

# Overview Data Scientist

# Virtual Internship Experience

# Disclaimer

“Dokumen ini memiliki hak cipta. Barang siapa yang menyebarluaskan atau menduplikasi tanpa izin dari instansi terkait dapat diproses sesuai dengan ketentuan hukum yang berlaku.”

# Outline

- 1 **Peran Data Scientist**
- 2 **Tugas Data Scientist**
- 3 **Alur Kerja**

# 1 Peran Data Scientist

Peran seorang data scientist adalah untuk mengumpulkan, menganalisis, dan menginterpretasikan data untuk menghasilkan wawasan bisnis yang berharga.

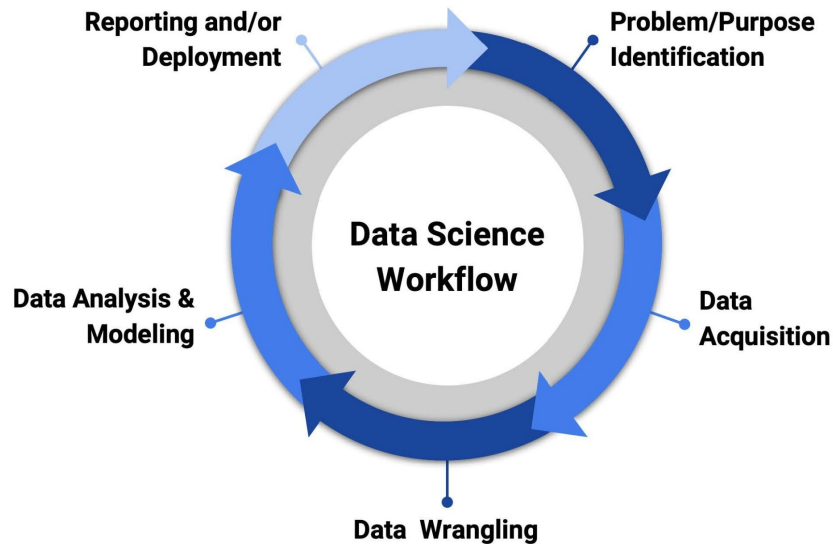
Mereka menggunakan berbagai teknik dan algoritma analisis data untuk memecahkan masalah, menemukan pola, dan membuat prediksi berdasarkan data yang ada



## **2** Tugas *Data Scientist*

- 1** Eksplorasi dan pembersihan data
- 2** Model dan analisis data
- 3** Pengembangan dan implementasi model
- 4** Komunikasi hasil analisis
- 5** Kolaborasi dengan tim lintas disiplin
- 6** Pelaporan dan pemantauan

### 3 Alur Kerja



Gambar 1. Alur kerja Data Scientist

Secara singkat, alur kerja data scientist memiliki pembagian sebagai berikut:

1. Sebelum memulai proyek Data Scientist bagian terpenting adalah mengidentifikasi dan **menyatakan pernyataan masalah** dengan jelas. Langkah ini berfungsi sebagai panduan untuk proyek Anda selanjutnya dan menginformasikan bagaimana langkah-langkah lainnya akan dilakukan.
2. Selanjutnya, **pengumpulan data** yang dapat diperoleh dari berbagai sumber. Baik itu file CSV, database, ekstraksi dari website, dll
3. Setelah data terkumpul, data perlu dibersihkan dan **diproses** terlebih dahulu untuk memastikan bahwa data tersebut dalam format yang sesuai untuk analisis. Langkah ini melibatkan *data cleansing*, *data integration*, *data transformation*, dan *feature engineering*

### 3 Alur Kerja

4. Ketika Data Scientist memiliki akses ke data yang dikumpulkan, waktu harus dihabiskan untuk membiasakan diri dengan data tersebut. Pada fase ini, sangat penting untuk merumuskan hipotesis sambil mencari pola dan anomali dalam data. Anda harus menyelesaikan masalah baik dengan menggunakan **model** seperti regresi, klasifikasi, dll
5. Data Scientist harus dapat **mengkomunikasikan** hasil mereka karena mereka akan sering melakukan ini. Data Scientist harus membagikan temuan, hasil, dan menceritakan dengan berbagai *stakeholders* selama fase ini. Para *stakeholders* biasanya tidak berpengalaman dalam data science, sehingga dapat mengubah pesan mereka dalam bentuk visualisasi yang menarik yang akan membantu pemahaman para stakeholders.



KALBE  
Nutritional



**How is it implemented in Kalbe Nutritional?**





# Thank You!



KALBE  
Nutritional