



CSGE603130 • Kecerdasan Artifisial dan Sains Data Dasar
Semester Genap 2021/2022
Fakultas Ilmu Komputer, Universitas Indonesia

Lab 1: Introduction to Data Science Tools
Tenggat Waktu: 19 Februari 2022, 23.55 WIB

Ketentuan:

1. Dokumen *template* lab dengan format .ipynb dan dataset (jika dibutuhkan) telah disediakan di SCeLe
2. Jalankan kode pada dokumen .ipynb dan perhatikan dengan saksama apa yang potongan kode tersebut lakukan beserta dengan keluarannya. Jawablah **pertanyaan** yang disisipkan pada potongan program yang diberikan.
3. Dokumen Jupyter Notebook yang telah dilengkapi dengan jawaban dikumpulkan dengan format penamaan **Kelas_LabX_NPM_Nama.ipynb**
Contoh: F_Lab1_1706979341_Lulu Ilmaknun Qurotaini.ipynb
4. Kumpulkan dokumen tersebut pada submisi yang telah disediakan di SCeLe sesuai dengan kelas masing-masing sebelum **XX Februari 2022, 23.55 WIB**. Keterlambatan pengumpulan akan dikenakan penalti.
5. Lab ini dirancang sebagai **tugas mandiri**. Plagiarisme tidak diperkenankan dalam bentuk apapun. Adapun kolaborasi berupa diskusi (tanpa menyalin maupun mengambil jawaban orang lain) dan literasi masih diperbolehkan dengan mencantumkan kolaborator dan sumber.

A. Jupyter Notebook

Pada mata kuliah Kecerdasan Artifisial dan Sains Data Dasar ini, kita akan menggunakan berkas *notebook* IPYNB untuk melakukan pemrosesan dan analisis data. Berkas IPYNB adalah dokumen yang digunakan pada **Jupyter Notebook** yang berisi *environment* interaktif untuk pemrograman menggunakan bahasa Python. Satu berkas IPYNB terdiri atas satu atau lebih cell yang dapat digunakan untuk menjalankan kode secara interaktif maupun menulis catatan dalam sintaks *markdown*.

Jupyter Notebook adalah aplikasi *open-source* yang digunakan untuk membuat dokumen *notebook* yang dapat berisi kode pemrograman, visualisasi, maupun teks. Aplikasi ini dapat digunakan dalam bidang sains data untuk melakukan berbagai pekerjaan, seperti membersihkan data, transformasi

data, memodelkan statistik data, membuat visualisasi data, menerapkan *machine learning*, dan masih banyak lagi.

Selain Jupyter Notebook, berkas IPYNB dapat dibuka dan dibuat menggunakan teknologi *cloud* (akses melalui internet), seperti [Google Colab](#) atau [Kaggle](#). Aplikasi tersebut memiliki fungsionalitas dasar yang sama dengan Jupyter Notebook. Namun, juga memiliki fitur-fitur tersendiri, seperti penyimpanan pada *cloud storage* (ex: Google Drive), dukungan kolaborasi, dan akses terhadap mesin komputasi yang mumpuni. Berikut adalah perbandingan antara aplikasi-aplikasi tersebut.

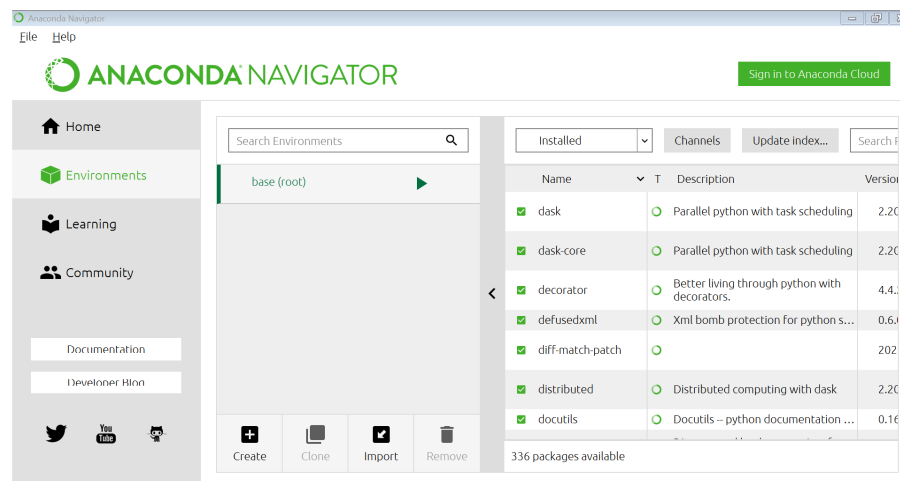
Aplikasi	Keunggulan	Kekurangan
Jupyter Notebook	<ul style="list-style-type: none">• Tidak memerlukan koneksi internet untuk membuka <i>notebook</i>.• Akses langsung ke <i>workspace</i> lokal.• Bisa dipasang beberapa extension tambahan.	<ul style="list-style-type: none">• Performa dibatasi oleh prosesor perangkat sehingga sulit melakukan komputasi besar (seperti <i>deep learning</i>) jika prosesor tidak memadai.• Perlu meng-<i>install</i> aplikasi Jupyter Notebook pada perangkat.• Perlu mengatur dan melakukan <i>update</i> pada <i>environment</i> lokal secara manual.• Tidak dapat digunakan secara kolaboratif.
Google Colab	<ul style="list-style-type: none">• Dapat diakses dari perangkat apapun, selama memiliki <i>web browser</i> dan koneksi internet.• <i>Notebook</i> tersimpan di Google Drive sehingga dapat diakses dan dibagikan dari dan ke mana pun.• Fitur <i>sharing</i> yang mudah digunakan, seperti pada Google Docs.• Mendapatkan akses GPU atau prosesor lain yang cukup mumpuni, bahkan pada versi gratis, sehingga mempermudah melakukan komputasi besar.• Fitur <i>version control/history</i> seperti pada Google Docs yang mempermudah penelusuran sejarah kolaborasi.• Beberapa <i>library</i>, seperti NumPy, Pandas, dan Matplotlib, telah tersedia	<ul style="list-style-type: none">• Responsivitas tampilan dipengaruhi oleh koneksi internet.• Tidak dapat diakses tanpa koneksi internet.• Akan kena <i>timeout</i> jika <i>idle</i> terlalu lama.

	<p>secara bawaan.</p> <ul style="list-style-type: none">• Fitur komentar/catatan seperti pada Google Docs.• Fitur <i>auto-save</i> selama terhubung dengan internet.	
--	---	--

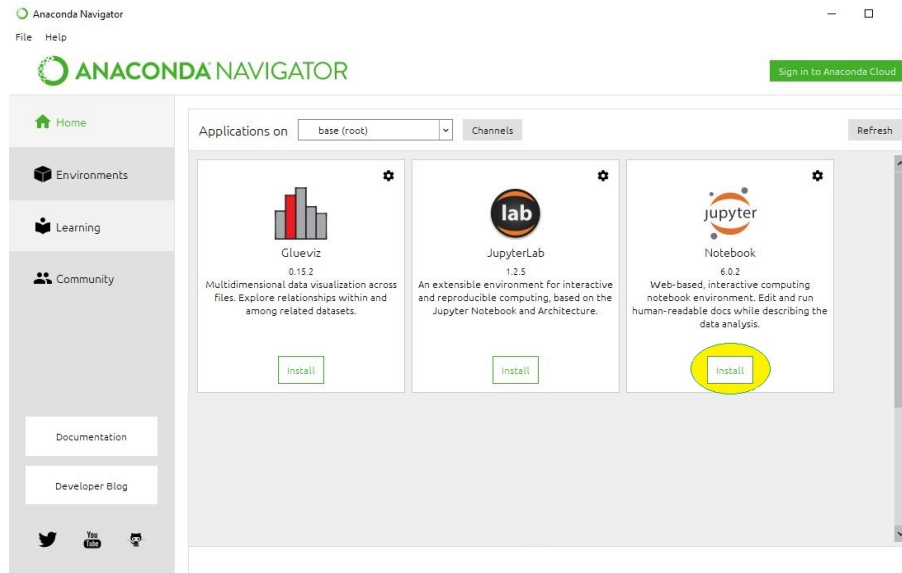
B. Instalasi Jupyter Notebook

Salah satu cara mudah untuk mengunduh dan menggunakan Jupyter Notebook adalah dengan menggunakan **Anaconda**. Kelebihan dari Anaconda adalah *library* yang sering digunakan pada tutorial mata kuliah ini (atau *data science* secara umum) sudah tersedia, seperti **Jupyter Notebook**, **Pandas**, **NumPy**, dan **Scikit-learn**, sehingga tidak diperlukan instalasi tambahan. Secara sederhana penggunaan Anaconda untuk Jupyter Notebook adalah sebagai berikut:

- Unduh Anaconda sesuai dengan sistem operasi pada link berikut <https://www.anaconda.com/products/individual>
- Lakukan langkah-langkah instalasi sesuai dengan perintah setiap langkah
- Jika ingin menggunakan *application interface* (umumnya di sistem operasi Windows), dapat langsung membuka aplikasi yang sudah di-*install*. Pada *interface* aplikasi, *environment* dan *library* dapat diatur pada tab “Environments”.



Kita juga dapat membuka Jupyter Notebook pada tab “Home” dengan memastikan *library jupyter* sudah di-*install*



- Jika ingin menggunakan *command line interface*, dapat menuju aplikasi Anaconda Prompt dan menggunakan beberapa *command*. Jika menggunakan terminal (bukan melalui Anaconda Prompt), gunakan *command* `conda activate` untuk aktivasi *environment* `conda`. Untuk dokumentasi lebih lengkap mengenai *command* yang digunakan dapat mengunjungi docs.anaconda.com atau [conda cheat sheet](#). Berikut beberapa *command* sederhana yang dapat digunakan.

- Untuk melihat library yang sudah di-*install*, gunakan *command*:

```
conda list
```

- Untuk melakukan instalasi package `jupyter`, gunakan *command*:

```
conda install jupyter
```

- Untuk menjalankan Jupyter Notebook setelah instalasi, gunakan *command*:

```
jupyter notebook
```

Alternatif lainnya, kita dapat mengikuti langkah-langkah instalasi **Jupyter Notebook menggunakan pip** sebagai berikut:

1. Instalasi **pip**

- a. Periksa apakah **pip** telah terpasang pada komputer dengan menjalankan perintah **pip** apapun, misal seperti berikut pada *command line interface*:

```
pip --version
```

*Jika **pip** belum terpasang maka akan muncul informasi bahwa perintah tersebut tidak dikenali.

- b. Unduh berkas **get-pip.py** dari <https://bootstrap.pypa.io/get-pip.py>.
- c. Buka *command line interface* pada direktori/folder tempat **get-pip.py** berada dan jalankan perintah berikut:

```
python get-pip.py
```

*Jika perintah **python** juga tidak dikenali oleh CLI maka **python** perlu diunduh dan di-*install* terlebih dahulu. Silakan merujuk pada [petunjuk platform masing-masing](#).

- d. Periksa keberhasilan instalasi dengan perintah **pip** seperti pada butir 1a.
2. Instalasi **Jupyter Notebook** menggunakan **pip**
 - a. Jika **pip** telah terpasang, jalankan perintah berikut pada *command line interface*:

```
pip install jupyter
```
 - b. Luncurkan **Jupyter Notebook** dengan menjalankan perintah berikut pada *command line interface*:

```
jupyter notebook
```
- *Aplikasi akan terbuka sebagai *tab* baru pada *web browser*.
3. Cara instalasi lain dapat dibaca pada [dokumentasi JupyterLab](#).

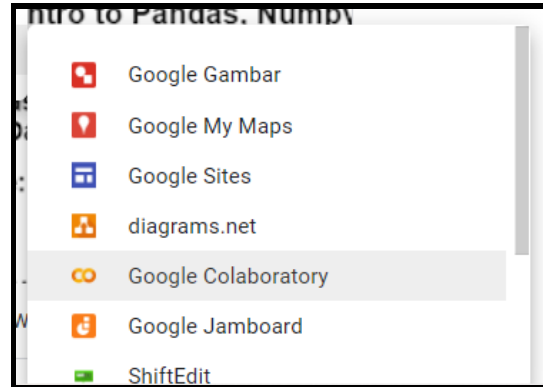
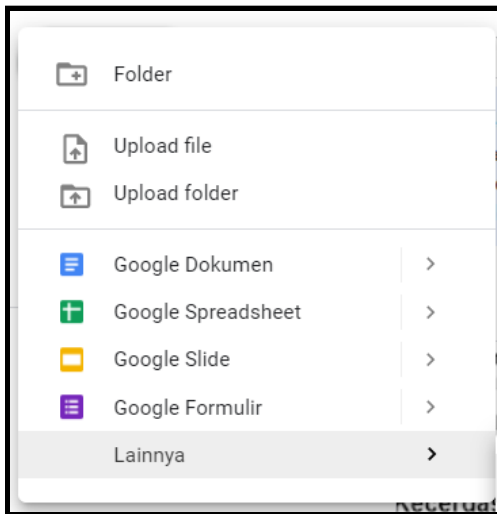
C. Google Colab

Opsi lain untuk membuat dan mengedit berkas IPYNB adalah dengan menggunakan aplikasi berteknologi cloud. Salah satu aplikasi yang umum digunakan adalah Google Colab.

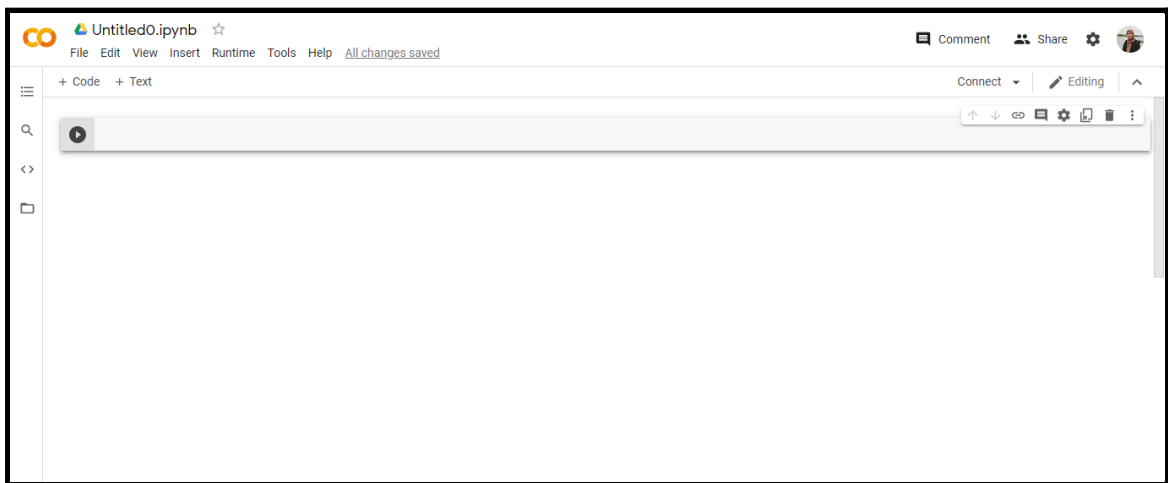
Berikut adalah langkah-langkah menggunakan Google Colab:

1. Membuat berkas IPYNB

Masuk ke [Google Drive](#), klik “Baru” → klik “Lainnya” → klik “Google Colaboratory”



Akan muncul tampilan berkas IPYNB pada Google Colab seperti berikut:

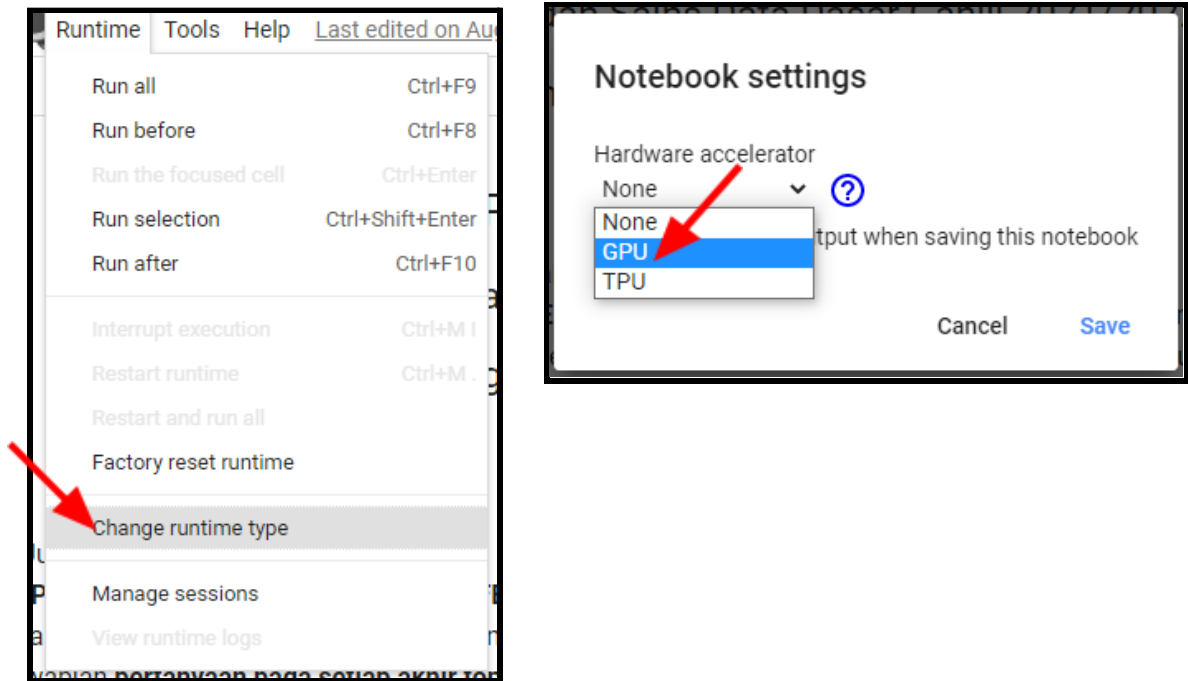


2. Mengubah tipe *runtime* (prosesor)

Saat proses *training* model *machine learning* terkadang dibutuhkan prosesor yang mumpuni seperti GPU agar pembuatan model dapat dieksekusi lebih cepat. Google Colab menyediakan tiga macam prosesor, yaitu CPU, GPU, dan TPU secara gratis dengan total waktu eksekusi selama 12 jam berturut-turut, jika lebih dari itu perlu dilakukan *restart runtime* dan menjalankan ulang kode. Fitur ini merupakan salah satu kelebihan yang tidak dimiliki oleh Jupyter Notebook lokal.

Berikut cara mengubah tipe prosesor:

Klik "*runtime*" → "*change runtime type*" → pilih tipe runtime yang dibutuhkan → klik "*Save*".

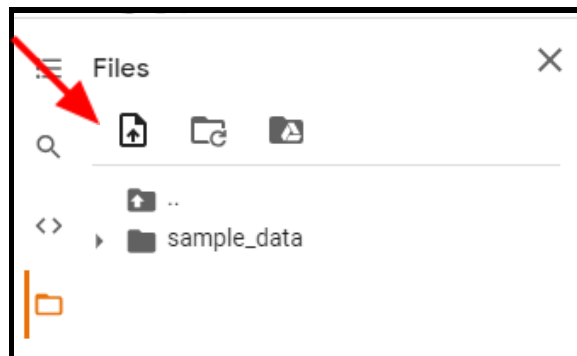


3. Mengunggah berkas

Untuk mengunggah berkas dapat dilakukan salah satu cara berikut:

- Melalui “*Upload to session storage*”

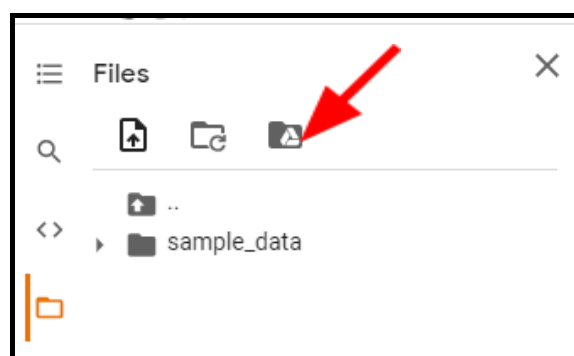
Klik ikon “*files*” → ikon “*Upload to session storage*” → pilih berkas yang dibutuhkan

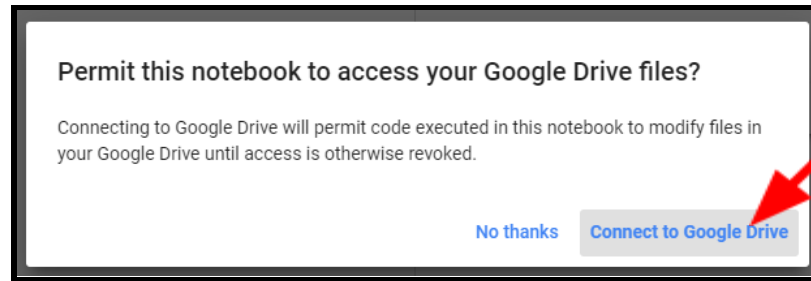


Dengan cara ini, berkas yang diunggah hanya akan tersedia hingga akhir *session*.

- Melalui “*Mount Drive*”

Klik ikon “*files*” → klik ikon “*Mount Drive*” → klik “*Connect to Google Drive*”

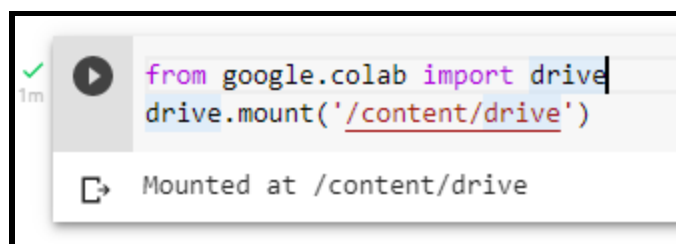




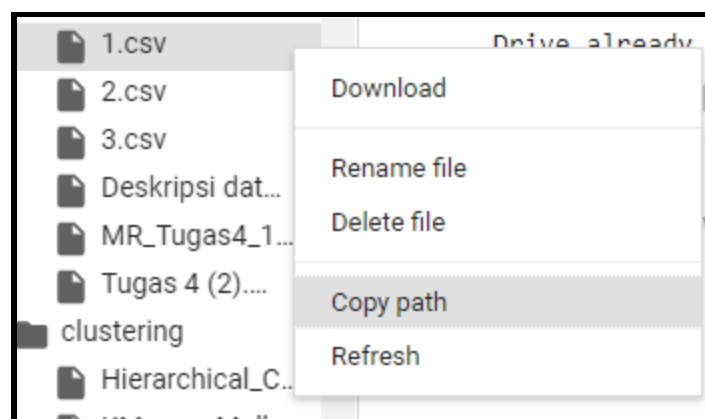
run kode yang muncul → klik link pada *output* → salin kode otorisasi → tempelkan kode pada *text box* → tekan *enter*.



Google Colab berhasil terhubung ke Google Drive.



Pada tab "*files*" akan muncul folder "*drive*" → pilih berkas yang dibutuhkan dari folder "*drive*" → klik kanan → klik "*Copy path*".



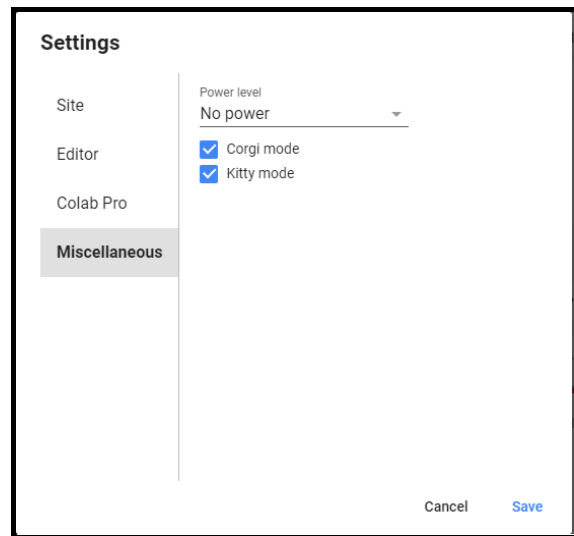
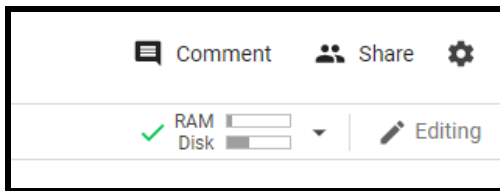
Path yang telah disalin, ditempel pada kode untuk membaca berkas.

4. [BONUS] Mengaktifkan Mode Corgi atau Kitty

Kita dapat mengaktifkan mode ini sebagai hiburan saat mengedit *notebook* di Google Colab.

Berikut langkahnya:

Klik ikon “*settings*” → klik “*Miscellaneous*” → centang “*Corgi mode*” atau “*Kitty mode*” → klik “*Save*”.



Mode Corgi atau Kitty akan aktif seperti berikut:



D. Pengenalan Pandas, Numpy, dan Scikit-learn

Silakan buka berkas **Lab1.ipynb** menggunakan Jupyter Notebook, Google Colab, atau aplikasi lain. Jalankan setiap *cell* untuk mencoba kode yang disediakan. Silakan modifikasi kode untuk eksplorasi mandiri.

Setelah selesai mencoba kode-kode yang ada, pada setiap bagian akhir topik terdapat beberapa pertanyaan yang perlu dijawab. Jawab pertanyaan pada *cell* di bawah setiap pertanyaan.

Catatan: Pada tutorial mata kuliah ini, kita akan menggunakan bahasa pemrograman Python. Jika kamu butuh penyegaran kembali mengenai Python, kamu dapat mengunjungi channel Youtube pada link berikut: <https://bit.ly/pymooc-id>. Selamat belajar!

- Selamat belajar dan mengerjakan! -