**Crawling Data dengan Selenium**

* Hal yang dibutuhkan:

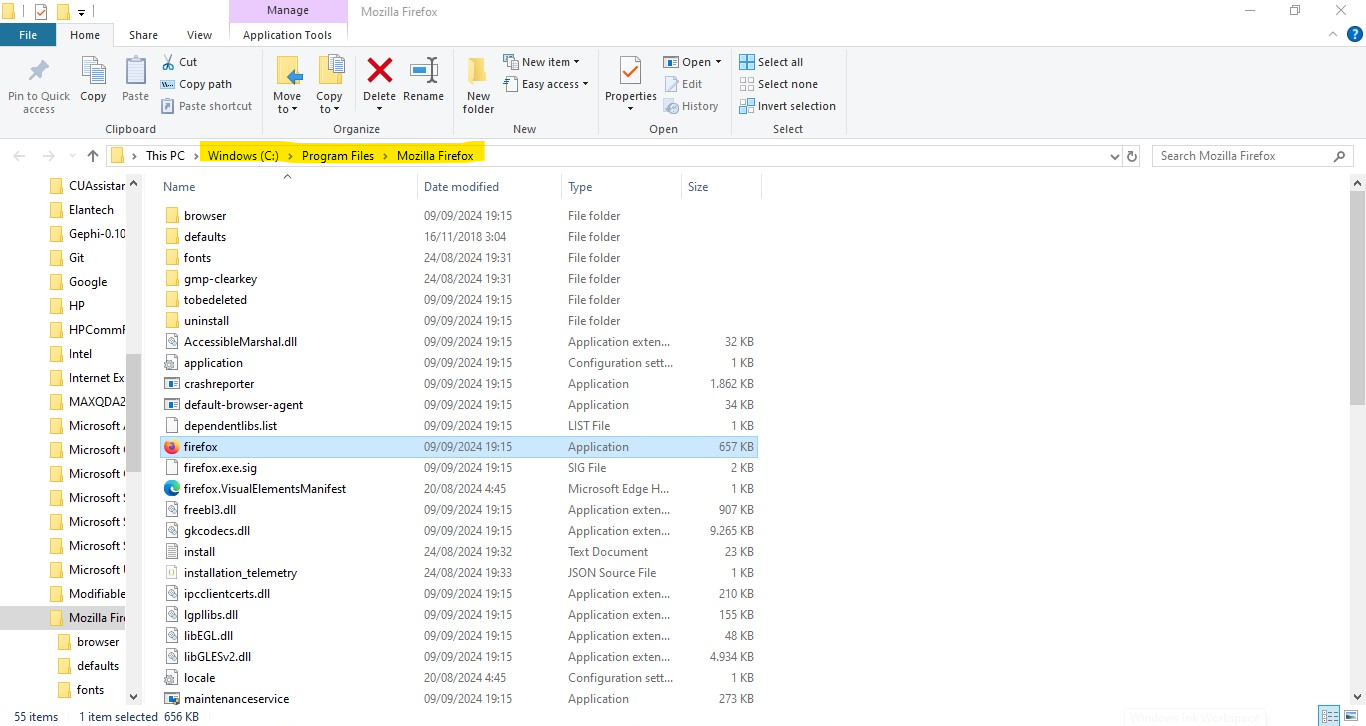
1. Jupyter Notebook
2. Web browser Mozilla Firefox versi terbaru.
3. Driver firefox yakni geckodriver.exe versi terbaru.
4. Akun Twitter aktif.

* Install Jupyter Notebook:

1. Download Python disini : <https://www.python.org/downloads/>
2. Install Python seperti biasa pada device anda. Mohon diperhatikan, sebelum klik `Install Now` mohon ceklis `Add pyhton.exe to PATH` terlebih dahulu, kemudian klik install.
3. Setelah Python terinstall, buka CMD dengan cara menekan pada keyboard `window + R` kemudian ketik cmd.
4. Jika sudah masuk CMD, ketik `pip install jupyter notebook` + enter, tunggu sampai penginstalan sukses.
5. Kemudian close cmd, dan buka lagi cmd yang baru.
6. Lanjut, ketik `python -m notebook` + enter, untuk membuka jupyter notebook. Kemudian akan diarahkan ke browser. Jika tampilannya berupa Home Page Jupyter, maka penginstalan sukses.

* Mendowload Mozilla Firefox & Geckodriver:
* Mozilla Firefox:

1. Download Mozilla Firefox versi terbaru disini : <https://www.mozilla.org/id/firefox/new/>
2. Install seperti biasa saja. Jika masih menggunakan versi terdahulu, mohon install ulang yang baru.
3. Jika sudah, klik kanan icon Firefox pada desktop, kemudian klik `Open file location`.
4. Nanti akan di arahkan kesini :



Disini yang ditandai kuning itu adalah alamat path firefox.exe yang tersimpan. Alamat ini akan dipanggil pada saat penarikan data pada code, maka dari itu penting untuk mengingat alamat path ini. Disini untuk alamat path nya pada gambar : `C:\\Program Files\\Mozilla Firefox\\firefox.exe`. Penting untuk menyesuaikan pada device masing-masing.

* Geckodriver :

1. Download geckodriver atau driver firefox disini : <https://github.com/mozilla/geckodriver/releases> (download yang versi terbaru)
2. File yang setalah di download akan berbentuk ZIP, maka dari itu ekstrak terlebih dahulu sampai hasilnya `geckodriver.exe`.
3. Kemudian sama seperti firefox.exe, geckodriver ini akan dipanggil didalam code. Maka dari itu penting untuk mengingat alamat path nya. Sebagai contoh, disini disimpan `D:\\geckodriver.exe`. Namun ini bisa disimpan dimana saja, asal nanti saat memanggil nya pada code alamat pathnya sesuai. Untuk mencari alamatnya ketentuannya sama seperti gambar pada saat mencari firefox.exe, yang bertanda kuning. *Catatan: untuk file exe geckocriver nya, tidak perlu diinstall. Cukup di ekstraksi saja di device anda.*

* Twitter:

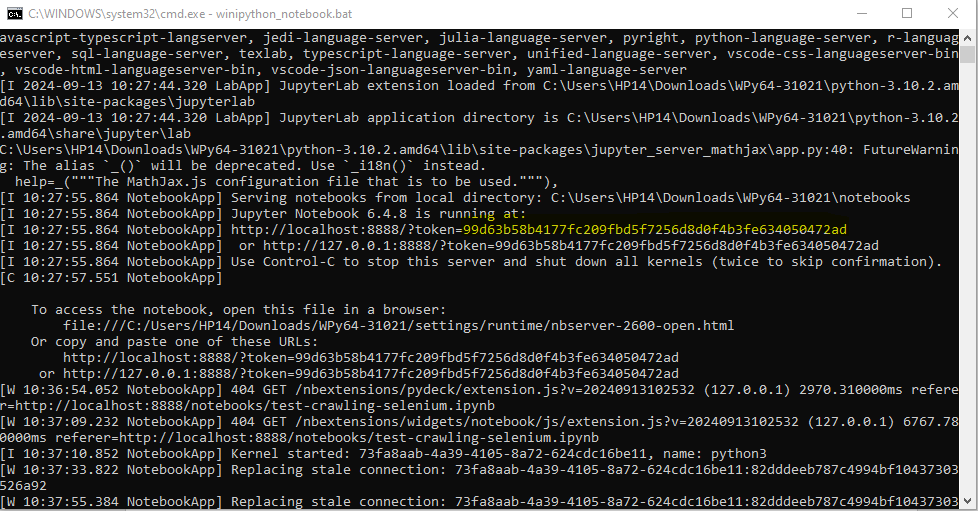
1. Buka https://twitter.com/login
2. Klik `daftar` jika belum punya akun, jika sudah ada klik `login`.
3. Kemudian jika sudah masuk, klik `Jelajahi` atau `Explore`. Disini kita cari kata kunci yang akan kita tarik datanya. Sebagai contoh disini akan mencari keyword `makan siang gratis`. Maka kita ketik di perncarian nya dengan format seperti ini `makan siang gratis lang:id until:2024-09-12 since:2024-01-01`. Yang artinya, mencari dengan keyword `makan siang gratis`, berbahasa `Indonesia`, dalam rentang waktu dari yang terbaru 12 Sepetember 2024 sampai yang terlampau 1 Januari 2024. Kemudian salin link web nya, contohnya seperti ini: `https://x.com/search?q=makan%20siang%20gratis%20lang%3Aid%20until%3A2024-09-12%20since%3A2024-01-01&src=typed\_query` ini untuk pencarian yang Top atau Populer. Jika ingin mencari dari yang Latest maka bisa di klik dulu `latest` nya kemudian salin link web nya, seperti ini: `https://x.com/search?q=makan%20siang%20gratis%20lang%3Aid%20until%3A2024-09-12%20since%3A2024-01-01&src=typed\_query&f=live`.

Untuk yang versi latest biasanya akan mendapatkan banyak data daripada versi yang top.

1. Simpan juga alamat link ini seperti alamat path sebelumnya karena ini akan dipanggil dalam code.

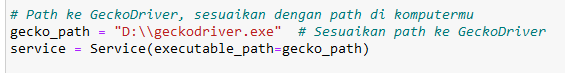
* Langkah Penarikan Data :

1. Pastikan bahan nya sudah terinstall dan ada (seperti alamat path dan alamat link).
2. Buka Jupyter Notebook pada Firefox. Jika case nya jupyter notebook ini terbuka pada browser lain, maka salin link alamat home page jupyter notebook, kemudian paste pada browser Firefox. Jika dimintai token maka masukan token ini yang ada di cmd jupyter:



Salin token yang ditandai kuning. Catatan, tiap device tokennya berbeda, jadi sesuaikan saja.

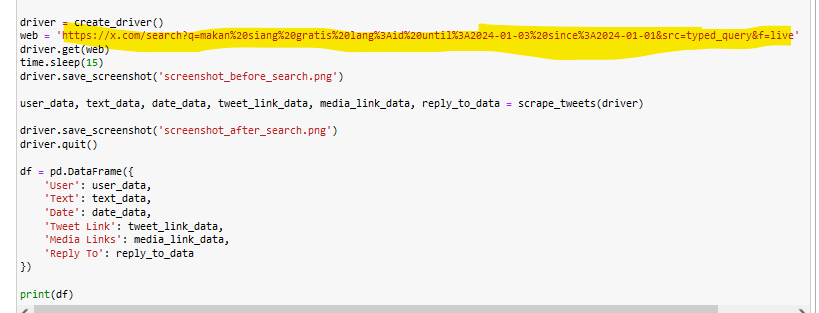
1. Jika sudah masuk home page, klik `upload` dan masukan code yang kami kirim yang dalam format `.ipynb`.
2. Kemudian sesuaikan link path dan link web twitter yang sudah anda miliki, contoh:



Yang tanda merah, hapus dan isi dengan link path geckodriver.exe anda.



Yang tanda merah, hapus dan isi dengan link path firefox.exe anda.



Yang tanda kuning, hapus dan isi dengan link web twitter yang akan anda Tarik datanya.

1. Kemudian klik `Run` code nya. Disini run nya satu per-satu dulu. Yang tadi sudah di isi dan di sesuaikan link nya, nanti akan otomatis membuat tab browser baru yang mana mengarah pada web twitter. Kemudian anda isi username dan password akun anda, klik login. Web akan otomatis meng-scroll sendiri, yang artinya sedang mengambil data. Jika sudah selesai, web akan close sendiri dan hasil akan tersimpan di notebook. Kemudian setelah itu run saja code yang tersisa. Dan penarikan sukses berhasil.

Catatan, saat tab baru pada web meng-scroll otomatis jangan pernah meng-close nya atau meng-scroll nya sendiri oleh anda, nanti akan terjadi error, maka dari itu biarkan saja sampai tab ke close otomatis.

1. Jika ingin melihat hasilnya, buka kembali Home page jupyter, kemudian cari hasil yang dalam format csv. Kemudian buka csv tersebut, nanti klik `File`, kemudian klik `Download`. Maka hasil csv akan tersimpan di device anda.