Dokumentasi Final: **Shopee Product Scraper**

**Link kode :** [**link copas kode**](https://github.com/MFirSyah/scrap_data_jual_terlaris_shopee/blob/main/scraper_shopee.py)

1. **Tujuan:**

Script ini digunakan untuk **mengambil (scrape)** data produk dari sebuah toko Shopee, baik produk yang **masih tersedia (ready stock)** maupun yang **sudah habis**. Data akan disimpan ke file Excel dalam dua sheet: **Produk\_Tersedia** dan **Produk\_Habis**.

1. **Fitur Utama**

|  |  |
| --- | --- |
| **Fitur** | **Keterangan** |
| **🔎 Scrape Produk Ready** | Mengambil nama, harga, dan jumlah terjual dari produk yang masih tersedia. |
| **🛑 Scrape Produk Habis** | Mengambil data produk yang stoknya habis, tidak termasuk rekomendasi Shopee. |
| **⏱️ Delay Otomatis** | Jeda acak (1–10 detik) untuk menyimulasikan perilaku manusia dan hindari deteksi bot. |
| **📁 Simpan Excel Otomatis** | Data disimpan di **hasil\_scraping\_shopee.xlsx** dengan 2 sheet. |
| **🧼 Anti-Duplikat Produk** | Menghindari produk yang sama masuk dua kali. |
| **👀 Preview di Terminal** | Menampilkan sebagian data hasil scraping (hingga batas **preview\_limit**). |

1. **Kebutuhan Sistem**

* **Python 3.8** atau lebih baru
* **Google Chrome** (dibuka dengan --remote-debugging-port=9222)
* **ChromeDriver** yang cocok dengan versi Chrome
* **Package Python**:
* Selenium
* beautifulsoup4
* pandas
* openpyxl

1. **Instalasi Package (sekali saja)**

“pip install selenium beautifulsoup4 pandas openpyxl”

1. **Cara Menjalankan**
2. **Buka Google Chrome dengan mode debugging**

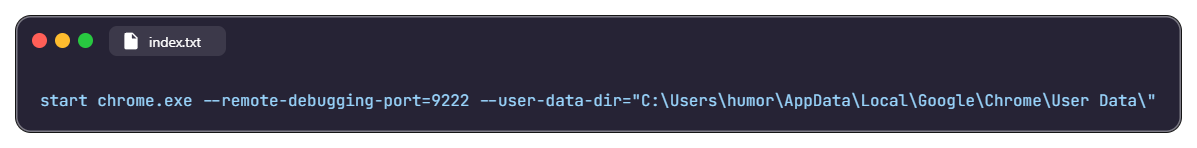
Buka Command Prompt, lalu jalankan:



**Note:**

Gantilah path Chrome sesuai lokasi Chrome di komputer Anda.

**Contoh:**



Jika ada beberapa profil, cek **dir** berada dimana, dengan cara:

1. Buka chrome
2. Ketik **chrome://version**
3. Lihat pada **Jalur Profil**, disitu tertera **dir** profil
4. Maka jadinya seperti ini:



1. **Login ke Shopee**
2. Di Chrome yang terbuka, login ke akun Shopee.
3. Buka halaman toko Shopee yang ingin di-scrape.
4. Klik tab produk
5. Klik urutkan berdasarkan apa, sesuai keinginan
6. Salin alamat websitenya
7. Jangan tutup chrome tampilan terbaru ini
8. **Jalankan script**
9. Buka visual studio code
10. Masuk ke direktori tempat menyimpan kode
11. Buka file-nya
12. Tekan CTRL+J
13. Ketik **python scraper\_shopee.py**
14. Tempel alamat website, lalu enter
15. Masukkan jumlah preview data yang di scrape, lalu enter
16. Tunggu proses scraping selesai

**Note:**

Jika gagal karna terkena sistem bot, ulangi dari nomor 5 hingga 8

1. **Input yang Diminta Saat Menjalankan**

|  |  |
| --- | --- |
| **Pertanyaan** | **Contoh Jawaban** |
| **🔗 Masukkan URL toko Shopee** | https://shopee.co.id/dbklik?page=0&sortBy=sales&tab=0 |
| **📊 Masukkan jumlah preview** | 30 |

1. **Output yang Dihasilkan**

* File: **hasil\_scraping\_shopee.xlsx**
* Sheet **Produk\_Tersedia** → berisi semua produk ready stock
* Sheet **Produk\_Habis** → berisi semua produk yang sudah habis

**Contoh kolom:**

* Nama Produk
* Harga (dalam Rupiah)
* Terjual per Bulan (jumlah penjualan estimasi)

1. **Contoh output pada terminal:**

**A screenshot of a computer

AI-generated content may be incorrect.**

**CATATAN PENTING:**

* Script ini **tidak membuka browser baru**, tapi menghubungkan ke Chrome yang sudah dibuka (debug mode).
* Pastikan **tidak menutup Chrome** saat scraping berlangsung.
* Shopee bisa membatasi akses jika scraping terlalu cepat → script ini sudah otomatis memperlambat proses.

**Tips Tambahan**

* Untuk toko besar, gunakan preview\_limit lebih besar (misal: 1000) agar hasil terlihat jelas.
* Jika muncul error koneksi, pastikan Chrome sudah terbuka dengan debugging.
* Jika ingin menyimpan juga ke format .csv, tinggal tambahkan:

A computer code with orange and blue text

AI-generated content may be incorrect.

**Batasan & Pengembangan Lanjut**

|  |  |
| --- | --- |
| **Batasan Saat Ini** | **Bisa Ditambah** |
| Tidak scrape gambar, rating, atau kategori | Tambahkan **div** dan **img** selector di fungsi **extract** |
| Belum ada filter diskon / promo | Bisa tambahkan field baru di scraping |
| Hanya Excel output | Bisa tambahkan CSV, JSON, atau upload ke Google Sheets |