

Menggunakan Aplikasi Data Science (Tableau and Anaconda)



William Yohanes Sirait, S.Si

- Pengenalan tentang “Tableau”
- Keunggulan dari “Tableau”
- Cara menginstall aplikasi “Tableau” dan “Anaconda”
- Memahami tools dan button pada aplikasi “Tableau”

Pengenalan tentang Tableau



Tableau merupakan *interactive data visualization* yang berdiri sejak tahun Januari 2003 di Seattle, Washington, Amerika Serikat yang berfokus pada *business Intelligence*. Tableau memiliki berbagai software products seperti Tableau Desktop, Tableau Server, Tableau Online, Tableau Reader dan Tableau Mobile. Tableau dapat digunakan untuk memvisualisasikan data dengan berbagai jenis seperti data histogram, data spasial, data kategorik, data numerik dengan sangat mudah, hanya dengan “drag and drop”.



Keunggulan Tableau

1. Sifatnya open source untuk keperluan akademis.
2. Dapat digunakan untuk mencleaning data secara otomatis
3. Terintegrasi dengan berbagai jenis data (excel, pdf, json, Sql, ms.access, dll).
4. Dapat terintegrasi dengan perangkat data science lain seperti “Phyton, R, dll”
5. Tampilannya menarik dan tidak rumit penggunaannya (Drag and Drop).



Instalasi Tableau

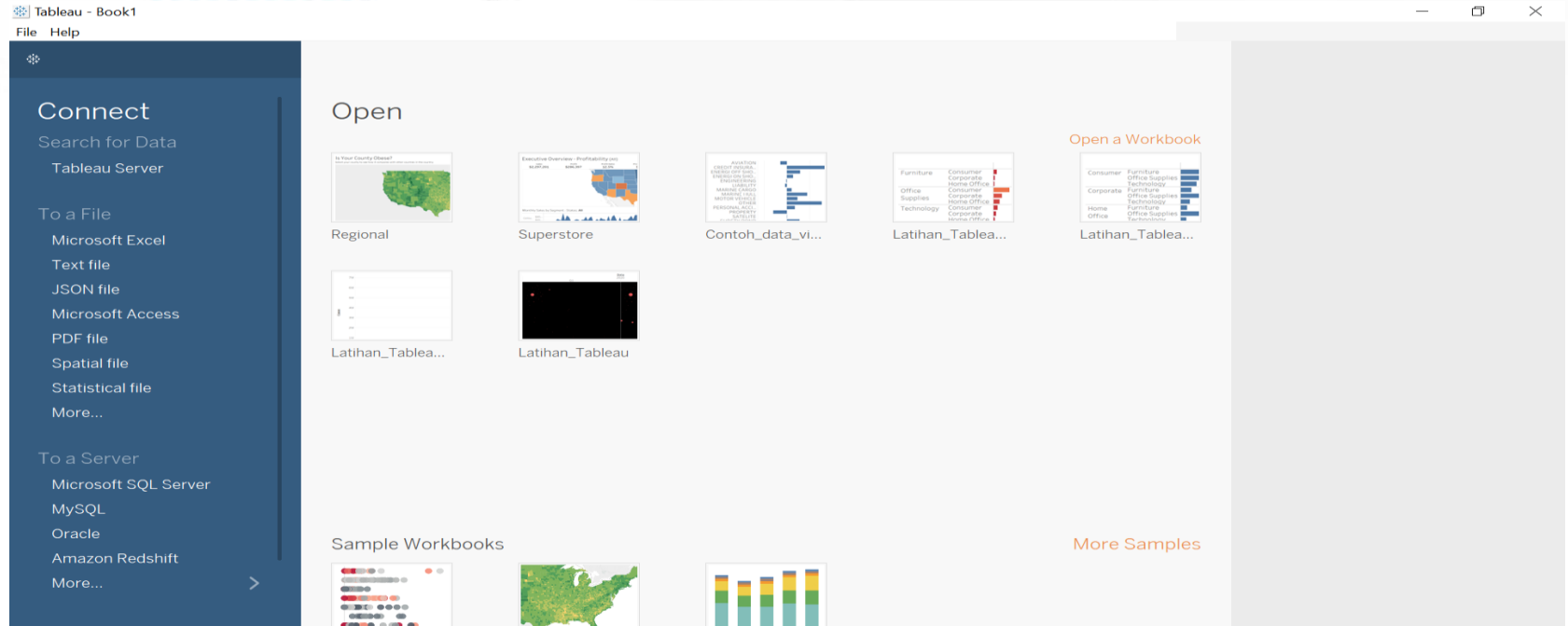
1. Buka Website <https://www.tableau.com/products/desktop/download>
2. Mempunyai email sekolah/lembaga pendidikan contoh (William.sirait@kanaan.sch.id)
3. <https://www.tableau.com/academic/students> (lalu pilih free student license)
4. Isi data diri dengan (“Country/Region, Personal Information, dan School Information) lalu “Verify Student Status” lewat email sekolah kalian



Instalasi “Anaconda”

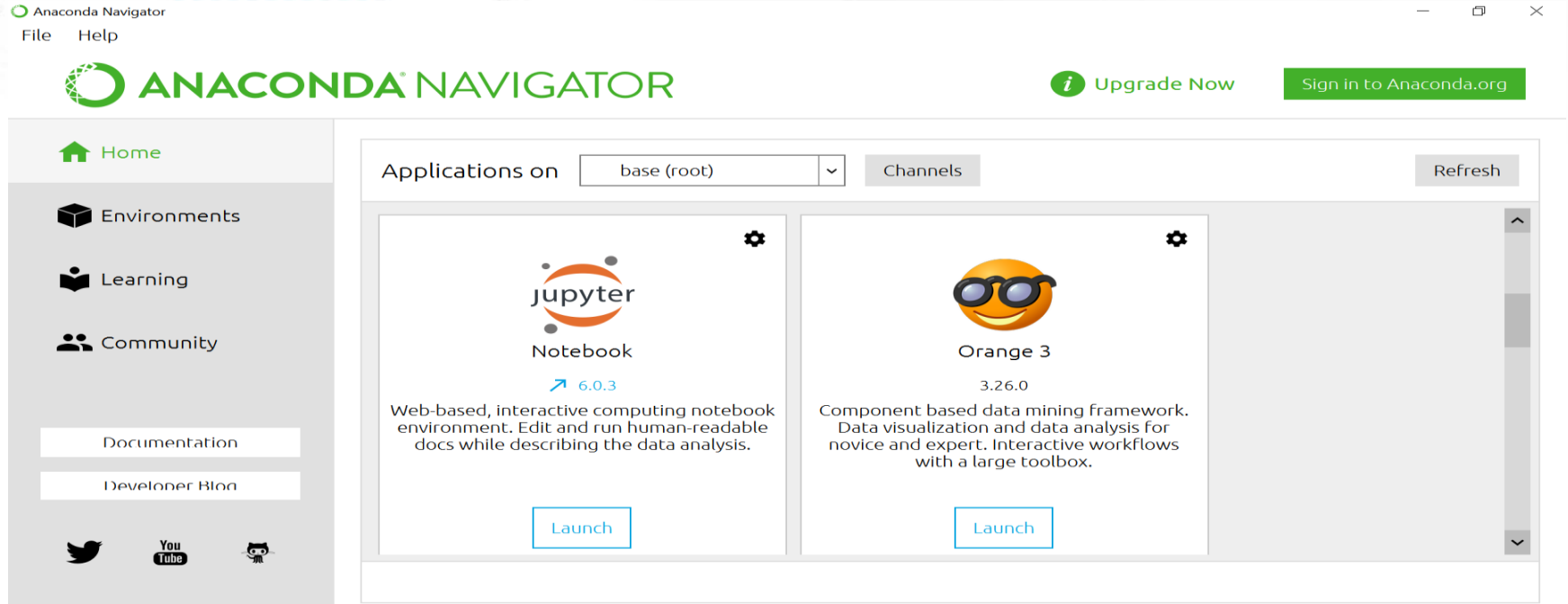
1. Buka web <https://www.anaconda.com/>
2. <https://www.anaconda.com/products/individual#Downloads> (pilih Sistem Operasi yang kalian pakai, Microsoft, Linux, atau MacOS)
3. Lalu launch pada “Interface Anaconda” pilih python.

Interface Tableau



Gambar. Interface awal Tableau

Interface Anaconda



Gambar. Interface awal "Anaconda"



Tableau Public - Book1

File Data Window Help

Connections

pelatihan_tab...covid19_dunia
Microsoft Excel

Sheets

Use Data Interpreter
Data Interpreter might be able to clean your Microsoft Excel workbook.

Sheet1

New Union

Sheet1 (pelatihan_tableau_covid19_dunia)

Filters
0 Add

Sheet1

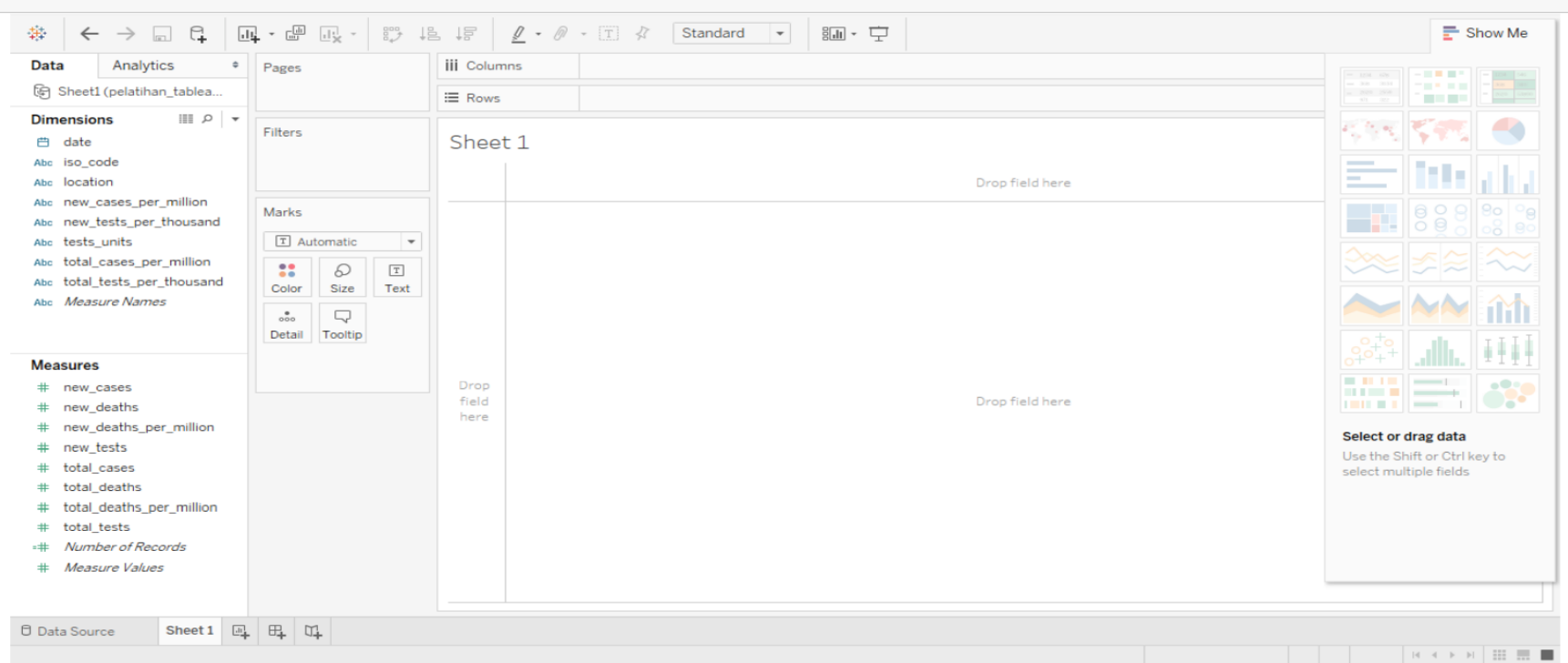
Sort fields Data source order

Show aliases Show hidden fields 1,000 rows

Sheet1 iso_code	Sheet1 location	Sheet1 date	Sheet1 total_cases	Sheet1 new_cases	Sheet1 total_deaths	Sheet1 new_deaths	Sheet1 total_cases_per_m...	Sheet1 new_cases_per_m...
ABW	Aruba	13/3/2020	2	2	0	0	18733.0	18,733
ABW	Aruba	20/3/2020	4	2	0	0	37465.0	18,733
ABW	Aruba	24/3/2020	12	8	0	0	112395.0	74,93
ABW	Aruba	25/3/2020	17	5	0	0	159227.0	46,831
ABW	Aruba	26/3/2020	19	2	0	0	177959.0	18,733
ABW	Aruba	27/3/2020	28	9	0	0	262256.0	84,296
ABW	Aruba	28/3/2020	28	0	0	0	262256.0	0
ABW	Aruba	29/3/2020	28	0	0	0	262256.0	0

Data Source Sheet1

Gambar. Tampilan Data Source Tableau



Gambar. Tampilan Workspace Tableau



Data Pane **Analytic Pane** **View Cards**

Tableau Public - Book1
File Data Worksheet Dashboard Story Analysis Map Format Window Help

Data Analytics

Sheet1 (pelatihan_tablea...)

Dimensions

- date
- iso_code
- location
- new_cases_per_million
- new_tests_per_thousand
- tests_units
- total_cases_per_million
- total_tests_per_thousand
- Measure Names

Measures

- new_cases
- new_deaths
- new_deaths_per_million
- new_tests
- total_cases
- total_deaths
- total_deaths_per_million
- total_tests
- Number of Records
- Measure Values

Filters

Marks

Automatic

Color Size Text

Detail Tooltip

Columns

Rows

Sheet 1

Drop field here

Drop field here

Drop field here

Drop field here

Show Me

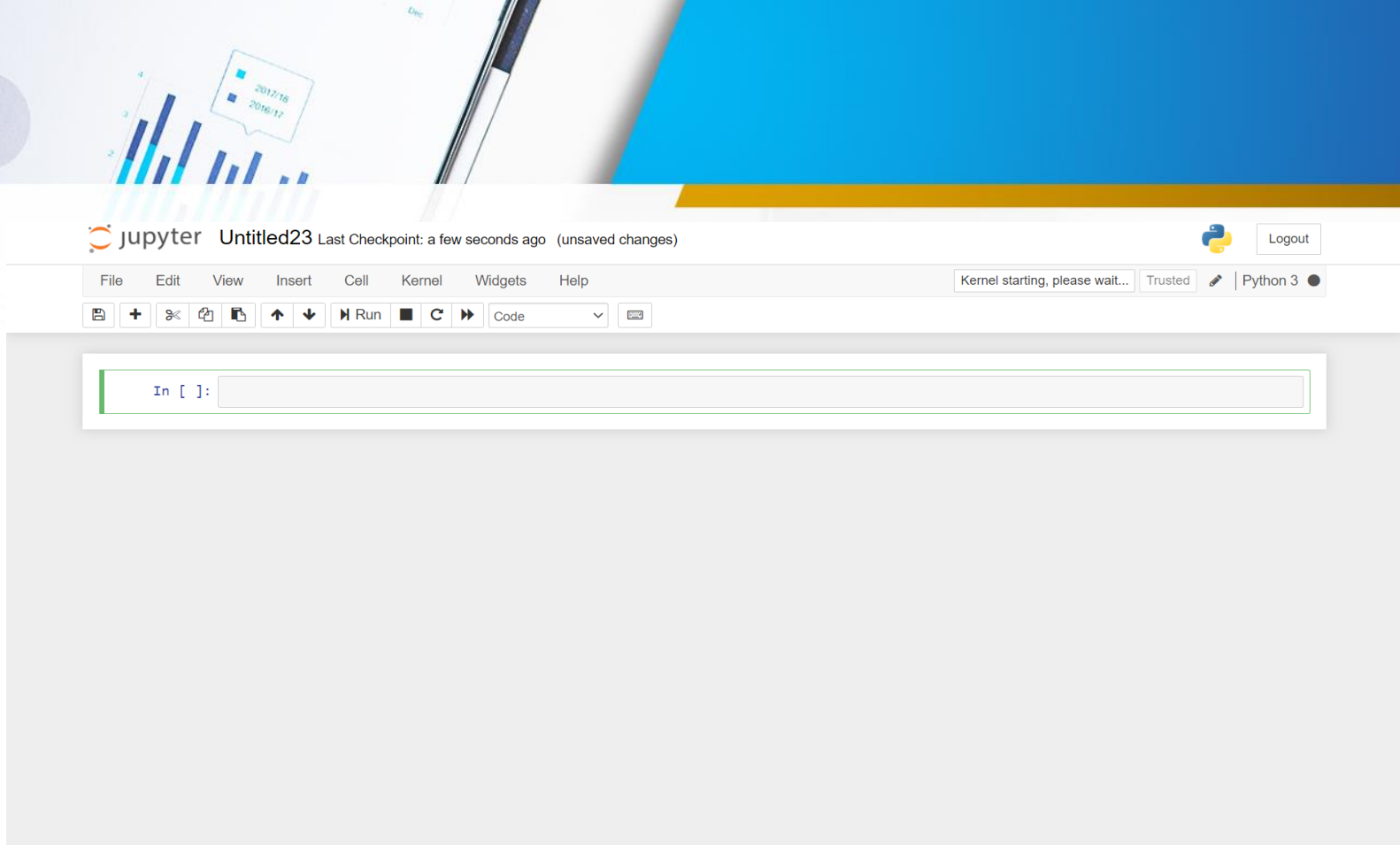
Select or drag data
Use the Shift or Ctrl key to select multiple fields

Data Source Sheet 1 Dashboard

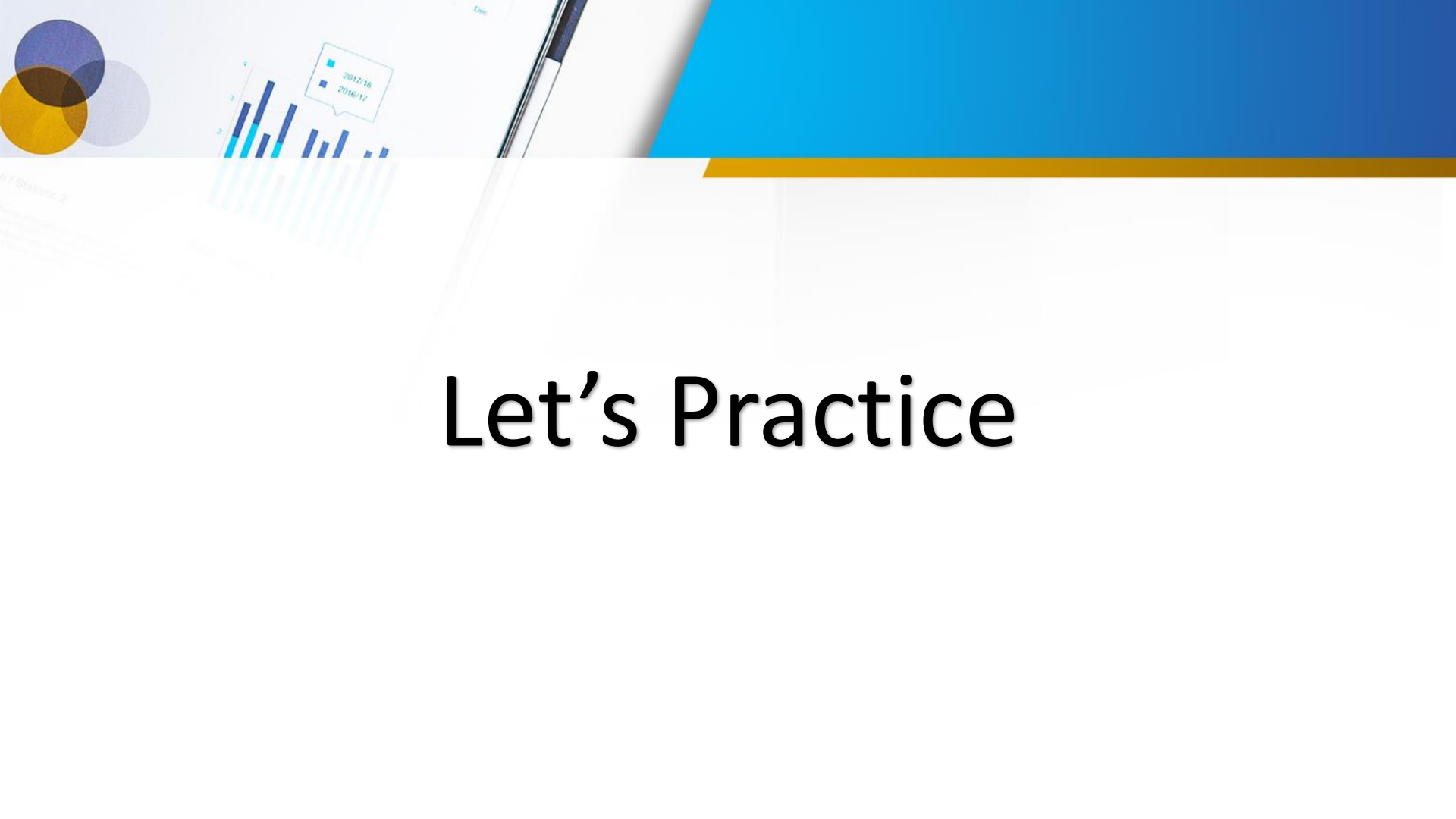
Worksheet Tabs



Term	Deskripsi
Start Page	Tombol untuk kembali ke halaman awal Tableau
Data Pane	Berisi Dimensions dan Measures dari data yang diupload ke dalam halaman utama
Analytics Pane	Terdapat berbagai pilihan analisis yang dapat digunakan untuk membuat trend lines, dan lain sebagainya
Workbook Name	Nama workbook
View Cards	Digunakan untuk memodifikasi workbook
Toolbar Icon	Akses cepat ke fitur favorit
Worksheet	Sheet yang digunakan untuk membuat visualisasi
Data Source	Kembali ke sumber data
Worksheet Tabs	Melihat worksheet, dashboard, atau story board
Dashboard	Untuk membuat dashboard visualisasi data



Gambar. Tampilan awal “Notebook” di Anaconda



Let's Practice



Tujuan Modeling Data : ingin melihat suatu penyebaran Kasus Data COVID-19 di suatu wilayah/negara


Metodologi Data Science


a. Business Understanding :

- Pengetahuan COVID-19, Proses Penularan COVID-19, Pemahaman akan geografis dan karakteristik suatu wilayah dan hubungannya dengan banyak/tidaknya penderita COVID-19.
- Data ini dibutuhkan oleh WHO, Negara, suatu pemerintah Provinsi (suatu wilayah).

b. Data Requirements :

- Sumber data dari <https://www.kaggle.com/sudalairajkumar/novel-corona-virus-2019-dataset>
- Pembuat data : sudalairajkumar
- Institusi : Johns Hopkins University untuk tujuan pendidikan dan penelitian
- Update Data : 22 January 2020

- 
- Data Collection :
 - Case Type (terkonfirmasi +, mengalami kematian)
 - Cases (Jumlah kasus)
 - Country Region (Negara)
 - Province State (Provinsi suatu Wilayah)
 - Iso2/Iso3 (Kode Abjad Suatu wilayah)
 - Population (Jumlah populasi dari suatu negara)
 - Lat (garis lintang)
 - Longitude (garis bujur)
 - Prep Flow Runtime(jadwal pengambilan data terakhir)

- 
- Data Preparation :
 - Membersihkan data yang memiliki entry data (Null, N/A, error) pada bagian kolom (People total Husted, Province State, People Hostipitalized Cummmulative Count).
 - Modeling Data
 - (Menggunakan kolom Country Region, Population, Cases, Lattitude, Longitude).



Challenge

1. Instalasi Aplikasi “Tableau” dan “Anaconda” di Komputer kalian
2. Buatlah sebuah model data yang tersedia di grup lalu buat model data dengan aplikasi “tableau” sesuai dengan konsep metodologi data science (Business Understanding, Data Requirements, Data Preparation, Data Collection, Data Preparation, Modeling Data) lalu kumpulkan di grup.



Reference

Satria, Dias. 2020. *Visulisasi Data dan Data Analitik dengan Tableau*. Malang: Brawijaya University