



Data Science

Pokok Bahasan : *Data Analytics Lifecycle*

Dosen Pengampu : Hary Sabita, ST., MTI

12 April 2022



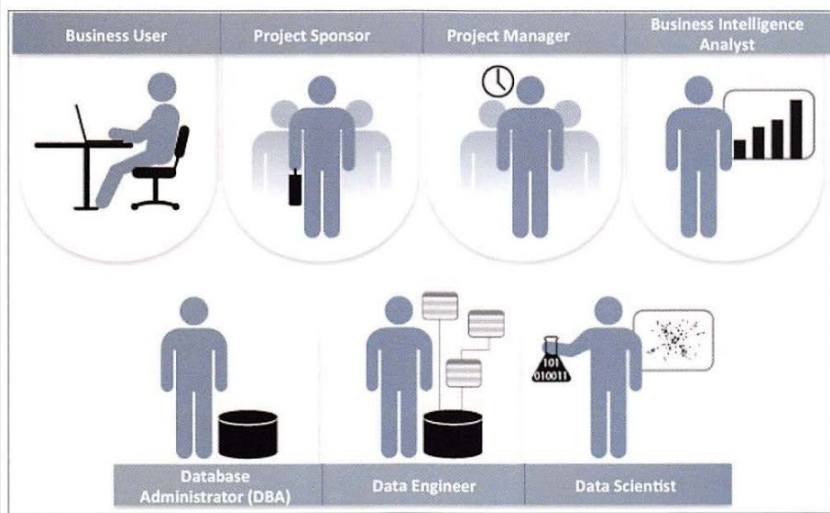
Outline :

1. *Discovery*
2. *Data Preparation*
3. *Model Planning*
4. *Model Execution*
5. *Communicate Results*
6. *Operationalize*

Data Analytics Lifecycle Overview

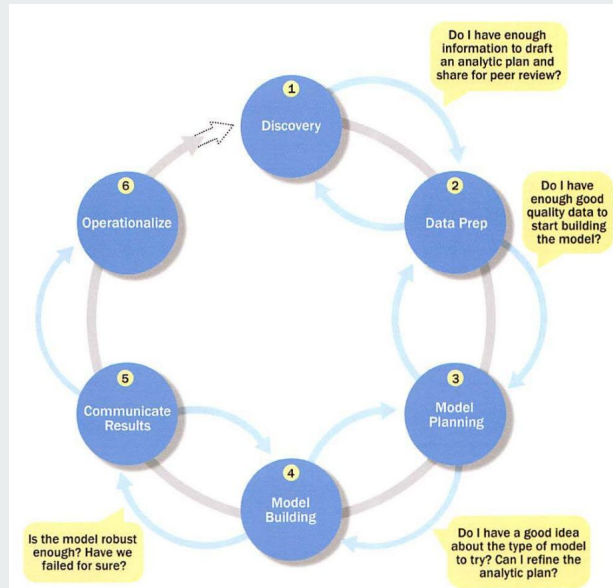
1. Didesain untuk *big data problems* dan *data science project*
2. Memiliki 6 fase
3. Setiap proyek dapat terjadi di beberapa fase dalam satu waktu
4. Umumnya fase dapat bergerak maju (*forward*) atau mundur (*backward*)
5. Bisa kembali ke fase sebelumnya bila ditemukan informasi baru ; agar dapat bergerak secara iteratif dan mendorong menuju operasional pekerjaan proyek

Peran peran Kunci dalam Kesuksesan Proyek Analisis



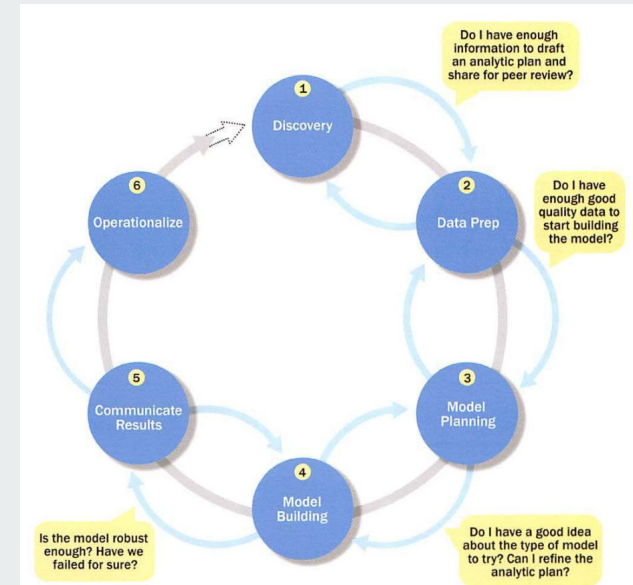
1. **Bisnis User**; Owner/ pemilik pekerjaan
2. **Project Sponsor**; Penyedia dana
3. **Project Manager**; Pemimpin proyek
4. **Business Intelligence Analyst**; Membuat dashboard laporan, KPI
5. **Database Administrator (DBA)**; Menyediakan akses ke database
6. **Data Engineer**; Membantu kemampuan teknis dalam query.
7. **Data Scientist**; Bertanggung jawab terhadap teknik analisis data, membuat model dan mengimplementasikannya

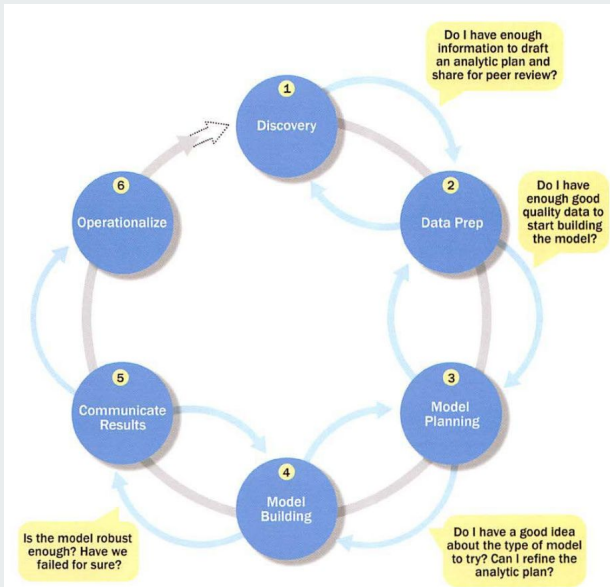
Data Analytics Lifecycles



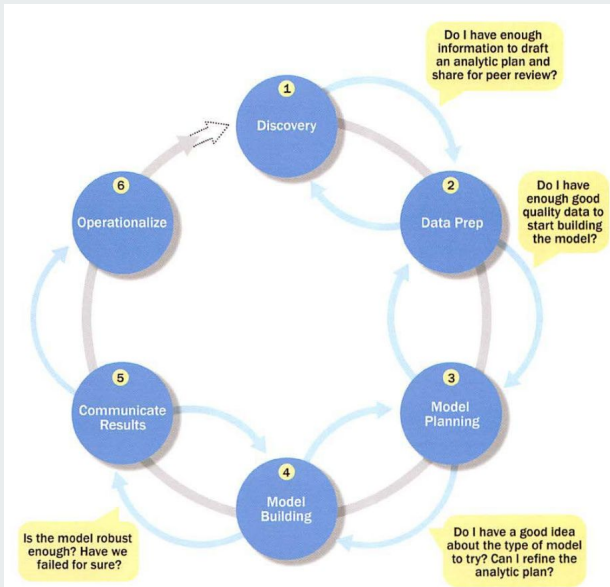
Fase 1 - Discovery; Mempelajari domain bisnis, Bagaimana organisasinya, Apakah ada proyek serupa dimasa lalu?, Menilai sumber daya yang ada (teknologi, waktu, data), Membingkai masalah bisnis sebagai tantangan analitik bisnis yang dapat diatasi, Merumuskan hipotesis awal untuk menguji dan mempelajari data agar dapat bergerak ke fase berikutnya

Fase 2 - Data Preparation; Membutuhkan keberadaan *analytic sandbox*, Dapat bekerja dengan data selama proyek berlangsung, Merubah data menjadi proses ETLT(Ext, Transform, Load, Transform) agar dapat bekerja dengan data dan menganalisanya, Memahami data secara menyeluruh dan mengkondisikannya.



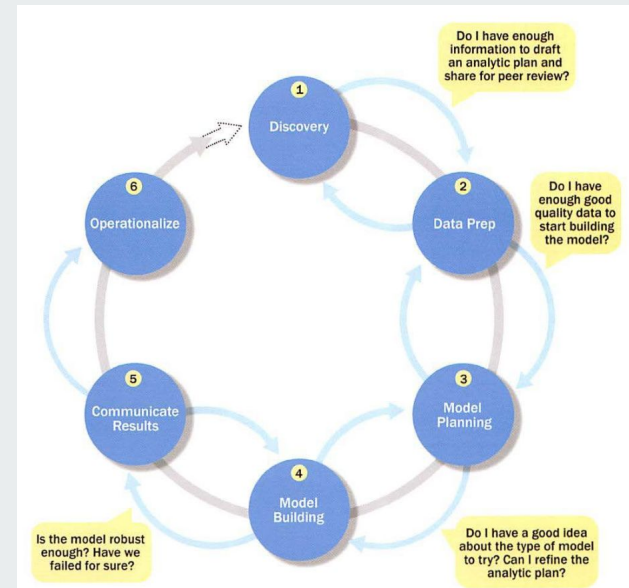


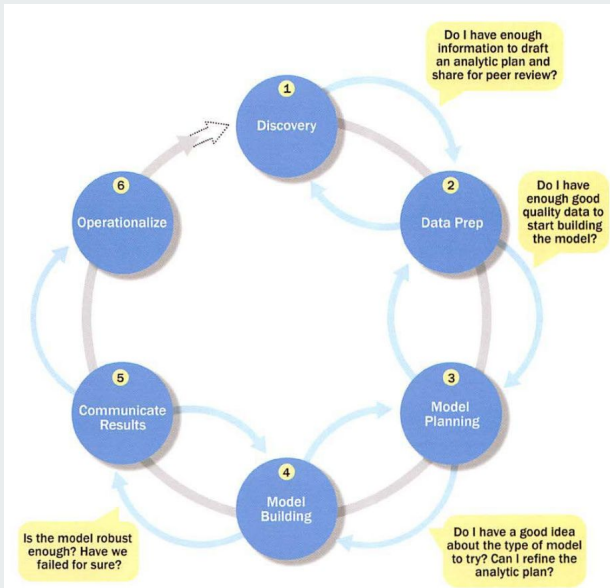
Fase 3 - Model Planning; Merencanakan model yang paling cocok yang akan digunakan, Menentukan metode, teknik untuk alur kerja berikutnya, Mengeksplorasi data, Menentukan variabel data



Fase 4 - Model Building; Mengembangkan data untuk diuji, dilatih dan diproduksi, Membangun dan mengeksekusi model, Mempertimbangkan apakah tool yang ada bisa untuk membangun model yang telah ditentukan, Mempertimbangkan apakah ada lingkungan luar yang paling mungkin untuk mengeksekusi model

Fase 5 - Communicate Results; Berkolaborasi dengan stakeholder apakah hasil proyek berhasil atau gagal sesuai dengan apa yang dikembangkan di fase 1, Mengidentifikasi temuan utama, Mengukur nilai bisnis, Memberikan laporan ringkas kepada stakeholder agar mudah dipahami.





Fase 6 - Operationalize; Memberikan laporan akhir proyek,
Menjalankan proyek model percontohan



Setelah anggota tim menjalankan model dan menghasilkan temuan, sangat penting untuk:

1. Membedakan hasil ini dengan cara yang disesuaikan dengan audiens yang melibatkan tim.
2. Membedakan hasil dari pekerjaan dengan cara yang menunjukkan nilai yang jelas.
3. Jika tim melakukan analisis yang akurat secara teknis tetapi gagal menerjemahkan hasilnya ke dalam bahasa yang sesuai dengan audiens, orang tidak akan melihat nilainya, dan sebagian besar waktu dan usaha untuk proyek akan terbuang percuma.

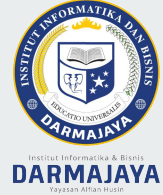


Next..

Pertemuan berikutnya akan membahas secara rinci bagaimana masing-masing dari enam fase bekerja dan sebuah tugas dalam bentuk studi kasus di sebuah perusahaan...



TERIMA KASIH



Tugas Kelompok (Tugas 3)

1. Di fase mana tim berharap untuk menginvestasikan sebagian besar waktu proyek?
Mengapa? Di mana tim berharap untuk menghabiskan waktu paling sedikit? Mengapa?