

Lab 1

Sains Data Genap 2020/2021

Deadline: 28 Februari 2021, 23.55 WIB

Petunjuk umum:

1. Buka dokumen Jupyter Notebook **Lab1.ipynb** yang telah disediakan di scele.
2. Lengkapi data **NPM**, **KOLABORATOR** (jika ada), dan **REFERENSI** (jika ada), pada *cell* di bawah Petunjuk Umum.
3. Jalankan kode yang telah disediakan dan perhatikan dengan saksama apa yang potongan kode tersebut lakukan beserta dengan keluarannya. Jawablah **pertanyaan pada bagian akhir dokumen** sesuai dengan perintah yang diberikan.
4. Dokumen Jupyter Notebook yang telah dijawab dikumpulkan dengan format penamaan **KodeAsdos_Lab1_NPM_Nama.ipynb**
Contoh: LI_Lab1_1706979341_Lulu Ilmaknun Qurotaini.ipynb
5. Kumpulkan pada submisi yang telah disediakan di scele sebelum **28 Februari 2021, 23.55 WIB**. Keterlambatan pengumpulan akan dikenakan pinalti.

A. Jupyter Notebook

Pada mata kuliah Sains Data 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. 1 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 aplikasi daring, seperti [Google Colab](#). Aplikasi tersebut memiliki fungsionalitas dasar yang sama dengan Jupyter Notebook, namun juga memiliki fitur-fitur tersendiri, seperti penyimpanan pada *cloud storage*, 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 	<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 secara bawaan. • Fitur komentar/catatan seperti pada Google Docs. • Fitur <i>auto-save</i> selama terhubung dengan internet. 	<ul style="list-style-type: none"> • Responsivitas tampilan dipengaruhi oleh koneksi internet. • Tidak dapat diakses tanpa koneksi internet.

B. Instalasi Jupyter Notebook

Berikut adalah langkah-langkah instalasi **Jupyter Notebook** secara umum:

1. Instalasi **pip**

- 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.

- Unduh berkas **get-pip.py** dari <https://bootstrap.pypa.io/get-pip.py>.
- 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](#).

- Periksa keberhasilan instalasi dengan perintah **pip** seperti pada butir 1a.

2. Instalasi **Jupyter Notebook** menggunakan **pip**

- Jika **pip** telah terpasang, jalankan perintah berikut pada *command line interface*:

```
pip install jupyter
```

- 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. Pengenalan Pandas

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

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

- Selamat belajar dan mengerjakan! -