

Introduction to Python for Data Science

Pertemuan 1


Tutor : Kak Vanya






Apa yang kita pelajari ?

1. Intro to the class
2. Algoritma
3. Flowchart
4. Syntax



Intro to the class (Pengenalan)





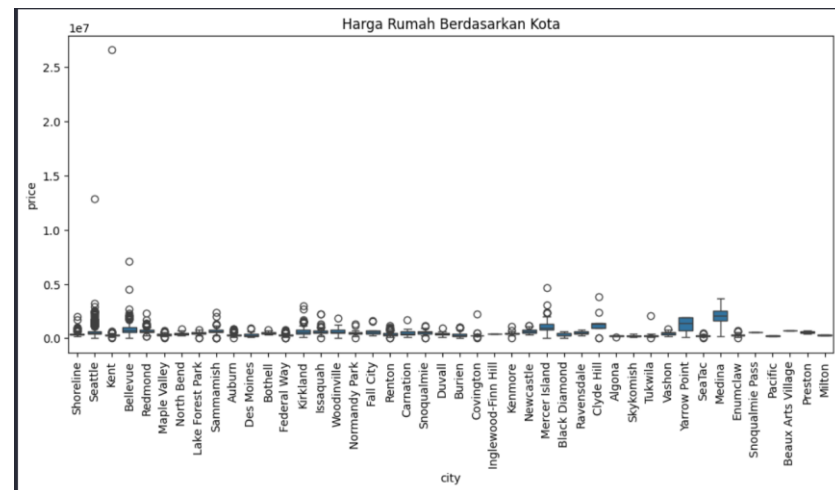
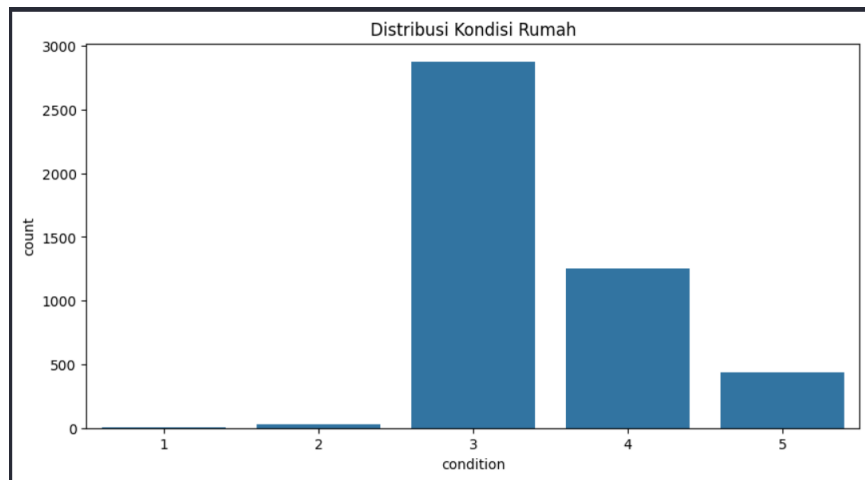
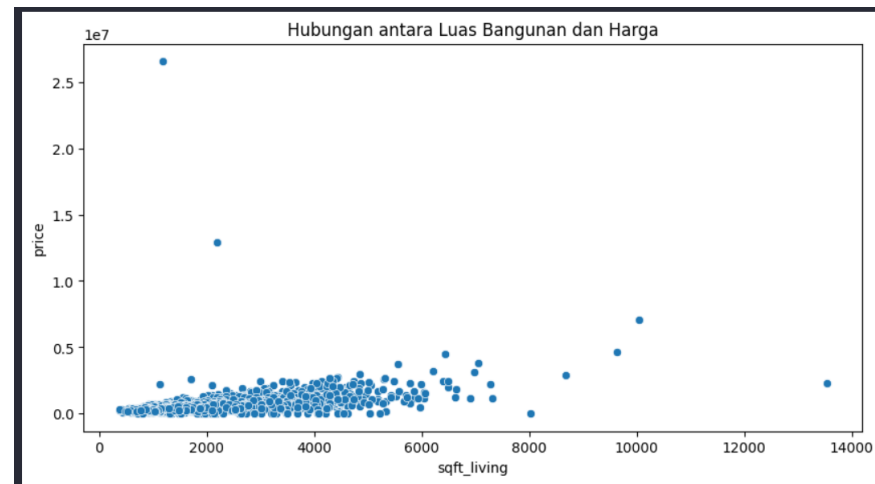
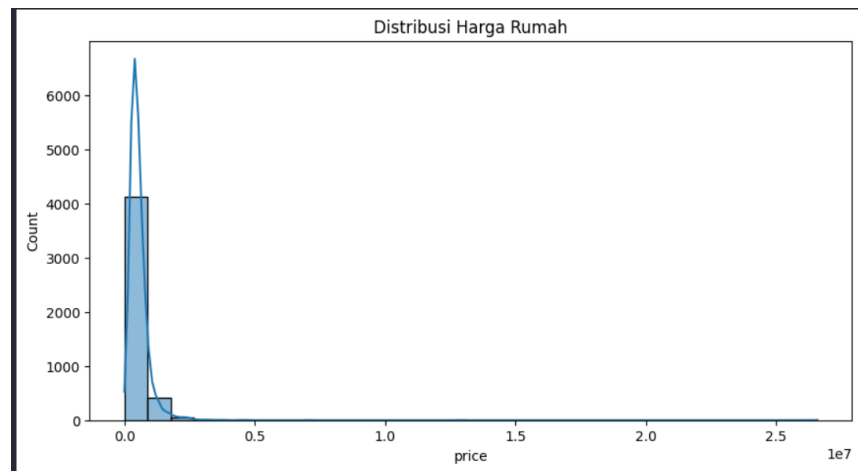
Data science

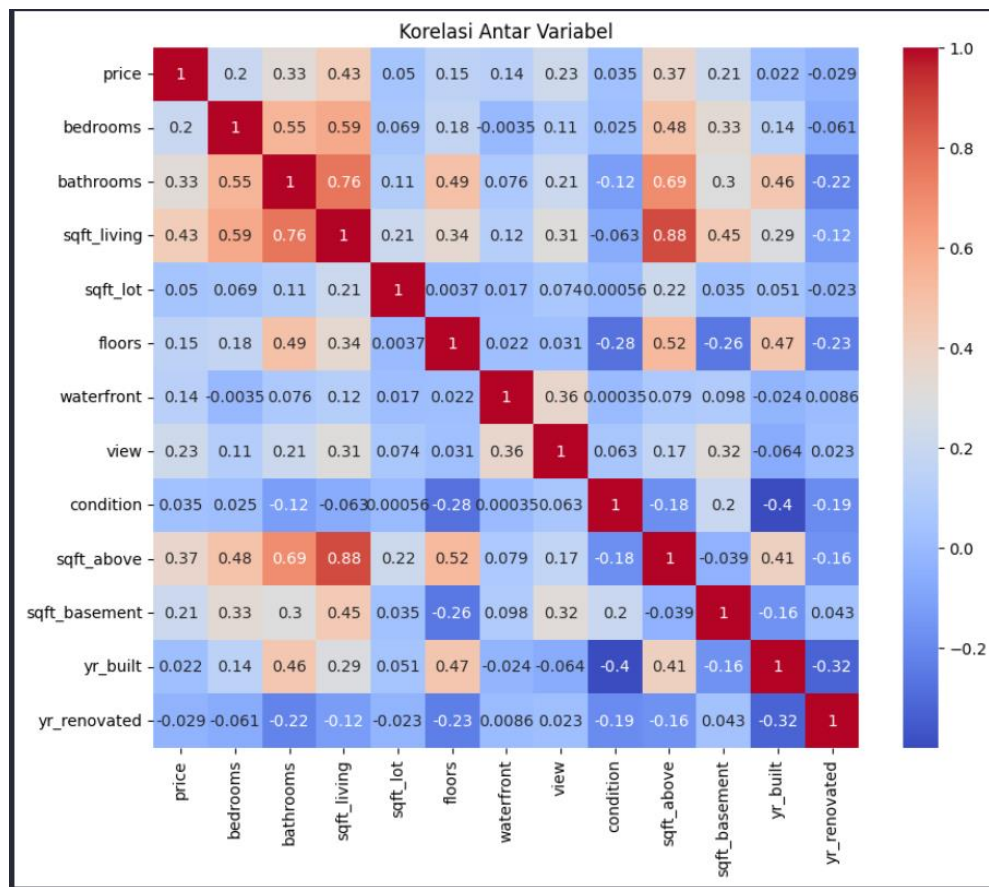
- Data Science adalah bidang ilmu yang fokus untuk mengolah data menjadi informasi yang berguna.
- Tujuan Data Science adalah mengubah data mentah menjadi **wawasan atau insight** yang digunakan untuk :
 - 1) Untuk mengambil Keputusan.
 - 2) Memprediksi masa depan atau insight.
 - 3) Process automation. (Pekerjaan yang repetitive yang dilakukan manusia tetapi dilakukan oleh computer)



Komponen Data science

- Data Collection – mengumpulkan data dari berbagai sumber (web, database, sensor, dsb.)
- Data Cleaning – membersihkan data agar rapi dan konsisten.
- Data Exploration – analisis awal (pakai Python, Pandas, dll).
- Data Analysis – mencari pola, tren, dan hubungan.
- Machine Learning – membuat model prediktif.
- Data Visualization – membuat grafik/chart agar hasil mudah dipahami





```
display(recommended_houses[['price', 'bedrooms', 'bathrooms', 'sqft_living', 'yr_built', 'yr_renovated', 'city']])
```

✓ 0.0s

	price	bedrooms	bathrooms	sqft_living	yr_built	yr_renovated	city
4356	0.0	4.0	2.75	2600	1960	2001	Seattle
4376	0.0	4.0	2.25	2170	1960	2001	Normandy Park
4408	0.0	5.0	2.25	2880	1990	0	Covington
4362	0.0	4.0	4.00	3680	1990	2009	Sammamish
4374	0.0	2.0	2.50	2200	2007	0	Enumclaw
4382	0.0	5.0	4.50	4630	2006	0	Snoqualmie
4383	0.0	5.0	4.00	4430	2013	1923	Bellevue
4385	0.0	4.0	4.50	5030	2008	0	Mercer Island
4389	0.0	4.0	3.50	4210	2006	0	Bellevue
4413	0.0	4.0	3.00	1990	1990	2009	Federal Way



Data dari smartphone!

- **Profil Transaksi**
 - berapa x pesan Gojek
 - berapa x pesan grabfood
 - jam berapa pesan gojek
 - berapa x menggunakan promo
- **Profil Session**
 - seberapa sering buka apps gojek? Kapan? Jam berapa?
- **Profil Location**
 - lokasi mana yang paling sering digunakan untuk antar? Jam berapa?
- **Profil Device**
 - device apa yang dipakai?

- Identifying Consumers
- Recommending Products
- Analyzing Reviews

E-commerce



- Predicting Potential Problems
- Monitoring Systems
- Automating Manufacturing Units
- Maintenance Scheduling
- Anomaly Detection

Manufacturing



- Fraud Detection
- Credit Risk Modeling
- Customer Lifetime Value

Banking



Healthcare

- Medical Image Analysis
- Drug Discovery
- Bioinformatics
- Virtual Assistants



Transport

- Self Driving Cars
- Enhanced Driving Experience
- Car Monitoring System
- Enhancing the safety of passengers



Finance

- Customer Segmentation
- Strategic Decision Making
- Algorithmic Trading
- Risk Analytics

Data Science Applications



Tools yang digunakan

- Python (dengan Pandas, Numpy, Matplotlib, Scikit-Learn).
- R, SQL.
- Jupyter Notebook, Google Colab (tanpa perlu install).
- Power BI, Tableau (visualisasi)



Data Science Job

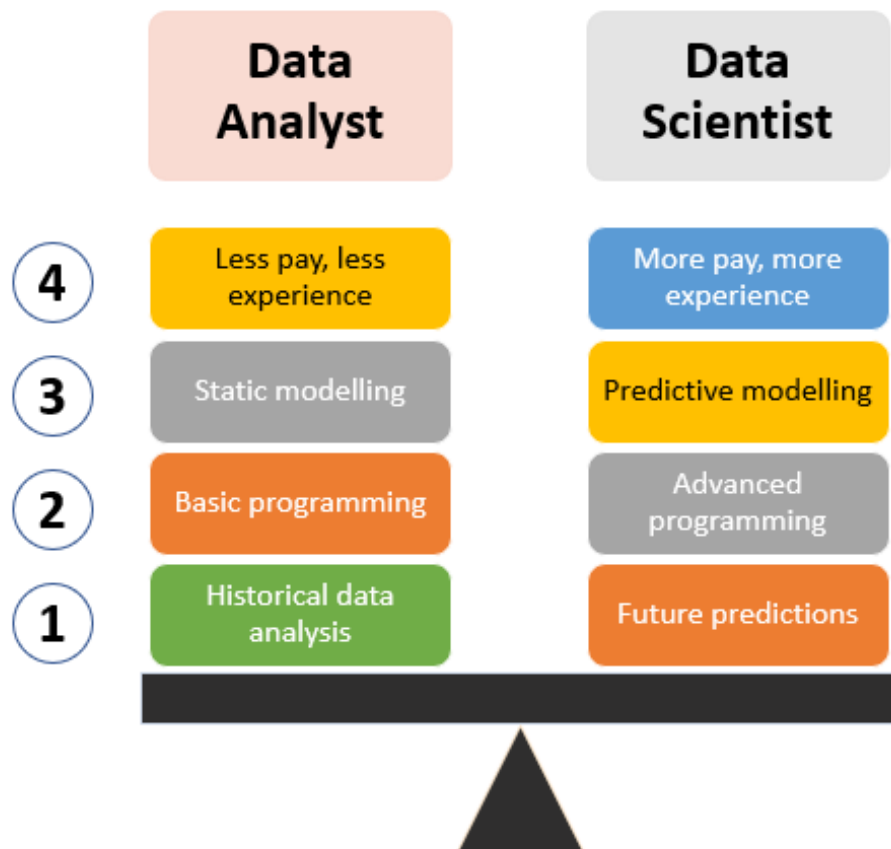
- Artificial Intelligence Specialist (#1)
- Data Scientist (#3)
- Data Analyst
- Business Intelligence

Skills

- Python/R (Programming)
- Matematika dan Statistik
- Komunikasi
- Database, data visualisasi
- dll

Role	Fokus	Tools	Output
AI Specialist	Kecerdasan buatan	ML, DL, Python	Sistem pintar
Data Scientist	Analisis + Prediksi	Python, ML	Model prediktif
Data Analyst	Menjelaskan data	Excel, SQL	Grafik, laporan
BI Analyst	Keputusan bisnis	Power BI, Tableau	Dashboard & insight

Peran	Ibaratnya	Fokus Utama	Contoh Tugas	Tools Umum
Data Engineer	Tukang saluran air	Menyiapkan dan mengalirkan data agar bersih dan bisa dipakai	Mengambil data dari database atau API, membersihkan, menyimpan di data warehouse	SQL, Python, Airflow, Cloud (BigQuery, AWS)
Data Analyst	Tukang baca peta	Membaca dan menjelaskan isi data	Buat grafik penjualan, temukan tren, buat laporan mingguan	Excel, SQL, Tableau, Power BI
Data Scientist	Peramal cerdas	Membuat model prediksi dari data	Prediksi pelanggan yang akan berhenti, deteksi penipuan, buat rekomendasi	Python, Pandas, Scikit-learn, Jupyter Notebook





Job Family	Posisi	Kota	Seniority Level	Rata2 Tahun Pengalaman	Gaji Per Bulan			
					Min	Mid	Max	Average
Data/Analytics	Data Engineer	DKI Jakarta	Entry Level	1,5	Rp5.890.000	Rp7.500.000	Rp30.000.000	Rp9.074.194
	Data Analyst		Entry Level	2,39	Rp3.000.000	Rp6.900.000	Rp13.860.000	Rp6.916.564
	Data Scientist		Entry Level	1,5	Rp6.000.000	Rp7.500.000	Rp13.500.000	Rp8.102.174
	Data Scientist	DKI Jakarta	Intermediate Level	2,4	Rp6.500.000	Rp8.500.000	Rp30.000.000	Rp11.110.690
	Data Analyst	DKI Jakarta	Intermediate Level	2,36	Rp6.900.000	Rp10.000.000	Rp18.000.000	Rp12.758.966
		Jawa Barat	Entry Level	1	Rp4.414.000	Rp6.950.000	Rp9.500.000	Rp6.798.357



Algoritma





Apa itu Algoritma ?

- Langkah-langkah yang tersusun secara logis untuk menyelesaikan suatu masalah.
- Ciri-ciri algoritma baik yaitu jelas, berurut, tidak ambigu dan efisien.
- Contoh sederhana membuat mie instant :
 - 1) Rebus air
 - 2) Masukkan mie ke dalam air mendidih
 - 3) Tunggu 3 menit
 - 4) Masukkan bumbu ke mangkuk
 - 5) Tuang mie dan airnya ke mangkuk
 - 6) Aduk, lalu sajikan



Flowchart










Apa itu Flowchart ?

- Flowchart adalah gambar atau diagram yang menggambarkan langkah-langkah dari suatu proses secara berurutan.
- Tujuannya adalah untuk mempermudah memahami alur logika dari sebuah program atau aktivitas.
- Untuk tools bisa Draw.io, Whimsical dan lainnya.
- Kenapa penting ?
 - 1) Membantu merancang program sebelum mulai coding.
 - 2) Membuat logika jadi lebih jelas dan terstruktur.
 - 3) Mudah untuk diskusi atau presentasi ke orang lain.
 - 4) Berguna untuk mendeteksi kesalahan logika sejak awal.

Symbol Flowchart



Symbol	Name	Function
	Start/end	An oval represents a start or end point
	Arrows	A line is a connector that shows relationships between the representative shapes
	Input/Output	A parallelogram represents input or output
	Process	A rectangle represents a process
	Decision	A diamond indicates a decision

Mulai

Buka kemasan

Tuangkan isi kemasan ke dalam cangkir

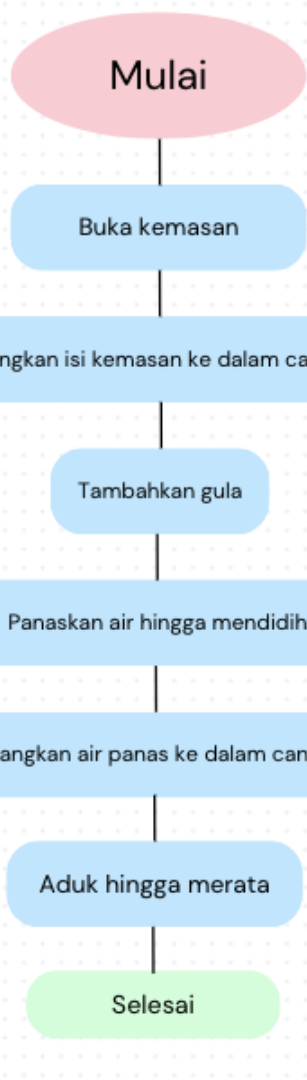
Tambahkan gula

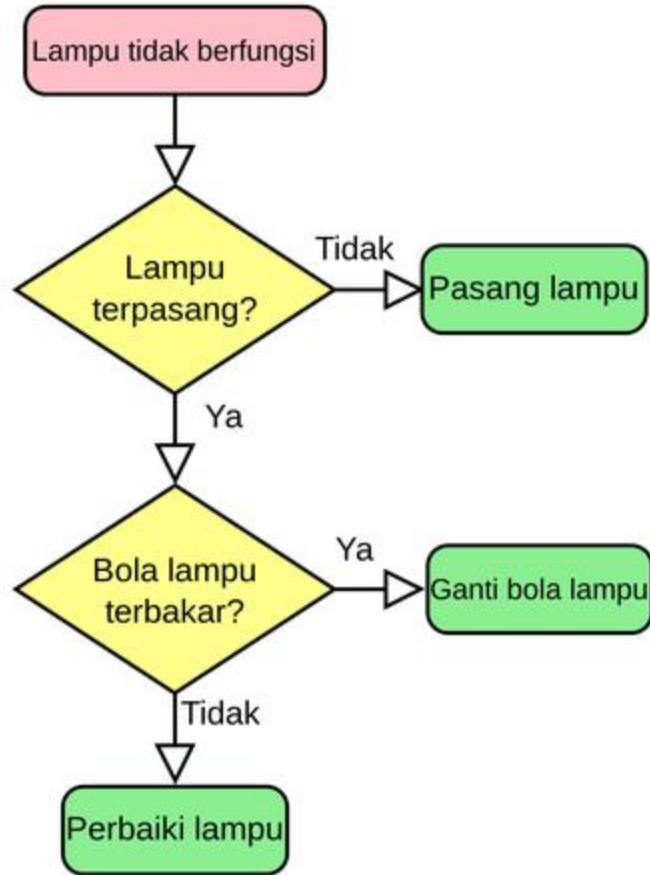
Panaskan air hingga mendidih

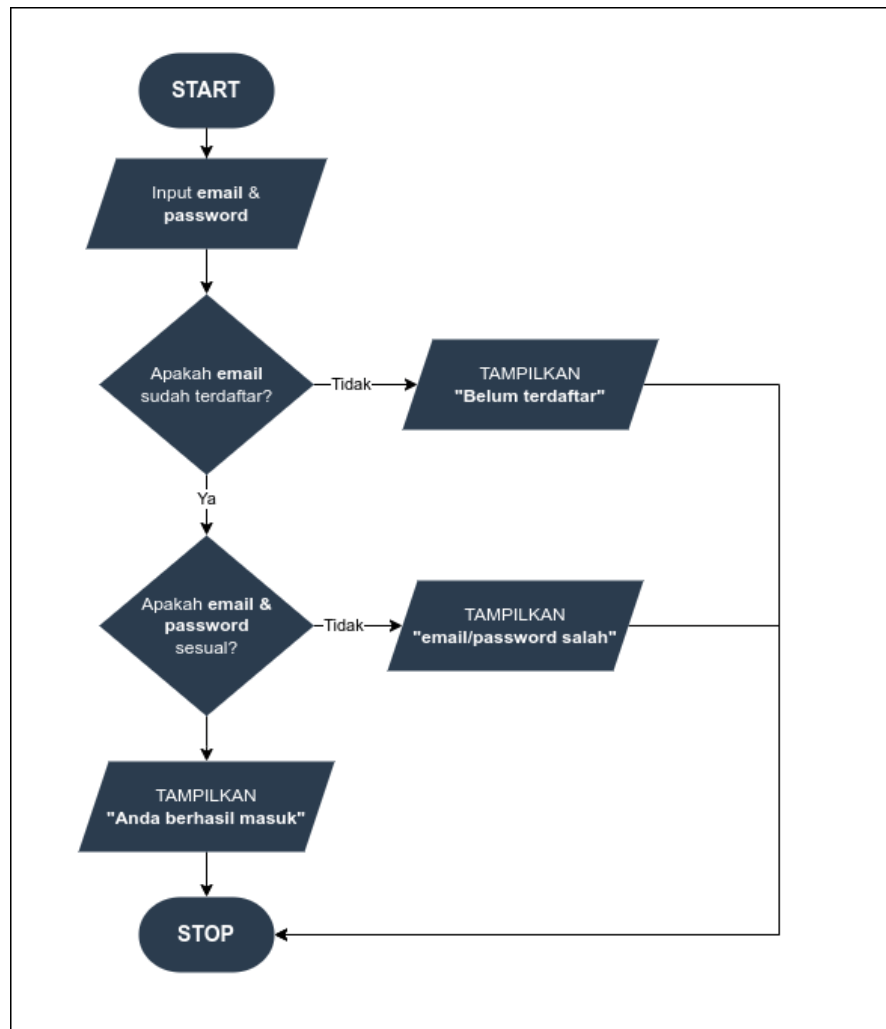
Tuangkan air panas ke dalam cangkir

Aduk hingga merata

Selesai









Syntax (Python)





Apa itu Syntax ?

- Sintaks (Syntax) adalah aturan penulisan kode dalam suatu bahasa pemrograman.
- Sama seperti tata bahasa dalam Bahasa Indonesia atau Inggris, setiap bahasa pemrograman punya aturan sendiri agar komputer bisa mengerti perintah kita.
- Kenapa Syntax penting ?
 - 1) Komputer itu "literal", hanya paham instruksi yang ditulis dengan benar
 - 2) Kalau salah tulis sedikit saja → bisa muncul error (contoh: lupa tanda titik dua : atau tanda kutip ")



```
1 # Contoh yang benar
2 print("Hello, World !")
3
4 # Contoh yang salah
5 Print("Halo, Dunia!")
6 # Error: huruf P harus kecil (Python sensitif huruf besar-kecil)
```



```
1  # Contoh yang benar
2  nama = "Vanya"
3  umur = 22
4
5  # Contoh yang salah
6  22 = umur
7  # Error: angka tidak bisa dijadikan nama variabel
```



Apa itu Python ?

- Python adalah bahasa pemrograman tingkat tinggi yang digunakan untuk membuat program atau aplikasi di berbagai bidang.
- Python for Data Science berarti menggunakan bahasa Python untuk mengolah, menganalisis, dan memvisualisasikan data agar bisa digunakan dalam pengambilan keputusan atau membuat model prediktif.



Setup Python

- Download Python : <https://www.python.org/>
- Download VSCode : <https://code.visualstudio.com/>
- Google Collab (tanpa install) : <https://colab.research.google.com/>
- How to install Python offline (Reference Only) :
<https://sanbercode.com/bootcamp/class/725/materi/persiapan-perangkat/1b3ec8be-0424-4ac1-95cc-14003264033b>

Ekstensi	Fungsi	Contoh
.py	File utama Python (kode program/script biasa)	program.py
.ipynb	Notebook interaktif (digunakan di Jupyter Notebook/Colab)	analisis.ipynb
.pyc	File Python yang sudah di-compile menjadi bytecode (otomatis dibuat)	main.pyc
.pyo	File Python yang di-optimize saat compile (jarang dipakai sekarang)	script.pyo
.pyw	Python tanpa terminal/console (khusus Windows GUI)	aplikasi.pyw

Alasan	Penjelasan
Mudah dibaca	Sintaks Python simpel, cocok untuk pemula
Punya banyak pustaka khusus data	Ada banyak library siap pakai (pandas, numpy, matplotlib, dsb.)
Dukungan machine learning & AI	Bisa dipakai untuk membuat model prediktif dan analisis canggih
Digunakan di industri	Perusahaan-perusahaan besar menggunakan Python untuk analisis data mereka

Library	Fungsi
NumPy	Operasi angka & array
Pandas	Mengolah data tabel (mirip Excel)
Matplotlib / Seaborn	Membuat grafik dan visualisasi
Scikit-learn	Machine learning (prediksi, klasifikasi)
Statsmodels	Analisis statistik lanjutan

C++



```
1  #include <iostream>
2  using namespace std;
3
4  int main() {
5      cout << "Hello, World!" << endl;
6      return 0;
7  }
8
```

VS

Python



```
1  print("Hello, World !")
```



Latihan

1. Buatlah contoh algoritme dalam kehidupan sehari2 dan jelaskan.
2. Buatlah Flowchart bebas, boleh flowchart suatu aplikasi atau kehidupan sehari2.
3. Buatlah program python sederhana dengan output : (Google Collab)
→ **“Halo nama saya dan saya tertarik untuk belajar Data Science”**
4. Buatlah program penjumlahan menggunakan python.
5. Buatlah program input nama dan usia lalu variable tersebut dipanggil kembali.
→ **Masukkan nama kamu: Vanya**
Masukkan usia kamu: 22
Halo, Vanya! Kamu berusia 22 tahun.