

PENERAPAN TEKNIK WRANGLING DALAM ANALISIS STATISTIK JUMLAH PEGAWAI NEGERI SIPIL SUMATERA BAGIAN SELATAN 2024

**Kelompok 5
Komputasi Statistik**

- 1. Keren Marito Lumban Gaol 123450020**
- 2. Rosalia Siregar 123450036**
- 3. Enggli Rahmadhani 123450043**
- 4. Akeyla Fairuz Shafi 123450119**





TEKNIK WRANGLING/MUNGING

Teknik wrangling atau Teknik munging adalah proses membersihkan dan mengubah data mentah yang masih berantakan menjadi data yang rapi dan siap dianalisis. Bayangkan seperti merapikan kamar yang berantakan sebelum dipakai - data juga perlu dibersihkan dulu sebelum bisa dianalisis. Prosesnya meliputi menghapus data yang hilang, memperbaiki duplikasi, menyeragamkan format, dan mengatur ulang struktur data





LATAR BELAKANG



Provinsi Sumatera Selatan memiliki ribuan PNS yang tersebar di 41 instansi pemerintahan. Data kepegawaian ini sangat penting sebagai dasar pengambilan keputusan dan kebijakan pemerintah, seperti perencanaan kebutuhan pegawai baru, distribusi pegawai, dan kebijakan kesetaraan gender. Namun data yang dikumpulkan dari BPS masih dalam kondisi mentah dengan berbagai masalah seperti data yang hilang, penulisan nama instansi tidak seragam, dan format yang belum siap dianalisis. Jika data ini langsung dianalisis tanpa dibersihkan, hasilnya bisa menyesatkan dan kebijakan yang dibuat bisa salah sasaran.





TUJUAN



Penerapan teknik ini bertujuan untuk menjawab tiga pertanyaan utama:

1. Tahapan wrangling apa saja yang diperlukan untuk mengolah data PNS?
2. Bagaimana cara menerapkan setiap tahapan tersebut?
3. Seperti apa hasil akhir data setelah melalui proses wrangling?

Tujuan akhirnya adalah mengaplikasikan 6 teknik wrangling untuk menghasilkan data yang terstruktur dan menghasilkan informasi yang berguna sebagai dasar kebijakan kepegawaian di Sumatera Selatan.





METODE



- **TEKNIK WRANGLING**

Ada enam teknik wrangling yang diterapkan dalam penelitian ini. Pertama, creating variables yaitu membuat variabel baru seperti rasio gender atau persentase. Kedua, subsetting untuk memilih bagian data yang relevan. Ketiga, sorting untuk mengurutkan data berdasarkan kriteria tertentu. Keempat, recoding untuk menyeragamkan kategori dan penulisan. Kelima, merging untuk menggabungkan data dari berbagai sumber. Keenam, reshaping untuk mengubah struktur data dari format wide ke long atau sebaliknya agar sesuai kebutuhan analisis.

- **METODE PENELITIAN**

Penelitian ini menggunakan metode kuantitatif deskriptif dengan data sekunder dari Badan Pusat Statistik tahun 2024. Data mencakup 41 instansi pemerintah Provinsi Sumatera Selatan dengan variabel yang diamati meliputi jumlah PNS pria, wanita, total PNS, dan nama instansi. Pengumpulan data dilakukan melalui studi dokumentasi dari publikasi resmi BPS, kemudian data tersebut diolah menggunakan teknik wrangling secara sistematis mulai dari input dataset hingga data siap dianalisis.





DATASET

Dataset mencakup 41 instansi pemerintah dengan total 13.219 PNS. Lima instansi dengan jumlah PNS terbanyak adalah Dinas Pendidikan dengan 7.775 pegawai, Dinas Kesehatan dengan 810 pegawai, Badan Pendapatan Daerah dengan 418 pegawai, Setda Provinsi dengan 334 pegawai, dan RS Dr. Ernaldi Bahar dengan 285 pegawai. Data awal memiliki berbagai masalah seperti missing value, inkonsistensi penulisan nama instansi, dan format yang belum terstruktur sehingga memerlukan proses wrangling yang sistematis.

Link

<https://sumsel.bps.go.id/id/statistics-table/2/OTY0IzI=/jumlah-pegawai-negeri-sipil-menurut-dinas-instansi-pemerintah-dan-jenis-kelamin.html>





HASIL & PEMBAHASAN



Penelitian ini menggunakan data sekunder jumlah PNS Provinsi Sumatera Selatan tahun 2024 dari BPS, mencakup 41 instansi dengan total 13.049 pegawai. Dataset awal berbentuk wide dan kemudian diproses melalui berbagai teknik data wrangling seperti pembuatan variabel baru (persentase gender, rasio gender, kategori dominasi), subsetting instansi berdasarkan ukuran dan dominasi gender, sorting instansi menurut jumlah PNS dan rasio gender, recoding ukuran instansi dan status gender, merging dengan data penduduk dan klasifikasi fungsi instansi, serta reshaping ke format long untuk kebutuhan visualisasi. Analisis deskriptif menunjukkan distribusi jumlah pegawai yang sangat tidak merata dan right-skewed, dengan sebagian besar instansi berukuran kecil namun hanya menampung sekitar 20% PNS, sementara Dinas Pendidikan menjadi outlier besar yang menyerap hampir 60% PNS. Variabilitas antarinstansi sangat tinggi dan komposisi gender cenderung didominasi pegawai wanita, terutama pada sektor pendidikan dan kesehatan. Keseluruhan proses ini memberikan gambaran menyeluruh mengenai distribusi kepegawaian dan ketimpangan gender, sekaligus menjadi dasar penting untuk perencanaan kebijakan kepegawaian yang lebih efektif, merata, dan responsif terhadap kebutuhan tiap instansi.





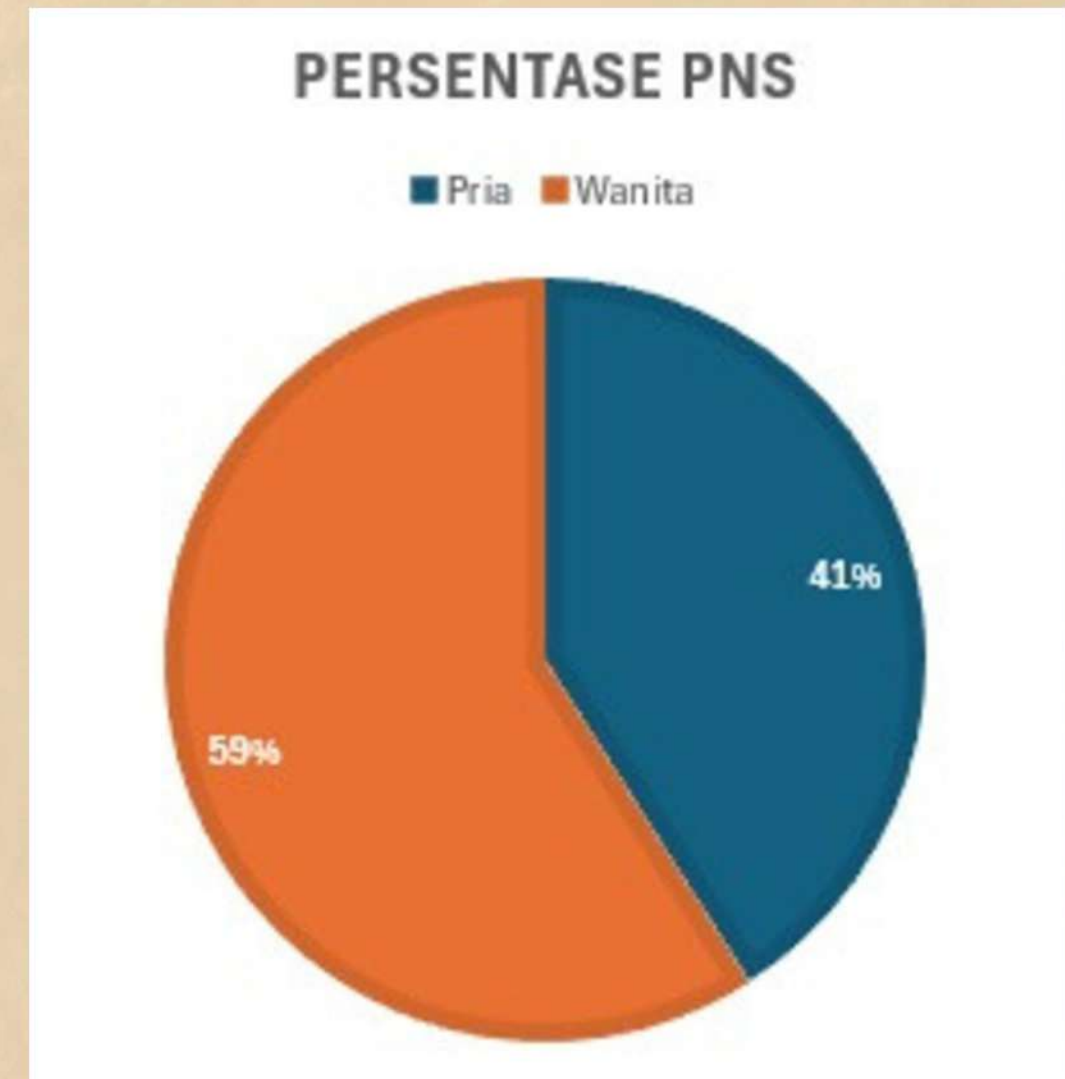
HASIL & PEMBAHASAN

• Visualisasi Data

Dibuat variabel tambahan seperti persentase gender dan sorting data untuk mengategorikan tiap instansi

Instansi <chr>	Total <int>	Persentase_Wanita <dbl>
1 Setda Provinsi Sumatera Selatan	334	44.610778
2 Sekretariat DPRD Provinsi Sumatera Selatan	89	38.202247
3 Inspektorat Provinsi Sumatera Selatan	96	47.916667
4 Satuan Polisi Pamong Praja	138	6.521739
5 RS.Dr. Ernaldi Bahar	285	75.789474

Kategori	Jumlah	%Instansi	%Total PNS
Sangat Besar	1	2,4%	59,6%
Besar	1	2,4%	6,2%
Sedang	6	14,6%	13,3%
Kecil	33	80,6%	20,9%





KESIMPULAN & SARAN

Kesimpulan

Data wrangling terbukti menjadi tahap krusial yang mengubah data mentah menjadi dataset berkualitas melalui enam tahapan sistematis. Hasil penelitian mengungkap bahwa Dinas Pendidikan menyerap hampir 60% total PNS provinsi, 80,6% instansi berukuran kecil, dan 46,3% instansi telah mencapai keseimbangan gender. Data wrangling merupakan langkah strategis untuk menghasilkan informasi akurat sebagai dasar pengambilan keputusan kepegawaian dan perencanaan kebijakan publik berbasis data.

Saran

Pemerintah daerah perlu menetapkan standarisasi sistem pengelolaan data kepegawaian yang seragam di seluruh instansi dan meningkatkan kapasitas SDM melalui pelatihan rutin tentang data wrangling. Data PNS yang telah dibersihkan harus dimanfaatkan secara aktif untuk evaluasi kebutuhan pegawai, pemerataan distribusi, penyusunan formasi baru, dan perumusan kebijakan kesetaraan gender. Diperlukan penelitian lanjutan untuk menganalisis efisiensi instansi berdasarkan rasio pegawai terhadap beban kerja guna mendukung perencanaan kepegawaian yang lebih tepat sasaran dan berkelanjutan.



Thank You

