**FOUNDATION DATA : DATA EVERYWHERE**

data adalah kumpulan fakta. Kumpulan ini bisa berupa angka, gambar, video, kata-kata, pengukuran, pengamatan, dan banyak lagi. Data adalah kumpulan fakta yang dapat digunakan untuk membuat kesimpulan, membuat prediksi, dan membantu dalam pengambilan keputusan.

Analisis data adalah pengumpulan, transformasi, dan pengorganisasian data untuk menarik kesimpulan, membuat prediksi, dan mendorong pengambilan keputusan yang tepat.

seorang analis data adalah seseorang yang mengumpulkan, mentransformasi, dan mengatur data untuk membantu membuat keputusan yang tepat.

Decision Intelligence adalah kombinasi dari ilmu data terapan dan ilmu sosial dan manajerial. Ini adalah tentang memanfaatkan kekuatan dan keindahan data.

Ekosistem data terdiri dari berbagai elemen yang berinteraksi satu sama lain untuk menghasilkan, mengelola, menyimpan, mengatur, menganalisis, dan berbagi data. Elemen-elemen ini mencakup perangkat keras dan perangkat lunak, serta orang-orang yang menggunakannya

Perbedaan data science dan data analyst:

* Data Science didefinisikan sebagai menciptakan cara-cara baru untuk memodelkan dan memahami yang tidak diketahui dengan menggunakan data mentah sedangkan Data analitics adalah pengumpulan, transformasi, dan organisasi data untuk menarik kesimpulan, membuat prediksi, dan mendorong pengambilan keputusan yang tepat. Data analytics dalam istilah yang paling sederhana adalah ilmu data.
* Data Science membuat pertanyaan baru menggunakan data, sementara Data analitics menemukan jawaban untuk pertanyaan yang ada dengan menciptakan wawasan dari sumber data

Data-driven decision-making didefinisikan sebagai menggunakan fakta untuk membimbing strategi bisnis.

Langkah data driven decision making :

1. Mencari tahu kebutuhan bisnis, setelah di definisikan, seorang data analis harus menemukan data, menganalisisnya dan menggunakannya untuk mengungkap tren, pola dan hubungan

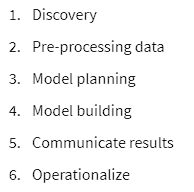
Untuk mendapatkan hasil maksimal dari pengambilan keputusan berbasis data, penting untuk memasukkan wawasan dari orang-orang yang akrab dengan masalah bisnis. Orang-orang ini disebut Subject Matter Expert, dan mereka memiliki kemampuan untuk melihat hasil analisis data dan mengidentifikasi ketidaksesuaian apa pun, memahami area abu-abu, dan akhirnya memvalidasi pilihan yang dibuat. Organisasi yang bekerja dengan cara ini menempatkan data di jantung setiap strategi bisnis, tetapi juga mendapat manfaat dari wawasan orang-orang mereka. Sebagai analis data, Anda memainkan peran kunci dalam memberdayakan organisasi-organisasi ini untuk membuat keputusan berbasis data, itulah sebabnya sangat penting bagi Anda untuk memahami bagaimana data berperan dalam proses pengambilan keputusan.

Para analis menggunakan pengambilan keputusan berdasarkan data dan mengikuti proses langkah demi langkah. Anda telah mempelajari bahwa ada enam langkah dalam proses ini:

1. Mengajukan pertanyaan dan mendefinisikan masalah. (Ask)
2. Menyiapkan data dengan mengumpulkan dan menyimpan informasi. (Prepare)
3. Memproses data dengan membersihkan dan memeriksa informasi. (Process)
4. Menganalisis data untuk menemukan pola, hubungan, dan tren. (Analyze)
5. Bagikan data dengan audiens Anda. (share)
6. Bertindak berdasarkan data dan menggunakan hasil analisis. (act)

Naluri adalah pemahaman intuitif tentang sesuatu dengan sedikit atau tanpa penjelasan. Hal ini tidak selalu merupakan sesuatu yang disadari; kita sering menangkap sinyal tanpa menyadarinya. Anda hanya memiliki "perasaan" bahwa itu benar.

## EMC's data analysis life cycle

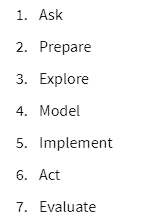
Cycle analitik data EMC Corporation adalah siklus dengan enam langkah:

1. Penemuan
2. Data pra-pemrosesan
3. Perencanaan model
4. Pembuatan model
5. Mengkomunikasikan hasil
6. Mengoperasionalkan

EMC Corporation sekarang menjadi Dell EMC. Model ini, yang dibuat oleh David Dietrich, mencerminkan sifat siklus dari proyek-proyek di dunia nyata. Fase-fase tersebut bukanlah tonggak sejarah yang statis; setiap langkah terhubung dan mengarah ke langkah berikutnya, dan pada akhirnya berulang. Pertanyaan-pertanyaan kunci membantu para analis menguji apakah mereka telah mencapai cukup banyak hal untuk melangkah maju dan memastikan bahwa tim telah menghabiskan waktu yang cukup untuk setiap fase dan tidak memulai pemodelan sebelum datanya siap.

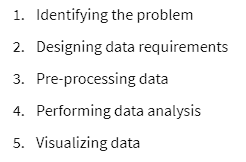
**SAS's iterative life cycle**

Siklus hidup berulang diciptakan oleh perusahaan bernama SAS, penyedia solusi analisis data terkemuka. Siklus hidup ini dapat digunakan untuk menghasilkan hasil yang dapat diulang, dapat diandalkan, dan dapat diprediksi:

1. Menanyakan
2. Mempersiapkan
3. Mengeksplorasi
4. Memodelkan
5. Menerapkan
6. Bertindak
7. Mengevaluasi

Model SAS menekankan sifat siklus dari model mereka dengan memvisualisasikannya sebagai simbol tak terhingga. Siklus hidup mereka memiliki tujuh langkah, yang sebagian besar telah kita lihat di model lainnya, seperti Ask, Prepare, Model, dan Act. Namun siklus hidup ini juga sedikit berbeda; siklus hidup ini memiliki langkah setelah fase act yang dirancang untuk membantu analis mengevaluasi solusi mereka dan berpotensi untuk kembali ke fase ask lagi.

## Project-based data analytics life cycle

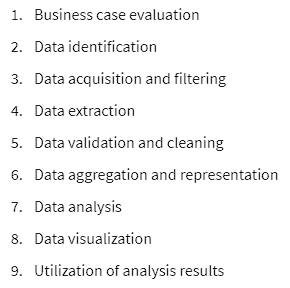
Siklus hidup analisis data berbasis proyek memiliki lima langkah sederhana:

1. Mengidentifikasi masalah
2. Merancang kebutuhan data
3. Pra-pemrosesan data
4. Melakukan analisis data
5. Memvisualisasikan data

Siklus hidup proyek analisis data ini dikembangkan oleh Vignesh Prajapati. Siklus hidup ini tidak mencakup fase keenam, atau yang kami sebut sebagai fase Act. Namun, siklus ini masih mencakup banyak langkah yang sama dengan siklus hidup yang telah kami jelaskan. Dimulai dengan mengidentifikasi masalah, menyiapkan dan memproses data sebelum analisis, dan diakhiri dengan visualisasi data.

## Big data analytics life cycle

Penulis Thomas Erl, Wajid Khattak, dan Paul Buhler mengusulkan siklus hidup analitik data besar dalam buku mereka, Big Data Fundamentals: Konsep, Pendorong & Teknik. Siklus hidup mereka menunjukkan fase-fase yang dibagi menjadi sembilan langkah:

1. Evaluasi kasus bisnis
2. Identifikasi data
3. Akuisisi dan penyaringan data
4. Ekstraksi data
5. Validasi dan pembersihan data
6. Agregasi dan representasi data
7. Analisis data
8. Visualisasi data
9. Pemanfaatan hasil analisis

Siklus hidup ini tampaknya memiliki tiga atau empat langkah lebih banyak daripada model siklus hidup sebelumnya. Namun pada kenyataannya, mereka baru saja memecah apa yang selama ini kita sebut sebagai Persiapan dan Proses menjadi langkah-langkah yang lebih kecil. Hal ini menekankan pada tugas-tugas individu yang diperlukan untuk mengumpulkan, mempersiapkan, dan membersihkan data sebelum tahap analisis.