel (2021)

#### **b. Interpretasi Hasil Analisis Data Kualitatif**

Interpretasi Data Kualitatif berarti bahwa peneliti menarik sebuah makna dari temuan dari analisis data (Creswell & Creswell, 2018). Makna ini dapat berupa pelajaran yang telah diperoleh, membandingkan informasi dengan literatur lainnya dan pengalaman pribadi. Secara umum analisis data kualitatif menghasilkan suatu narasi yang bersifat deskriptif yang telah dikelompokkan kedalam tema-tema tertentu.

**Tabel 7.** Interpretasi Pada Jenis Data Kuantitatif

| No | Jenis Data | Cara Interpretasi                                                                                                                                                                                                         |
|----|------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1  | Observasi  | Interpretasi dilakukan dengan menajamkan pola-pola perilaku yang terjadi dalam suatu kelompok. Pola tersebut dapat dijelaskan lebih dalam terhadap waktu melakukan, jenis perilaku maupun metode komunikasi yang dipakai. |
| 2  | Dokumen    | Interpretasi dilakukan dengan memetakan kode-kode yang muncul dari berbagai dokumen yang dikumpulkan.                                                                                                                     |
| 3  | Wawancara  | Interpretasi dilakukan melalui tema, topik dan kategori berdasarkan jawaban pertanyaan wawancara.                                                                                                                         |

Sumber: Islam et al. (2022)

## **2. Kesalahan Dalam Interpretasi**

### **a. Korelasi yang dianggap sebagai kausalitas**

Kesalahan pertama yang rentan terjadi dalam interpretasi adalah kecenderungan analis data dalam menyampuradukkan suatu fenomena dalam korelasi. Jika dua variabel memiliki korelasi yang tinggi bukan berarti satu variabel memengaruhi variabel yang lainnya. Memaksakan nilai korelasi akan menghasilkan interpretasi yang tidak bermakna. Suatu kausalitas perlu didukung dengan suatu temuan-temuan terdahulu, kajian ilmiah maupun masukan dari pakar.

### **b. Bias Konfirmasi**

Kesalahan interpretasi yang kedua dipicu adanya bias konfirmasi. Seorang analis harus menginterpretasi suatu hasil analisis secara objektif dan terbuka terhadap segala kemungkinan yang terjadi. Sebagai contoh seorang analis adalah sosok yang peduli akan lingkungan tetapi tidak peduli terhadap perkembangan teknologi sehingga saat menganalisis data kesehatan hanya akan membahas peran atau pengaruh lingkungan terhadap kesehatan dan menutup diri akan peran teknologi dalam mendukung kesehatan.

**c. Penggunaan data yang tidak relevan**

Kesalahan ketiga yang rentan terjadi dalam interpretasi data adalah penggunaan data yang tidak relevan. Relevansi penggunaan data harus ditentukan sejak dalam masa perumusan tujuan sehingga temuan yang dihasilkan akan secara utuh mewakili fenomena yang akan dianalisis.

**B. Tahap Interpretasi Hasil Analisis**

Sebagai suatu tahap tersendiri dalam proses analisis data, dibutuhkan sebuah tahapan yang dapat menjadi rujukan dalam melakukan interpretasi data sehingga tercipta pemahaman yang lebih mendalam terhadap suatu temuan. Baik data kuantitatif maupun kualitatif, suatu hasil analisis perlu diinterpretasi menggunakan sebuah informasi dalam konteks yang relevan. Terdapat 4 langkah yang dapat digunakan untuk menghasilkan interpretasi hasil analisis yang mendalam. Berikut langkah-langkah interpretasi data yang disampaikan oleh Dates (2021):

**1. Langkah-langkah melakukan interpretasi hasil analisis data**

Dalam mengembangkan suatu interpretasi kita perlu mengumpulkan dan menyatukan seluruh informasi yang relevan terhadap temuan/hasil analisis data. Sebagai contoh, informasi tersebut dapat berupa: standar dalam suatu pengukuran, hasil observasi umum, data atau hasil analisis lainnya yang serupa. Dates (2021) turut menyampaikan tiga jenis informasi yang dapat dijadikan referensi dalam menciptakan interpretasi yang lebih mendalam dan komprehensif:

**a. Sensitivitas dari metode dan instrumen yang digunakan dalam analisis data**

Instrumen dan metode yang digunakan dapat memberikan batasan terhadap temuan dari suatu data. Kita dapat memberikan informasi mengenai berbagai potensi yang bisa didapat jika menggunakan instrumen atau metode lainnya.

**b. Tingkat perubahan nilai dalam suatu indikator**

Perbedaan nilai indikator maupun dimensi dalam suatu pengukuran bisa menjadi informasi yang dapat dielaborasi terhadap hasil analisis yang diperoleh. Suatu hasil pengukuran dapat dibandingkan dengan standar yang telah ada atau pengukuran-pengukuran sebelumnya.

**c. Tingkat kepercayaan terhadap pengambilan data**

Kendala-kendala yang dialami dalam pengambilan data bisa menjadi informasi tambahan dalam melengkapi hasil analisis. Pernyataan kita terhadap kendala yang dialami akan memperkaya konteks dari temuan yang diperoleh.

**2. Kembangkan Temuan**

Temuan merupakan suatu hasil observasi dari data yang kita miliki dan suatu pernyataan yang merangkum seluruh aspek penting. Interpretasi akan membuat sebuah temuan menjadi jembatan dalam memutuskan kesimpulan yang tepat. Terdapat tiga sumber yang dapat digunakan untuk mengembangkan temuan yakni: membandingkan dengan standar, membandingkan temuan dengan data yang kita miliki, dan membandingkan temuan dengan data atau hasil analisis yang lain. Berikut penjelasan dari tiga sumber interpretasi tersebut:

**a. Bandingkan temuan dengan standar yang diketahui**

Hasil pengukuran suatu indikator maupun dimensi dapat dibandingkan dengan standar yang sudah ditentukan. Jika suatu sumber tidak ditemukan maka perlu berdiskusi dengan seorang pakar untuk memberikan interpretasi mengenai keseusian nilai yang kita ukur. Perbedaan nilai pengukuran dan suatu standar dapat menjadi informasi tambahan terhadap suatu fenomena seperti seberapa jauh nilai pengukuran menyimpang dari standar.

**b. Bandingkan temuan dengan data yang kita miliki**

Membandingkan temuan dengan data yang dimiliki berujuan untuk memperdalam hasil temuan. Sebagai contoh, kita dapat membandingkan pengamatan mana yang memiliki nilai lebih rendah atau lebih tinggi dari rata-rata, apakah ada pencilan/*outliers* dalam data, ditanggal berapa suatu

pengamatan memiliki nilai lebih tinggi dari rata-rata, dan lain sebagainya.

**c. Bandingkan temuan dengan data atau temuan lainnya**

Memantau data dari sumber yang lainnya akan memperluas perspektif kita terhadap suatu fenomena. Suatu temuan yang kita hasilkan dalam data perlu diberikan sebuah konteks sehingga dapat dibandingkan dengan temuan lainnya. Kita perlu memberikan informasi mengenai seluruh kondisi yang terjadi pada saat pengumpulan maupun analisis data sehingga dapat ditemukan perbedaan dengan hasil analisis yang lainnya.

**3. Kembangkan Kesimpulan**

Kesimpulan adalah penjelasan mengenai mengapa suatu temuan tercipta seperti itu. Kesimpulan yang ditarik harus sesuai dengan dan menjawab pertanyaan yang diajukan pada awal proses analisis data. Proses interpretasi data dalam hal ini memiliki peran untuk memberikan penjelasan lebih mendalam mengenai berbagai konteks dari suatu proses analisis data sehingga kita memiliki argumentasi untuk menarik sebuah kesimpulan sesuai dengan kondisi yang kita alami.

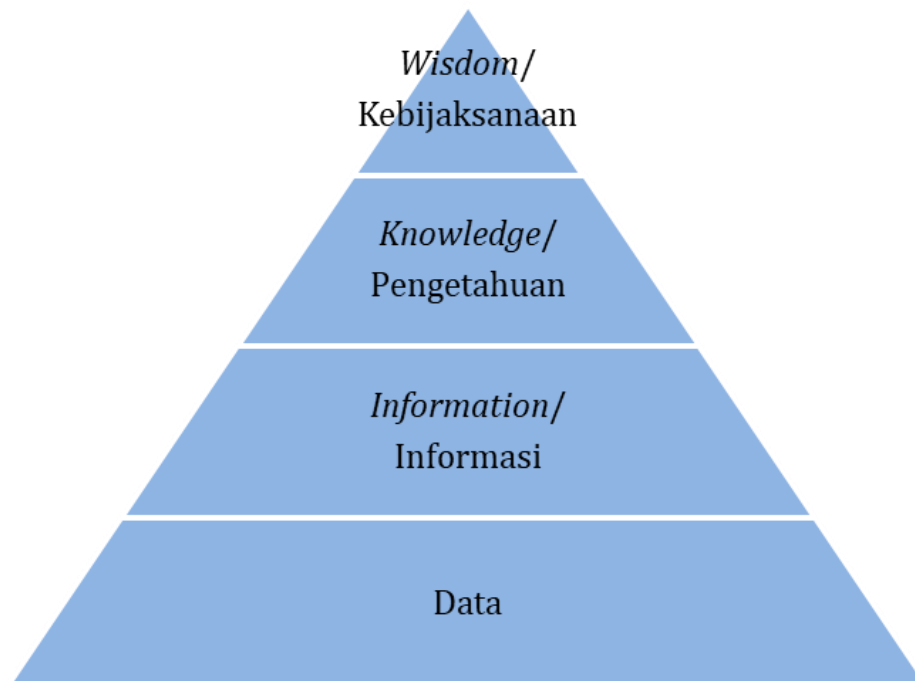
**4. Kembangkan Rekomendasi**

Rekomendasi diperoleh berdasarkan temuan dan kesimpulan yang ditarik. Terdapat dua jenis kesimpulan yakni: aksi yang harus dikerjakan serta informasi selanjut yang dapat atau harus digali lebih dalam. Contoh rekomendasi aksi sebagai berikut: Dinas Kesehatan Pemerintah Provinsi X harus membuat program peningkatan kebersihan lingkungan dalam rangka menekan kasus DBD yang tinggi. Adapun rekomendasi informasi sebagai berikut: Diperlukan pemahaman lebih mendalam mengenai pemahaman anggota rumah tangga di Provinsi X terhadap penanganan jentik nyamuk di lingkungan rumahnya.

**C. Hirarki Pengetahuan**

Sebagai sebuah proses, analisis data merupakan suatu kegiatan yang memerlukan *input*/masukan dan *output*/keluaran. Secara singkat proses tersebut menggunakan data sebagai *input* dan menghasilkan temuan/*findings*

sebagai *output*. Namun begitu diperlukan sebuah pemahaman lebih dalam mengenai konsep *input* dan *output* tersebut sehingga diperoleh pemahaman lebih mendalam mengenai fungsi dari suatu analisis data.



Sumber: Rowley (2007)

**Gambar 31.** Piramida Data-Information-Knowledge-Wisdom (DIKW)

Piramida DIKW merupakan sebuah model untuk menggambarkan bagaimana perjalanan dari pemanfaatan suatu data sehingga memiliki nilai guna. Model tersebut turut memberikan definisi dari tiap-tiap komponennya sehingga dapat dibedakan masing-masing fungsinya. Berikut definisi dari tiap komponen Piramida DIKW:

- Data didefinisikan sebagai simbol yang mewakili sifat suatu objek, kejadian dan lingkungannya. Data adalah produk dari suatu kegiatan observasi. Perbedaan data dan informasi bersifat fungsional dimana data tidak memiliki nilai guna sebelum dilakukan pengolahan atau analisis dengan metode tertentu untuk menjadi informasi.
- Informasi berisikan deskripsi yang berisikan jawaban untuk pertanyaan-pertanyaan seperti siapa, apa, kapan dan seberapa banyak. Informasi merupakan suatu bentuk data yang dapat dibaca oleh manusia dimana

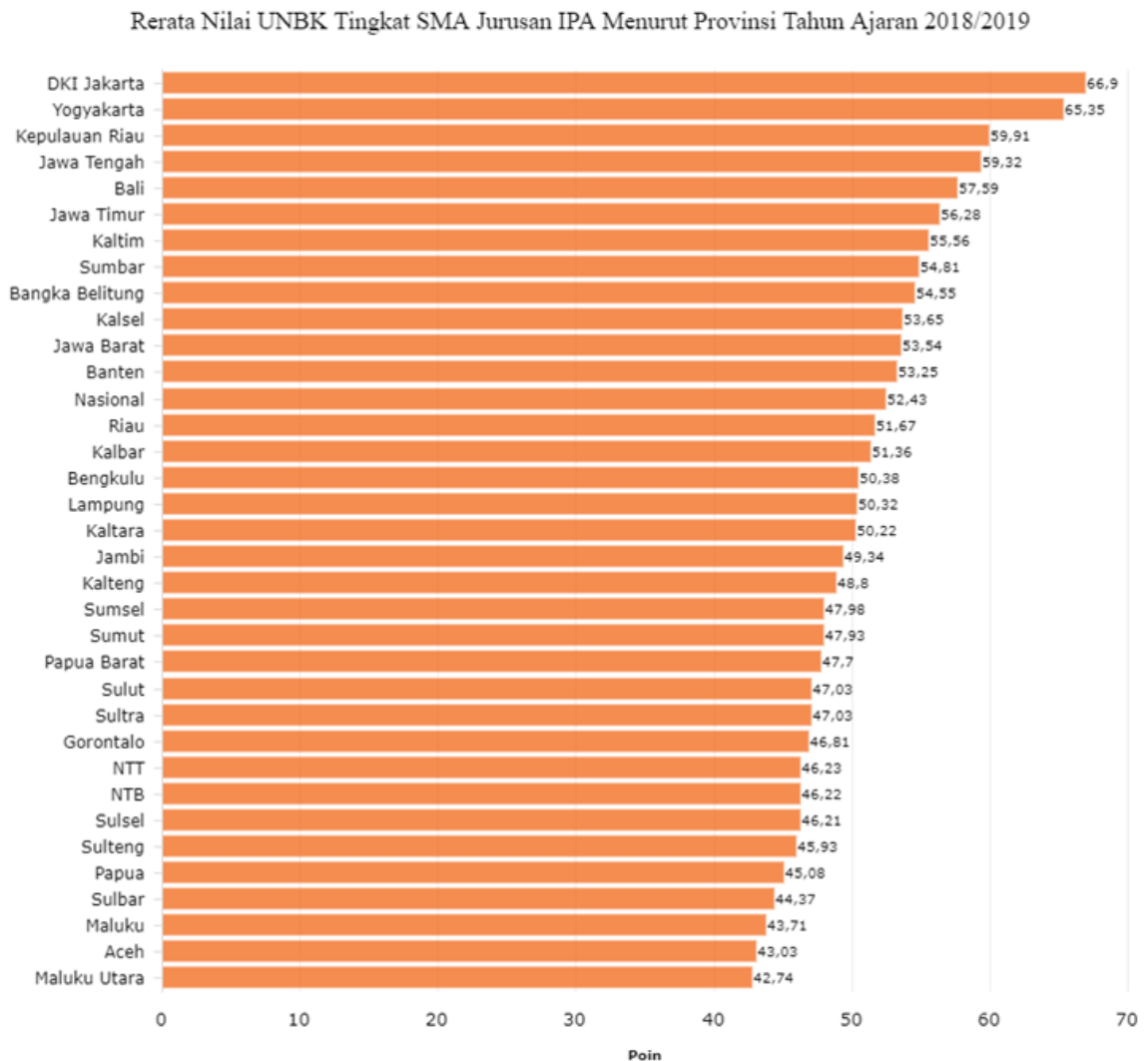
secara umum digambarkan melalui visualisasi seperti grafik, tren, diagram dan bentuk lainnya.

- Pengetahuan adalah kemampuan untuk mengubah informasi yang dapat menjadi sebuah instruksi/aksi. Dalam menciptakan sebuah pengetahuan diperlukan pengalaman, opini serta keahlian dari seorang *expert*. Sesuai dengan bentuk diagram piramida DIKW, tidak semua informasi dapat diubah menjadi sebuah instruksi.
- Kebijakan adalah kemampuan untuk mengenali pilihan yang relevan dan dapat diterapkan untuk situasi yang dihadapi, membandingkan pro dan kontra dari pendekatan yang dipilih serta membuat keputusan yang optimal dan efektif bagi suatu pekerjaan.

Sebagai sebuah proses, analisis data merupakan sebuah perjalanan dimana data diolah sedemikian rupa sehingga dapat menjadi sebuah informasi. Piramida DIKW merupakan model konseptual yang dapat digunakan untuk menggambarkan bahwa suatu analisis data yang menghasilkan sebuah temuan/*findings* masih perlu dielaborasi dengan pengalaman dan keahlian dari seorang *expert* dalam menciptakan sebuah aksi. Rowley (2007) menyampaikan bahwa semakin tinggi piramida maka akan semakin mudah suatu makna disampaikan dan diaplikasikan kepada manusia. Kebijakan dalam hal ini adalah suatu tingkat yang paling tinggi dalam pendayagunaan data dan merupakan proses pengaplikasian temuan dari suatu data yang paling baik. Melalui piramida ini dapat disampaikan bahwa kemampuan dalam meninterpretasi hasil menjadi proses yang penting dalam proses analisis data karena suatu informasi pada akhirnya harus dapat dikonversi, diterima dan dimengerti oleh manusia.

#### **D. Studi Kasus Interpretasi Hasil Analisis Data**

Studi kasus yang akan dibahas merupakan hasil analisis data dari bidang pendidikan.



Dkatadata.co.id

databoks

Sumber: Kusnandar (2019)

**Gambar 32 .** Piramida Data Information-Knowledge-Wisdom (DIKW)

Gambar 30 merupakan visualisasi dari rata-rata nilai UNBK tingkat SMA Jurusan IPA per Provinsi pada Tahun Ajaran 2018/2019. Dari visualisasi tersebut kita dapat melihat mana provinsi yang memiliki nilai tertinggi, terendah, maupun variansi nilai dari rata-ratanya. Selain itu kita bisa melihat provinsi mana saja yang memiliki nilai rata-rata di atas maupun dibawah rata-rata secara nasional. Dalam studi kasus ini, dimisalkan Kementerian Pendidikan Pendidikan, Kebudayaan, Riset dan Teknologi/Kemendikbudristek ingin melihat lima provinsi mana yang memiliki nilai rata-rata UNBK yang paling rendah

sehingga program kementerian selanjutnya difokuskan kepada daerah-daerah yang memiliki nilai rata-rata UN rendah. Berikut contoh interpretasi data yang dapat disampaikan dalam memperkaya hasil analisis di gambar 7 dalam perspektif infrastruktur sekolah:

“Lima Provinsi yang memiliki nilai rata-rata UN paling rendah adalah Papua, Sulawesi Barat, Maluku, Aceh dan Maluku Utara. Jika dilihat dari statistik mengenai kondisi infrastruktur sekolah, lima provinsi tersebut memiliki persentase jumlah sekolah berkondisi baik yang relatif rendah sehingga kurang memungkinkannya terjadi proses belajar dan mengajar yang efektif. Begitu pula jika kita melihat dari faktor guru. Pengawasan terhadap perbaikan kualitas infrastruktur sekolah harus ditingkatkan oleh pihak-pihak yang berwenang sehingga semakin banyak sekolah-sekolah di lima provinsi tersebut yang berstatus baik. Penggunaan anggaran perbaikan sekolah perlu di analisis untuk mendapatkan gambaran yang lebih luas”

## **E. Rangkuman**

Interpretasi data merupakan sebuah tahapan dalam proses analisis data yang bertujuan untuk memberikan makna lebih mendalam pada hasil analisis data sehingga dapat mengarahkan pada suatu kesimpulan maupun rekomendasi yang komprehensif. Melalui tahap interpretasi ini kita dapat menguji hipotesis maupun asumsi yang dibangun pada awal proses analisis data. Secara umum terdapat perbedaan cara menginterpretasi hasil analisis jenis data kuantitatif dan kualitatif. Dalam proses interpretasi, kita dapat membandingkan hasil analisis data dengan suatu standar, temuan lainnya yang serupa maupun data yang kita miliki. Proses interpretasi menjadi penting karena pada dasarnya suatu informasi harus dapat dimaknai dengan tepat dan dimengerti oleh khalayak yang membacanya agar menjadi rekomendasi aksi yang memiliki nilai.

## **F. Evaluasi**

1. Jelaskan perbedaan analisis dan interpretasi data
2. Sebutkan dan jelaskan 4 langkah untuk menghasilkan interpretasi yang mendalam
3. Pilihlah satu topik yang Anda kuasai dan lakukanlah sebuah analisis terhadap satu data. Interpretasikan hasil analisis tersebut menggunakan informasi-informasi yang relevan sehingga membentuk kesimpulan yang kuat.

## MATERI POKOK 5:

# PRAKTIK VISUALISASI DATA

### **Indikator Hasil Belajar:**

Peserta mampu **membuat visualisasi data dengan aplikasi tableau, dengan benar, melalui:**

1. Mengetahui Aplikasi Tableau untuk Analisis dan Visualisasi Data dengan mudah
2. Memahami Analisis dan Interpretasi Data Univariat, Bivariat, Multivariat, dengan tepat
3. Mempraktikkan Analisis dan Visualisasi Data, sesuai dengan tujuan

### **A. Pengenalan Aplikasi Tableau untuk Analisis dan Visualisasi Data**

Salah satu perangkat lunak (*software*) dalam visualisasi data yaitu tableau. Tableau merupakan *tools* yang membuat visualisasi data yang kita rancang agar mudah dipahami, dapat dianalisa, dan tampilannya lebih interaktif. Proses visualisasi data pada tableau dengan mengolah data mentah atau kaku menjadi data bentuk diagram, grafik, map, dan lainnya sehingga menunjukkan perbedaan atau perubahan data tersebut menjadi lebih jelas dan interaktif.

Tableau juga menyediakan *dashboard* interaktif untuk menghasilkan tampilan data dari analisa statistik menjadi lebih menarik dan secara visual yang *eye catching* dengan beragam tampilan tipe grafik, tipe diagram, dan *colour*.

Kemampuan tableau dalam memvisualisasikan data tidak diragukan sehingga data yang dihasilkan menjadi lebih mudah dianalisis, data lebih mudah dibaca dan juga interaktif. Selain digunakan untuk visualisasi data, tableau juga dipergunakan dalam analisis data dan pembuatan laporan. Teknik visualisasi dalam tableau yakni membuat data dalam bentuk tabel baik sederhana maupun kompleks menjadi sebuah informasi dalam bentuk diagram, peta (map) maupun grafik, sehingga data menjadi informasi terlihat lebih menarik secara visual, lebih jelas, dan informatif. Tableau juga menyediakan dashboard untuk menyusun data dan informasi serta analisis statistik yang akan disajikan dalam satu halaman dashboard interaktif juga tampilan yang menarik

secara visual. Sistem *drag and drop* dalam *software* tableau mempermudah user dalam menggunakan tableau untuk menunjang kegiatannya. Tableau bisa mengakomodir dan menggabungkan/mengintegrasikan data dalam berbagai *source* seperti ms.excel/spreadsheet, berbagai jenis database, dan cloud storage. Melalui kemampuan dan kelebihan yang dimiliki Tableau tersebut maka aplikasi tersebut akan menjadi alat yang digunakan dalam melakukan analisis deskriptif pada modul ini.

Secara umum aplikasi Tableau memiliki dua jenis versi, yaitu Tableau *Public* dan Tableau Desktop. Kedua jenis ini masing-masing memiliki perbedaan bila dilihat dari kelebihan dan kekurangan. Dibawah ini tabel 8 penjelasan perbedaan Tableau *Public* dan Tableau Desktop:

**Tabel 8 .** Perbedaan Tableau Public dan Tableau Desktop

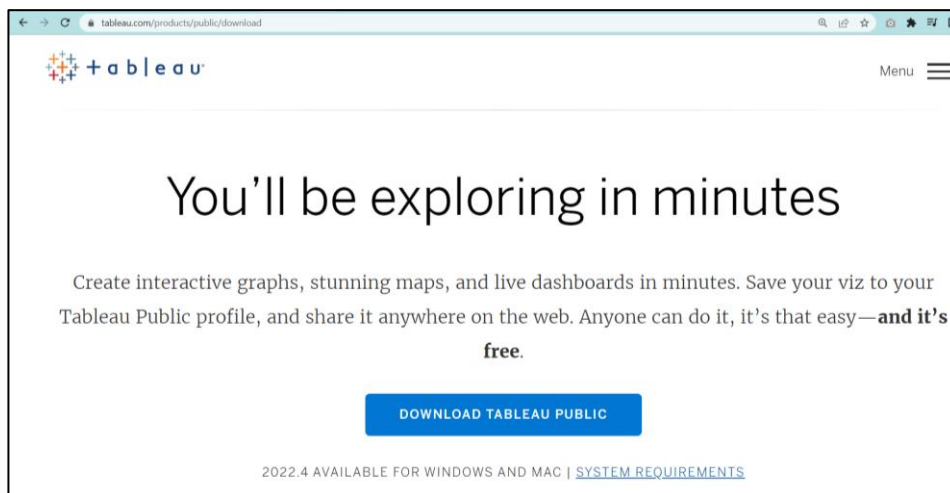
| <b>Perbedaan</b>                                                       | <b>Tableau <i>Public</i></b>                                                                                                                                             | <b>Tableau Desktop</b>                                                                                                                                       |
|------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <i>Open Source</i>                                                     | Gratis ( <i>open source</i> )                                                                                                                                            | Berbayar ( <i>bukan open source</i> )                                                                                                                        |
| Jumlah Baris Data yang dapat diunggah                                  | Terbatas (100 baris data dalam sekali koneksi <i>file</i> yang dilakukan)                                                                                                | Tidak terbatas                                                                                                                                               |
| Konektivitas Sumber Data/Jenis file sumber dataset yang bisa digunakan | Ms. Excel, Text, Ms. Access, TDE                                                                                                                                         | Semua jenis yang tersedia di tableau (termasuk ke database Amazon Redshift, Postgres, SQL Server, Google BigQuery dll)                                       |
| Ruang Kerja                                                            | <i>Cloud</i><br>(Tidak bisa menyimpan <i>workbook</i> hasil kerja secara <i>local</i> namun akan disimpan di <i>Tableau public server</i> yang dapat diunduh siapa saja) | <i>Local</i><br>(Hasil kerja di Tableau Tableau <i>Workbook/TWB</i> dan Tableau Data<br><br><i>Extract/TDE</i> disimpan secara <i>local</i> di desktop kita) |
| Kerja                                                                  | Online                                                                                                                                                                   | Offline and online                                                                                                                                           |
| Keamanan Data                                                          | Tidak terjamin<br>(Semua hasil kerja tersimpan secara online dan tidak bisa disimpan secara <i>local</i> )                                                               | Terjamin<br>(Semua hasil kerja disimpan secara <i>local</i> dan dapat diatur <i>password</i> serta username)                                                 |

Tableau Public merupakan *software* yang disediakan secara gratis (*open source*) yang digunakan dalam membuat visualisasi data menjadi lebih mudah dianalisis, data lebih mudah dibaca dan interaktif. Tableau public dapat terhubung dengan data atau *file* yang kita miliki berupa spreadsheet atau Microsoft Excel.

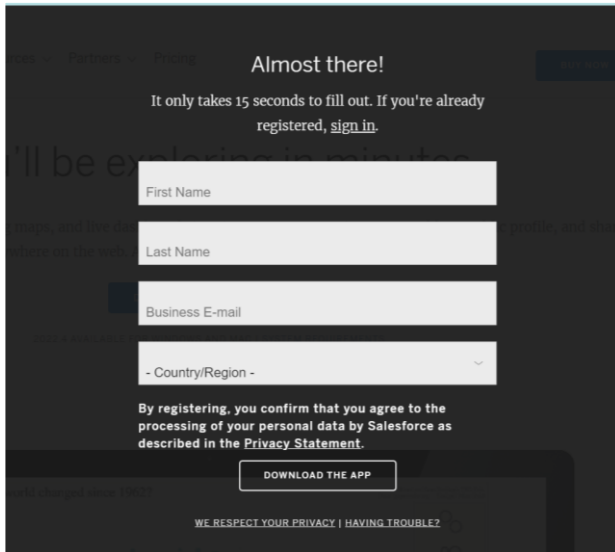
Tutorial atau langkah-langkah yang dapat dipergunakan dalam menggunakan Tableau Public adalah sebagai berikut:

Tahap pertama, download *software* Tableau Public yang dapat diakses secara gratis:

1. Buka halaman URL berikut <https://public.tableau.com/en-us/s/download> selanjutnya klik download tableau public.



2. Setelah itu, langkah selanjutnya input nama, alamat email, dan negara, lalu klik tombol download the app.



Almost there!

It only takes 15 seconds to fill out. If you're already registered, [sign in](#).

First Name

Last Name

Business E-mail

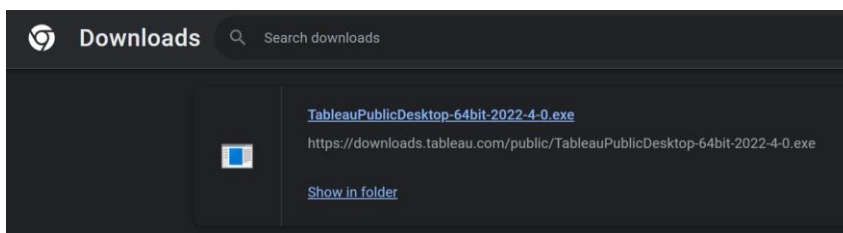
- Country/Region -

By registering, you confirm that you agree to the processing of your personal data by Salesforce as described in the [Privacy Statement](#).

[DOWNLOAD THE APP](#)

[WE RESPECT YOUR PRIVACY](#) | [HAVING TROUBLE?](#)

3. Proses selanjutnya, *software* Tableau Public akan terdownload dan tunggu hingga proses download selesai.



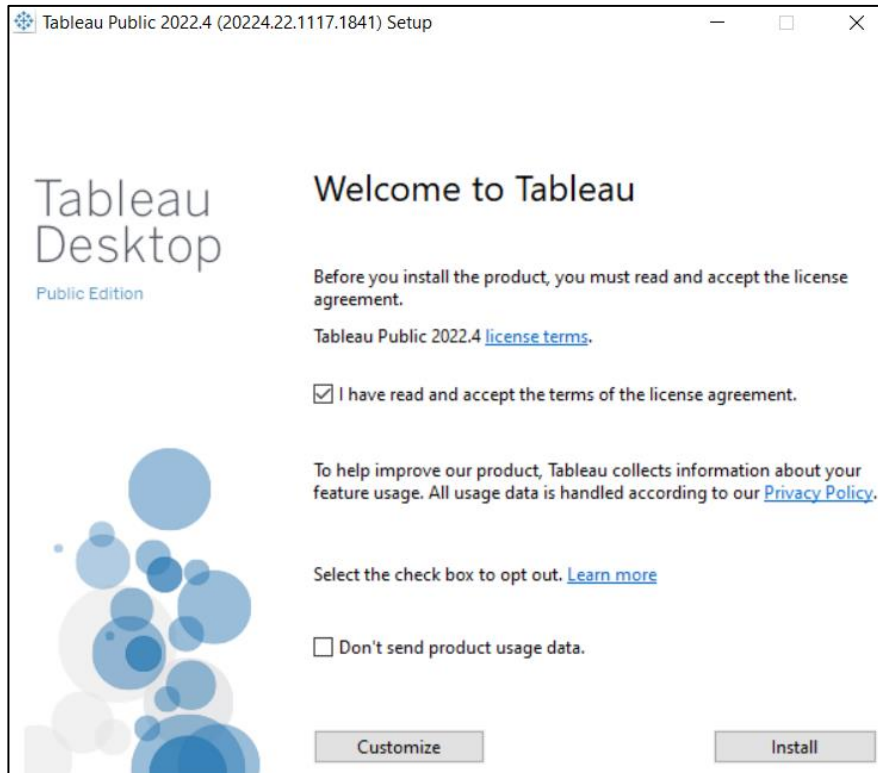
Setelah selesai

mendownload, lalu tahap selanjutnya dengan menginstall Tableau Public ke device atau perangkat komputer yang kita miliki dengan cara sebagai berikut:

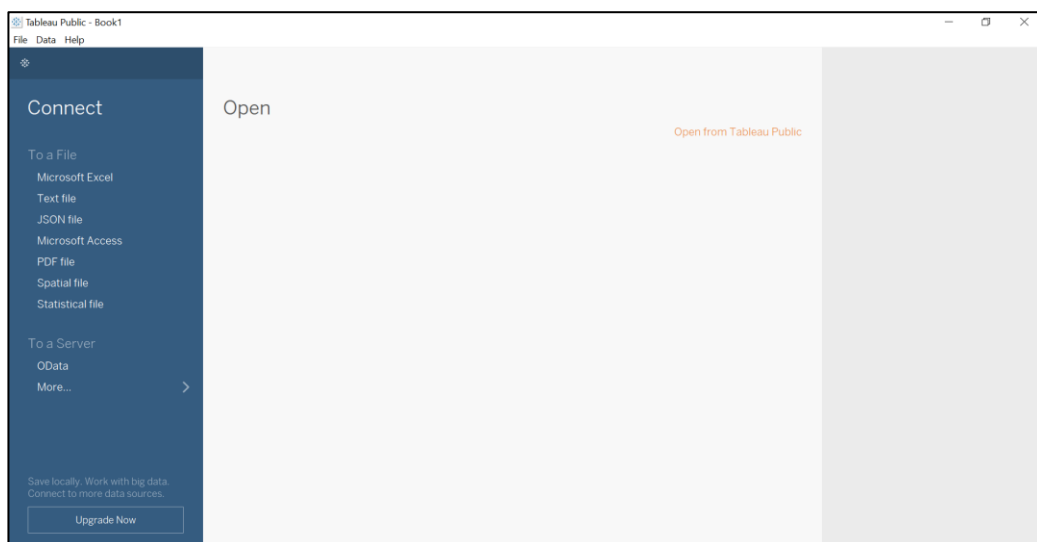
4. Pertama, klik aplikasi dengan nama Tableau Public Desktop yang telah didownload dengan cara klik lalu pilih run.



5. Selanjutnya, check list *license agreement* dan klik install pada bagian kolom kanan bawah. Maka tunggu proses instalasi hingga selesai.



6. Setelah instalasi selesai, maka Tableau Public otomatis terbuka dengan tampilan dibawah ini.



Setelah dipastikan *software* Tableau Public terinstall dengan baik, maka selanjutnya bisa memulai untuk membuat visualisasi data dengan langkah sebagai berikut:

Tahap-tahap membuat visualisasi data dengan Tableau Public.  
Visualisasi Data dengan mengkoneksikan dataset.

1. Pertama, pilih dataset yang akan dikoneksikan dan dipergunakan untuk visualisasi data, misalnya format Microsoft Excel. Untuk contoh dataset bisa diakses dan diunduh pada URL

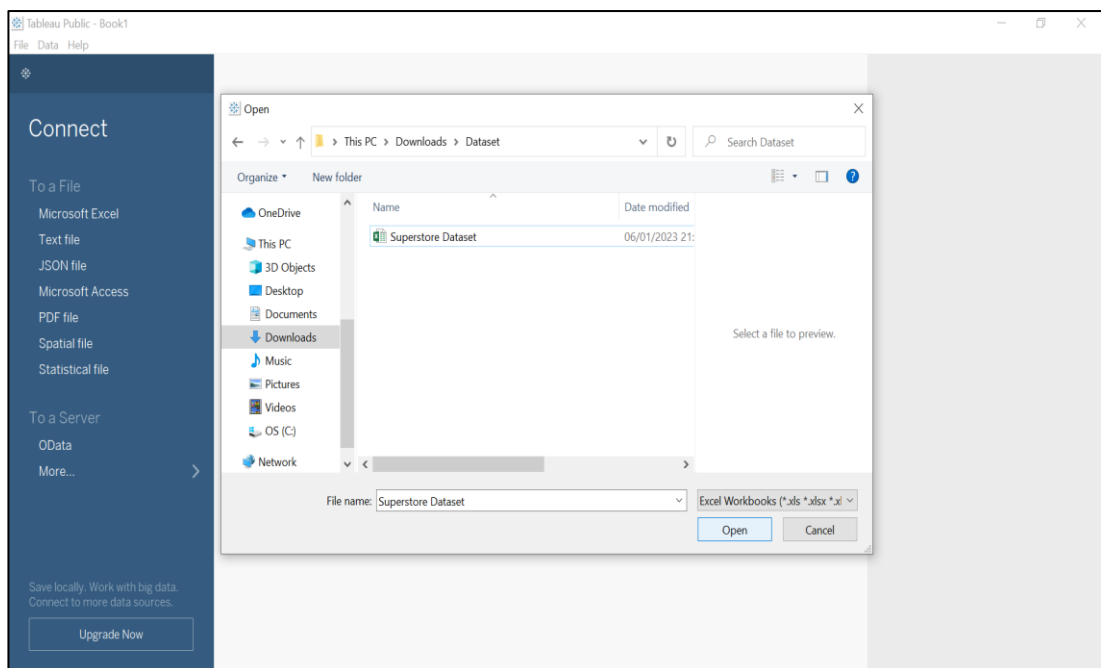
berikut <https://www.kaggle.com/datasets/addhyay/superstore-dataset>



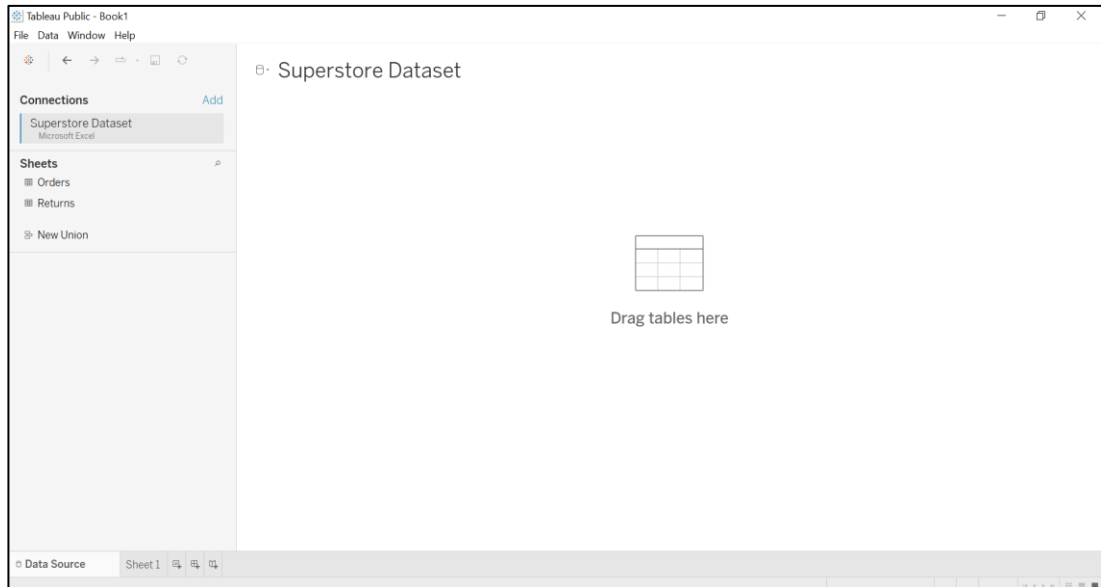
Superstore Dataset

2. Kedua, buka Tableau Public dan pilih file yang akan

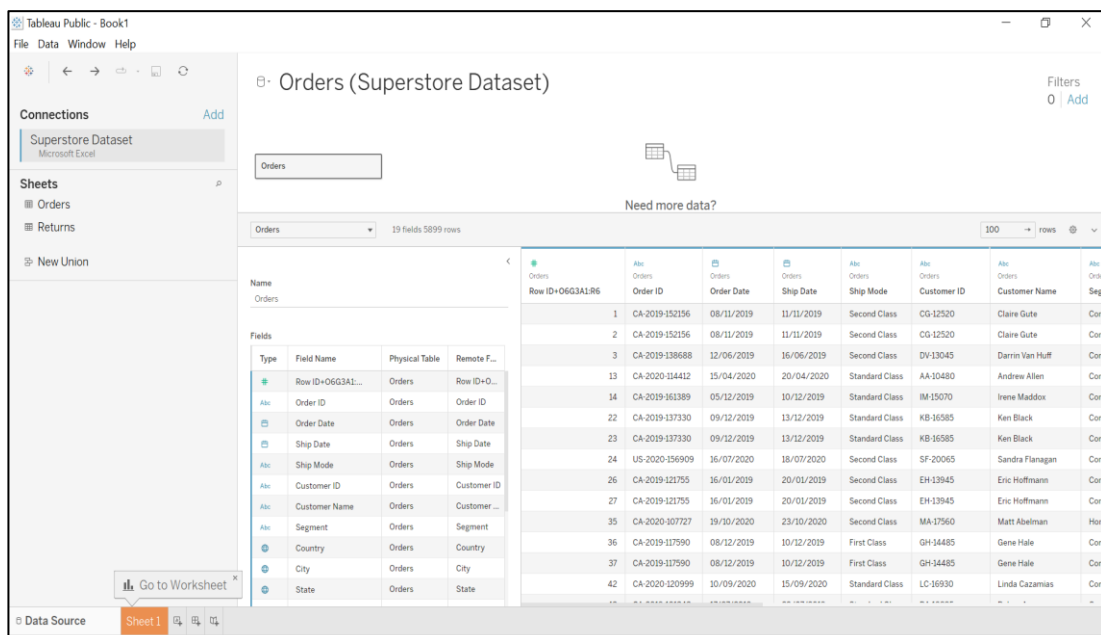
dikoneksikan lalu klik open



3. Ketiga, setelah file terbuka maka pilih sheet yang akan dipergunakan untuk visualisasi data. Pada Superstore Dataset terdapat dua sheet, yaitu sheet Orders dan sheet Returns. Saat ini yang akan dipergunakan sheet Orders, maka klik drag and drop sheet Orders ke dalam tables here pada Tableau Public.



4. Keempat, jika proses mengkoneksikan data berhasil, maka isian dataset akan tampil seperti tampilan dibawah ini.



Pembelajaran tentang instalasi Tableau Public dan mengkoneksikan Dataset telah selesai. Selanjutnya akan dibahas Analisis Deskriptif Univariat dan Analisis Deskriptif Lebih dari Satu Variabel.

## B. Analisis Deskriptif Univariat

Analisis deskriptif univariat merupakan analisis yang menggunakan satu variabel. Analisis ini dilakukan terhadap setiap variabel masing-masing dari hasil penelitian. Fungsi analisa univariat dengan merangkum bermacam data

dari output pengukuran, maka bermacam data tersebut menjadi informasi yang bermanfaat, seperti grafik, tabel dan statistik yang terukur. Tujuan analisis deskriptif univariat adalah memahami karakteristik maupun bentuk distribusi data.

Data yang digunakan dalam praktik analisis data deskriptif ini adalah data set dengan judul *superstore* yang dapat diunduh dan digunakan secara publik pada website kaggle. Adapun data tersebut dapat diunduh pada link dibawah ini:

<https://www.kaggle.com/datasets/addhyay/superstore-dataset>

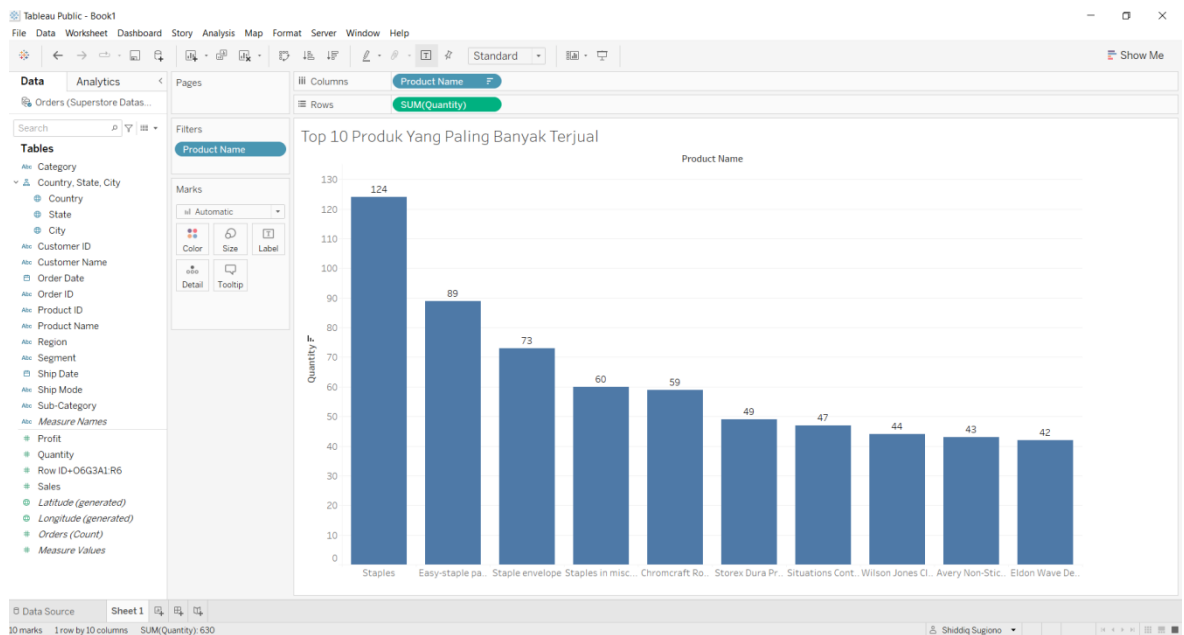
Secara umum data tersebut menyajikan hasil penjualan berbagai produk dari sebuah toko dari tahun 2019-2020. Data tersebut terdiri atas variabel-variabel sebagai berikut:

- Row ID => ID Unik pada setiap baris.
- Order ID => ID Order pada setiap customer.
- Order Date => Tanggal order pada setiap produk.
- Ship Date => Tanggal pengiriman suatu produk.
- Ship Mode=> Jenis pengiriman tertentu yang ditentukan customer.
- Customer ID => ID unik untuk mengidentifikasi customer.
- Customer Name => Nama dari seorang customer.
- Segment => segmen dimana customer berada.
- Country => Negara tempat tinggal customer.
- City => Kota tempat tinggal customer.
- State => Negara bagian customer tinggal.
- Postal Code => Kode Pos dari setiap pelanggan.
- Region => Wilayah dimana customer berada.
- Product ID => ID Unik dari setiap produk.
- Category => Kategori dari produk yang dipesan.
- Sub-Category => Sub-Kategori dari produk yang dipesan.
- Product Name => Nama dari produk.
- Sales => Harga Penjualan dari produk.
- Quantity => Kuantitas dari produk yang dipesan.
- Discount => Diskon yang diberikan.
- Profit => keuntungan/kerugian yang didapat.

Peserta pelatihan dapat memetakan jenis dari setiap variabel yang tersedia di dataset superstore dalam bentuk kategorikal atau numerik. Adapun peserta pelatihan perlu memahami definisi dari tiap variabel perlu dipahami sebelum melakukan analisis lebih lanjut.

## 1. Praktik Membuat Grafik Batang/Kolom

Jenis grafik batang/kolom dapat digunakan untuk memvisualisasikan data dari sebuah variabel. Pada bagian ini, analisis yang dilakukan adalah untuk menentukan 10 jenis produk yang paling banyak terjual antara tahun 2019-2020. Dalam hal ini kita membutuhkan total/sum dari kuantitas setiap produk yang dipesan oleh customer. Selain itu kita perlu melakukan filter untuk menampilkan 10 produk yang menempati urutan teratas. Gambar dibawah merupakan hasil proses penggunaan tableau:



Sumber: hasil analisis data superstore

**Gambar 33.** Hasil Analisis Deskriptif berupa grafik batang

Berikut langkah dari analisis diatas:

1. Memasukan variabel Product Name ke dalam Columns dan variabel Quantity ke dalam Rows.
2. Variabel Quantity diagregasi menjadi penjumlahan/sum sehingga tampilannya berubah menjadi sum(Quantity)
3. Pilih bentuk bars pada tab "Show Me"

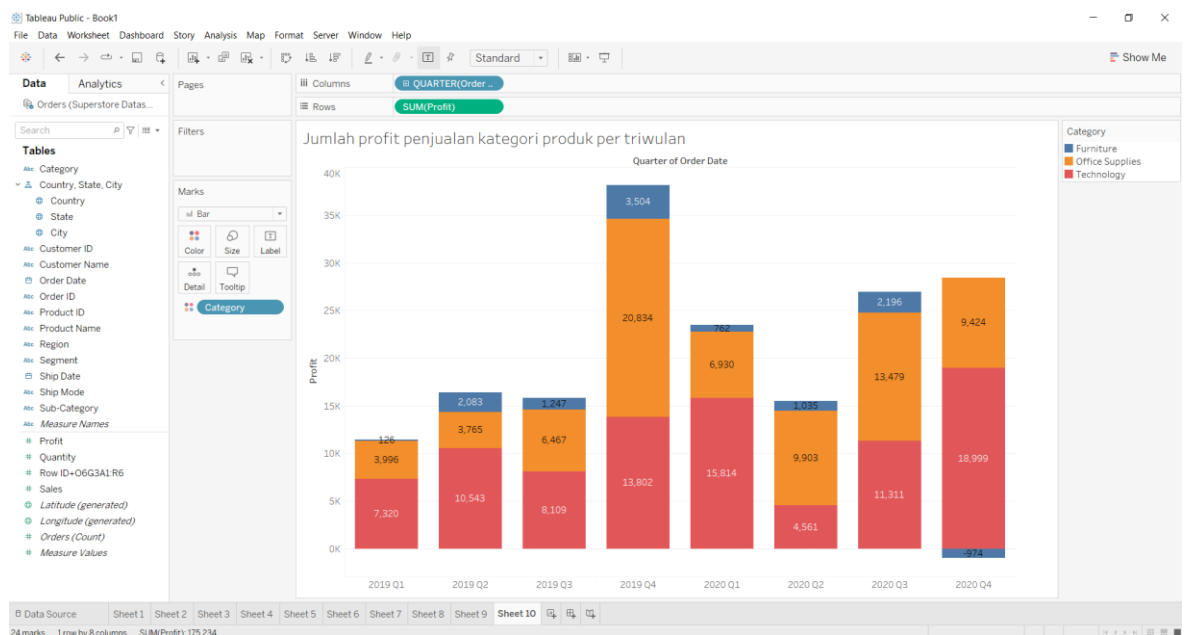
4. Menginput Product Name ke bagian filter untuk menentukan top 10 nilai terbesar.

Hasil analisis menunjukkan bahwa Staples adalah nama barang yang paling banyak dipesan oleh customer. Terlebih top 4 ditempati oleh barang jenis staples meskipun dengan nama yang berbeda.

Peserta pelatihan dapat membuat sebuah evaluasi, terutama efektivitas dari grafik yang ditampilkan.

## 2. Praktik Membuat stacked bars

Jenis grafik stacked bars digunakan untuk menganalisis lebih dalam suatu grafik batang melalui kategori yang ada di dalamnya. Analisis kali ini bertujuan untuk menampilkan profit pada tiap kategori produk per triwulan. Variabel yang akan digunakan adalah Profit, Category dan Order Date. Berikut hasil analisis menggunakan aplikasi Tableau:



Sumber: hasil analisis data superstore

**Gambar 34.** Hasil Analisis Deskriptif berupa stacked bars

Berikut langkah dari analisis diatas:

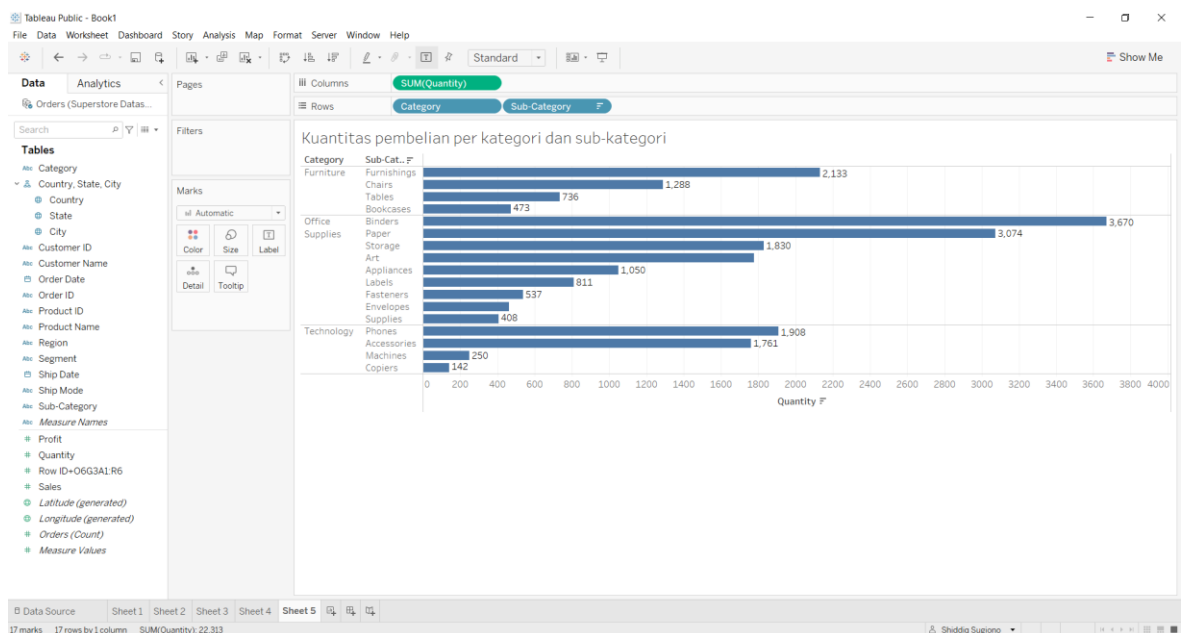
1. Masukan variabel Order Date ke Columns dan variabel profit dan category ke rows.

- Ubah satuan order date menjadi triwulan dan variabel profit menjadi agregat jumlah.
- Pada tab “Show Me” pilih grafik stacked bars.

Hasil analisis diatas menunjukkan nilai profit per triwulan yang dipecah melalui kategori. Secara visual dapat ditunjukkan bahwa kategori office supplies dan technology memberikan lebih banyak profit dibandingkan furniture. Interpretasi lainnya dapat melibatkan total profit per triwulan.

### 3. Praktik Membuat grafik batang/kolom terkelompok

Selain dapat digunakan untuk menganalisis satu variabel, grafik batang/kolom dapat digunakan untuk menganalisis lebih dalam suatu variabel yang memiliki kategori tertentu. Dalam bagian ini, analisis ditujukan untuk melihat kuantitas penjualan sub-kategori terbanyak pada masing-masing kategori. Analisis ini menggunakan 2 variabel yakni kuantitas penjualan subkategori dan kuantitas per kategori. Berikut hasil analisis menggunakan tableau:



Sumber: hasil analisis data superstore

**Gambar 35.** Hasil Analisis Deskriptif berupa grafik grouped bar/column chart

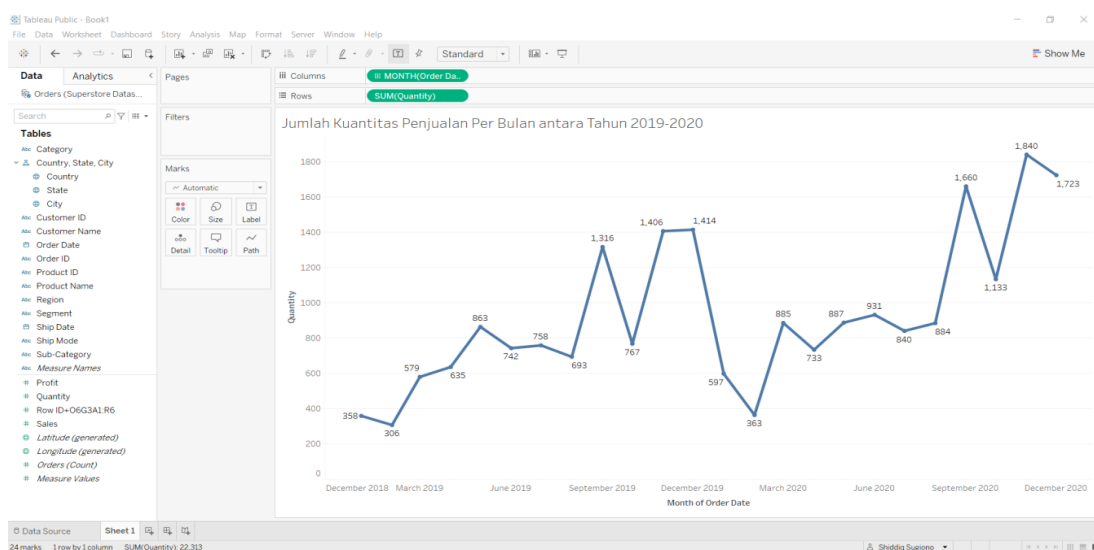
Berikut langkah dari analisis diatas:

1. Masukkan variabel category ke dalam row lalu diikuti variabel sub-category ke dalam row.
2. Masukkan variabel SUM(Quantity) ke dalam column.
3. Pilih icon sort descending untuk menyortir data dari yang paling tinggi ke yang paling rendah.

Melalui analisis ini kita dapat memiliki dua temuan. Pertama, Subkategori Furnishing, Binder dan Phones adalah subkategori produk yang paling banyak dijual dalam hal kuantitas pada setiap kateogrinya. Kedua, kita dapat melihat juga bahwa office supplies merupakan subkategori yang paling banyak dipesan. Peserta pelatihan dapat melakukan evaluasi terhadap bentuk evaluasi serta memberikan interpretasi yang lebih mendalam. Proses interpretasi dapat menggunakan hasil analisis pada grafik lainnya.

#### 4. Praktik Membuat Grafik Garis

Grafik garis merupakan jenis grafik yang digunakan untuk menjelaskan satu variabel yang bergantung terhadap waktu/*time dependent*. Pada bagian ini, analisis yang akan dilakukan bertujuan untuk melihat jumlah barang yang dipesan per bulan selama tahun 2019-2020. Dalam analisis ini kita membuat variabel penunjuk waktu serta jumlah barang yang dipesan dalam suatu bulan tertentu. Gambar dibawah merupakan proses analisis pembuatan grafik garis:



Sumber: hasil analisis data superstore

**Gambar 36.** Hasil Analisis Deskriptif berupa grafik garis

Berikut langkah dari analisis diatas:

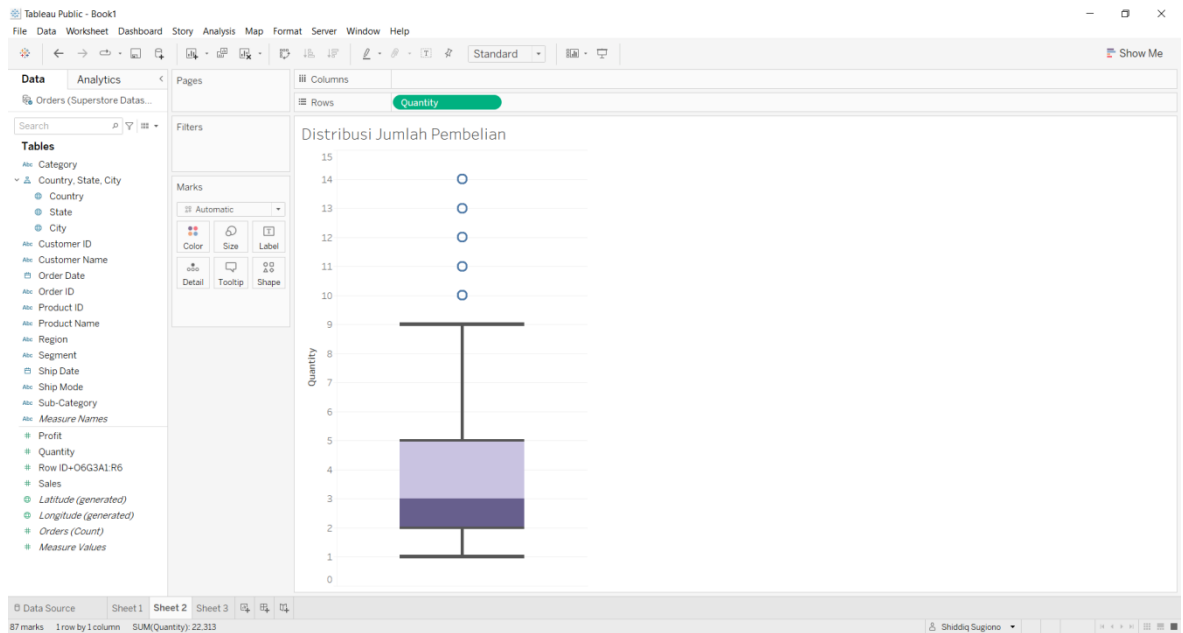
1. Memasukan variabel Order Date ke dalam kolom dan variabel Quantity ke dalam Row.
2. Mengganti jenis satuan waktu menjadi tanggal-tahun dan mengagregasi nilai kuantitas menjadi jumlah/sum.
3. Pilih bentuk lines pada tab “show me”
4. Kita dapat memunculkan label dengan memiliki kotak yang bertuliskan label pada bagian “Mark”.

Hasil analisis pada grafik batang ini dapat bervariasi tergantung dengan fokus yang ingin dilihat, kita dapat menunjukkan bahwa kuantitas penjualan tertinggi berada pada bulan November 2020 dengan total kuantitas penjualan produk adalah 1840 unit. Selain itu kita dapat melihat bahwa terjadi penurunan tajam di bulan Februari 2019 atau dalam hal ini bisa jadi disebabkan karena adanya pandemi COVID-19.

Peserta pelatihan dapat mencoba mengevaluasi serta menginterpretasi lebih dalam hasil grafik batang tersebut.

## **5. Praktik Membuat Boxplot**

Grafik boxplot digunakan untuk melihat distribusi dari suatu variabel tertentu. Pada bagian ini, analisis bertujuan untuk melihat distribusi dari kuantitas penjualan selama tahun 2019-2020. Pemilik toko ingin mengetahui bagaimana pola jumlah barang yang dipesan oleh customer. Dalam analisis ini kita membutuhkan variabel Quantity untuk menggambarkan distribusi kuantitas barang pada setiap pembelian. Berikut adalah hasil proses analisis menggunakan tableau:



Sumber: hasil analisis data superstore

**Gambar 37 .** Hasil Analisis Deskriptif berupa grafik boxplot

Berikut langkah dari analisis diatas:

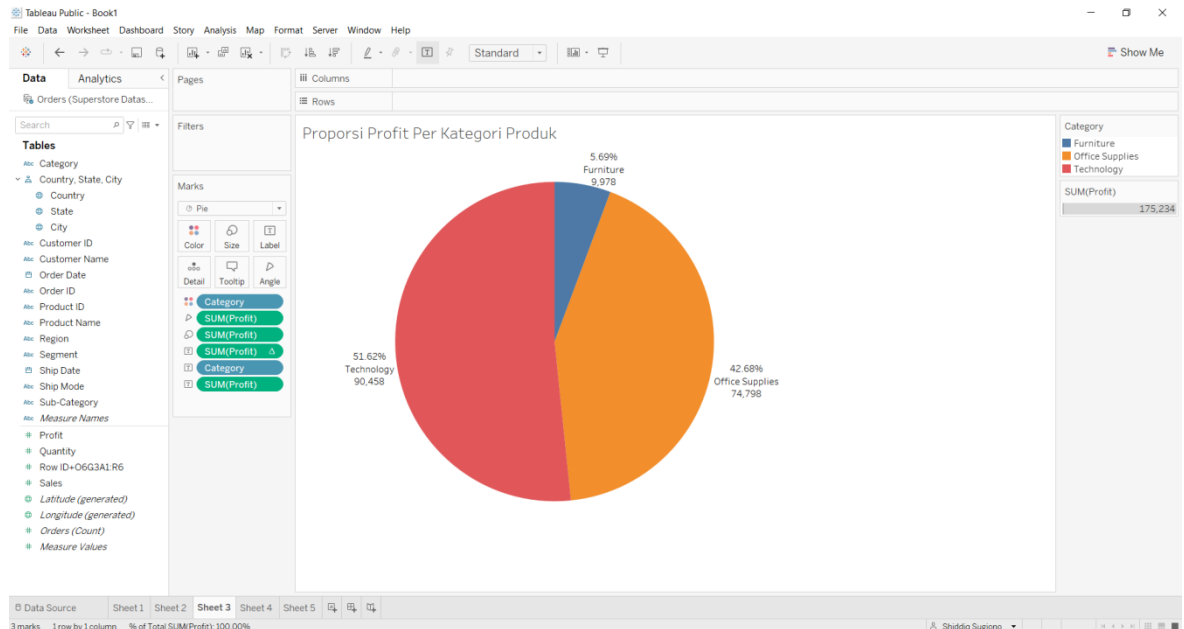
1. Memasukan variabel quantity kedalam Row, dalam kasus ini kita menonaktifkan pengukuran agregat dengan cara pilih menu: Analysis > (uncheck) Aggregate Measure.
2. Klik kanan sumbu Y (Quantity) > Add Reference Line > Boxplot.
3. Lakukan setting dalam visualisasi boxplot lalu tekan OK.

Hasil analisis tersebut menunjukkan bahwa nilai median dari jumlah pembelian customer berada di angka 3. Pembelian paling banyak adalah 14 unit dan paling rendah adalah 1 unit. Pembelian dengan jumlah 10, 11, 12, 13, dan 14 unit adalah jenis pembelian outlier dimana sangat sedikit orang yang membeli dengan jumlah tersebut. Adapun 75% jumlah pembelian terpusat pada 2-5 unit. Peserta pelatihan dapat memberikan interpretasi yang berbeda dari hasil analisis tersebut.

## 6. Praktik Membuat Pie Chart

Pie chart merupakan salah satu bentuk visualisasi data yang menunjukkan persentase dari sebuah kategori dalam suatu distribusi. Pada bagian ini,

analisis akan ditujukan untuk melihat kategori produk apa yang memberikan profit paling banyak pada toko. Pada analisis ini kita membutuhkan data terkait variabel produk dan jumlah profit yang dihasilkan tiap produknya. Berikut hasil analisis menggunakan tableau:



Sumber: hasil analisis data superstore

**Gambar 38.** Hasil Analisis Deskriptif berupa grafik boxplot

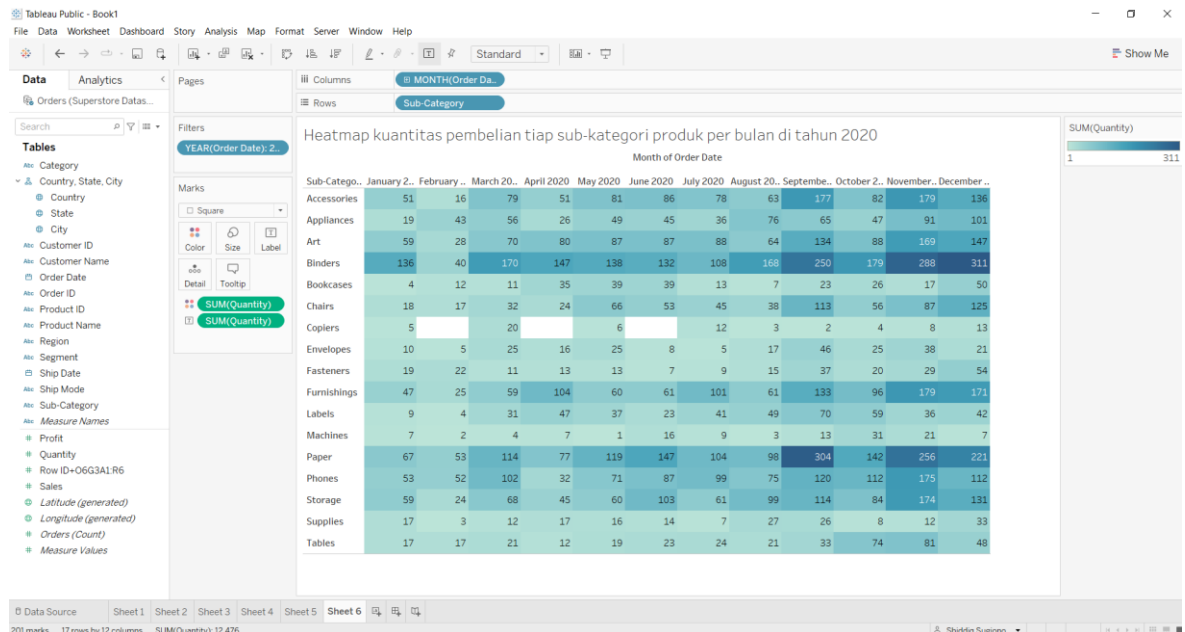
Berikut langkah dari analisis diatas:

1. Memasukan Sum(Profit) ke dalam row dan memasukan Category ke dalam columns.
2. Dalam Tab Show Me pilih bentuk Pie
3. Drag and drop category dan quantity ke dalam kotak "Label"
4. Untuk memunculkan persentase, pilih Sum(Profit) pada tab Marks lalu pilih opsi "Quick Table Calculation" > Percent of Total.

Hasil analisis menunjukkan bahwa produk dengan kategori "Technology" memiliki persentase profit/keuntungan terbesar yakni 90.458 USD. Disusul dengan kategori office supplies dan furniture. Peserta pelatihan dapat memberikan evaluasi maupun interpretasi mengenai hal apa saja yang dapat dilakukan selanjutnya setelah mengentahui temuan ini.

## 7. Praktik Membuat Heatmap

Diagram Heatmap memiliki manfaat untuk memberikan gambaran mengenai derajat tinggi atau rendahnya suatu nilai dalam bentuk tabel. Analisis pada bagian ini bertujuan ingin melihat sub-kategori apa yang memiliki nilai kuantitas tertinggi pada bulan tertentu di tahun 2020. Pada analisis dibutuhkan variabel sub-category, order date serta quantity. Berikut hasil analisisnya:



Sumber: hasil analisis data superstore

**Gambar 39.** Hasil Analisis Deskriptif berupa diagram heatmap.

Berikut langkah dari analisis diatas:

1. Masukkan variabel sub-category ke rows dan order date ke columns.
2. Pilih opsi MONTH pada variabel order date sehingga menampilkan bulan saja.
3. Drag and drop order date ke dalam filters untuk memfilter sehingga yang ditampilkan hanya tahun 2020 saja.
4. Masukkan SUM(Quantity) ke tab colors dan tab labels.
5. Dalam tab "Show Me" pilih highlight tables

Hasil analisis menunjukkan bahwa kuantitas penjualan cenderung meningkat pada akhir tahun. Adapun produk yang kuantitas pembeliannya tinggi adalah binders dan papers. Melalui temuan ini maka toko tersebut perlu memiliki stok yang banyak pada barang-barang tersebut khususnya di akhir

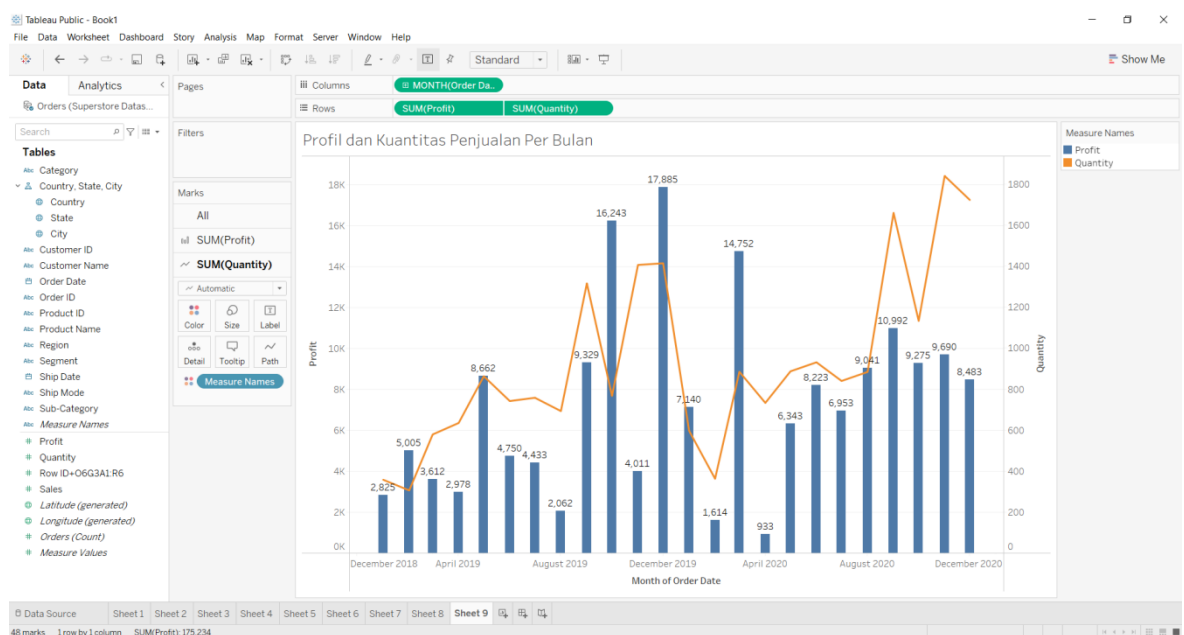
tahun sehingga tetap bisa melayani pembeli. Peserta pelatihan dapat melakukan interpretasi lebih dalam mengenai barang-barang yang memiliki kuantitas penjualan rendah.

### C. Analisis Deskriptif Lebih dari Satu Variabel

Analisis deskriptif lebih dari dua variabel adalah analisis yang memperlihatkan hubungan lebih dari satu variabel secara bersama dengan mengontrol variabel lain. Jenis analisis deskriptif ini memberikan temuan lebih mendalam jika dibandingkan dengan analisis deskriptif univariat. Dalam bagian ini kita akan melakukan praktik visualisasi data dengan menggunakan dataset yang sama yakni *superstore*.

#### 1. Praktik membuat Diagram Dual Combination

Dual combination merupakan salah satu bentuk visualisasi yang menggabungkan 2 jenis grafik. Salah satu jenis dual combination yang dapat digunakan adalah gabungan antara grafik garis dengan grafik batang. Pada bagian ini, tujuan analisis adalah untuk melihat pergerakan profit dan kuantitas penjualan perbulan secara bersamaan. Dalam analisis ini dibutuhkan variabel quantity dan profit. Berikut hasil analisis menggunakan aplikasi tableau:



Sumber: hasil analisis data superstore

**Gambar 40.** Hasil Analisis Deskriptif berupa grafik dual combination

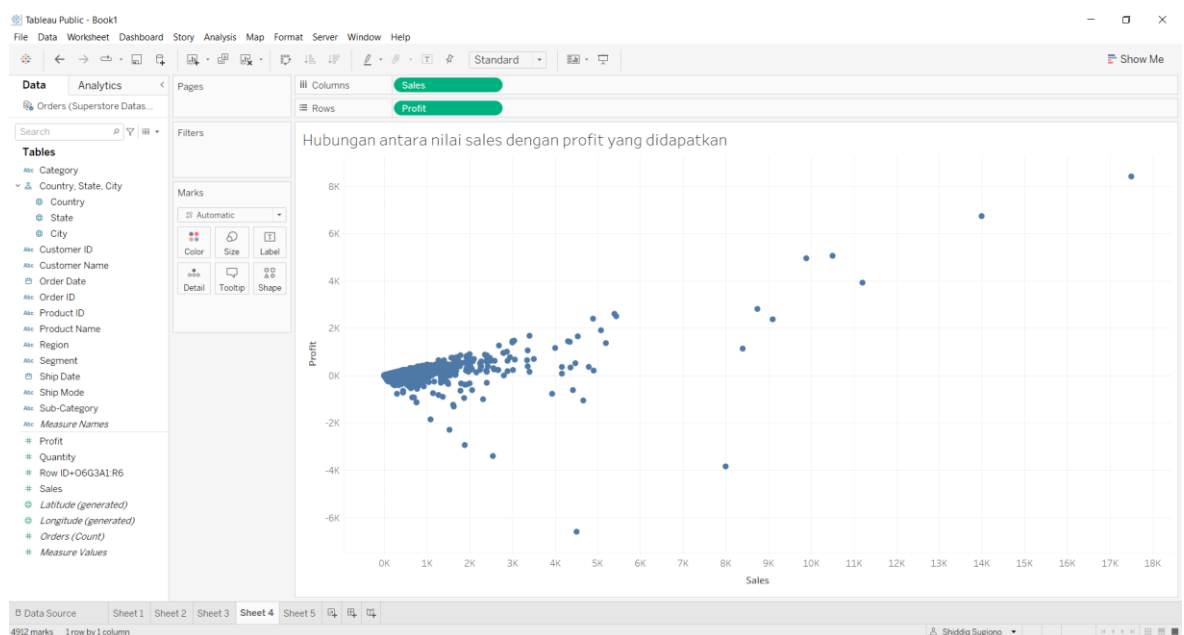
Berikut langkah dari analisis diatas:

1. Masukkan variabel Order Date ke columns dan variabel profit serta quantity ke rows.
2. Ubah jenis tanggal menjadi satuan bulan, variabel profit dan quantity menjadi agregat sum.
3. Pada tab “Show Me” pilih jenis diagram dual combination.

Grafik tersebut memiliki 2 jenis pengukuran, sumbu y sebelah kiri adalah profit dengan satuan USD dan sumbu Y sebelah kanan adalah quantity dengan satuan unit. Melalui grafik ini kita dapat bagaimana perbandingan antara jumlah unit yang terjual dengan profit yang didapatkan. Peserta pelatihan diharapkan dapat menyampaikan beberapa hasil analisis yang diperoleh dari grafik diatas.

## 2. Praktik membuat Diagram Pencar/Scatter Plot

Diagram pencar/scatter plot merupakan salah satu bentuk visualisasi yang bertujuan untuk memperlihatkan bentuk keterhubungan antara dua variabel. Dalam bagian ini, tujuan dari analisis adalah melihat hubungan antara variabel nilai penjualan (sales) dengan profit yang didapatkan. Terdapat 2 variabel yang digunakan dalam analisis ini yakni nilai sales dan nilai profit. Berikut hasil dari analisis menggunakan tableau:



Sumber: hasil analisis data superstore

**Gambar 41** Hasil Analisis Deskriptif berupa diagram pencar/scatter plot

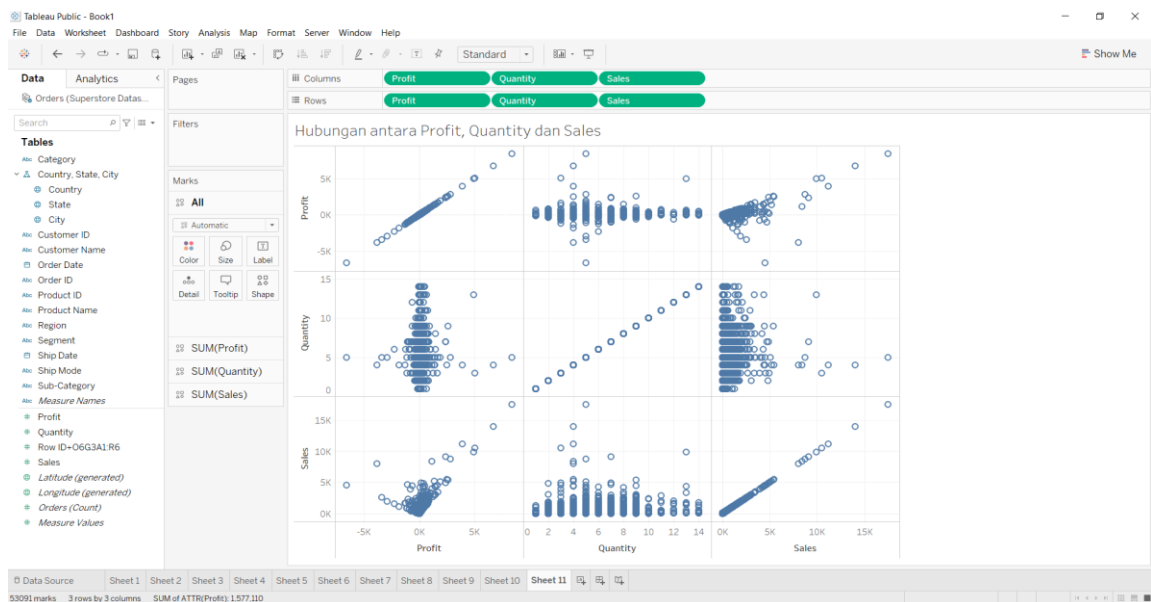
Berikut langkah dari analisis diatas:

1. Masukkan variabel profit ke rows dan sales ke columns. Pilih bentuk dimension untuk kedua variabel tersebut.
2. Pilih bentuk scatter pada tab “Show Me”

Hasil analisis diatas menunjukkan bahwa ada kecenderungan dimana jika nilai sales tinggi maka akan meningkatkan profit. Namun begitu ada beberapa kasus yang menunjukkan bahwa semakin tinggi sales maka akan semakin tinggi pula kerugiannya. Meskipun begitu, semakin tinggi nilai sales maka semakin rendah kemungkinan untuk rugi. Peserta pelatihan diharapkan bisa melakukan evaluasi maupun interpretasi terhadap hasil analisis.

### 3. Praktik Membuat Matrix Scatter Plot

Matrix Scatter Plot pada dasarnya memiliki bentuk dan tujuan yang sama seperti scatter plot tetapi dalam bentuk matriks ini terdapat beberapa variabel yang dapat di analisis secara sekaligus. Analisis kali ini bertujuan untuk melihat hubungan Profit, Quantity serta Sales secara sekaligus. Berikut hasil analisis menggunakan tableau:



Sumber: hasil analisis data superstore

**Gambar 42.** Hasil Analisis Deskriptif berupa matrix scatter plot

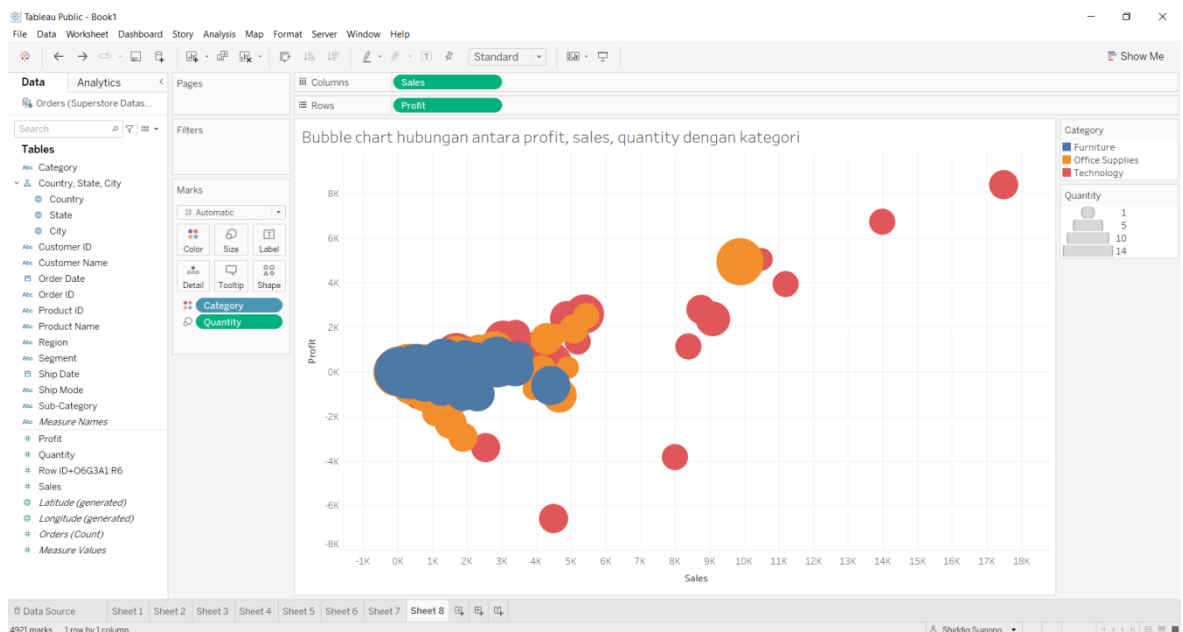
Berikut langkah dari analisis diatas:

1. Masukan variabel Profit, Quantity, dan Sales ke dalam row dan columns.
2. Buat semua variabel berbentuk dimension/bukan agregat dengan cara pilih menu Analysis > Aggregate Measures

Melalui gambar diatas, maka dapat dilihat bagaimana bentuk hubungan antar variabel. Pada garis diagonal adalah hubungan antara masing-masing variabel, sehingga hanya berbentuk garis lurus. Peserta pelatihan diharapkan dapat memberikan analisis pada masing-masing bentuk hubungan diatas.

#### 4. Praktik Membuat Bubble Chart

Bubble chart merupakan salah satu bentuk visualisasi yang memiliki kemampuan dalam menampilkan empat variabel. Sama seperti diagram pencar/scatter plot, jenis grafik ini memiliki tujuan untuk melihat hubungan antar variabel. Dalam analisis kali ini, kita akan melihat bagaimana hubungan antara kuantitas harga sales, profit, kategori produk dan kuantitas penjualan. Berikut hasil analisis dengan menggunakan



Sumber: hasil analisis data superstore

**Gambar 43 .** Hasil Analisis Deskriptif berupa bubble chart

Berikut langkah dari analisis diatas:

1. Masukkan variabel profit ke row dan variabel sales ke columns, pastikan

semua data berjenis dimension/bukan hasil agregat.

2. Drag and drop variabel category ke colors dan quantity ke size.
3. Jika bentuk lingkaran terlihat kecil maka sesuaikan ukurannya di menu size.

Hasil analisis tersebut menghasilkan berbagai macam temuan dua diantaranya adalah sebagai berikut: Pertama, kategori teknologi cenderung memiliki harga sales yang paling bervariasi dibanding kategori lainya dan cenderung memiliki profit yang lebih tinggi dibandingkan dengan kategori produk lainnya. Kedua, Furnitur dan ATK cenderung memiliki harga sales rendah namun memiliki kuantitas penjualan yang tinggi. Peserta penelitian diminta untuk menemukan temuan-temuan lainnya dari hasil analisis tersebut.

#### **D. Rangkuman**

Tableau merupakan salah satu aplikasi yang dapat digunakan untuk menganalisis data secara deskriptif, baik satu variabel maupun lebih dari satu variabel. Aplikasi tersebut turut mendukung berbagai macam bentuk visualisasi sehingga dapat memberikan berbagai kemungkinan untuk memberikan berbagai macam temuan dari suatu data. Dalam menggunakan aplikasi tableau, kita harus mampu memahami bagaimana bentuk agregasi data (sum/count) sehingga dapat secara tepat menampilkan hasil analisis. Pemilihan visualisasi pada analisis deskriptif harus dilakukan berdasarkan tujuan yang ingin dicapai serta jumlah variabel yang akan dianalisis sehingga hasil yang disampaikan dapat menampilkan informasi secara efektif.

#### **E. Evaluasi**

1. Pilih satu dataset yang bersifat open public, lalu lakukan satu jenis analisis deskriptif untuk satu variabel dan satu jenis analisis deskriptif lebih dari satu variabel. Gunakan jenis visualisasi dari salah satu contoh yang telah disampaikan dalam Bab 5.
2. Berikan interpretasi yang mendalam terhadap hasil analisis yang telah dilakukan pada nomor 1. Jika membutuhkan analisis lainnya silahkan tampilkan visualisasi tambahan.

## DAFTAR PUSTAKA

Agresti, A., Franklin, C., Klingenberg, B. (2017). *Statistics: The Art and Science of Learning from Data, Fourth Edition*. Pearson Education. ISBN 10: 1-292-16477-8

Bai, S., Zhou, X., Lyu, Y., Wang, J., & Pan, C. (2019). Data Visualization Model Methods and Techniques. *IOP Conference Series: Earth and Environmental Science*, 252(5). <https://doi.org/10.1088/1755-1315/252/5/052063>

Banerjee, A., Bandyopadhyay, T., & Acharya, P. (2013). Data Analytics: Hyped Up Aspirations or True Potential? *Vikalpa*, 38(4), 1–12. <https://doi.org/10.1177/0256090920130401>

Bilgin, Y. (2017). Qualitative Method Versus Quantitative Method in Marketing Research: An Application Example at Oba Restaurant. In *Qualitative versus Quantitative Research* (Issue November). <https://doi.org/10.5772/67848>

Creswell, J. W., & Creswell, J. D. (2018). *Research and Design Qualitative, Quantitative and Mixed Methods Approaches* (Fifth Edit). SAGE Publications Inc.

Dates, G. (2021). Step 1: Assemble the Information You ' ll Need. *Massachusetts Water Watch Partnership*, 1(2), 1–8. <https://www.umass.edu/mwwp/pdf/intmanl.pdf>

Delen, D., & Ram, S. (2018). Research challenges and opportunities in business analytics. *Journal of Business Analytics*, 1(1), 2–12. <https://doi.org/10.1080/2573234X.2018.1507324>

Hennink, M., Hutter, I., & Bailey, A. (2020). *Qualitative Research Methods*. SAGE.

Humasofdata. (2018). *Your Guide to Qualitative and Quantitative Data Analysis Methods*. <https://humansofdata.atlan.com/2018/09/qualitative-quantitative->

data-analysis-methods/

<https://medium.com/@rndayala/data-levels-of-measurement-4af33d9ab51a>

indonesiabaik.id. (2021). *Waspada Demam Berdarah Dengue*.  
<https://indonesiabaik.id/infografis/waspada-demam-berdarah-dengue>

Islam, M., Hossain, A., Hasan, R., & Soton, A. R. (2022). *The Methods, Benetifs and Problems of The Interpretation of Data*.

Kusnandar, V. B. (2019). *Nilai Ujian Nasional SMA Jurusan IPS 24 Provinsi di Bawah Rerata Nasional*.  
<https://databoks.katadata.co.id/datapublish/2019/06/29/nilai-ujian-nasional-sma-jurusan-ips-24-provinsi-di-bawah-rerata-nasional-20182019>

Li, R., Tao, Q., Luo, Y., & Yan, L. (2020). Research on practical application of data analysis and visualization. *Proceedings - 2020 International Conference on Virtual Reality and Intelligent Systems, ICVRIS 2020*, 78–81.  
<https://doi.org/10.1109/ICVRIS51417.2020.00026>

Ott, R. L., Longnecker, M. (2010) *An Introduction to Statistical Methods and Data Analysis*, Sixth Edition. Books/Cole Cengage Learning.

Richmond, B. (2006). Introduction to Data Analysis Handbook. *Academy for Educational Development*, 1–27. <http://eric.ed.gov/?id=ED536788>

Rowley, J. (2007). The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of Information Science*, 33(2), 163–180.  
<https://doi.org/10.1177/0165551506070706>

Sharma, B. (2018). Processing of data and analysis. *Biostatistics and Epidemiology International Journal*, 1(1), 3–5.  
<https://doi.org/10.30881/beij.00003>

Sheng, J., Amankwah-Amoah, J., Khan, Z., & Wang, X. (2021). COVID-19 Pandemic in the New Era of Big Data Analytics: Methodological Innovations and Future Research Directions. *British Journal of Management*, 32(4), 1164–1183.

<https://doi.org/10.1111/1467-8551.12441>

Sirait, E. R. E. (2016). Implementasi Teknologi Big Data Di Lembaga Pemerintahan Indonesia. *Jurnal Penelitian Pos Dan Informatika*, 6(2), 113. <https://doi.org/10.17933/jppi.2016.060201>

Terrel, S. R. (2021). *Statistics Translated: A Step-by-Step Guide to Analyzing and Interpreting Data*. The Guilford Press. <https://www.ptonline.com/articles/how-to-get-better-mfi-results>

Vali, S. M., & Rajasekaran, S. (2020). Effective use of Data Analysis and Data Visualization in E-Learning. *Solid State Technology*, Approximately 15000. [https://www.scihorizon.com/cdn/pdf/1627547144\\_67b3bcf5e42c88388ff6.pdf](https://www.scihorizon.com/cdn/pdf/1627547144_67b3bcf5e42c88388ff6.pdf)

Woodsong, C. (2005). *Qualitative* (Issue January).