



Data Science Workflow Essentials: A Practical Demonstration in R




Rajius Idzalika - Data Scientist @UN Global Pulse Asia Pacific

Disajikan pada acara R-Ladies meet up 26 Agustus 2023

Alur kerja data science: dari mendefinisikan tujuan hingga penyebaran hasil.

	Alur	Deskripsi
1	Mendefinisikan tujuan ↓	Contoh: <ul style="list-style-type: none">• Untuk mendapatkan wawasan mengenai xxx• Untuk mengeksplorasi data yyy
2	Identifikasi data dan akses ↓	Apa tipe data yang tersedia? Angka, teks, gambar, data spasial dll Apakah data tersebut data primer atau sekunder? Jika data sekunder, bagaimana cara mengakses data? Mendapatkan akses VS mengumpulkan data
3	Pemrosesan data ↓	Pembersihan data, analisa data, visualisasi data
4	Diseminasi hasil	Laporan, presentasi, blog dll

Demonstrasi alur kerja data science untuk project #LNOB.

	Alur	Deskripsi
1	Mendefinisikan tujuan 	Untuk mendapatkan wawasan mengenai sebaran Perguruan Tinggi di Indonesia yang mengakomodir kebutuhan khusus/penyandang disabilitas.
2	Identifikasi data dan akses 	Sumber data: Dikti . Akses data dilakukan dengan web scraping.
3	Pemrosesan data 	Pembersihan data, analisa data, visualisasi data
4	Diseminasi hasil	<ul style="list-style-type: none">• Tabel: ditampilkan ke dalam platform• Chart atau peta

Mengumpulkan data lewat web scraping apakah boleh?

Argumen boleh	Argumen tidak boleh
Data merupakan informasi publik	Sudah tertulis di Term of Service
Untuk tujuan yang menguntungkan publik seperti riset akademis, jurnalisme data	Dapat menguntungkan kompetitor
Tidak ada larangan eksplisit (lihat robot.txt)	Robot.txt secara eksplisit mendeskripsikan apa yang boleh dan tidak boleh
Data bersifat non-komersial, tidak memperlambat kerja server, tidak melibatkan hak cipta atau data pribadi	Data bersifat komersial, memperlambat kerja server, melibatkan hak cipta atau data pribadi

Data yang tersedia di situs Dikti merupakan data semi terstruktur

ptinklusif.kemdikbud.go.id/s/5/pt-siap-menerima-mhs-disabilitas

Education Links for shared d... Resilience My trainings Administration Warung makan Literature summar... The tricks propag...

Jln. Raya Jend Sudirman Pintu I, Gedung D Depdiknas, Senayan Jakarta 10270 (021) 57946100 ptinklusif@ken

Imc
Layanan Mahasiswa Disabilitas

HOME MAIN PAGE BERITA BERITA ME PER

DAFTAR PERGURUAN TINGGI YANG MENERIMA MAHASISWA BERKEBUTUHAN KHUSUS

No	Perguruan Tinggi	Program Studi
1.	Akademi Kesejahteraan Sosial Ibu Kartini Jalan Sultan Agung No 77 Candibaru Semarang	D3 Desain Busana <i>B - Disabilitas Rungu</i>
2.	Akademi Pekerjaan Sosial Kupang Jalan Jambu No 10 Oepura (Kompleks Sosial) Oepura Kupang	D3 Pekerja Sosial <i>B - Disabilitas Rungu, F - Gangguan Komunikasi, A - Disabilitas Netra</i>
3.	Akademi Tata Boga Bandung Jalan Raya Cinunuk No 186	1. D3 Tata Boga <i>Q - Autis, C - Disabilitas Intelektual, B - Disabilitas Rungu</i> 2. D1 Tata Hidangan <i>Q - Autis</i>

Data menjadi terstruktur melalui proses pembersihan data

No	Perguruan Tinggi	Program Studi
1.	Akademi Kesejahteraan Sosial Ibu Kartini Jalan Sultan Agung No 77 Candibaru Semarang	D3 Desain Busana <i>B - Disabilitas Rungu</i>
2.	Akademi Pekerjaan Sosial Kupang Jalan Jambu No 10 Oepura (Kompleks Sosial) Oepura Kupang	D3 Pekerja Sosial <i>B - Disabilitas Rungu, F - Gangguan Komunikasi, A - Disabilitas Netra</i>
3.	Akademi Tata Boga Bandung Jalan Raya Cinunuk No 186	1. D3 Tata Boga <i>Q - Autis, C - Disabilitas Intelektual, B - Disabilitas Rungu</i> 2. D1 Tata Hidangan <i>Q - Autis</i>



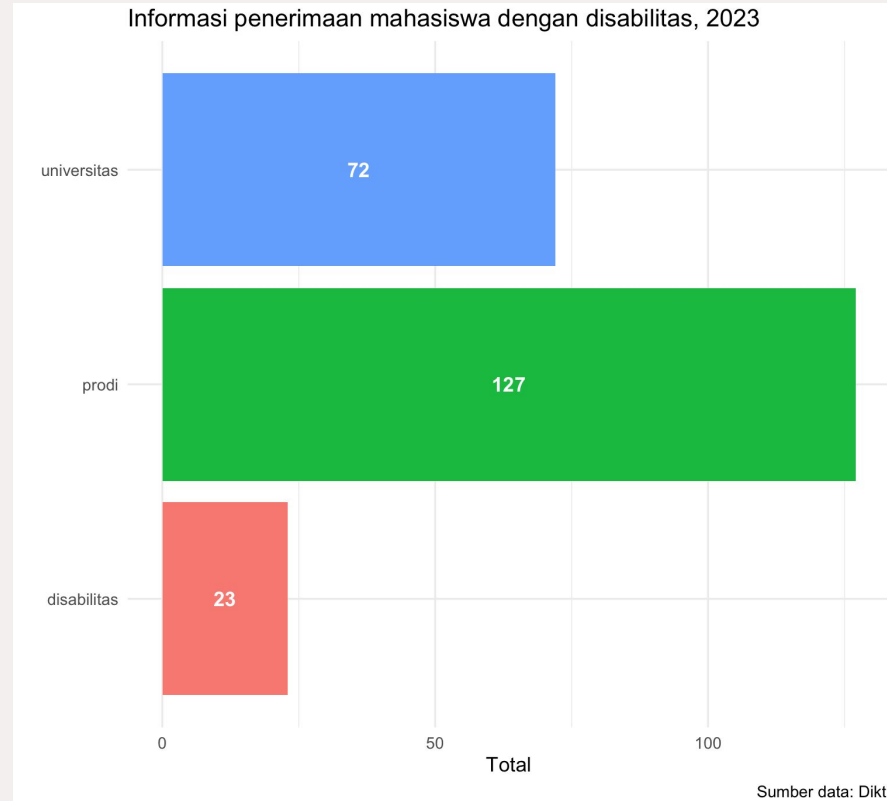
	A	B	C	D	E
1		perguruan_tinggi	alamat	prodi	disabilitas
2	1	Akademi Kesejahteraan Sosial Ibu Kartini	Jalan Sultan Agung No 77 Candibaru Semarang	D3 Desain Busana	B - Tuna runggu
3	2	Akademi Pekerjaan Sosial Kupang	Jalan Jambu No 10 Oepura (Kompleks Sosial) Oepura Kupang	D3 Pekerja Sosial	B - Tuna runggu
4	3	Akademi Pekerjaan Sosial Kupang	Jalan Jambu No 10 Oepura (Kompleks Sosial) Oepura Kupang	D3 Pekerja Sosial	F - Tuna wicara
5	4	Akademi Pekerjaan Sosial Kupang	Jalan Jambu No 10 Oepura (Kompleks Sosial) Oepura Kupang	D3 Pekerja Sosial	A - Tuna netra
6	5	Akademi Tata Boga Bandung	Jalan Raya Cinunuk No 186	D3 Tata Boga	Q - Autis
7	6	Akademi Tata Boga Bandung	Jalan Raya Cinunuk No 186	D3 Tata Boga	C - Tuna grahita ringan
8	7	Akademi Tata Boga Bandung	Jalan Raya Cinunuk No 186	D3 Tata Boga	B - Tuna runggu
9	8	Akademi Tata Boga Bandung	Jalan Raya Cinunuk No 186	D1 Tata Hidangan	Q - Autis

Pembersihan data bertujuan untuk mempersiapkan data agar dapat dianalisa.

Langkah-langkah:

1. Memisahkan kolom **perguruan tinggi** menjadi **perguruan tinggi** dan **alamat**.
2. Memisahkan kolom **program studi** menjadi **program studi** dan **tipe disabilitas**.

Hasil visualisasi pertama

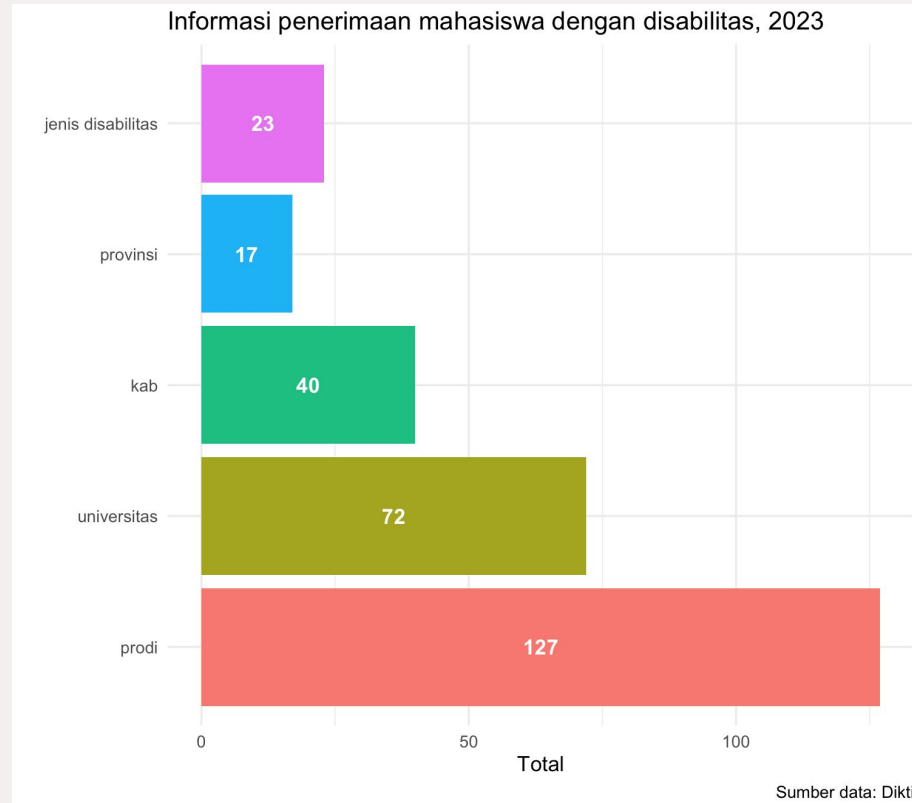


Pembersihan data bertujuan untuk mempersiapkan data agar dapat dianalisa.

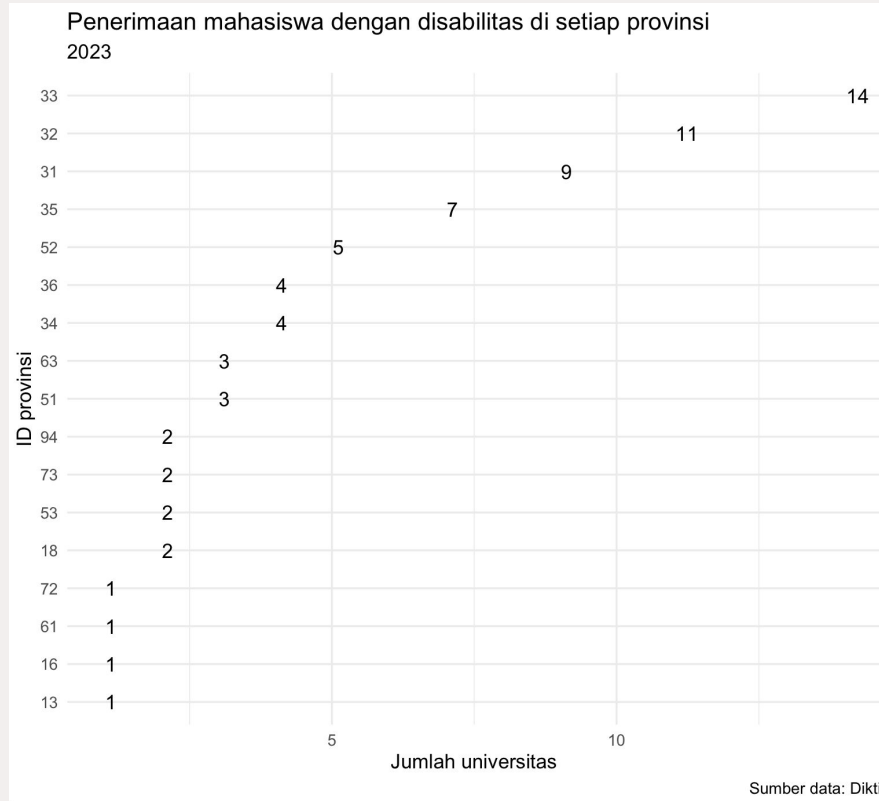
Langkah-langkah:

1. Memisahkan kolom **perguruan tinggi** menjadi **perguruan tinggi** dan **alamat**. ✓
2. Memisahkan kolom **program studi** menjadi **program studi** dan **tipe disabilitas**. ✓
3. Menambahkan kolom **ID kabupaten/kota** dan **provinsi**

Hasil visualisasi kedua



Hasil visualisasi ketiga



Setelah visualisasi:

Langkah-langkah:

1. **Membuat interpretasi** hasil visualisasi
2. **Merapikan script**