**ANALISIS FAKTOR DEMOGRAFI DAN SOSIAL-EKONOMI TERHADAP TINGKAT KERENTANAN EKONOMI DAN SOSIAL DI INDONESIA**

**(STUDI KASUS DATA SUSENAS 2017)**



**Disusun Oleh:**

**Fadzilla Kusuma Ningrum**

**222313071**

**2KS3**

**Dosen Pengampu:**

**Yuliagnis Tranver Wijaya, S.ST., M.Sc.**

**PROGRAM STUDI D4 KOMPUTASI STATISTIK**

**POLITEKNIK STATISTIKA STIS**

**2024/2025**

Tahapan Pengolahan Data

1. Identifikasi Kebutuhan Data/Business Understanding
2. Masalah

Kerentanan sosial menunjukkan seberapa rentan individu atau wilayah terhadap dampak negatif dari guncangan, baik alamiah seperti bencana alam maupun non-alamiah seperti krisis ekonomi atau wabah penyakit. Dampak ini bisa memperburuk kondisi sosial-ekonomi masyarakat, memperluas kesenjangan, dan menghambat pembangunan. Indeks Kerentanan Sosial (Social Vulnerability Index – SVI) dikembangkan sebagai alat untuk mengukur, memetakan, dan memahami dimensi-dimensi kerentanan ini. SVI mengintegrasikan berbagai indikator demografi, sosial, ekonomi, dan lingkungan untuk memberikan gambaran tentang tingkat kerentanan suatu wilayah secara menyeluruh.

Berdasarkan permasalahan mengenai SVI di atas, topik "Analisis Faktor Demografi dan Sosial-Ekonomi terhadap Tingkat Kerentanan Ekonomi dan Sosial di Indonesia (Studi Kasus Data SUSENAS 2017)" berfokus pada tiga hal utama:

1. Identifikasi Faktor-Faktor yang Mendorong Kerentanan

Menganalisis bagaimana indikator seperti usia anak/lansia, pendidikan, kemiskinan, dan akses dasar berkontribusi terhadap kerentanan di tingkat kabupaten/kota.

1. Menyediakan Bukti Empiris untuk Pengambilan Kebijakan

Memberikan jawaban berbasis data terhadap pertanyaan, seperti apakah kabupaten dengan proporsi lansia yang tinggi cenderung lebih rentan? Apakah rendahnya tingkat pendidikan berpengaruh signifikan terhadap kemiskinan?

1. Melihat Keragaman Wilayah di Indonesia

Menyajikan gambaran yang lebih jelas dan mendalam tentang kerentanan sosial dan ekonomi di berbagai kabupaten/kota di Indonesia melalui visualisasi data.

Beberapa permasalahan yang ingin dianalisis dalam proyek ini sebagai berikut.

1. Bagaimana sebaran dan karakteristik indikator demografi, sosial-ekonomi, serta tingkat kemiskinan di berbagai kabupaten/kota di Indonesia?
2. Apakah data yang digunakan telah memenuhi asumsi dasar statistik untuk analisis lanjutan?
3. Apakah terdapat perbedaan atau hubungan signifikan secara statistik antar kelompok berdasarkan tingkat kerentanan ekonomi dan indikator lainnya?
4. Seberapa besar pengaruh faktor demografi dan sosial-ekonomi terhadap tingkat kerentanan ekonomi (kemiskinan)?
5. Tujuan proyek

Tujuan proyek ini adalah untuk:

1. Memberikan gambaran awal karakteristik indikator demografi, sosial-ekonomi, serta tingkat kemiskinan berdasarkan variabel-variabel seperti proporsi anak-anak, lansia, tingkat pendidikan, kepala rumah tangga perempuan, dan kemiskinan.
2. Memeriksa asumsi-asumsi statistik dasar seperti normalitas sebaran data dan homogenitas varians antar kelompok agar hasil analisis dapat dipercaya.
3. Melihat apakah perbedaan kondisi tersebut memang secara statistik berhubungan dengan tingkat kemiskinan atau kerentanan sosial lainnya.
4. Melihat hubungan antar berbagai variabel independen (misalnya proporsi lansia, tingkat pendidikan, dan akses infrastruktur) dengan variabel dependen yaitu tingkat kemiskinan.
5. Rancangan solusi dan tahapan untuk mencapai tujuan

Untuk mencapai tujuan proyek, tahapan-tahapan yang akan dilakukan adalah sebagai berikut.

1. Menggunakan visualisasi grafik, seperti diagram batang dan boxplot untuk mengetahui gambaran awal karakteristik berbagai variabel.
2. Melakukan uji normalitas dan homogenitas sebagai langkah awal sebelum melakukan statistik inferensia seperti uji t dan regresi linear. Jika asumsi tidak terpenuhi, maka perlu dipertimbangkan penggunaan transformasi data.
3. Membandingkan rata-rata, proporsi, atau varians dari kelompok-kelompok demografi atau sosial-ekonomi yang berbeda, seperti kelompok kabupaten dengan pendidikan tinggi dan rendah.
4. Melakukan analisis regresi linear berganda untuk melihat hubungan antar berbagai variabel independen (misalnya proporsi lansia, tingkat pendidikan, dan akses infrastruktur) dengan variabel dependen yaitu tingkat kemiskinan. Hasil dari model ini dievaluasi melalui signifikansi masing-masing variabel serta validitas model berdasarkan uji asumsi dan kebaikan model.
5. Mengambil Data
6. Mengunduh data

Data diunduh melalui tautan: <https://raw.githubusercontent.com/bmlmcmc/naspaclust/main/data/sovi_data.csv>.

1. Membuka file

Menggunakan RStudio

1. Mengintegritaskan Data

Data yang diperoleh dapat dipertanggungjawabkan karena berasal dari sumber yang terpercaya, yaitu jurnal “Revisiting social vulnerability analysis in Indonesia data”.

1. Menelaah Data
2. Identifikasi tipe data

Data SUSENAS 2017 memiliki 17 variabel dengan tambahan 2 variabel nama provinsi dan nama kabupaten/kota sesuai kode distrik. Berikut variabel yang ada di Data SUSENAS 2017 beserta tipe datanya.

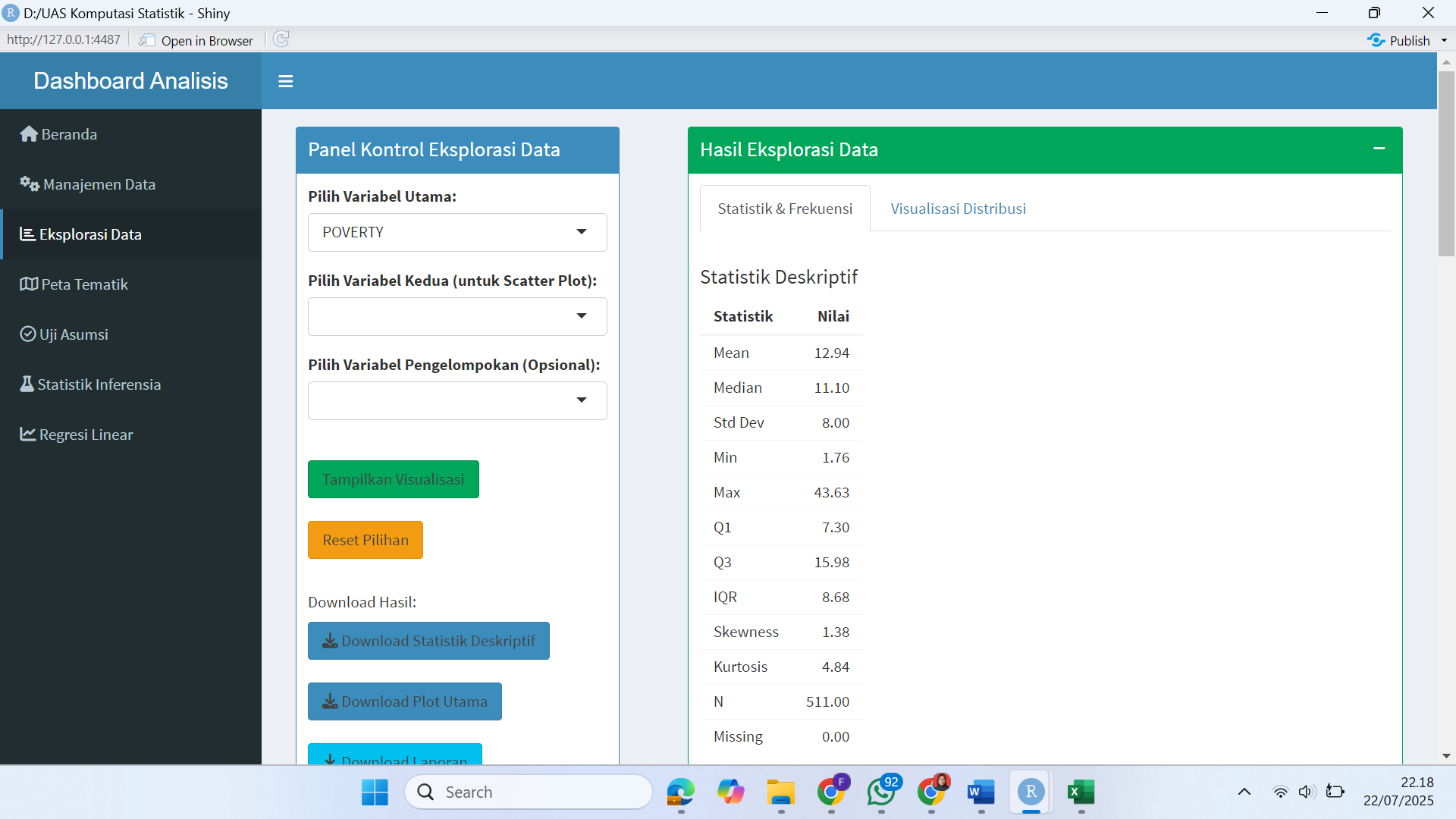


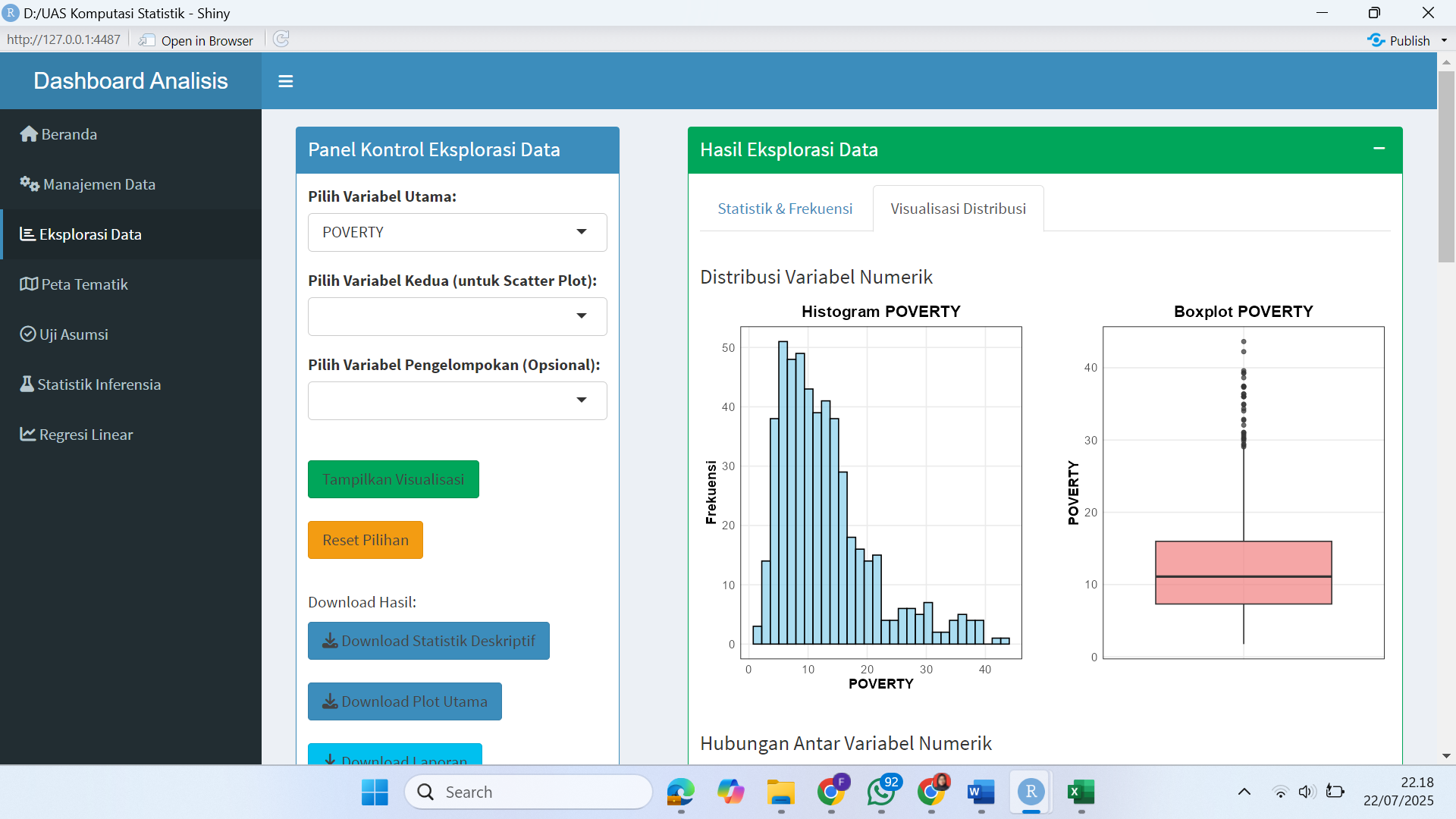
1. Cleaning data
2. Pengecekan *Missing Values*

Berdasarkan pemeriksaan yang dilakukan melalui dashboard interaktif, ditemukan bahwa tidak ada missing values pada seluruh variabel dalam dataset Indeks SVI SUSENAS 2017. Hal ini mengindikasikan bahwa data yang tersedia lengkap dan tidak memerlukan proses imputasi atau penghapusan data akibat nilai yang hilang. Kelengkapan data ini memastikan bahwa analisis statistik yang akan dilakukan dapat menggunakan seluruh observasi tanpa mengurangi ukuran sampel atau memperkenalkan bias akibat penanganan missing values. Jumlah total record (baris) dalam data adalah 511 baris.

1. Pengecekan *Outlier*

Pengecekan outlier (nilai pencilan) dilakukan untuk mengidentifikasi observasi yang secara signifikan menyimpang dari pola umum data, yang dapat memengaruhi asumsi statistik dan hasil model. Metode yang digunakan untuk deteksi *outlier* antara lain adalah metode Interquartile Range (IQR) dan visualisasi Boxplot.





1. Penanganan *Outlier*

Setelah tidak terdapat *missing values*, dapat dilakukan pengecekan kembali data *outlier*. Untuk menanganinya, dilakukan dengan menghapus nilai *outlier* extreme.

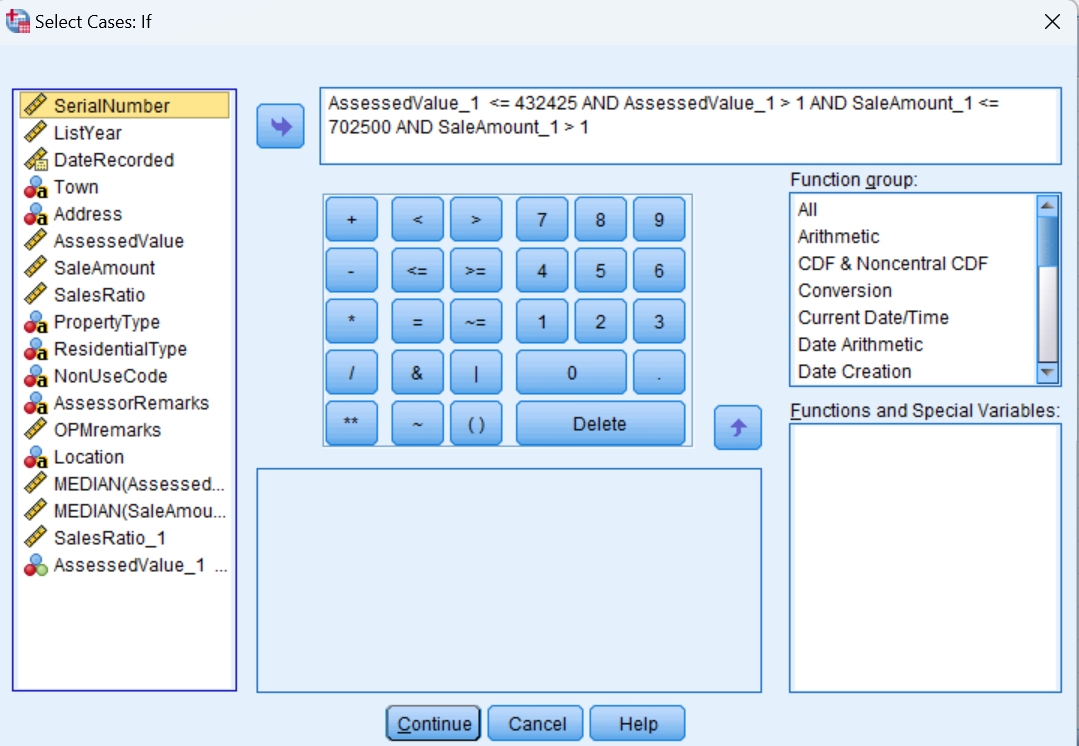
Langkah:

1. Mencari batas-batas *outlier* dengan menggunakan nilai Q1, Q3, dan IQR

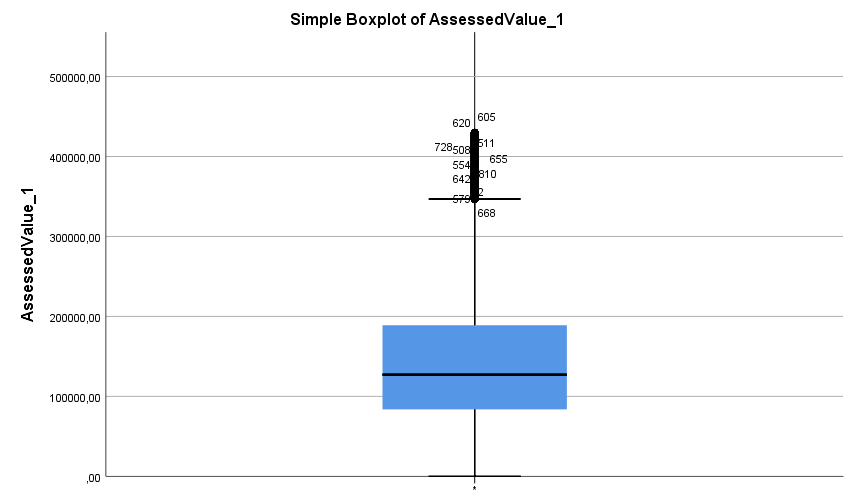
Rumus:



1. Menghapus nilai *outlier* atas dan bawah menggunakan Select Cases

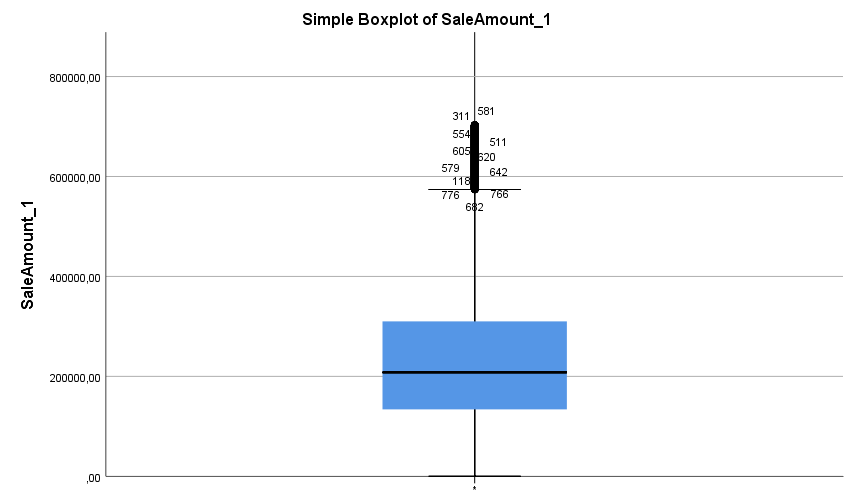


1. Visualisasi data awal dalam bentuk grafik
2. Boxplot variabel Assessed Value\_1



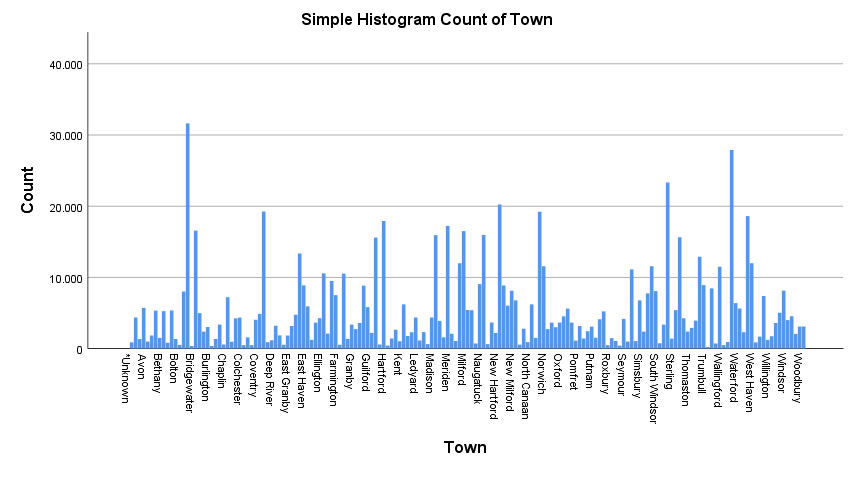
Dari boxplot terlihat bahwa variabel AssessedValue\_1 cenderung berdistribusi miring ke kanan (positively skewed) karena banyak outlier dengan nilai tinggi.

1. Boxplot variabel Sale Amount\_1



Dari boxplot terlihat bahwa variabel SaleAmount\_1 cenderung berdistribusi miring ke kanan (positively skewed) karena banyak outlier dengan nilai tinggi.

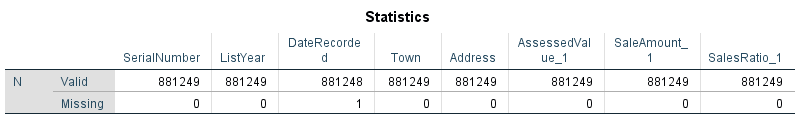
1. Histogram variabel Town



Dari histogram terlihat beberapa kota memiliki inputan terbanyak, yaitu Kota Bridgeport, Danbury, Norwalk, Hartford, dan West Hartford. Kota yang memiliki inputan paling banyak adalah Kota Bridgeport.

1. Memvalidasi Data

Sebelum melakukan analisis data lebih lanjut, dilakukan validasi data untuk memastikan bahwa data sudah lengkap.



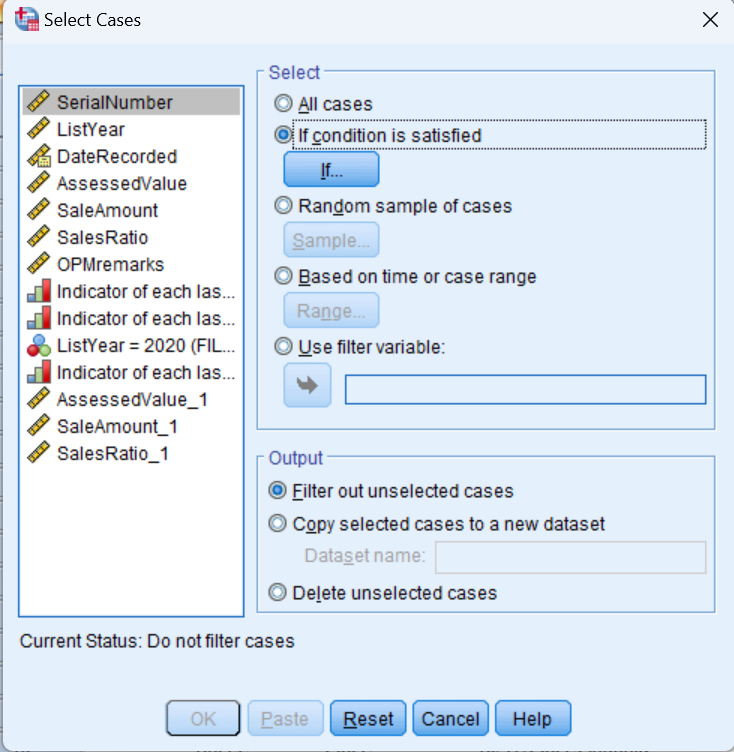
Dari tabel di atas, terlihat data sudah lengkap karena tidak ada *missing values*.

1. Menentukan Objek Data

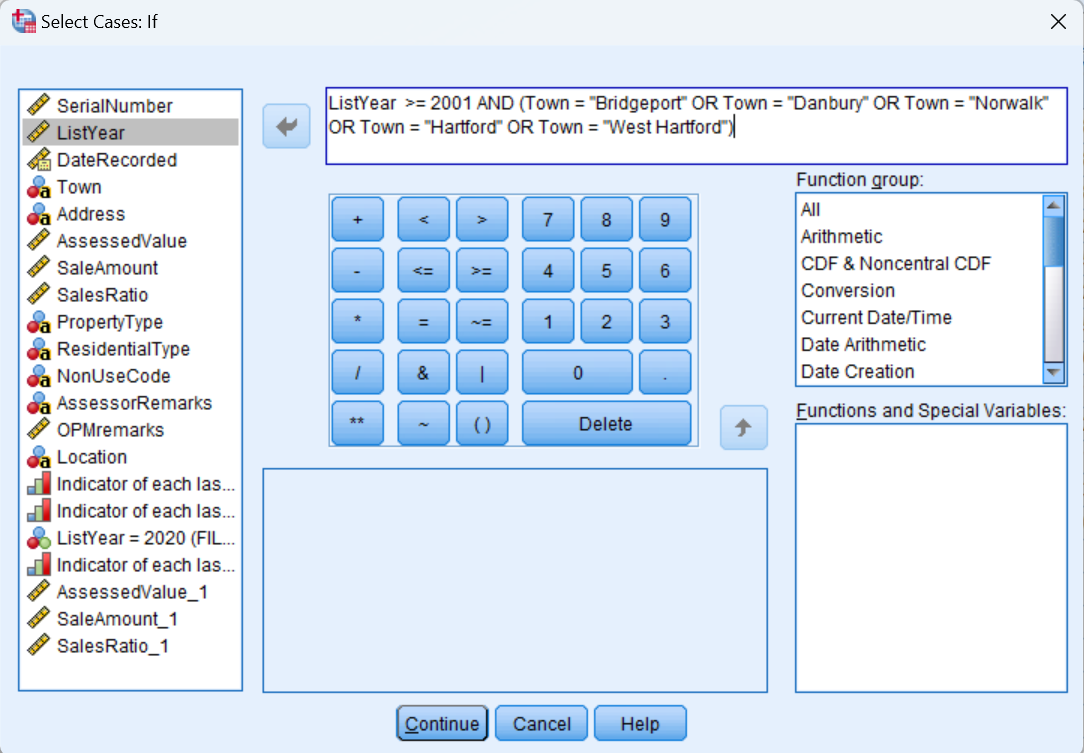
Proyek ini bertujuan untuk menganalisis penjualan properti di lima kota terpilih di Amerika Serikat, yaitu Kota Bridgeport, Danburry, Northwalk, Hartford, dan West Hartford. Data yang digunakan dimulai dari tahun 2001 sampai dengan 2020.

Langkah:

Klik Data 🡪 Select Cases. Pada kotak Select, pilih “If condition is satisfied” dan klik If.

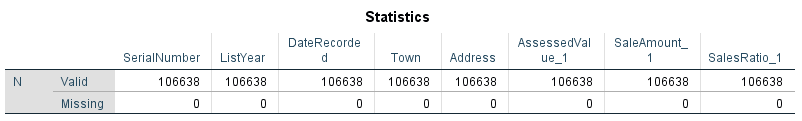


Isikan pada jendela Select Cases: If kasus *record* yang ingin dipih. Lalu, klik Continue.

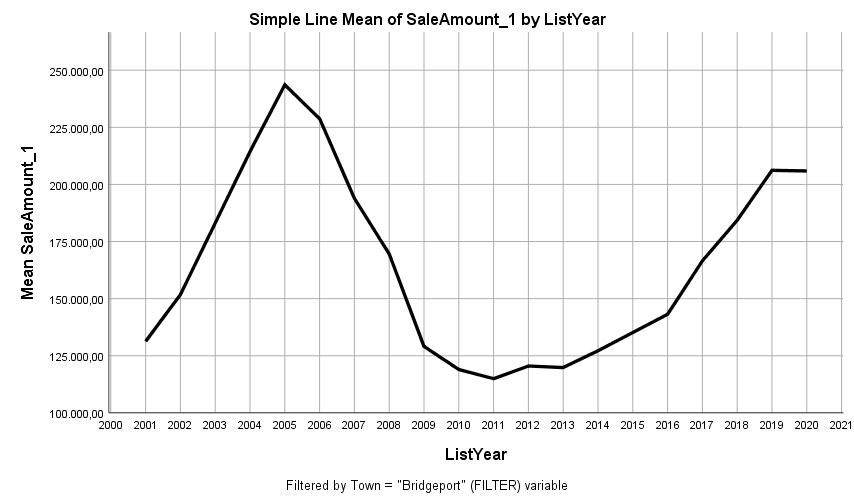


Pada kotak Output, pilih Copy selected cases to a new dataset untuk membuka data yang telah difilter di lembar kerja yang baru.

Tabel berikut menunjukkan validasi terhadap variabel-variabel yang akan digunakan untuk analisis data.



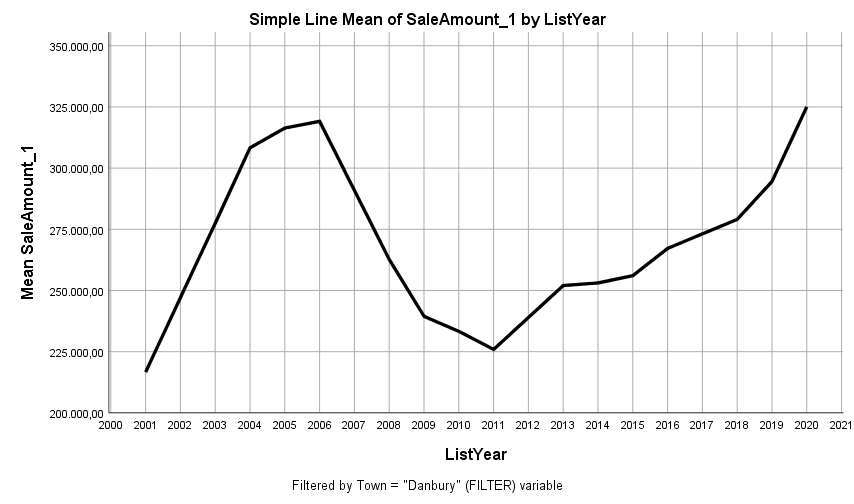
1. Membuat Business Intelligence
2. Visualisasi rata-rata penjualan properti tahun 2001 – 2020 di setiap kota
3. Kota Bridgeport



Interpretasi:

Grafik ini menunjukkan tren rata-rata nilai penjualan properti di Kota Bridgeport pada tahun 2001 – 2020. Pada awal periode tahun 2001, rata-rata nilai penjualan berada di kisaran 135.000 dolar. Nilai ini mengalami peningkatan yang cukup tajam hingga mencapai puncaknya pada tahun 2005, dengan nilai penjualan mendekati 250.000 dolar. Kenaikan ini mencerminkan kondisi pasar properti yang sedang tumbuh. Setelah tahun 2005, terlihat penurunan yang cukup tajam hingga tahun 2011 dan rata-rata nilai penjualan turun sampai di bawah 125.000 dolar. Penurunan ini kemungkinan terkait dengan dampak krisis keuangan global tahun 2008 yang menyebabkan melemahnya pasar properti di berbagai wilayah, termasuk Bridgeport. Mulai tahun 2012, tren mulai menunjukkan sedikit peningkatan, meskipun masih fluktuatif pada awalnya. Kenaikan yang lebih stabil terlihat sejak tahun 2015 dan terus berlangsung hingga mencapai lebih dari 200.000 dolar pada tahun 2019. Setelah itu, tren cenderung datar pada tahun 2020.

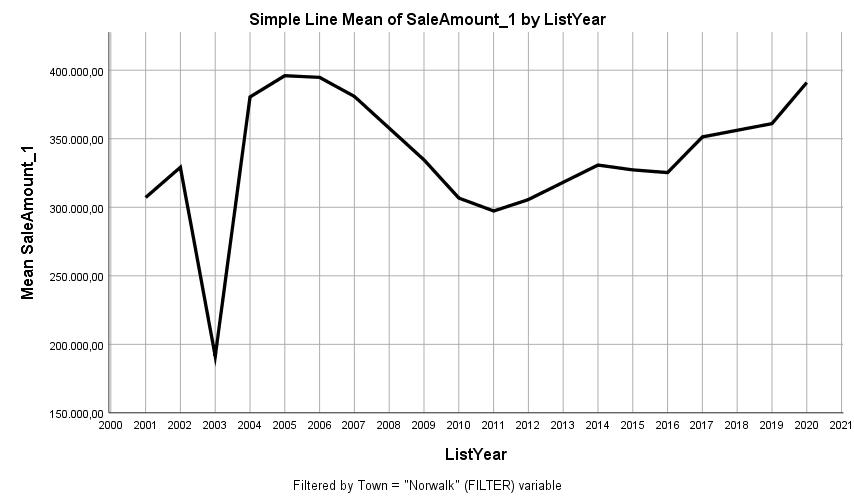
1. Kota Danbury



Interpretasi:

Grafik ini menunjukkan tren rata-rata nilai penjualan properti di Kota Danbury pada tahun 2001 – 2020. Pada tahun 2001, rata-rata nilai penjualan di Kota Danbury berada di kisaran 225.000 dolar. Angka ini kemudian meningkat tajam hingga mencapai puncaknya pada tahun 2006, dengan nilai mendekati 325.000 dolar. Kenaikan ini mencerminkan kondisi pasar properti yang sedang tumbuh. Setelah mencapai puncaknya, terjadi penurunan yang signifikan mulai tahun 2007 hingga mencapai titik terendah pada tahun 2011, dengan rata-rata nilai penjualan menyentuh sekitar 225.000 dolar. Penurunan ini sejalan dengan dampak krisis ekonomi global yang memengaruhi sektor properti, termasuk di Kota Danbury. Mulai tahun 2012, pasar mulai menunjukkan tanda-tanda pemulihan. Rata-rata nilai penjualan meningkat secara perlahan dan konsisten, dengan kenaikan yang semakin kuat dari tahun 2016 ke atas. Tren naik ini berlanjut hingga tahun 2020, di mana nilai rata-rata penjualan properti kembali mendekati 325.000 dolar.

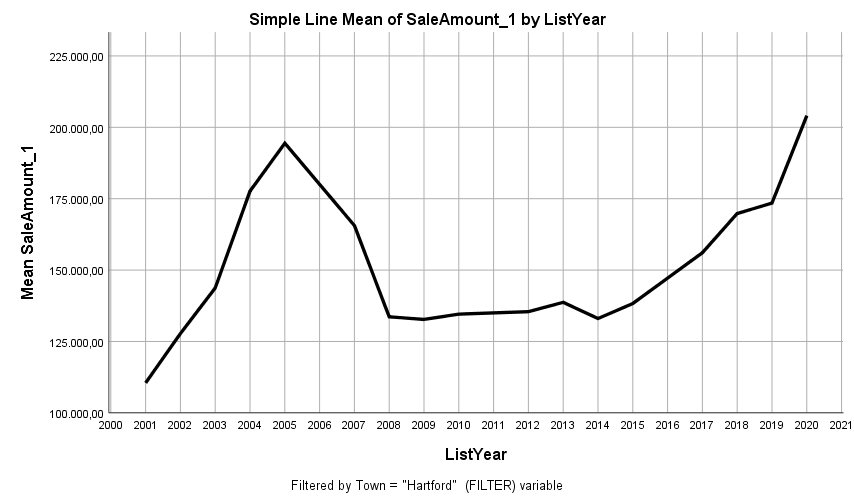
1. Kota Norwalk



Interpretasi:

Grafik ini menunjukkan tren rata-rata nilai penjualan properti di Kota Norwalk pada tahun 2001 – 2020. Pada tahun 2001, rata-rata nilai penjualan berada di atas 300.000 dolar dan sempat meningkat di tahun 2002. Namun, pada tahun 2003 terlihat penurunan yang sangat tajam, dengan rata-rata nilai penjualan turun hingga di bawah 200.000 dolar. Penurunan ekstrem ini disebabkan oleh jumlah data yang sangat sedikit pada tahun tersebut (hanya 4 data), sehingga tidak merepresentasikan kondisi pasar secara akurat. Setelah itu, nilai penjualan melonjak pada tahun 2004 dan mencapai puncaknya pada tahun 2005 dan 2006, dengan rata-rata mendekati 400.000 dolar. Mulai tahun 2007 hingga 2011, terjadi penurunan dalam nilai penjualan yang kemungkinan dampak dari krisis keuangan global. Titik terendah berada di tahun 2011 dengan rata-rata nilai penjualan menyentuh 300.000 dolar. Mulai tahun 2012, tren menunjukkan pemulihan secara bertahap. Peningkatan nilai penjualan berlangsung cukup konsisten, meskipun dengan sedikit penurunan di tahun 2016. Namun, pasar kembali menguat dari tahun 2017 dan terus naik hingga tahun 2020, dengan rata-rata nilai penjualan kembali mendekati angka 400.000 dolar.

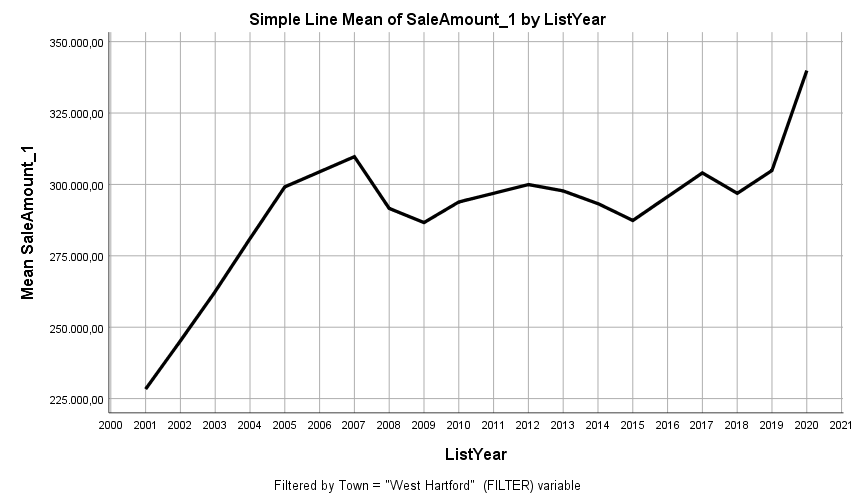
1. Kota Hartford



Interpretasi:

Grafik ini menunjukkan tren rata-rata nilai penjualan properti di Kota Hartford pada tahun 2001 – 2020. Pada tahun 2001, rata-rata nilai penjualan properti di Hartford berada di kisaran 110.000 dolar. Angka ini kemudian menunjukkan peningkatan dan mencapai puncaknya pada tahun 2005 dengan nilai mendekati 200.000 dolar. Kenaikan ini mencerminkan kondisi pasar properti yang sedang tumbuh. Namun, setelah tahun 2005, nilai penjualan mulai menurun tajam, terutama dari tahun 2008 hingga 2009, yang kemungkinan merupakan dampak dari krisis keuangan global. Pada tahun 2009, rata-rata nilai penjualan turun ke kisaran 130.000 dolar dan relatif stabil hingga tahun 2015. Pemulihan pasar mulai terlihat pada tahun 2016, di mana nilai rata-rata penjualan mulai meningkat secara bertahap setiap tahunnya. Kenaikan signifikan terlihat pada tahun 2020, dengan rata-rata nilai penjualan mencapai angka 200.000 dolar yang menunjukkan pemulihan pasar yang kuat dan peningkatan minat terhadap properti di Hartford.

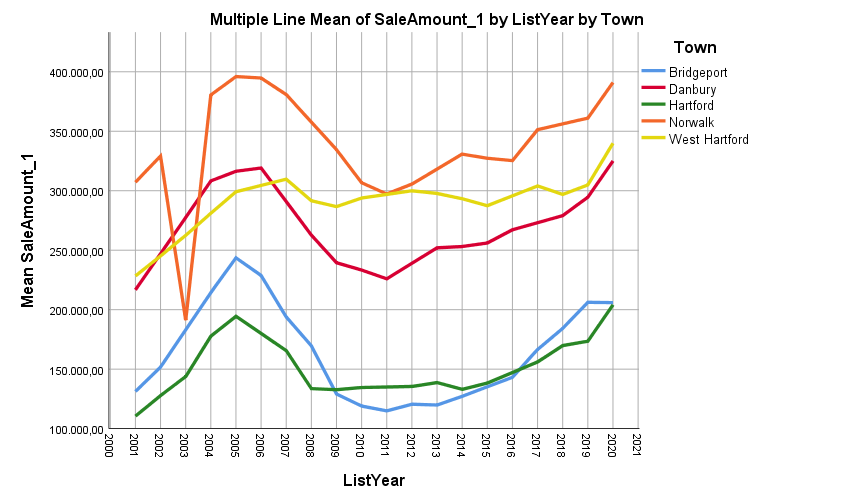
1. Kota West Hartford



Interpretasi:

Grafik ini menunjukkan tren rata-rata nilai penjualan properti di Kota West Hartford pada tahun 2001 – 2020. Pada tahun 2001, rata-rata nilai penjualan berada di kisaran 225.000 dolar. Selama lima tahun berikutnya, nilai ini meningkat secara konsisten dan signifikan, mencapai sekitar 310.000 dolar pada tahun 2006. Kenaikan ini mencerminkan kondisi pasar properti yang sedang tumbuh. Pada tahun 2007 hingga 2008, terlihat penurunan pada nilai rata-rata penjualan, tetapi tidak se-ekstrem kota-kota lain. Ini menunjukkan bahwa pasar properti West Hartford memiliki ketahanan yang relatif baik terhadap gejolak ekonomi. Nilai penjualan mulai naik secara bertahap dari 2010 hingga 2013. Meskipun terdapat fluktuasi antara tahun 2014 hingga 2019, pergerakan harga tetap berada dalam rentang yang relatif tinggi, antara 285.000 hingga 310.000 dolar. Kenaikan yang tajam terlihat pada tahun 2020, di mana rata-rata nilai penjualan melonjak hingga menembus angka 340.000 dolar. Lonjakan ini kemungkinan mencerminkan dampak dari tren pasar properti selama pandemi COVID-19, di mana permintaan rumah tinggal meningkat tajam di banyak wilayah pinggiran.

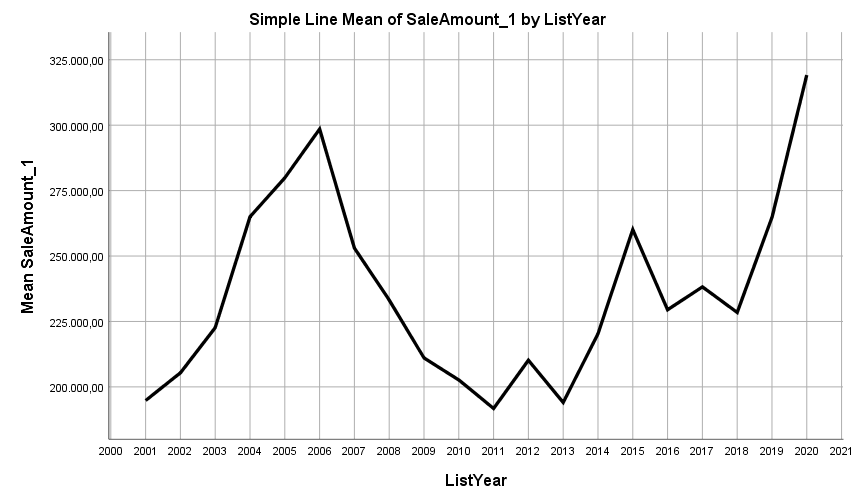
1. Visualisasi rata-rata penjualan properti di semua kota tahun 2001 – 2020



Interpretasi:

Grafik tersebut menunjukkan perbandingan rata-rata penjualan properti di Kota Bridgeport, Danbury, Norwalk, Hartford, dan West Hartford pada tahun 2001 – 2020. Secara umum, terlihat adanya pola fluktuasi yang serupa di hampir semua kota, dengan kecenderungan peningkatan nilai penjualan pada awal 2000-an, mencapai puncak pada tahun 2005 – 2006, penurunan penjualan sekitar tahun 2008 hingga awal 2010-an, dan pemulihan yang berlangsung secara bertahap pada tahun berikutnya. Kota Bridgeport, Danbury, dan Hartford memiliki pola rata-rata penjualan yang hampir sama. Kota Norwalk memiliki rata-rata penjualan properti yang relatif tinggi dibandingkan kota lain, meskipun mengalami penurunan tajam di tahun 2003 yang disebabkan oleh keterbatasan data. Kota West Hartford memiliki rata-rata yang cukup stabil dengan fluktuasi yang tidak terlalu ekstrem. Meskipun tiap kota memiliki dinamika tersendiri, pola yang terbentuk menunjukkan adanya pengaruh faktor eksternal, seperti krisis keuangan global. Menjelang akhir periode, khususnya tahun 2019 dan 2020, sebagian besar kota mengalami peningkatan rata-rata penjualan yang cukup signifikan. Hal ini mencerminkan adanya perkembangan pasar properti di kelima kota tersebut.

1. Visualisasi rata-rata penjualan properti tahun 2001 – 2020

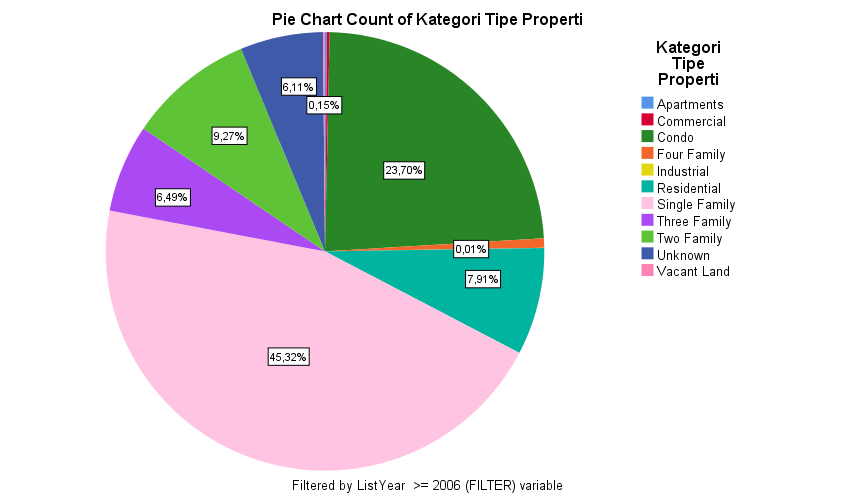


Interpretasi:

Grafik ini menunjukkan tren rata-rata nilai penjualan properti di Amerika Serikat pada tahun 2000 – 2020. Pada tahun 2001, nilai rata-rata penjualan berada di kisaran 195.000 dolar. Nilai ini mengalami kenaikan hingga mencapai puncaknya pada tahun 2006, mendekati angka 300.000 dolar. Namun, mulai tahun 2007 hingga 2011, terjadi penurunan signifikan dalam nilai penjualan yang disebabkan oleh krisis ekonomi global. Setelah periode krisis, grafik memperlihatkan pergerakan yang cenderung meningkat meski ada fluktuasi di beberapa tahun. Pada tahun 2016 hingga 2018 tampak sebagai periode transisi dengan variasi kenaikan dan penurunan yang tidak terlalu ekstrem. Pada tahun 2019 dan 2020, grafik menunjukkan lonjakan yang cukup mencolok hingga mencapai hampir 325.000 dolar. Pada tahun 2020 penjualan properti juga tetap menunjukkan rata-rata yang tinggi, meskipun tahun tersebut merupakan awal masa pandemi COVID-19. Pandemi tersebut ternyata tidak langsung menurunkan nilai pasar properti di Amerika Serikat, utamanya di kelima kota tersebut.

1. Visualisasi tipe properti terjual pada tahun 2006 – 2020

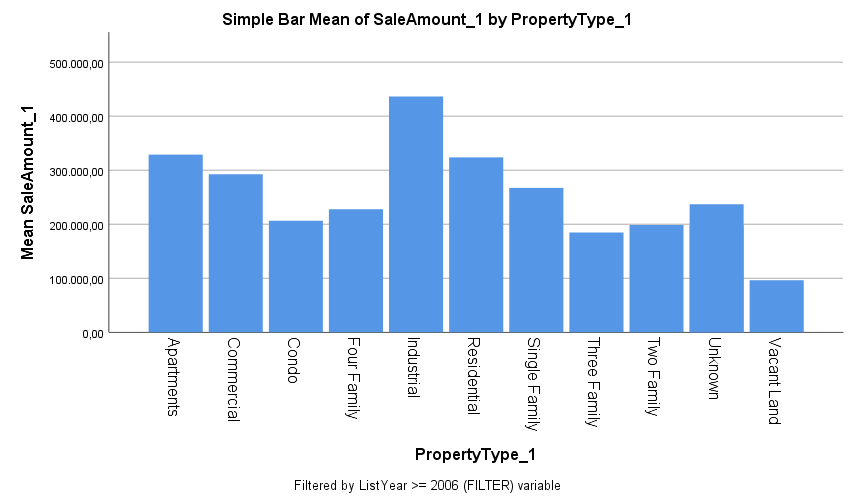
Analisis berdasarkan tipe properti dilakukan pada tahun 2006 hingga 2020 karena belum tersedia data untuk variabel ini pada tahun 2001 – 2005.



Interpretasi:

Grafik pie chart menunjukkan bahwa tipe properti yang paling dominan adalah Single Family (rumah tinggal satu keluarga) dengan proporsi sebesar 45,32% dari total data. Ini menunjukkan bahwa rumah tinggal satu keluarga menjadi properti yang umum dan paling banyak dimiliki di kelima kota tersebut. Tipe properti terbesar kedua adalah Condo (kondominium), sebesar 23,70% dari keseluruhan data. Ini menunjukkan bahwa unit hunian vertikal yang dimiliki secara individu juga cukup populer. Properti terbanyak selanjutnya adalah Two Family sebesar 9,27%, Residential sebesar 7,91%, dan Three Family sebesar 6,49%. Sebanyak 6,11% tidak diketahui jenis properti yang dimiliki dan jenis-jenis properti yang lain memiliki proporsi yang sangat kecil.

1. Visualisasi rata-rata penjualan berdasarkan tipe properti tahun 2006 – 2020



Interpretasi:

Berdasarkan grafik batang mengenai rata-rata nilai penjualan properti menurut tipe properti tahun 2006 – 2020, terlihat bahwa properti dengan tipe Industrial memiliki nilai penjualan rata-rata tertinggi dibandingkan tipe-tipe lainnya. Hal ini menunjukkan bahwa properti industri cenderung memiliki nilai yang lebih besar dibandingkan tipe lain. Sementara itu, tipe Residential, Apartments, dan Commercial juga menunjukkan nilai penjualan rata-rata yang relatif tinggi, mencerminkan permintaan yang cukup stabil dan potensial untuk kategori hunian dan komersial. Di sisi lain, tipe Single Family, Two Family, dan Three Family berada pada kisaran nilai menengah ke bawah, yang mencerminkan sifat pasar hunian yang lebih terjangkau. Tipe properti dengan rata-rata nilai penjualan paling rendah adalah Vacant Land atau tanah kosong. Nilai yang lebih rendah ini cukup wajar karena belum terdapat bangunan atau fungsi spesifik pada lahan tersebut. Selanjutnya, kategori *Unknown* menunjukkan rata-rata nilai penjualan yang cukup tinggi, kemungkinan karena adanya beberapa properti bernilai tinggi yang tidak tercatat dengan benar.

1. Uji kenormalan Kolmogorov-Smirnov

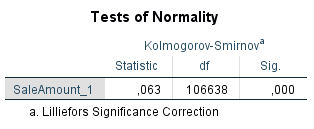
Tingkat signifikansi yang digunakan 5%

Hipotesis

Wilayah kritis

Tolak ketika

Hasil pengujian



Keputusan: Karena maka tolak

Kesimpulan: Dengan tingkat signifikansi 5%, cukup bukti untuk menyatakan bahwa data penjualan properti tidak berdistribusi normal.

1. Untuk mengetahui apakah terdapat perbedaan rata-rata penjualan properti di setiap kota, dapat dilakukan uji Kruskall-Wallis. Hal ini karena asumsi data berdistribusi normal tidak dapat dipenuhi sehingga tidak dapat menggunakan uji ANOVA.

Uji Kruskal Wallis

Tingkat signifikansi = 5%

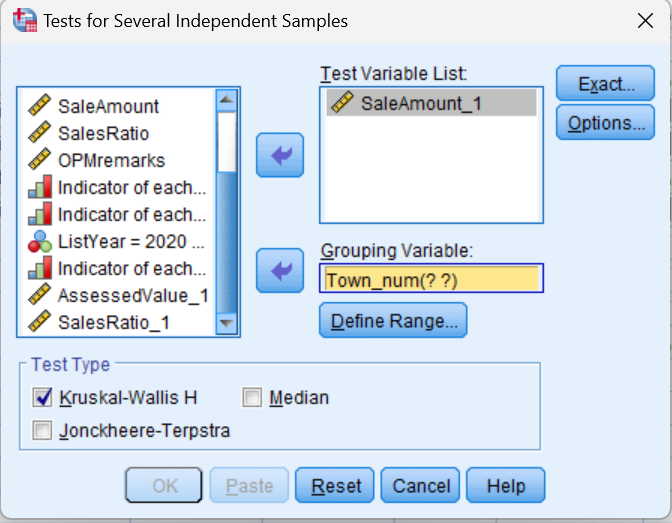
Hipotesis:

Langkah pengujian

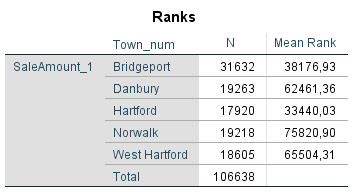
Ubah variabel Town menjadi bertipe numerik dengan label.

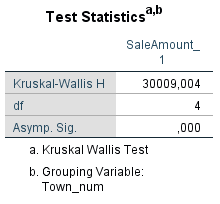
Klik Analyze 🡪 Non Parametric Test 🡪 Legacy Dialogs 🡪 K independent sample.

Masukkan Sale Amount ke Test Variabel List dan Town ke Grouping Variabel Klik Define Range 🡪 Max = 5 dan Min = 1 🡪 Ceklis Kruskal-Wallis H 🡪 OK.



Hasil pengujian





Keputusan: Karena maka tolak

Kesimpulan: Dengan tingkat signifikansi 5%, cukup bukti untuk menyatakan bahwa minimal terdapat satu kota yang memiliki rata-rata penjualan properti yang berbeda secara signifikan.

1. Untuk mengetahui hubungan antara nilai taksiran dengan harga penjualan properti, dapat dilakukan uji Korelasi Spearman. Hal ini karena asumsi data berdistribusi normal bivariat tidak dapat dipenuhi sehingga tidak dapat menggunakan uji Korelasi Pearson.

Uji Korelasi Spearman

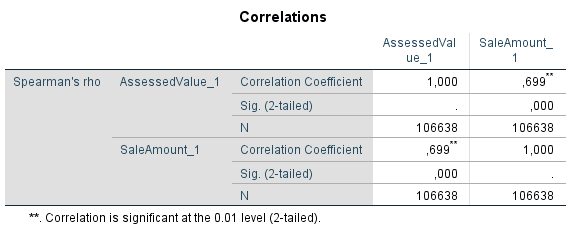
Tingkat signifikansi = 5%

Hipotesis:

(Tidak terdapat hubungan antara harga penjualan properti dengan nilai taksiran)

(Terdapat hubungan antara harga penjualan properti dengan nilai taksiran)

Hasil pengujian



Dari output di atas diperoleh hasil angka koefisien korelasi spearman sebesar 0,699. Artinya tingkat hubungan (korelasi) antara variabel nilai taksiran dan harga penjualan adalah kuat. Nilai positif menunjukkan bahwa arah hubungan kedua variabel tersebut searah artinya jika nilai taksiran tinggi maka harga penjualan properti juga tinggi. Dari hasil output diperolah nilai p-value (2 tailed) sebesar 0,000 yang berarti nilai ini lebih kecil dari tingkat signifikansi 5%. Oleh karena itu, dapat disimpulkan bahwa ada hubungan yang signifikan antara nilai taksiran dan harga penjualan.