**Pendahuluan**

* Visualisasi data merupakan cara mengomunikasikan sebuah informasi atau data dalam bentuk visual seperti diagram, grafik, atau representasi visual lainnya.
* Pentingnya penggunaan visualisasi data:
  + Komunikasi lebih efektif
  + Memantau data dengan lebih mudah
* Macam-macam media visualisasi data
* Tabel: Penyusunan data ke dalam baris dan kolom.
  + Diagram batang: Visualisasi grafik dalam bentuk batang.
  + Diagram garis: Visualisasi grafik dalam bentuk garis. Biasanya menyajikan perubahan data dalam periode waktu tertentu. Secara umum, diagram garis digunakan untuk melihat perkembangan data tertentu yang berlangsung secara terus menerus atau berkelanjutan.
  + Diagram lingkaran: Diagram lingkaran mirip seperti sebuah makanan pizza yang diiris dengan porsi tertentu. Pada konteks data, irisan pizza tersebut menggambarkan persentase data nilai atau kuantitas. Apabila irisan tersebut dijumlahkan nilainya, maka seharusnya menghasilkan 100 persen atau 360 derajat.
* Visualisasi data yang dapat diterapkan di bidang bisnis
  + Scorecard: Scorecard merupakan sebuah visualisasi yang fokus pada sebuah jenis data yang spesifik. Sehingga biasanya terdiri dari 1 bentuk visualisasi saja.
  + Dashboard: Kumpulan dari berbagai macam visualisasi yang menggabungkan dan merangkum informasi atau data bisnis.
  + Report: Bentuk penyampaian berita, keterangan, pemberitahuan, ataupun pertanggungjawaban sebagai bentuk pelaksanaan komunikasi dari pihak yang satu kepada pihak yang lainnya.
  + Analytic Report: Laporan yang berfokus pada analisis yang digunakan untuk menentukan keputusan.
* Tools yang dapat digunakan untuk membuat visualisasi data.
  + Tableau Public: Layanan gratis yang memungkinkan siapa saja dapat mempublikasikan visualisasi data ke dalam web. Aplikasi perlu didownload untuk menggunakannya.
  + Google Sheets: Layanan gratis yang semuanya tersedia secara online sehingga tidak perlu instalasi aplikasi spreadsheet ke komputer atau laptop.
  + Microsoft Excel: Aplikasi spreadsheet berbayar buatan Microsoft yang memerlukan instalasi ke komputer atau laptop.
* Konteks (makna inti) data dibutuhkan agar data lebih mudah dipahami. Apabila hanya memiliki data tanpa memahami konteks, maka data tersebut tidak bisa tersampaikan dengan baik.
* Penting untuk mengetahui audiens (siapa yang nanti membaca data atau visualisasi) untuk menentukan bagaimana cara menyampaikan data.
  + Kepada siapa data ini akan disampaikan: Anda harus bisa membedakan penyampaian data antara audiens umum dan yang sudah profesional (sudah ahli di bidangnya).
  + Data apa yang ingin disampaikan: Supaya data relevan dengan audiens, Anda harus menyaring terlebih dahulu data yang dimiliki.
  + Bagaimana cara menyampaikan data tersebut: Metode penyampaian perlu disesuaikan dengan audiens. Sebagai contoh jika sebagian besar audiens tidak terbiasa dengan data dalam bentuk angka maka kita dapat menggunakan metode seolah sedang bercerita (tidak hanya sebut angka saja).

**Persiapan Data Pra-Visualisasi**

* Spreadsheet merupakan sebuah lembaran berisi baris dan kolom yang dapat kita gunakan untuk analisis data.
* Kegunaan spreadsheet:
  + Menyimpan informasi secara detail.
  + Membuat tabel data dengan lebih mudah.
  + Membantu dalam perhitungan statistik.
  + Memvisualisasikan data ke dalam bentuk grafik.
  + Lengkap dengan rumus-rumus yang membantu dalam perhitungan data.
* Macam-macam aplikasi spreadsheet
  + Microsoft Excel
  + LibreOffice Calc
  + Google Sheets
* Spreadsheet pada dasarnya baris dan kolom. Baris dan kolom merupakan kumpulan dari sel. Data yang dapat dimasukkan dalam sel dapat berupa:
  + Angka : Anda dapat memasukkan angka yang nantinya dapat dihitung secara matematis.
  + Teks : Anda dapat memasukkan data atau informasi dalam bentuk teks seperti “Jumlah barang” atau “Nama lengkap”.
  + Rumus : Digunakan untuk melakukan perhitungan atau kalkulasi dari sebuah atau sekelompok sel. Misal, kita dapat menggunakan rumus untuk menjumlahkan sebuah deretan sel yang berisi data dalam bentuk angka.
* Elemen pada menu bar Google Sheets
  + Title bar: Bagian yang menunjukkan judul berkas.
  + Menu bar: Terletak di bawah judul berkas. Tersedia berbagai menu yang disediakan oleh Google Sheets mulai dari File hingga Help.
* Opsi ketika melakukan import file pada Google Sheets
  + Create new spreadsheet : Menambahkan berkas pada spreadsheet baru.
  + Insert new sheet(s) : Menambahkan berkas menjadi sheet baru dalam spreadsheet yang sedang aktif.
  + Replace spreadsheet : Menggantikan spreadsheet yang sedang aktif dengan berkas yang ditambahkan.
* Opsi pengolahan data pada berkas spreadsheet aplikasi Google Sheets:
  + Sortir data: Fitur untuk mengurutkan data supaya lebih tertata rapi.
  + Penyaringan data: Cara agar hanya menampilkan data tertentu berdasarkan kondisi yang kita tentukan. Pada saat penyaringan data dilakukan, maka data yang tidak memenuhi kriteria kondisi untuk dimunculkan akan secara otomatis disembunyikan. Namun data yang disembunyikan tersebut tidak terhapus dari Google Sheets.
  + Formula: Langkah atau rumus untuk melakukan perhitungan matematis sehingga mendapatkan nilai tertentu. Sebuah formula pasti merujuk pada sebuah alamat sel tertentu yang dapat disebut sebagai referensi sel.
  + Pivot table: Fitur untuk mengumpulkan dan mengelompokkan data berdasarkan parameter tertentu sesuai kebutuhan, tanpa perlu menggunakan penulisan rumus. Hasil akan disajikan dalam bentuk tabel.
* Operator adalah simbol atau tanda yang memiliki fungsi tertentu dan digunakan untuk melakukan pemrosesan atau perhitungan data.
  + Operator aritmatika: Operator aritmatika atau matematika digunakan pada aplikasi spreadsheet untuk melakukan operasi matematika dasar, seperti penambahan, pengurangan, perkalian, atau pembagian. Selain itu dapat digunakan untuk menghitung nilai pangkat atau persen.
  + Operator perbandingan: Dalam spreadsheet biasanya kita menggunakan jenis operator perbandingan untuk fungsi-fungsi logika seperti IF, OR, AND, dan NOT. Hasilnya adalah kondisi nilai True atau False.
  + Operator text: Contoh adalah ampersand (“&”) yang berfungsi untuk menggabungkan beberapa string text menjadi satu string tunggal
  + Operator referensi: Operator untuk menunjukkan lokasi sel yang digunakan dalam penerapan sebuah rumus atau range data.
* Elemen formula dalam spreadsheet aplikasi Google Sheets
  + Sama dengan (“=”): Tanda sama dengan “=” merupakan elemen yang paling awal ditulis dalam sebuah formula apa pun sebelum lanjut ke elemen lainnya.
  + Fungsi: Fungsi merupakan sebuah penamaan yang telah ditentukan untuk melakukan kalkulasi data berdasarkan susunan argumen dalam aplikasi spreadsheet.
  + Referensi Sel: Acuan sebuah sel atau range dalam aplikasi spreadsheet, mulai dari sheet yang sama hingga berbeda berkas atau workbook.
  + Operator: Simbol atau tanda yang memiliki fungsi tertentu dan digunakan untuk melakukan pemrosesan atau perhitungan data.
  + Konstanta: Nilai masukan yang bukan berasal dari perhitungan karena nilainya selalu sama dan tidak pernah berubah. Bentuk dari konstanta bisa berupa teks atau angka.
* Beberapa fungsi pada spreadsheet aplikasi Google Sheets
  + Sum: Mendapatkan nilai total dari rentang sel yang dipilih. Cara penulisan: =SUM(data ke-1, data ke-2, … , data ke-n)
  + SumIF: Menjumlahkan data yang sesuai kriteria tertentu. Hanya terbatas untuk satu kriteria. Cara penulisan: =SUMIF(range, ”kriteria”,sum\_range)
  + SumIFS: Menjumlahkan data yang sesuai kriteria tertentu. Menggunakan lebih dari satu kriteria. Cara penulisan: =SUMIFS(sum\_range, kriteria\_range1, “kriteria1”, kriteria\_range2, “kriteria2”, dan seterusnya)
  + Average: Rata-rata aritmatika yang menjumlahkan semua data kemudian dibagi dengan jumlah data yang ada. Cara penulisan: =AVERAGE(data ke-1, data ke-2, … , data ke-n)
  + Count: Menghitung banyaknya sel terpilih dalam rentang tertentu yang berisi nilai numerik. Cara penulisan: =COUNT(data ke-1, data ke-2, … , data ke-n)
  + CountA: Menghitung banyaknya sel terpilih dalam rentang tertentu, tak peduli apa pun jenis data yang ada di dalamnya (angka, teks, tanggal, kondisi benar atau salah, hingga kesalahan perhitungan), namun tidak menghitung sel yang kosong. Cara penulisan: =COUNTA(data ke-1, data ke-2, … , data ke-n)
  + CountIF: Menghitung banyaknya sel terpilih dalam rentang tertentu, yang sesuai kriteria tertentu. Hanya terbatas untuk satu kriteria. Cara penulisan: =COUNTIF(range, “kriteria”)
  + CountIFS: Menghitung banyaknya sel terpilih dalam rentang tertentu, yang sesuai kriteria tertentu. Menggunakan lebih dari satu kriteria. Cara penulisan: =COUNTIFS(kriteria\_range1, “kriteria1”, kriteria\_range2, “kriteria2”, dan seterusnya)
  + Min: Menentukan nilai terendah dalam rentang tertentu. Contoh penulisan: =MIN(data ke-1, data ke-2, … , data ke-n)
  + Max: Menentukan nilai tertinggi dalam rentang tertentu. Contoh penulisan: =MAX(data ke-1, data ke-2, … , data ke-n)
  + Trim: Menghilangkan ruang kosong yang tidak dibutuhkan pada sebuah teks. Fungsi ini hanya berjalan pada sel tunggal bukan pada rentang sel. Cara penulisan: =TRIM(teks)
  + Replace: Mengganti string baik berupa angka maupun teks biasa. Cara penulisan: =REPLACE(text, position, length, new\_text)
  + Unique: Mencari data yang unik dari sebuah rentang data yang mungkin saja ada yang sama. Cara penulisan: =UNIQUE(range)
  + If: Membuat perbandingan logis antara sebuah data dengan kondisi penguji yang yang diberikan. Cara penulisan untuk If sederhana: =IF(sel yang ingin diuji, [nilai jika benar]. [nilai jika salah])

**Transformasi Data ke Visual**

* Diagram adalah sebuah representatif visual terhadap informasi. Diagram digunakan untuk mempermudah kita memahami informasi dari data yang kita miliki.
* Beberapa bentuk grafik dalam visualisasi data:
  + Bar/Column Chart: Jenis grafik di mana setiap kategori diwakili oleh persegi panjang, dengan panjang perseginya yang sebanding dengan nilai agregasi data.
  + Line Chart: Jenis grafik yang menampilkan informasi sebagai serangkaian titik data yang disebut 'penanda' yang dihubungkan oleh segmen garis lurus.
  + Area Chart: Mirip dengan jenis grafik Line Chart. Namun, area di bawah garis yang menghubungkan titik data akan terisi dengan warna atau bayangan.
  + Pie Chart: Jenis grafik yang menampilkan data dalam grafik lingkaran. Seluruh "kue pai" mewakili 100% dari keseluruhan, sedangkan pai "irisan" mewakili bagian dari keseluruhan.
  + Scatter Chart: Menggunakan titik untuk merepresentasikan nilai numerik pada kategori yang berbeda. Posisi setiap titik pada sumbu horizontal dan vertikal menunjukkan nilai untuk titik data individual.
  + Map Chart: Memungkinkan kita untuk memvisualisasikan hubungan spasial dalam data dengan menunjukkan data pada peta geografis. Umumnya pada peta seperti ini, semakin tebal warnanya, semakin intens pula kepadatan datanya.
* Kesalahan umum dalam membuat visualisasi data:
  + Nilai persentase tidak sesuai (tidak sepenuhnya sampai 100%) pada grafik pai.
  + Terlalu banyak data yang ingin ditunjukkan.
  + Tidak mengikuti standar penulisan grafik seperti menaruh nilai minimum di sumbu Y paling atas sehingga grafik terbalik.
  + Terdapat sumbu yang terpotong seperti ketika sumbu Y tidak dimulai dari 0%.
  + Penggunaan grafik 3D yang kurang sesuai sehingga dapat terjadi bias pada audiens. Sebagai contoh sebuah irisan pai dapat terlihat lebih besar dari seharusnya dan sebuah batang terlihat lebih tinggi dari seharusnya.
  + Grafik susah dibandingkan karena tidak jelas tujuan dan informasi yang ingin disampaikan.
* Data provenance (dokumentasi data) adalah sebuah langkah untuk melihat sumber data yang kita peroleh.
  + Dokumentasi data dibutuhkan untuk tahu segala proses dalam pembuatan, perubahan, penanganan, mau pun hal lainnya yang berpengaruh kepada data.
  + Dokumentasi data penting untuk mengetahui keabsahan data dan memungkinkan kita menggunakannya kembali di waktu yang lain.
  + Dokumentasi data penting untuk dapat menilai keaslian dan menumbuhkan kepercayaan kepada data, sekaligus untuk mereproduksi data tersebut.
  + Contoh dari dokumentasi data dapat berupa metadata atau version history,
* Pendekatan yang dapat dilakukan dalam melakukan dokumentasi data sebagai berikut:
  + Dicatat dalam bentuk teks, bisa menggunakan skema penulisan umum atau bisa juga dengan skema khusus dalam data provenance.
  + Dicatat dan disimpan secara internal menggunakan program perangkat lunak atau dalam sistem eksternal lainnya.
  + Dituliskan dalam bentuk yang dapat dibaca oleh mesin atau yang bisa dibaca oleh manusia.
* Contoh bentuk penulisan dokumentasi data:
  + Berkas README
  + Data Dictionaries (kamus data)
  + Data Paper
* Tips dalam menjalankan dokumentasi data:
  + Gunakan Alur Ilmiah yang Terstruktur: Gunakan alur terstruktur yang terdiri dari pencatatan, eksekusi, pemrosesan, dan urutan secara ilmiah, supaya pembaca paham sumber asal dan teori yang mendukung dokumentasi data tersebut.
  + Informasi Harus Jelas dan Sedetail Mungkin: Tautan ke data sumber asli harus jelas dan sebutkan dari mana mendapatkannya.
* Exploratory analysis adalah proses penyampaian di mana Anda membimbing dan memandu audiens dalam melihat (menjelajahi) data yang telah dikumpulkan.
* Explanatory analysis adalah proses penyampaian di mana Anda tidak meminta audiens untuk menjelajahi data namun langsung ke poin-poin utama dan implikasi dari data tersebut.
* 3 poin penting dalam melakukan explanatory analysis:
  + Kepada siapa kita berkomunikasi: Berkomunikasi dengan terlalu banyak orang pada saat sama sekaligus, cenderung tidak tepat sasaran karena kebutuhan masing-masing orang dapat berbeda. Jika Anda sebelumnya tidak pernah bertemu dengan audiens, berarti Anda terlebih dahulu harus membangun reputasi bahwa Anda itu benar-benar mengerti topik yang akan diberikan. Biasanya hal ini dilakukan dengan cara perkenalan diri.
  + Bagaimana kita berkomunikasi dengan audiens: Pada proses presentasi secara langsung dan tatap muka, kita dapat menanggapi audiens jika terdapat hal yang kurang jelas, sehingga tidak semua yang kita sampaikan harus ditulis secara detail pada dokumen presentasi dan visualisasi data. Jika dalam bentuk dokumen yang dibaca sendiri-sendiri maka tingkat detail yang diperlukan pada penulisan dokumen biasanya lebih tinggi. Jika kita adalah orang yang menganalisis dan mengomunikasikan data, maka kita harus percaya diri.
  + Bagaimana kita bisa menggunakan data untuk membantu menegaskan maksud kita: Salah satu cara yang dapat dilakukan adalah menggunakan metode 5W (What, Who, When, Why, Where) untuk menegaskan informasi penting yang ingin kita sampaikan.
* Contoh visualisasi data yang efektif
  + Berbentuk Teks: Ketika kita hanya ingin menampilkan satu atau dua angka saja, gunakan teks sederhana.
  + Tabel: Tabel sangat baik untuk ketika audiens Anda beragam karena mereka dapat membaca baris tertentu sesuai kepentingan masing-masing. Jika kita berhadapan dengan banyak unit pengukuran, tabel lebih memudahkan saat dibaca.
  + Heatmap: Heatmap menggabungkan detail angka pada tabel dengan visual agar membantu audiens untuk lebih cepat fokus pada informasi yang penting pada tabel.
  + Grafik: Grafik lebih cepat diproses dan mudah dipahami dalam mendapatkan informasi.
    - Grafik scatterplot berguna untuk menunjukkan hubungan antara dua hal yang terdapat di sumbu x dan y karena grafik ini membuat kita memproses data serentak untuk melihat hubungan apa yang muncul.
    - Grafik garis biasanya digunakan untuk melihat data yang bersifat kontinu (berkelanjutan). Karena titik dihubungkan melalui garis, grafik ini memperlihatkan perubahan (pertumbuhan atau pengurangan) yang terjadi antar titik, sehingga tidak sesuai untuk data yang bersifat kategoris seperti data yang membandingkan jenis kelamin (berapa banyak yang laki-laki dan perempuan).
    - Slopegraph merupakan grafik berkategori garis yang hanya memperlihatkan perubahan antara dua titik saja agar audiens fokus pada perubahan kenaikan atau penurunan yang terjadi.
    - Grafik batang (vertikal dan horizontal) dapat membandingkan beberapa data yang digambarkan dalam bentuk batang secara cepat sehingga audiens bisa dengan cepat tahu data yang terbesar, terkecil, dan perbandingan antara kategori data.
    - Grafik batang bertumpuk (vertikal dan horizontal) memungkinkan untuk membandingkan antar kategori utama dan juga antar sub komponen dalam tiap kategori.
    - Grafik air terjun (waterfall chart) dapat digunakan untuk menampilkan bagian tertentu dari grafik batang bertumpuk sehingga lebih fokus menunjukkan data awal, kenaikan atau penurunan yang terjadi, serta data akhir.
    - Grafik area memungkinkan untuk menampilkan data secara lebih ringkas dari grafik lainnya namun tidak cocok memvisualisasikan data yang besaran perbedaan angkanya tidak tinggi.
* Contoh prinsip Gestalt dalam persepsi visual
  + Proximity (Kedekatan): Objek yang jaraknya berdekatan merupakan kelompok atau golongan yang sama.
  + Similarity (Kesamaan): Objek yang memiliki warna, bentuk, ukuran, dan arah yang sama dianggap terkait atau termasuk bagian dari kelompok yang sama.
  + Enclosure (Pembeda): Objek yang memiliki batas fisik atau border yang sama merupakan satu golongan sama.
  + Closure (Penutupan Bentuk): Manusia cenderung memandang sekumpulan objek yang terpecah-pecah sebagai bagian dari satu objek yang lengkap, daripada memandang kumpulan objek tersebut sebagai sesuatu yang benar-benar terpisah satu sama lain
  + Continuity (Kesinambungan Pola): Manusia cenderung secara alami membentuk garis pola walaupun tidak terlihat secara eksplisit.
  + Connection (Koneksi): Objek yang secara fisik terhubung adalah bagian dari satu golongan atau kesatuan.
* Cognitive load (beban kognitif) adalah usaha mental yang harus dilakukan untuk dapat memproses dan mempelajari informasi.
  + Salah satu penyebab utama yang menimbulkan cognitive load yang berlebihan adalah sebuah kerumitan atau disebut juga dengan clutter. Clutter merupakan elemen visual yang tidak menambah pemahaman.
  + Agar audiens dapat dengan efektif mengerti visual yang diberikan, maka Anda perlu mengurangi cognitive load. Secara umum, identifikasilah dan hapuslah elemen yang kurang efektif.
* Preattentive attributes adalah komponen atau atribut visual yang langsung tertangkap perhatian kita secara tanpa sadar.
  + Preattentive attributes dapat dimanfaatkan untuk membantu mengarahkan perhatian pada fokus utama yang ingin disampaikan.
  + Atribut ini dapat digunakan untuk membuat hierarki elemen visual yang mengarahkan perhatian pada informasi yang hendak diinformasikan sesuai dengan proses yang diharapkan.
  + Contoh preattentive attributes adalah penggunaan warna yang berbeda, ukuran font yang lebih besar, menggunakan font cetak tebal, dan memilih posisi yang lebih umum dilihat.
* 3 contoh konsep desain dalam berkomunikasi dengan data
  + Affordances: Dalam istilah desain, semua benda memiliki fungsinya masing-masing. Karakteristik ini menunjukkan bagaimana sebuah objek seharusnya berinteraksi. Cara penerapan:
    - Gunakan warna yang umum
    - Gunakan warna yang konsisten
    - Gunakan posisi yang konsisten
  + Accessibility: Visual dan desain seharusnya bisa dimengerti dan digunakan oleh orang dari berbagai latar belakang atau kemampuan. Cara penerapan:
    - Gunakan bahasa yang sederhana
    - Visual harus dapat menjelaskan dirinya sendiri
  + Aesthetics: Kebiasaan manusia untuk mendapatkan kesan negatif apabila melihat sesuatu yang kurang menarik atau tidak indah.
    - Membuat pilihan warna, bentuk, dan posisi menjadi satu kesatuan yang menarik