

# Data Science and Python Venv

Diki Fernandi - 20124023

May 2025

## 1 Penjelasan secara singkat tentang data science

Data science (ilmu data) adalah bidang interdisipliner yang menggunakan metode ilmiah, algoritma, dan sistem untuk mengekstraksi pengetahuan atau wawasan dari data, baik yang terstruktur maupun tidak terstruktur. Ilmu ini menggabungkan statistik, ilmu komputer, dan pemahaman tentang domain (bidang aplikasi) untuk menyelesaikan masalah nyata melalui analisis dan pemodelan data.

## 2 Perbedaan antara Data, Data Science, dan Data Scientist

- **Data:** Data adalah kumpulan fakta mentah yang belum memiliki makna. Contohnya berupa angka, teks, gambar, atau sinyal yang belum diolah.
- **Data Science:** Ilmu yang mempelajari cara mengumpulkan, mengolah, menganalisis, dan menginterpretasi data untuk menghasilkan informasi atau wawasan yang berguna. Data science menggabungkan statistika, ilmu komputer, dan pemahaman domain.
- **Data Scientist:** Profesional yang bekerja di bidang data science. Mereka menggunakan algoritma, statistik, dan pemrograman untuk menganalisis data dan memberikan solusi berbasis data.

## 3 Empat Aspek Dasar Data Science

Berikut ini adalah empat fondasi utama dalam data science:

1. **Pengetahuan Domain (Domain Expertise):** Pemahaman mendalam tentang bidang aplikasi tempat data digunakan, seperti kesehatan, keuangan, atau teknologi.
2. **Matematika dan Statistika:** Digunakan untuk membuat model prediktif dan melakukan inferensi dari data, termasuk konsep seperti regresi, probabilitas, dan uji hipotesis.
3. **Ilmu Komputer (Computer Science):** Kemampuan dalam bahasa pemrograman seperti Python atau R, serta penggunaan pustaka (library) dan framework untuk analisis data dan machine learning.
4. **Komunikasi dan Visualisasi:** Kemampuan menyampaikan hasil analisis dalam bentuk grafik, tabel, atau narasi yang mudah dipahami oleh pihak non-teknis.