LAPORAN UAS KOMPUTASI STATISTIK PEMBUATAN DASHBOARD

Disusun Untuk Memenuhi UAS SEMESTER GENAP TA. 2024/2025 Mata Kuliah Komputasi Statistik



Dosen Pengampu:

Yuliagnis Transver Wijaya, S.ST, M.Sc.

Disusun Oleh:

Muhammad Imaddudin Zaki 222313244 2KS3

D-IV KOMPUTASI STATISTIK POLITEKNIK STATISTIKA STIS TAHUN AJARAN 2024/2025

PAKTA INTEGRITAS UJIAN AKHIR SEMESTER

Mata Kuliah: Komputasi Statistik

Prodi: DIV Komputasi Statistik

Saya yang bertanda tangan di bawah ini:

NIM: 222313244

Nama: Muhammad Imaddudin Zaki

Kelas: 2KS3

Dengan ini menyatakan bahwa dashboard dan laporan dashboard Komstat yang saya buat dan kumpulkan untuk Ujian Akhir Semester Mata Kuliah Komputasi Statistik adalah hasil karya saya sendiri. Saya menyatakan tidak melakukan tindakan kecurangan, plagiat, atau pelanggaran akademik lainnya.

Apabila di kemudian hari terbukti melakukan pelanggaran terhadap Pakta Integritas ini, saya bersedia menerima sanksi akademik sesuai ketentuan yang berlaku di Politeknik Statistika STIS.

Demikian Pakta Integritas ini saya buat dengan sebenar-benarnya dan tanpa paksaan dari pihak manapun.

Jakarta, 23 Juli 2025

ttd

Muhammad Imaddudin Zaki

NIM. 222313244

LANGKAH 1 IDENTIFIKASI KEBUTUHAN DATA

Kerentanan sosial di Indonesia pada tahun 2017 mencerminkan kondisi ketimpangan yang kompleks, dipengaruhi oleh faktor demografis, ekonomi, dan kesiapsiagaan terhadap bencana di berbagai wilayah. Analisis ini menggunakan dataset sovi_data.csv yang mencakup 511 kabupaten/kota, dengan variabel utama seperti tingkat kemiskinan (POVERTY), tingkat pendidikan rendah (LOWEDU), dan proporsi penduduk yang tinggal di area rawan bencana (DPRONE). Data ini bersumber dari artikel ilmiah "Revisiting Social Vulnerability Analysis in Indonesia" dengan data primer berasal dari SUSENAS 2017. Tujuan dari pembuatan dashboard diharapkan dapat digunakan untuk mengeksplorasi pola kerentanan, melakukan uji statistik seperti ANOVA dan uji-t untuk membandingkan rata-rata antar kelompok, serta memodelkan hubungan antar variabel melalui regresi linear.

Dalam proses analisis, data yang digunakan tidak menunjukkan keberadaan outlier yang signifikan, sehingga memudahkan dalam penerapan metode statistik tanpa perlu penanganan khusus terhadap nilai ekstrem. Hal ini memungkinkan analisis dilakukan secara langsung pada data mentah tanpa transformasi yang kompleks. Sebagian besar variabel dalam dataset berskala rasio dan bersifat kontinu. Untuk analisis yang memperlukan variabel kategorik, pada dashboard sudah terdapat fitur yang memungkinkan untuk mengubah variabel numerik ke kategorik.

Dengan menggabungkan pendekatan statistik inferensial dan eksplorasi visual, analisis ini tidak hanya menyelesaikan tantangan teknis dalam pengolahan data, tetapi juga berupaya menghasilkan insight strategis. Hasil analisis diharapkan dapat memberikan kontribusi nyata dalam mendukung kebijakan mitigasi bencana dan pemerataan pembangunan berbasis bukti, terutama dengan identifikasi wilayah-wilayah yang menunjukkan tingkat kerentanan tinggi. Pendekatan ini menegaskan pentingnya komputasi statistik dalam memahami dinamika sosial Indonesia dan merancang intervensi kebijakan yang lebih tepat sasaran. Seluruh proses analisis mengikuti kerangka kerja Ujian Akhir Semester (UAS), dimulai dari manajemen data, eksplorasi visual, pengujian hipotesis, hingga pemodelan statistik.

MENGAMBIL DATA

1. Download data pada link di bawah ini

URL Data:

https://raw.githubusercontent.com/bmlmcmc/naspaclust/main/data/sovi data.csv

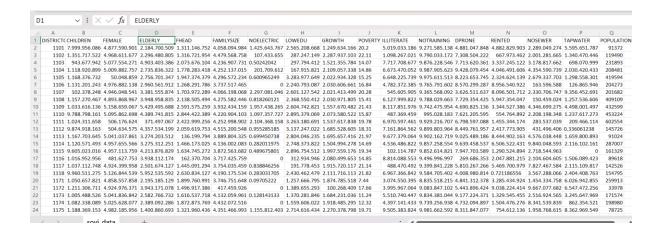
URL Matrik Penimbang Jarak:

 $\underline{https://raw.githubusercontent.com/bmlmcmc/naspaclust/main/data/distance.csv}$

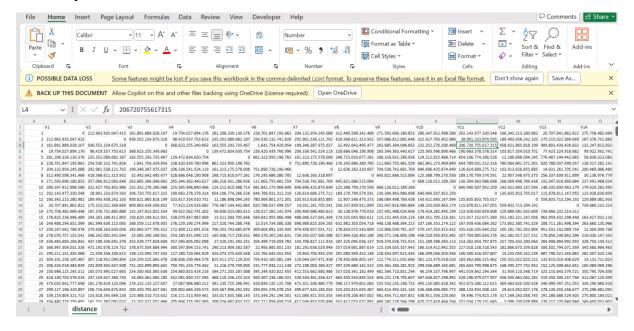
Kedua data akan bertipe CSV

MENGINTEGRITASKAN DATA

Pada tahap ini, proses pemeriksaan integritas data dilakukan untuk memastikan bahwa data yang digunakan dalam analisis telah memenuhi standar kelayakan dan dapat dipertanggungjawabkan. Pada data yang tersedia terdapat variabel variabel di antaranya DISTRICTCODE, CHILDREN, FEMALE, ELDERLY, FHEAD, FAMILYSIZE, NOELECTRIC, LOWEDU, GROWTH, POVERTY, ILLITERATE, NOTRAINING, DPRONE, RENTED, NOSEWER, TAPWATER, POPULATION. yang pastinya semua data itu dapat dipertanggungjawabkan dan sumber datanya dapat dipercaya dengan begitu pasti layak data data ini dijadikan sebagai bahan untuk analisis



Selain data sovi terdapat data lain yaitu matriks penimbang jarak yang bisa digunakan sebagai analisis spasial

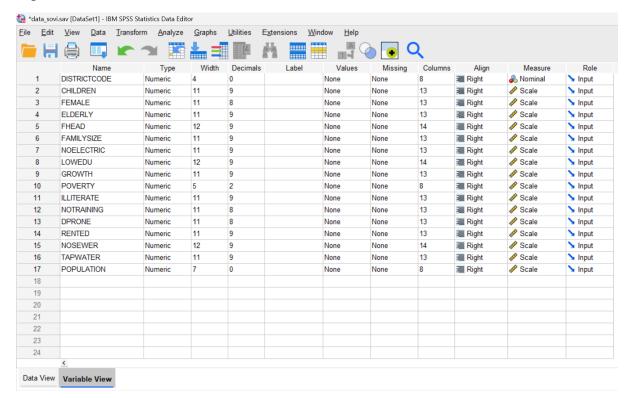


Disini saya tidak menggunakan matrik penimbang jarak untuk membuat visualisasi peta saya menggunakan data lain yaitu geojson wilayah Indonesia yang sudah di gabungkan datanya dengan data sovi yang terdapat data geometry yang menyimpan batas batas wilayah Indonesia



MENELAAH DATA

4.1 Tipe data



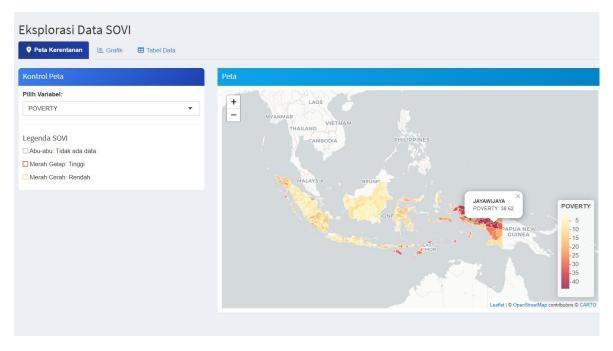
- a. DISTRICTCODE Kategorik, skala Nominal
- b. CHILDREN Numerik (Kontinu), skala Rasio
- c. DPRONE Numerik (Kontinu), skala Rasio
- d. ELDERLY Numerik (Kontinu), skala Rasio
- e. FAMILYSIZE Numerik (Kontinu), skala Rasio
- f. FEMALE Numerik (Kontinu), skala Rasio
- g. FHEAD Numerik (Kontinu), skala Rasio
- h. GROWTH Numerik (Kontinu), skala Rasio
- i. ILLITERATE Numerik (Kontinu), skala Rasio
- j. LOWEDU Numerik (Kontinu), skala Rasio
- k. NOELECTRIC Numerik (Kontinu), skala Rasio
- 1. NOSEWER Numerik (Kontinu), skala Rasio
- m. NOTRAINING Numerik (Kontinu), skala Rasio
- n. POPULATION Numerik (Diskrit), skala Rasio
- o. POVERTY Numerik (Kontinu), skala Rasio
- p. RENTED Numerik (Kontinu), skala Rasio
- q. TAPWATER Numerik (Kontinu), skala Rasio

4.2 Cleaning data

Data yang digunakan telah diberikan dalam kondisi bersih, tanpa adanya nilai hilang (missing values), duplikasi, maupun pencilan (outlier) yang signifikan. Seluruh variabel telah terdefinisi dengan jelas sesuai dengan tipe dan skala pengukuran masing-masing, baik kategorik maupun numerik, sehingga tidak memerlukan proses *data cleaning* tambahan. Dengan demikian, data siap langsung digunakan untuk keperluan eksplorasi, analisis statistik, dan pemodelan sesuai dengan tujuan studi.

4.3 Visualisasi

Dengan memanfaatkan fitur visualisasi yang tersedia di dashboard, seperti peta, histogram, dan scatter plot, pengguna dapat mengeksplorasi pola kerentanan sosial antar daerah secara interaktif. Peta tematik memungkinkan identifikasi spasial wilayah-wilayah dengan tingkat kerentanan tinggi berdasarkan variabel seperti POVERTY atau DPRONE, sedangkan histogram digunakan untuk melihat distribusi frekuensi dari masing-masing indikator kerentanan. Selain itu, scatter plot memfasilitasi analisis hubungan antar variabel, seperti keterkaitan antara tingkat pendidikan rendah (LOWEDU) dengan kemiskinan (POVERTY). Kombinasi visualisasi ini memberikan gambaran menyeluruh terhadap karakteristik data dan menjadi dasar kuat untuk analisis statistik yang lebih lanjut.







LANGKAH 5 MEMVALIDASI DATA

Proses validasi data telah dilakukan untuk menjamin integritas data yang digunakan dalam analisis. Berikut adalah hasil dari pemeriksaan tersebut:

- Ketepatan Data (Accuracy): Data telah divalidasi dan dipastikan akurat, merujuk pada sumber aslinya dari BPS SUSENAS 2017.
- Konsistensi Data: Pemeriksaan internal tidak menemukan adanya data yang kontradiktif. Semua informasi untuk setiap kabupaten/kota konsisten di seluruh variabel.
- Kelengkapan Data: Data yang digunakan dalam analisis akhir ini telah lengkap 100% dan bebas dari *missing values* pada kolom-kolom yang relevan.
- Keunikan Data: Telah dipastikan bahwa setiap baris data merepresentasikan satu entitas unik (satu kabupaten/kota) dan tidak ditemukan adanya duplikasi data.
- Format Data: Semua kolom telah memiliki format yang sesuai. Variabel numerik seperti (POVERTY) disimpan sebagai tipe numerik, dan variabel kategorik (seperti DISTRICTCODE) disimpan sebagai faktor/karakter, sesuai peruntukannya.

MENENTUKAN OBJEK DATA

Berdasarkan variabel-variabel yang tersedia, dapat dilakukan beberapa analisis untuk membuat *Business Intelligence* dan menghasilkan *insight*. Namun, langkah krusial sebelum analisis adalah menentukan objek data, di mana tidak semua data akan digunakan. Proses seleksi ini dilakukan secara strategis dengan mempertimbangkan beberapa kriteria utama

Berberapa analisis yang akan dilakukan:

6.1. Pengaruh Tingkat Pendidikan dan Ukuran Rumah Tangga terhadap Tingkat Kemiskinan

- Variabel yang Digunakan:
 - 1. POVERTY (sebagai variabel dependen/yang dipengaruhi)
 - 2. LOWEDU (sebagai variabel independen/yang memengaruhi)
 - 3. FAMILYSIZE (sebagai variabel independen/yang memengaruhi)
- Metode Analisis: Regresi Linear Berganda.

6.2. Kesenjangan Sosial-Ekonomi Berdasarkan Tingkat Risiko Bencana

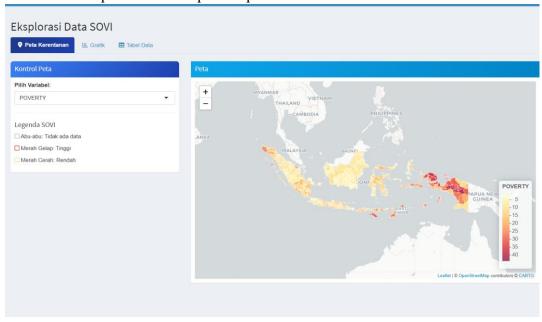
- Variabel yang digunakan:
 - 1. DPRONE
 - 2. FHEAD
 - 3. LOWEDU
 - 4. POVERTY
- Metode Analisis : Kategorasi berdasarkan DPRONE, Uji-t-sample independen, Analisis regresi linear berganda

MEMBUAT BUSINESS INTELLIGENCE

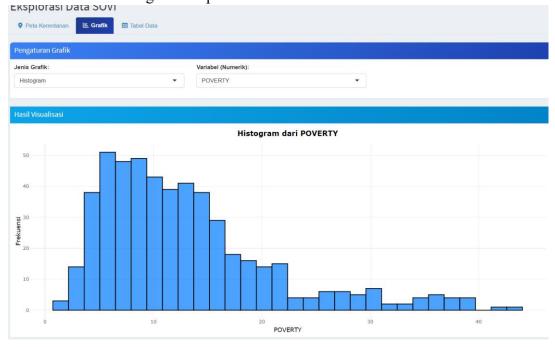
7.1 Pengaruh Tingkat Pendidikan dan Ukuran Rumah Tangga terhadap Tingkat Kemiskinan

Langkah – langkah:

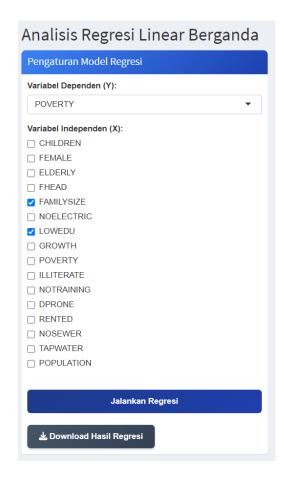
1. Klik menu Eksplorasi data -> peta -> pilih var POVERTY



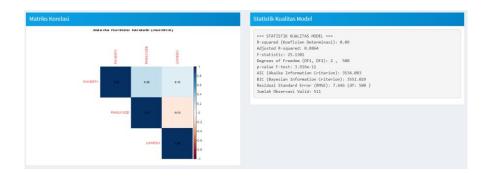
2. Klik menu sub menu grafik -> pilih variabel POVERTY

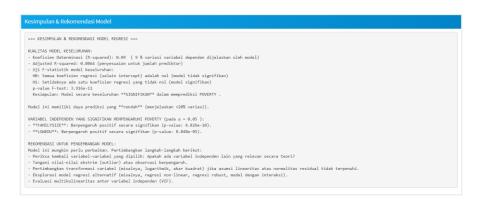


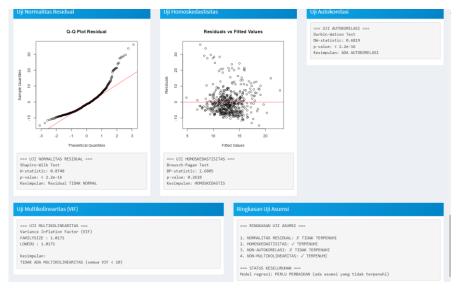
3. Klik menu regresi -> atur var dependen menjadi POVERTY dan pilih var independen FAMILYSIZE dan LOWEDU











Kesimpulan:

- 1. Pendidikan dan Struktur Keluarga adalah Faktor Kunci yang Terbukti berdasarkan output dapat dilihat bahwa dua pilar sosial pendidikan dan struktur keluarga memiliki hubungan langsung dengan tingkat kemiskinan di tingkat daerah.
 - o Insight A (Pendidikan): Setiap kenaikan 1% pada populasi berpendidikan rendah (LOWEDU) akan meningkatkan angka kemiskinan. dalam hal ini dapat diketahui bahwa *intervensi* kebijakan

yang paling mendasar dan terbukti untuk menekan kemiskinan adalah melalui peningkatan akses dan kualitas pendidikan.

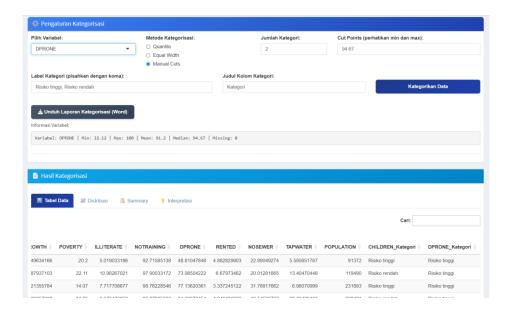
- o Insight B (Keluarga): Semakin besar rata-rata ukuran keluarga (FAMILYSIZE), semakin tinggi pula tingkat kemiskinan. Ini menyoroti adanya beban ekonomi yang lebih besar pada rumah tangga dengan banyak tanggungan. Yang mana hal ini dapat disimpulkan bahwa program pengentasan kemiskinan mungkin perlu dipadukan dengan program keluarga berencana (KB) atau bantuan spesifik untuk keluarga besar.
- 2. Keterbatasan Model yanag mana model ini hanya bisa menjelaskan 9% dari kemiskinan

Ini memberitahu para pengambil keputusan bahwa untuk benar-benar mengatasi kemiskinan, mereka tidak bisa hanya fokus pada pendidikan dan ukuran keluarga saja. Pasti ada faktor-faktor lain yang jauh lebih kuat pengaruhnya.

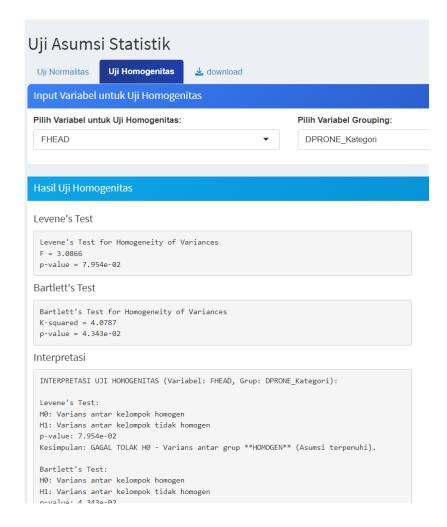
7.2. Kesenjangan Sosial-Ekonomi Berdasarkan Tingkat Risiko Bencana

Langkah – langkah:

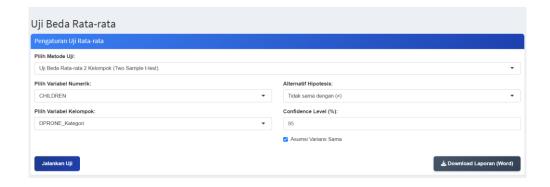
1. Manajemen data -> atur variabel -> pilih metode manual -> atur cut point berdasarkan nilai median



2. Lakukan uji asumsi homogenitas

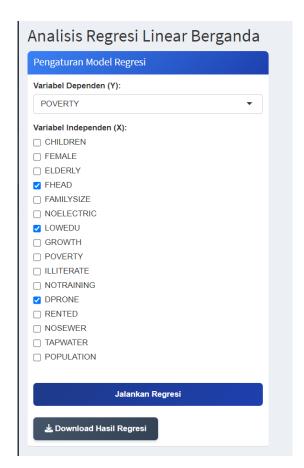


3. Lakukan uji rata rata dua kelompok





4. Lakukan analisis regresi



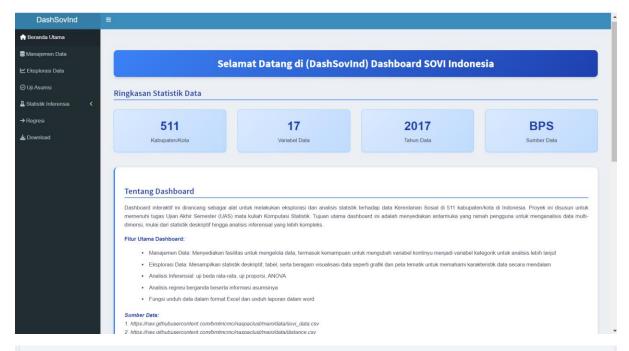
```
Call:
    Informula = a.formula(formula_str), data = reg_data)

Residuals:
    Rin 1Q Residuals
    Residuals
    Rin 1Residuals
    Residuals
    Residuals tandard error; 7.841 on 507 degrees of freedom
    Residuals standard error; 7.841 on 507 degrees of freedom
    Rultiple Residuals standard error; 7.841 on 507 degrees of freedom
    Rultiple Residuals standard error; 7.841 on 507 degrees of freedom
    Rultiple Residuals standard error; 7.841 on 507 degrees of freedom
    Rultiple Residuals standard error; 7.841 on 507 degrees of freedom
    Rultiple Residuals standard error; 7.841 on 507 degrees of freedom
    Rultiple Residuals standard error; 7.841 on 507 degrees of freedom
    Rultiple Residuals standard error; 7.841 on 507 degrees of freedom
    Rultiple Residuals standard error; 7.841 on 507 degrees of freedom
    Rultiple Residuals standard error; 7.841 on 507 degrees of freedom
    Rultiple Residuals standard error; 7.841 on 507 degrees of freedom
    Rultiple Residuals standard error; 7.841 on 507 degrees of freedom
    Rultiple Residuals standard error; 7.842 on 3 and 507 DF, province: 0.89334
    F-statistic: 7.962 on 3 and 507 DF, province: 0.89334
    F-statistic: 7.962 on 3 and 507 DF, province: 0.89334
    F-statistic: 7.962 on 3 and 507 DF, province: 0.89334
    F-statistic: 7.962 on 3 and 507 DF, province: 0.89334
    F-statistic standard error; 7.841 on 507 degrees degree degr
```

Kesimpulan:

Berdasarkan hasil pada gambar, model regresi linear berganda secara keseluruhan signifikan secara statistik (p-value F-statistik = 3.391e-05), yang menunjukkan bahwa kombinasi variabel independen ini mampu menjelaskan variasi kemiskinan lebih baik namun nilai R-squared (0.045) sangat rendah, artinya model ini hanya mampu menjelaskan sekitar 4.5% dari total variasi tingkat kemiskinan (POVERTY).

LAMPIRAN



Variabel	Nama	Deskripsi
CHILDREN	Anak-anak	Persentase penduduk usia di bawah lima tahun
WOMEN	Perempuan	Persentase penduduk perempuan
ELDERLY	Lansia	Persentase penduduk usia 65 tahun ke atas
FEMALE_HEAD	Kepala RT Perempuan	Persentase rumah tangga dengan kepala rumah tangga perempuan
FAMILY_SIZE	Ukuran Keluarga	Rata-rata jumlah anggota rumah tangga
NO_ELECTRICITY	Tanpa Listrik	Persentase rumah tangga yang tidak menggunakan listrik
LOW_EDU	Pendidikan Rendah	Persentase penduduk usia 15+ dengan pendidikan rendah
GROWTH	Pertumbuhan	Persentase perubahan populasi
POVERTY	Kemiskinan	Persentase penduduk miskin
ILLITERATE	Buta Huruf	Persentase penduduk yang tidak bisa membaca dan menulis
NO_TRAINING	Tanpa Pelatihan	Persentase rumah tangga tanpa pelatihan bencana
DISASTER_PRONE	Rawan Bencana	Persentase rumah tangga di daerah rawan bencana
RENT	Sewa Rumah	Persentase rumah tangga yang menyewa rumah
NO_DRAINAGE	Tanpa Drainase	Persentase rumah tangga tanpa sistem drainase
TAP_WATER	Air Keran	Persentase rumah tangga yang menggunakan air keran
POPULATION	Populasi	Jumlah penduduk

