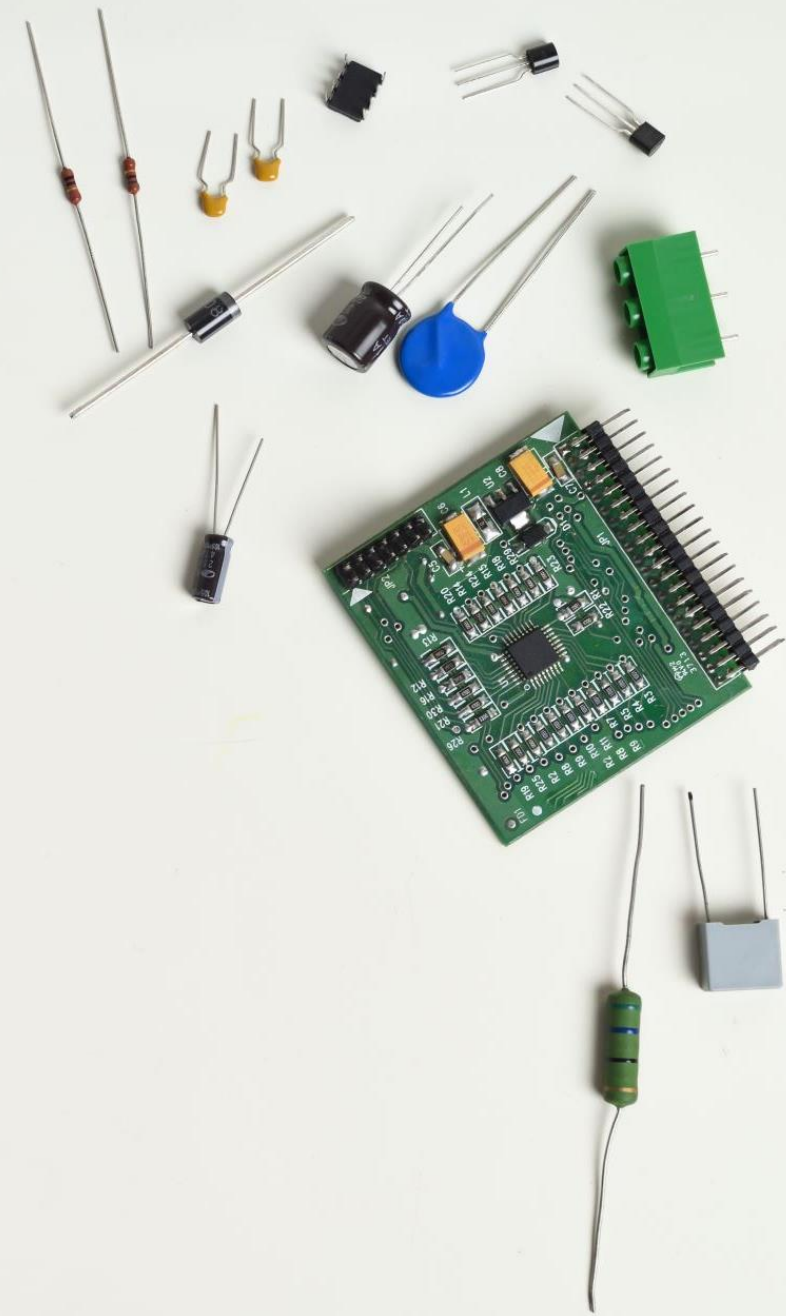


Pengenalan Kaggle, Github, dan Kedro



- Alat-alat yang dapat membantu ber data scientist ria: Kaggle, github dan kedro



Kaggle

- <https://www.kaggle.com/>
- Adalah platform untuk mendapatkan dataset, berkompetisi dan melatih kemampuan mengolah data hingga memnuat machine learning
- Dataset yang tersedia di Kaggle dapat diunduh dengan metode:
 1. Unduh dari website
 2. Undung dengan API (Aplication Platform Interface) biasanya berbasis website: <https://www.kaggle.com/docs/api> bisa menggunakan command (Kaggle datasets download -d[DATASET])

Menjalakan notebook menggunakan Kaggle

Dua metode dapat menjalankan notebook dengan kaggle



The screenshot shows a Jupyter Notebook titled "Airbnb NYC Price Prediction - Ajifi...". The notebook contains a Python script with the following content:

```
# This Python 2 environment comes with many helpful analytics libraries installed
# It is defined by the kaggle/python Docker image: https://github.com/kaggle/docker-python
# For example, here's several helpful packages to load

# Input data files are available in the read-only "/input/" directory
# For example, running this (by clicking run or pressing Shift+Enter) will list all files under the input directory

import os
for dirname, _, filenames in os.walk('/kaggle/input'):
    for filename in filenames:
        print(os.path.join(dirname, filename))

# You can write up to 500MB to the current directory (/kaggle/working/) that gets preserved as
# You can also write temporary files to /kaggle/temp/, but they won't be saved outside of the
session
```

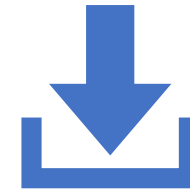
Below the code cell, there are two buttons: "+ Code" and "+ Markdown". The notebook also shows a sidebar on the right with sections like "Data", "Input", "Output (800B / 18.80B)", "Settings", "Schedule a notebook run", and "Code Help".



Praktikum (akses Kaggle
untuk login)



Menjelaskan isi dari Kaggle
ada kompetisi , ada dataset



Koding mendapatkan
dataset, bisa download juga

Mengenal Github



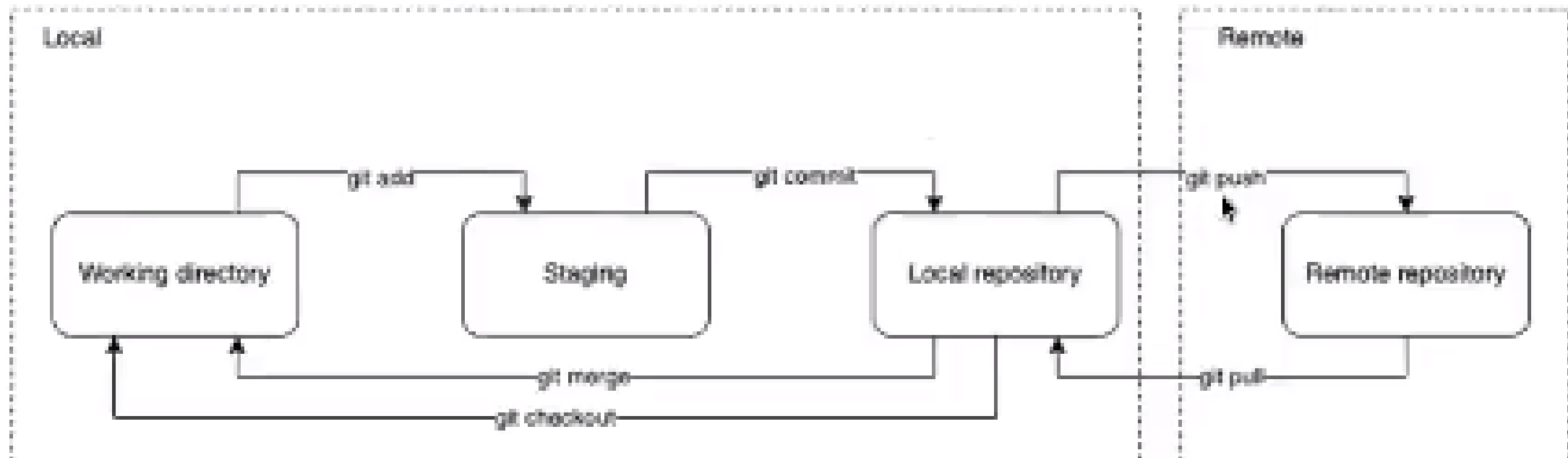
GITHUB.COM



GITHUB ADALAH TEMPAT UNTUK
BERKOLABORASI, MENYIMPAN, DAN
MENGELOLA KODE PROGRAM

Keutamaan menggunakan Github

- Memudahkan kolaborasi pengerjaan project
- Mencegah perubahan kode yang bisa merusak kode asli
- Sebagai portfolio bagi Developer



Mengenai Kedro

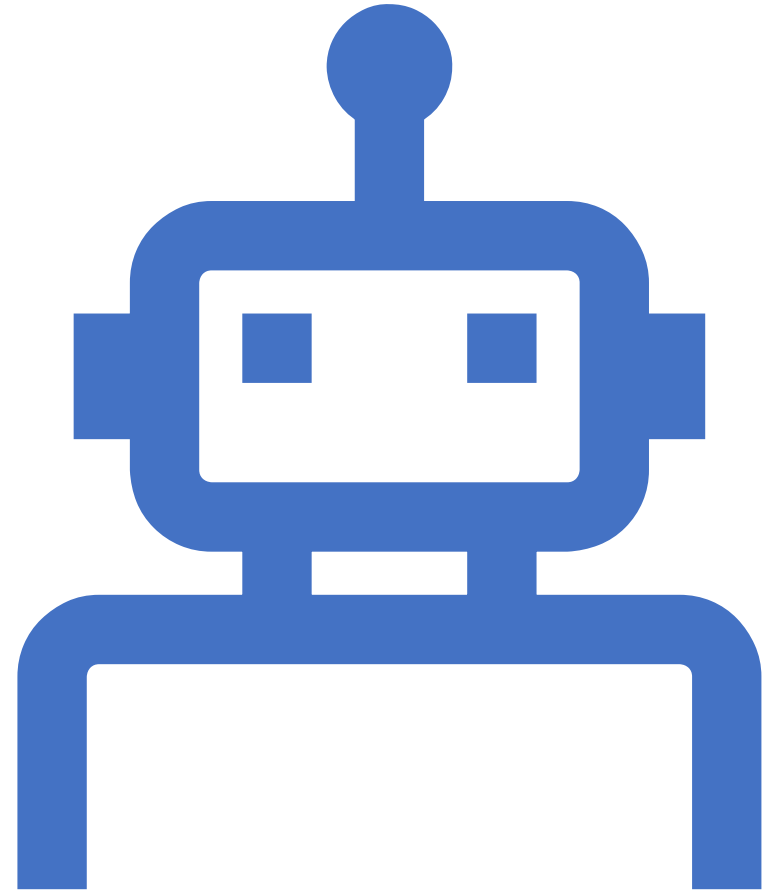


Kedro.org



Kedro adalah framework atau kerangka yang digunakan untuk mengatur alur jalannya kode dari data science

- Catalog: tempat menyimpan definisi data yang akan digunakan dalam proyek
- Node: tempat definisi fungsi, library, class dan komponen kode lainnya
- Pipeline: tempat untuk mengatur input output, parameter, serta fungsi apa yang digunakan untuk pekerjaan tertentu
- Pipeline registry: tempat untuk mengatur alur pipeline



Demo
kedro

- Terima kasih