



DATA ANALITIK

DATA ANALYSIS (UTS)

**ANALISIS DESKRIPTIF DAN VISUALISASI DATA BISNIS
MENGGUNAKAN TABLEAU DENGAN PENDEKATAN
CRISP-DM PADA DATASET VIDEO GAME SALES 2024**

Understanding data analysis helps businesses make informed decisions, optimize performance, and drive growth.



KELOMPOK 02



Our Team



11423006 - Samuel Alex Bonaparte Sirait



11423011 - Ruth Heppi Evelin Sinambela



11423028 Grace Advryanti Tampubolon



KELOMPOK 02

LATAR BELAKANG ANALISIS

Dalam era digital yang kompetitif, perusahaan game perlu memahami tren pasar dan preferensi pemain secara mendalam agar tetap unggul. Proyek ini menerapkan metodologi CRISP-DM dan Descriptive Analytics menggunakan Tableau untuk menganalisis dataset Video Game Sales 2024, yang mencakup data penjualan global berdasarkan genre, platform, wilayah, dan tahun rilis. Melalui visualisasi interaktif, diperoleh insight mengenai faktor-faktor yang memengaruhi keberhasilan game di pasar, seperti genre populer, wilayah dengan penjualan tertinggi, dan performa antar platform. Hasil analisis ini mendukung pengambilan keputusan berbasis data (data-driven decision making) dalam strategi produksi dan pemasaran industri game.



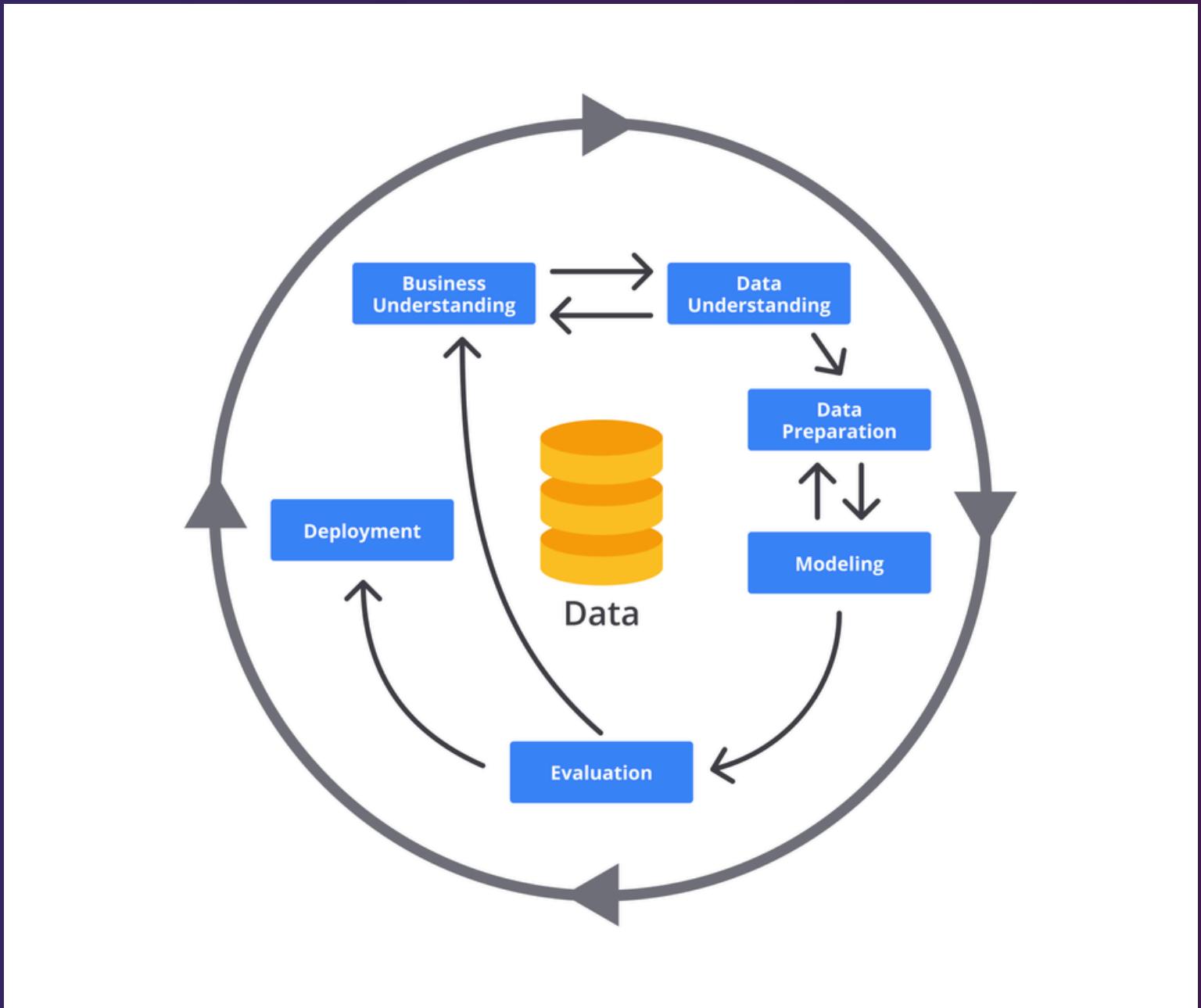
KELOMPOK 02



02



CRISPM-DM METHODOLOGY





BUSINESS UNDERSTANDING



- **Konteks Bisnis:**

- Industri video game berkembang pesat secara global, dengan persaingan yang tinggi dan preferensi pemain yang sangat beragam.
- Permasalahan Utama:
- Perusahaan game perlu memahami faktor-faktor yang memengaruhi penjualan agar strategi produksi, distribusi, dan pemasaran menjadi lebih efektif.

- **Tujuan Analisis:**

- Mengidentifikasi tren penjualan video game berdasarkan genre, console, tahun rilis untuk menemukan faktor yang berkontribusi terhadap keberhasilan penjualan.

- **Pertanyaan Analitik:**

- Genre apa yang memiliki rata-rata dan median penjualan tertinggi di pasar global?
- Bagaimana perkembangan rata-rata skor kritikus (AVG Critic Score) terhadap waktu rilis game dari tahun ke tahun?
- Genre atau console apa yang paling sering muncul dan paling populer dalam industri game?
- Seberapa besar variasi nilai kritik (Variance Critic Score) tiap tahun rilis, dan apa artinya terhadap konsistensi kualitas game yang dirilis?
- Apakah terdapat hubungan antara Critic Score dan Total Sales suatu game?

- **Hasil yang Diharapkan:**

- Insight bisnis yang membantu perusahaan dalam menyusun strategi berbasis data (data-driven decision making) di industri game.





DATA UNDERSTANDING

Sumber Data:

Dataset berasal dari Kaggle dengan judul “Video Game Sales 2024” oleh @asaniczka. Dataset ini berisi 39.798 baris dan 14 kolom yang mencakup informasi penjualan video game di berbagai wilayah dunia.

Atribut Utama:

- **Title:** Nama game
- **Console:** Platform atau konsol game (PS, Xbox, PC, Switch, dll.)
- **Release_Date:** Tanggal rilis game
- **Genre:** Jenis atau kategori game (Action, Adventure, Sports, dll.)
- **Publisher / Developer:** Pihak penerbit dan pengembang game
- **NA_Sales, PAL_Sales, JP_Sales, Other_Sales:** Penjualan per wilayah (dalam juta unit)
- **Total_Sales:** Total penjualan global
- **Critic_Score:** Nilai dari kritikus profesional



DATA UNDERSTANDING

Hasil Eksplorasi Awal:

- Terdapat missing value pada kolom Critic_Score, JP_Sales, dan Last_Update.

#	Critic Score
8.50000	8.80000
8.50000	8.50000
8.50000	8.50000
null	8.50000
8.50000	9.10000
8.50000	8.50000
8.60000	8.60000
8.30000	8.30000
null	9.00000
9.00000	9.50000
9.50000	9.50000

JP Sales
null
0.24000
0.14000
null
0.02000
0.06000
0.29000
0.35000
0.08000
0.08000

Last Update
null
null
3/21/2018
3/21/2018
10/11/2018
null
null
11/7/2018
9/12/2018





DATA UNDERSTANDING

Hasil Eksplorasi Awal:

- Ditemukan data duplikat pada beberapa judul game seperti “2 Games in 1 Double Pack: Finding Nemo / Monsters Inc.” dan “2 Days to Vegas”.

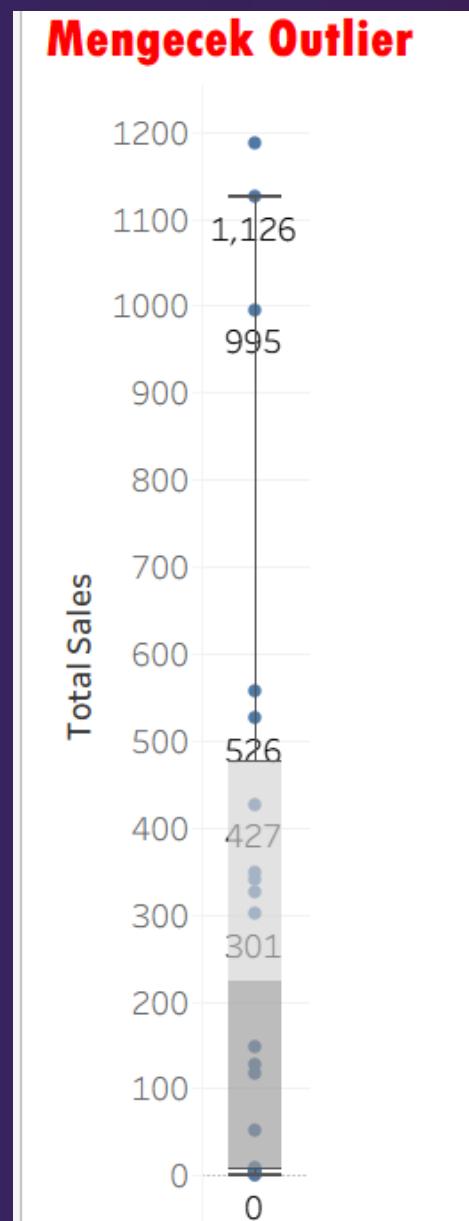
Mengecek Duplikat	
Title	Count
1/2 Summer +	1
2 Games in 1	1
Double Pack: Fin..	2
2 Games in 1	1
Double Pack: Hot..	1
2 Games in 1	1
Double Pack: The..	1
2 Games in 1:	1
Disney Princess ..	1
2 Games in 1:	1
Disney's Brother..	1
2 Games in 1:	1
Sonic Advance & ..	1
2 Games in 1:	1
Sonic Battle & Ch..	1
2 Games in 1:	1
Sonic Pinball Par..	1
2 Games in 1:	1
SpongeBob Squa..	2
2 in 1 Combo	1
Pack: Sonic Hero..	2
2 Games in 1	1



DATA UNDERSTANDING

Hasil Eksplorasi Awal:

- Beberapa game memiliki outlier penjualan ekstrem pada genre tertentu (terlihat di Boxplot Tableau). Dataset siap untuk tahap Data Preparation, setelah dilakukan pembersihan (data cleaning) dan pemeriksaan tipe data agar sesuai untuk proses analisis dan visualisasi di Tableau.



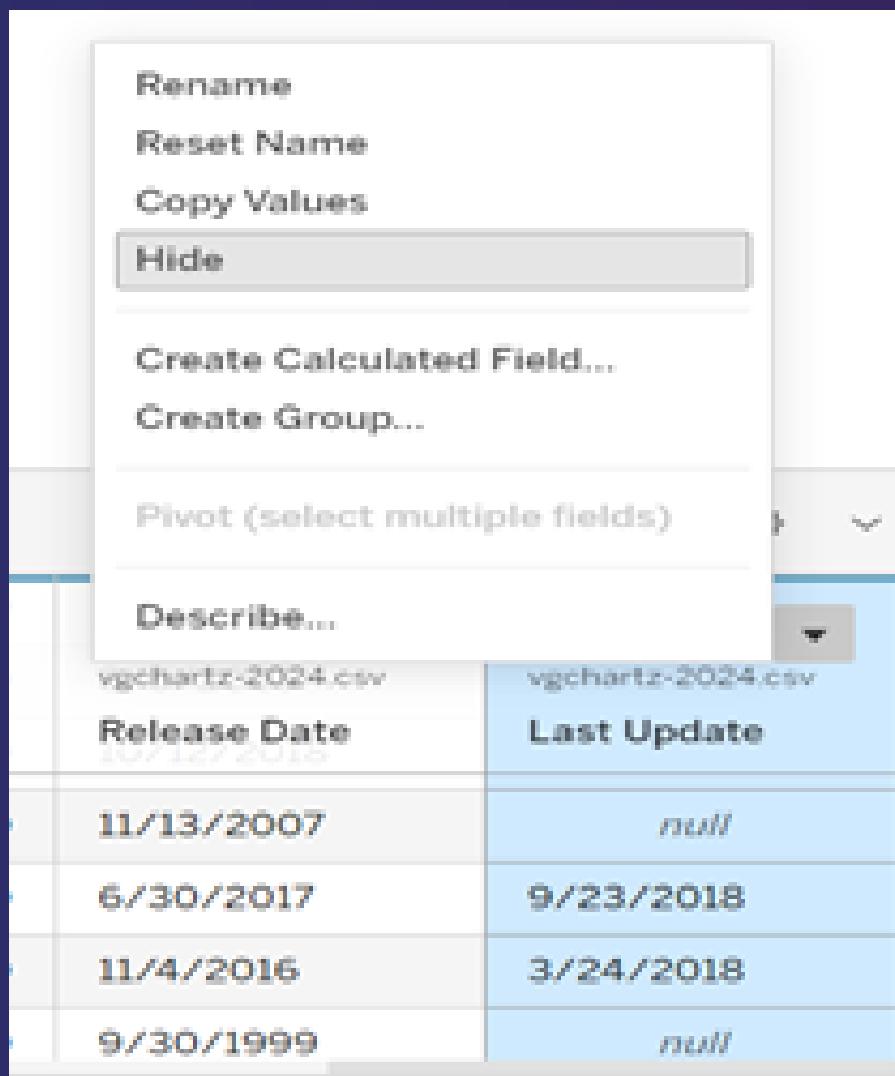


DATA PREPARATION

- **PEMBERSIHAN DATA :**



Menghapus Kolom yang Tidak Diperlukan
Beberapa kolom mungkin tidak relevan untuk analisis penjualan atau menyebabkan noise.



A screenshot of a data analysis interface showing a context menu open over a table. The menu includes options like 'Rename', 'Reset Name', 'Copy Values', and 'Hide'. Below the menu, there are buttons for 'Create Calculated Field...' and 'Create Group...'. The table has two columns: 'Release Date' and 'Last Update'. The 'Last Update' column contains several null values. The 'Last Update' column header is highlighted with a blue selection bar.

Release Date	Last Update
11/13/2007	null
6/30/2017	9/23/2018
11/4/2016	3/24/2018
9/30/1999	null

Last_Update kami sembunyikan karena tidak relevan dengan tujuan analisis. Kolom itu hanya berisi tanggal terakhir pembaruan data, bukan informasi yang menggambarkan performa atau karakteristik penjualan game. Selain itu, sebagian besar nilainya kosong, sehingga jika tetap digunakan bisa menurunkan kualitas data dan menambah noise dalam analisis.

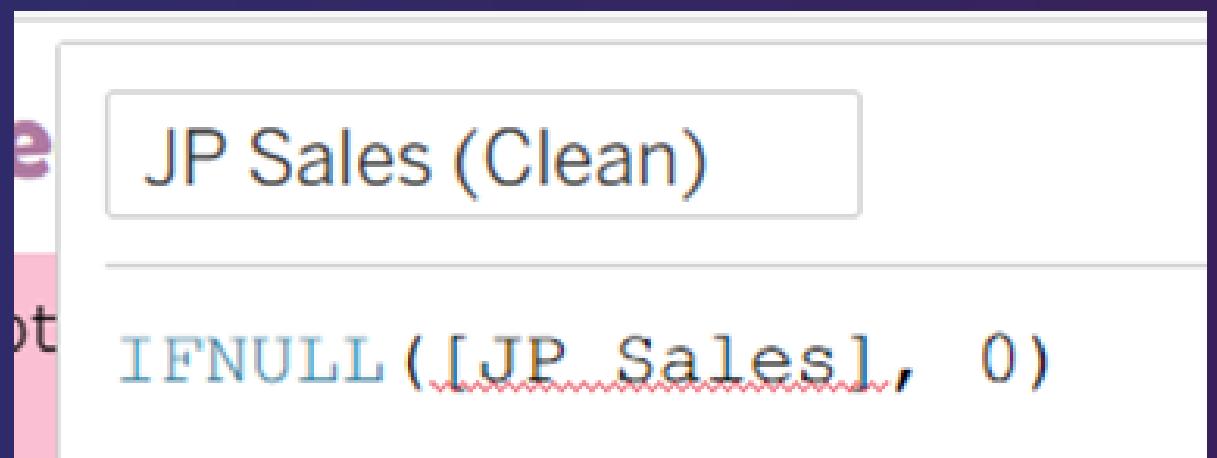




DATA PREPARATION

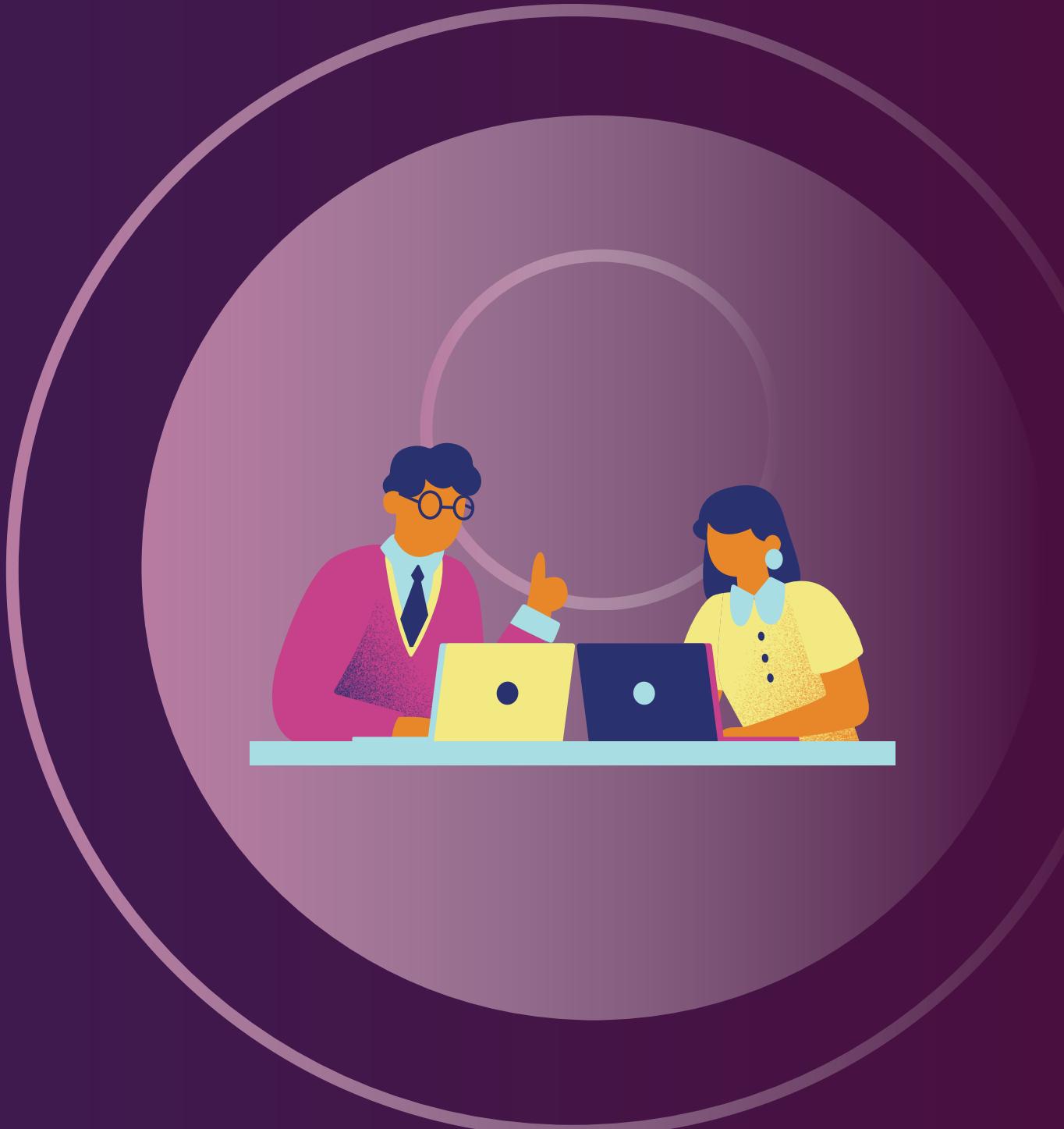
- Pembersihan Data :

-  Memeriksa dan Isi Nilai Kosong di Jp Sales dengan 0



```
JP Sales (Clean)
+-----+
| IFNULL([JP Sales], 0) |
+-----+
```

JP Sales (clean) ini yang akan digunakan untuk mengganti nilai null di JP Sales





DATA PREPARATION

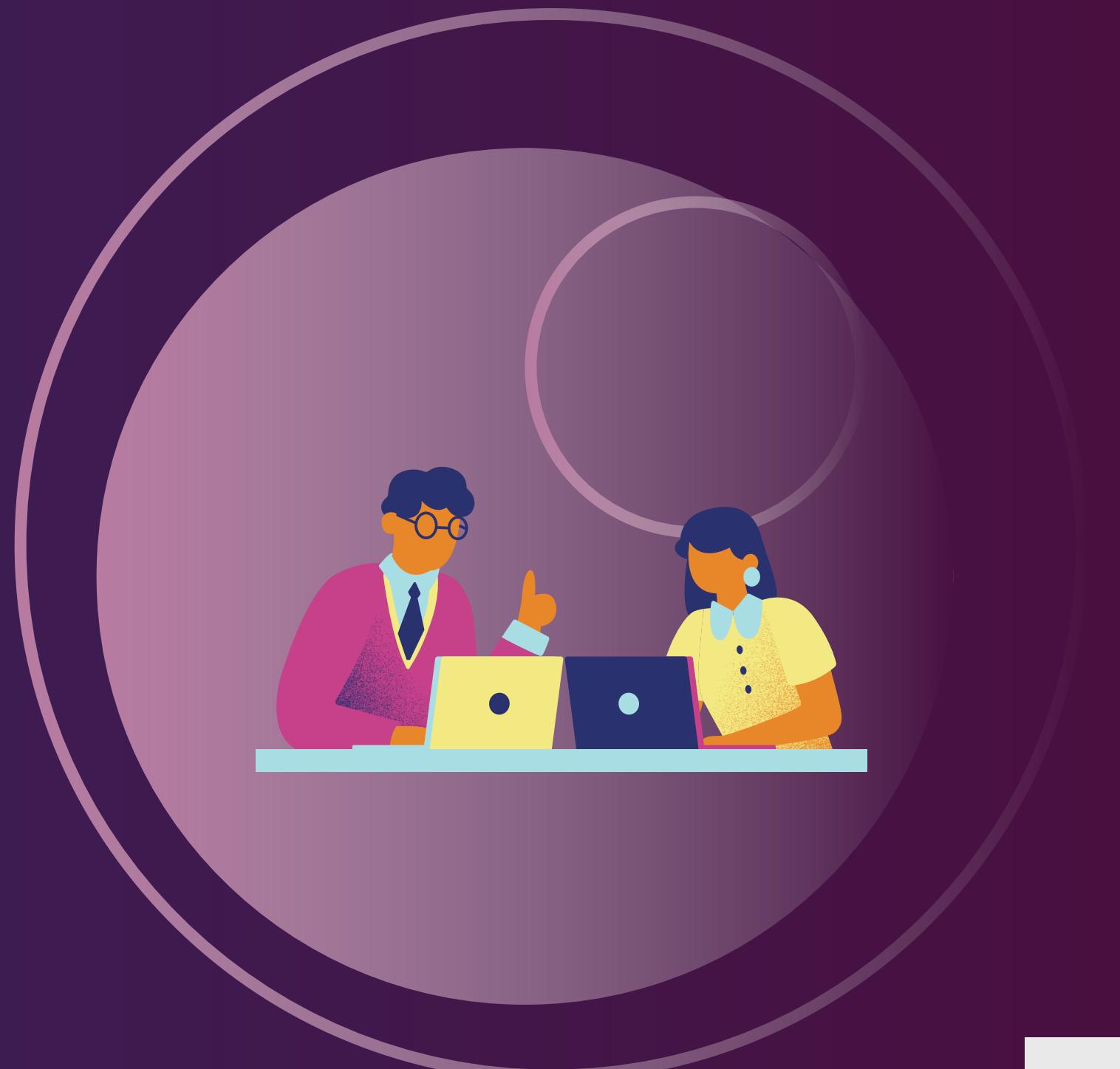
- **PEMBERSIHAN DATA :**



Memeriksa dan Isi Nilai Kosong di Critic Score dengan 0 atau “Unknown” dengan menggunakan calculated field → critic score (clean)

```
Critic Score (Clean)  
IFNULL([Critic Score], 0)
```

Critic Score (clean) ini yang akan digunakan untuk mengganti critic score yang berisi nilai null

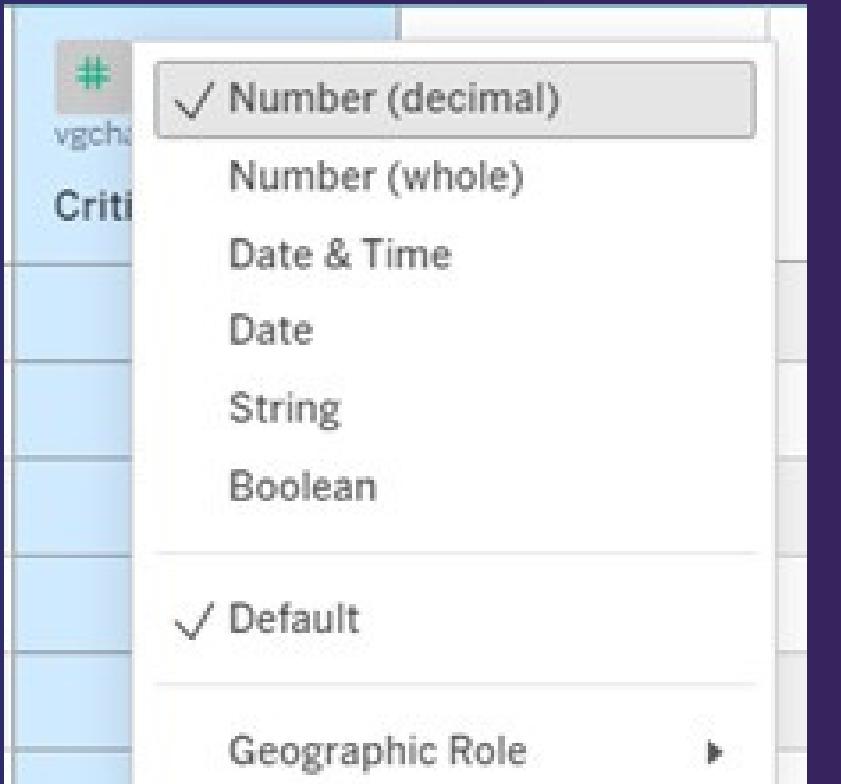




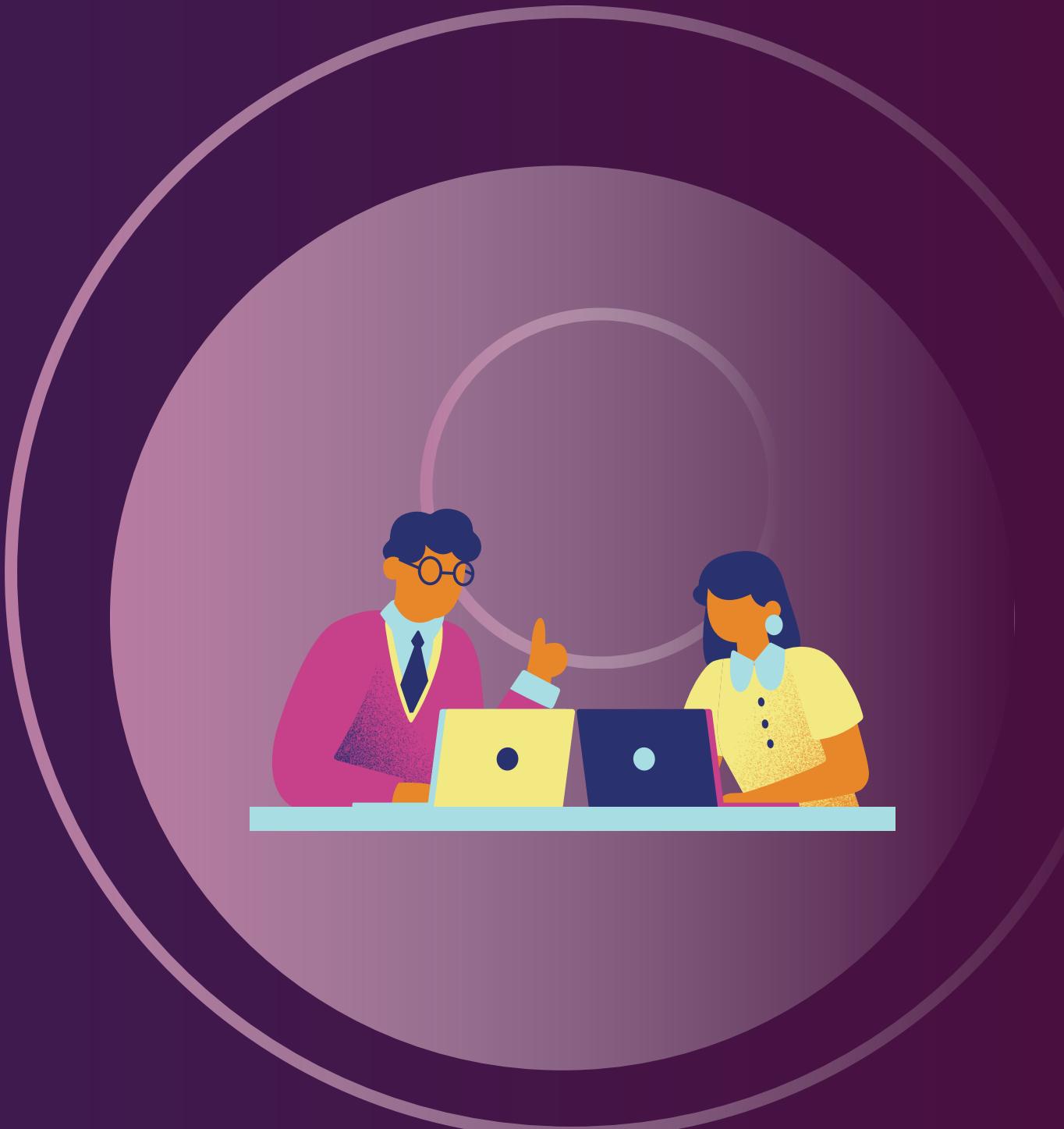
DATA PREPARATION

- **TRANSFORMATION DATA :**

- Menstandarkan Format Data



Menstandarkan format data dilakukan untuk memastikan setiap kolom memiliki tipe data yang sesuai, seperti angka pada kolom penjualan, teks pada kolom kategori

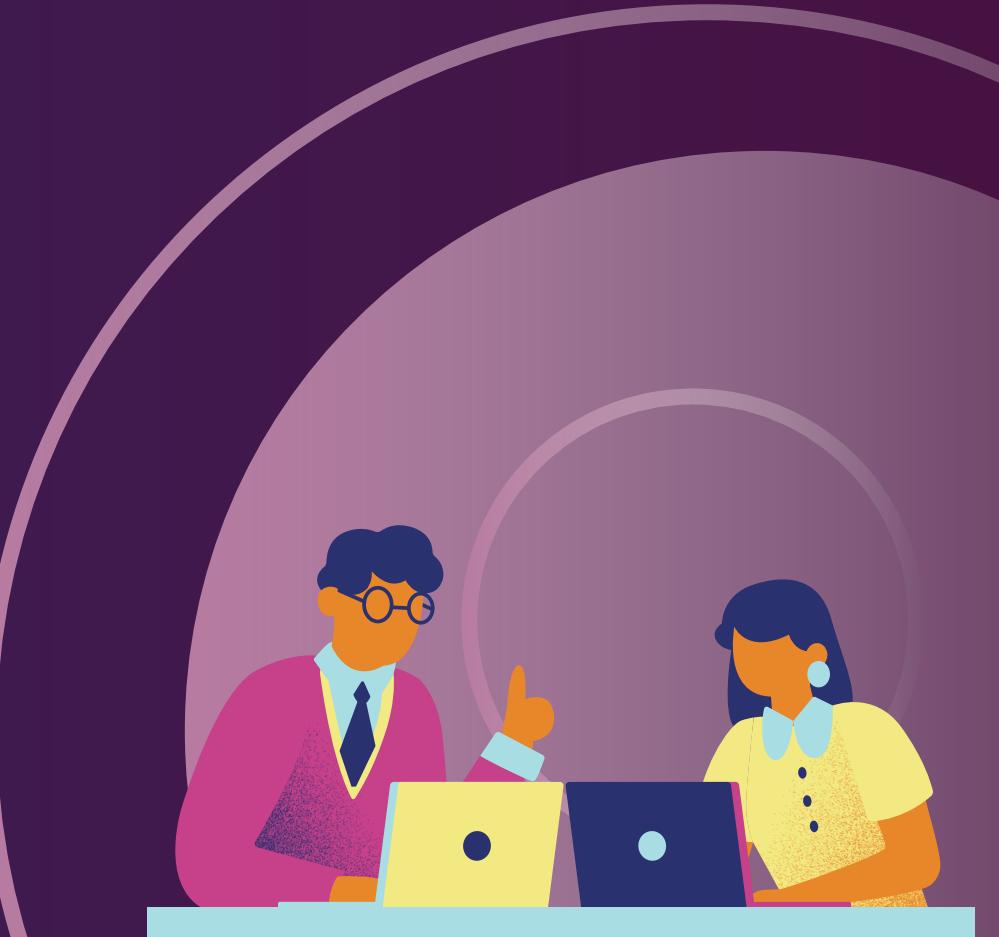




DATA PREPARATION

Pada tahap integrasi data, dilakukan pemeriksaan untuk memastikan apakah data perlu digabungkan dari beberapa sumber. Karena dataset yang digunakan hanya berasal dari satu sumber (satu file CSV), maka proses integrasi tidak diperlukan. Namun, pemeriksaan struktur dan konsistensi data tetap dilakukan untuk memastikan dataset siap dianalisis.

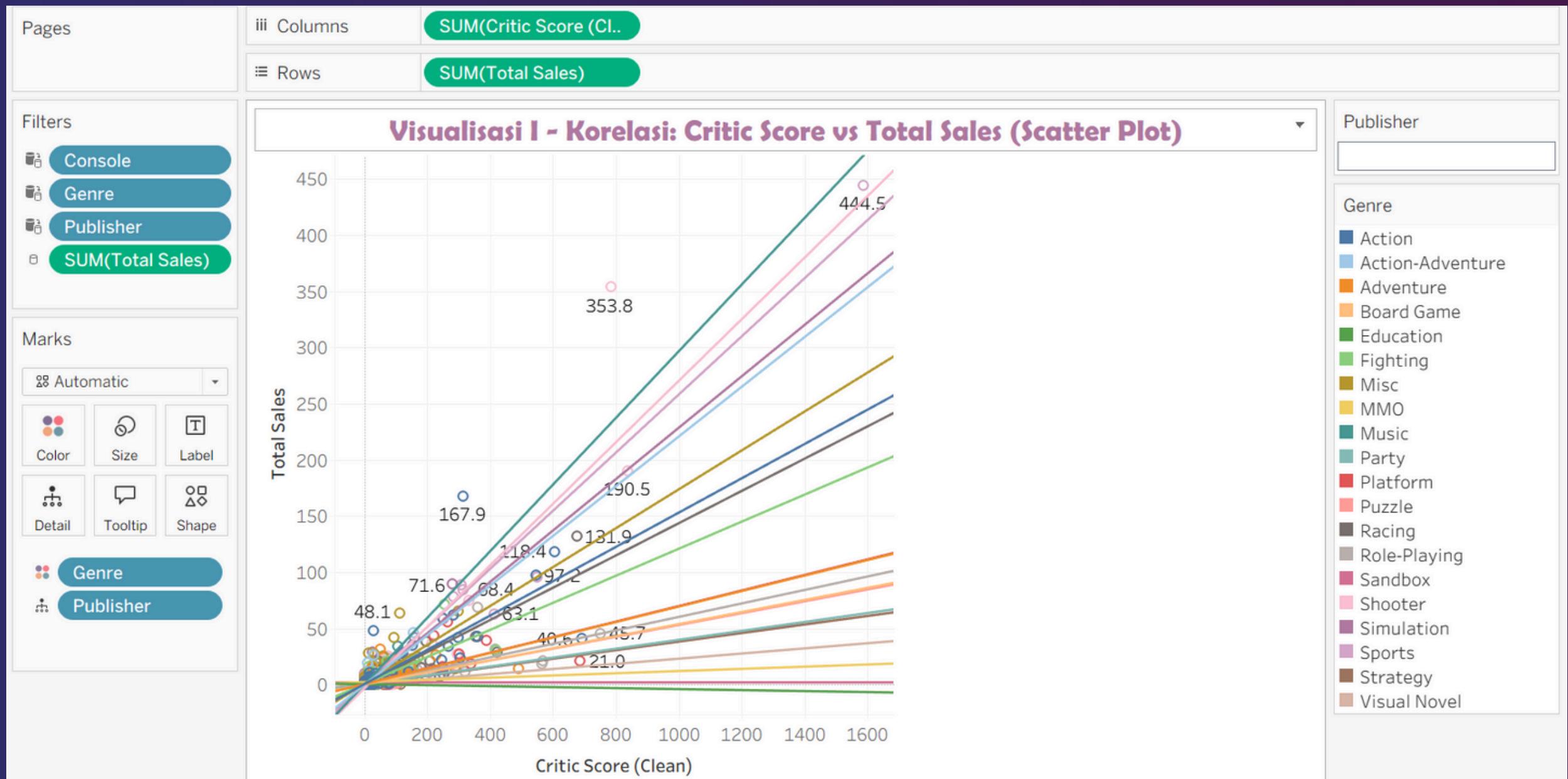
dan disimpan dalam file csv sebagai data yang sudah di bersihkan



DATA MODELING

Penerapan Descriptive Analytics:

- Analisis statistik sederhana: mean, median, mode, variance, correlation
 - Visualisasi I – Korelasi: Critic Score (Clean) vs Total Sales (Scatter Plot)



DATA MODELING

Penerapan Descriptive Analytics:

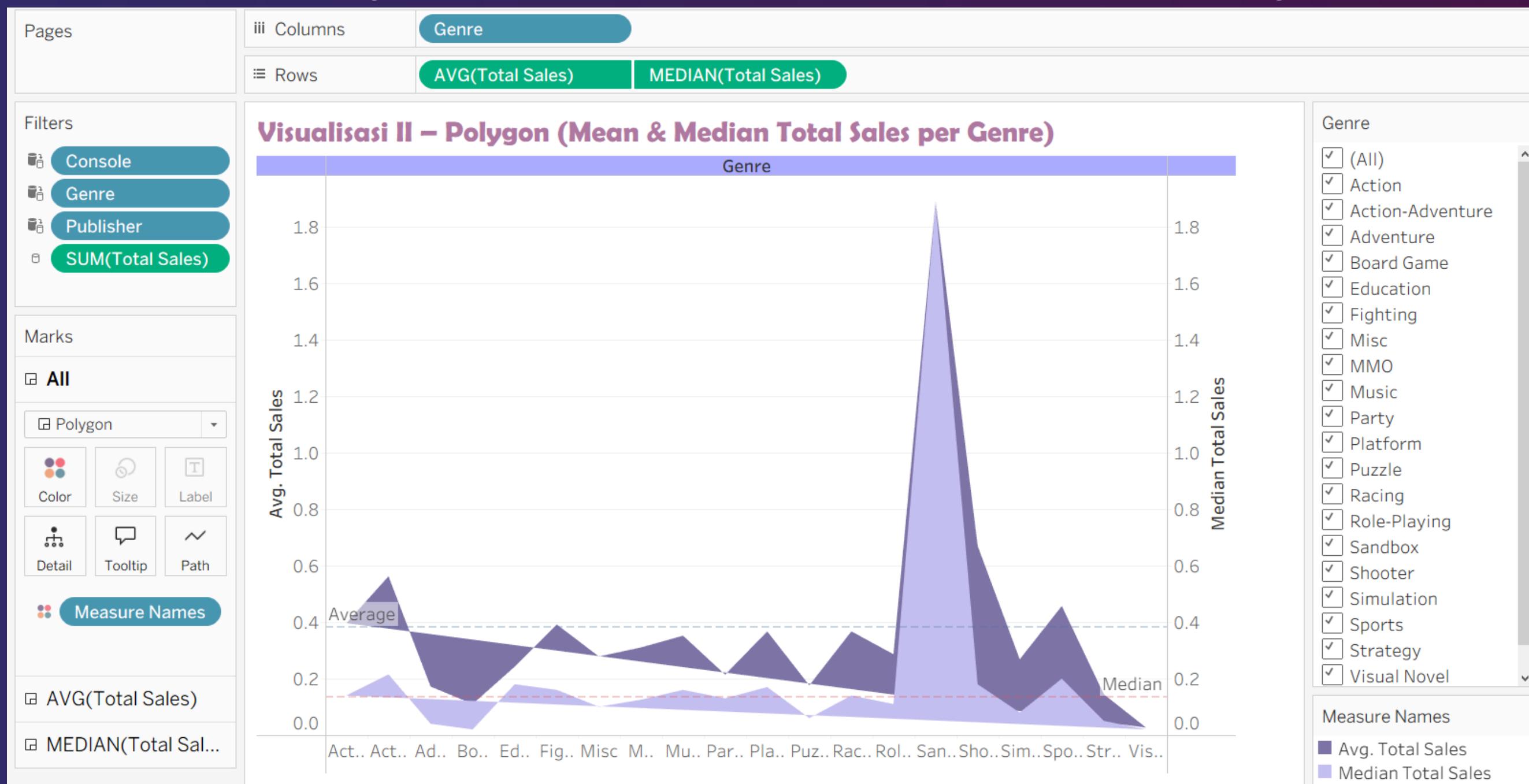
- Analisis statistik sederhana: mean, median, mode, variance, correlation
 - **Visualisasi I – Korelasi: Critic Score (Clean) vs Total Sales (Scatter Plot)**
- **Hasil Analisis Deskriptif**
 - Setelah data dibersihkan, dilakukan analisis korelasi antara Critic Score dan Total Sales menggunakan Scatter Plot di Tableau.
 - Setiap titik mewakili game, dengan warna berdasarkan genre.
 - Hasil menunjukkan pola korelasi positif: semakin tinggi skor kritik, semakin besar total penjualan.
 - Genre tertentu seperti Shooter dan Sports menunjukkan hubungan yang paling kuat.
 - Analisis ini menunjukkan pentingnya skor review terhadap performa penjualan game.

Scatter Plot sangat efektif untuk analisis korelasi karena memberikan gambaran langsung tentang hubungan antar variabel dan membantu dalam pengambilan keputusan berbasis data.

DATA MODELING

Penerapan Descriptive Analytics:

- Analisis statistik sederhana: mean, median, mode, variance, correlation
 - Visualisasi II – Polygon (Mean & Median Total Sales per Genre) dengan dual Axis



DATA MODELING

Penerapan Descriptive Analytics:

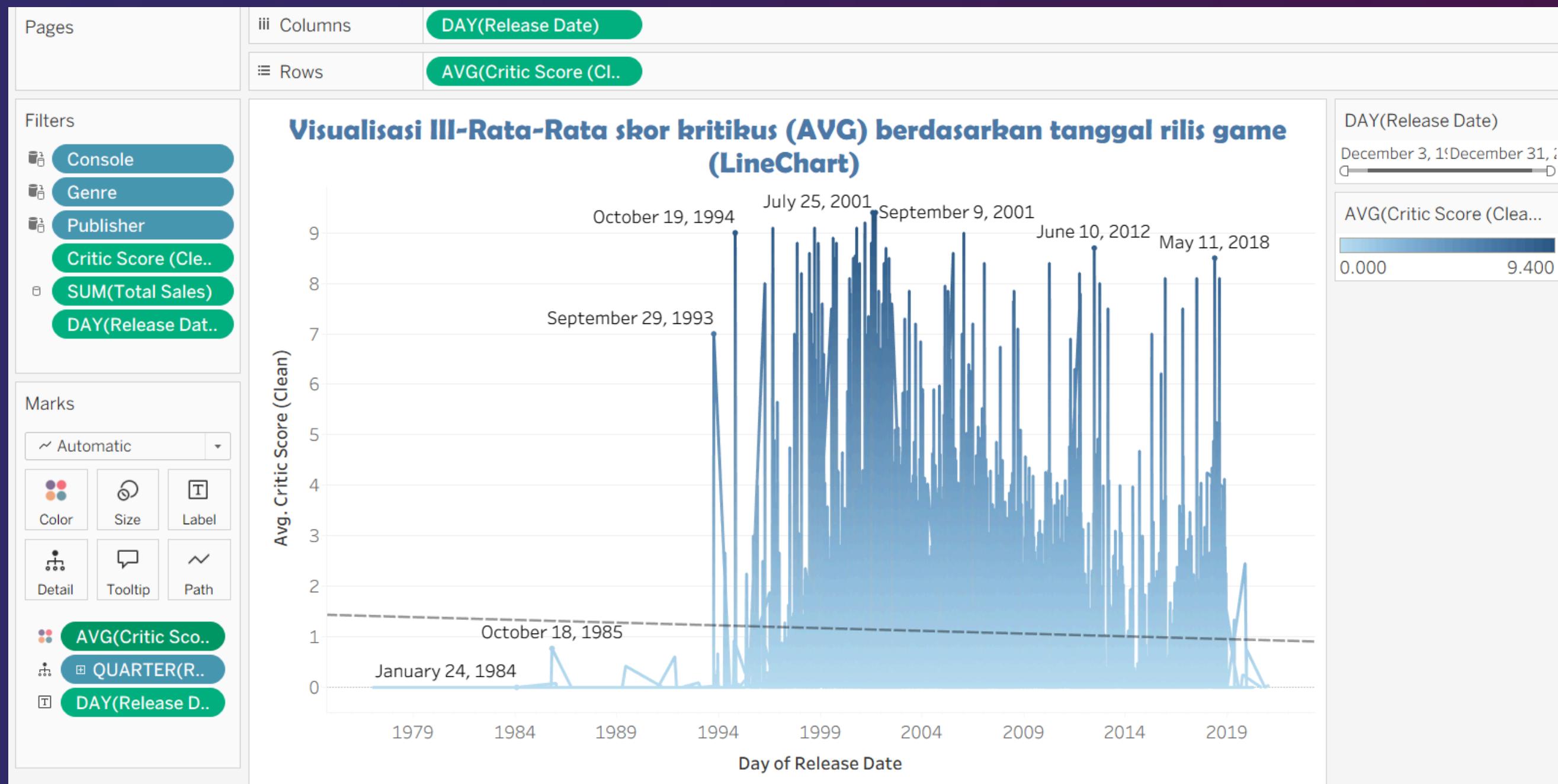
- Analisis statistik sederhana: mean, median, mode, variance, correlation
 - **Visualisasi II – Polygon (Mean & Median Total Sales per Genre)**
 - **Hasil Analisis Deskriptif**
 - Setelah data dibersihkan, dilakukan analisis perbandingan rata-rata (mean) dan nilai tengah (median) dari Total Sales tiap genre menggunakan Polygon Chart di Tableau.
 - Genre digunakan sebagai dimensi, sedangkan Total Sales sebagai measure utama.
 - Hasil menunjukkan bahwa genre Puzzle memiliki rata-rata penjualan tertinggi, tetapi median yang jauh lebih rendah, menandakan adanya game dengan penjualan sangat tinggi dibandingkan genre lainnya.
 - Genre lain seperti Action dan Sports menunjukkan perbedaan kecil antara mean dan median, yang berarti distribusi penjualannya lebih stabil.

Polygon Chart sangat efektif untuk analisis penjualan karena menyajikan informasi yang kaya dalam format yang ringkas dan mudah dibaca.

DATA MODELING

Penerapan Descriptive Analytics:

- Analisis statistik sederhana: mean, median, mode, variance, correlation
 - Visualisasi III-Rata-Rata skor kritikus (AVG) berdasarkan tanggal rilis game (LineChart)



DATA MODELING

Penerapan Descriptive Analytics:

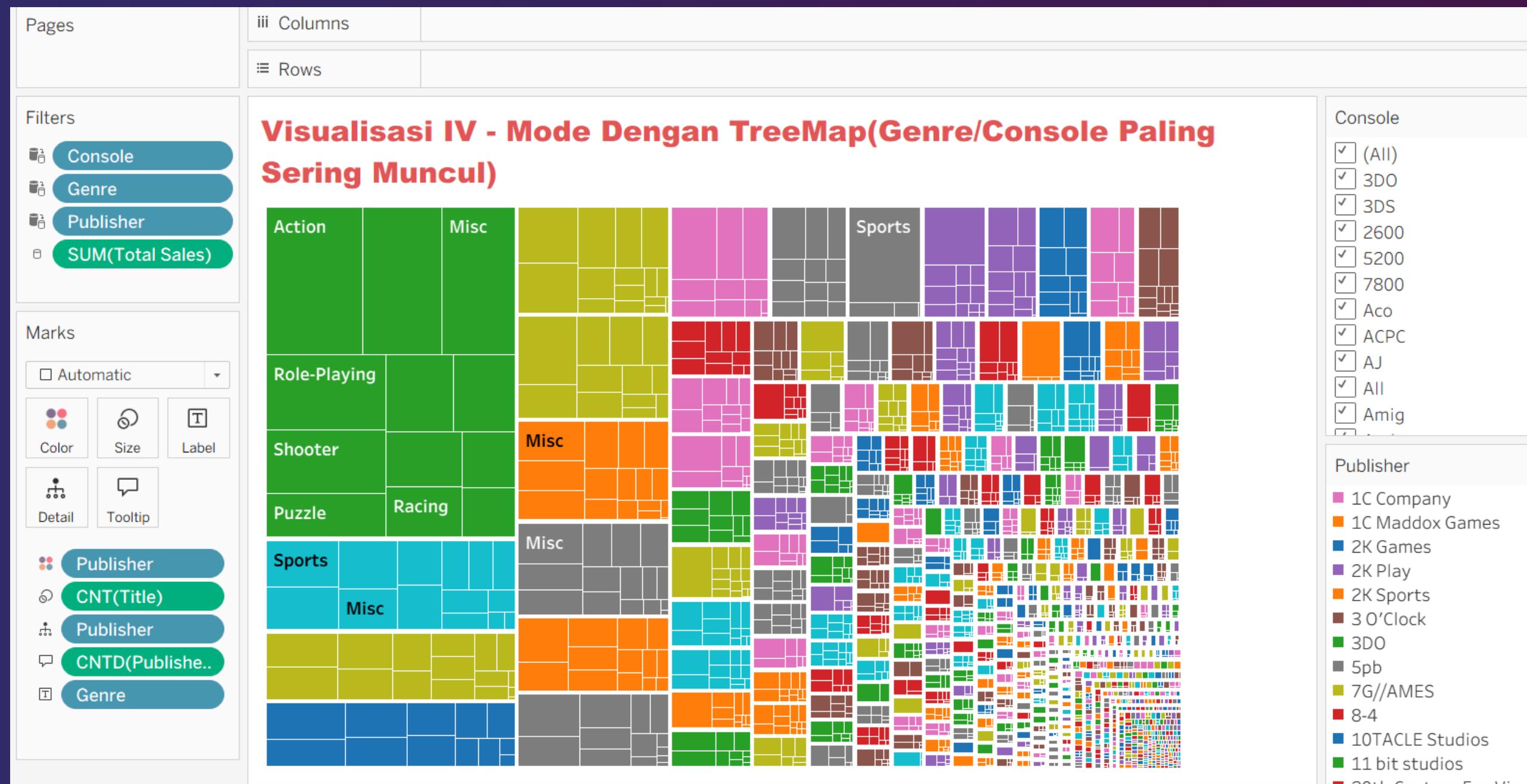
- Analisis statistik sederhana: mean, median, mode, variance, correlation
 - **Visualisasi III–Rata–Rata skor kritikus (AVG) berdasarkan tanggal rilis game (LineChart)**
 - **Hasil Analisis Deskriptif**
 - Analisis ini menggunakan Line Chart untuk melihat tren rata-rata skor kritikus berdasarkan tanggal rilis game.
 - Data Release Date dan Critic Score telah dibersihkan dari nilai kosong dan diformat dengan benar.
 - Visualisasi menunjukkan fluktuasi besar pada periode 1995–2015 dengan puncak sekitar tahun 2001 dan 2012.
 - Garis tren menunjukkan kecenderungan menurun, menandakan penilaian kritikus terhadap game secara rata-rata sedikit menurun dari tahun ke tahun.
 - Insight: standar penilaian yang meningkat dan ekspektasi pasar yang lebih tinggi dapat menjadi faktor penyebab tren ini.

LineChart sangat cocok untuk analisis mean karena mampu menyoroti perubahan nilai rata-rata secara kronologis dan membantu dalam pengambilan keputusan berbasis waktu.

DATA MODELING

Penerapan Descriptive Analytics:

- Analisis statistik sederhana: mean, median, mode, variance, correlation
 - Visualisasi IV - Mode Dengan TreeMap(Genre/Console Paling Sering Muncul)



DATA MODELING

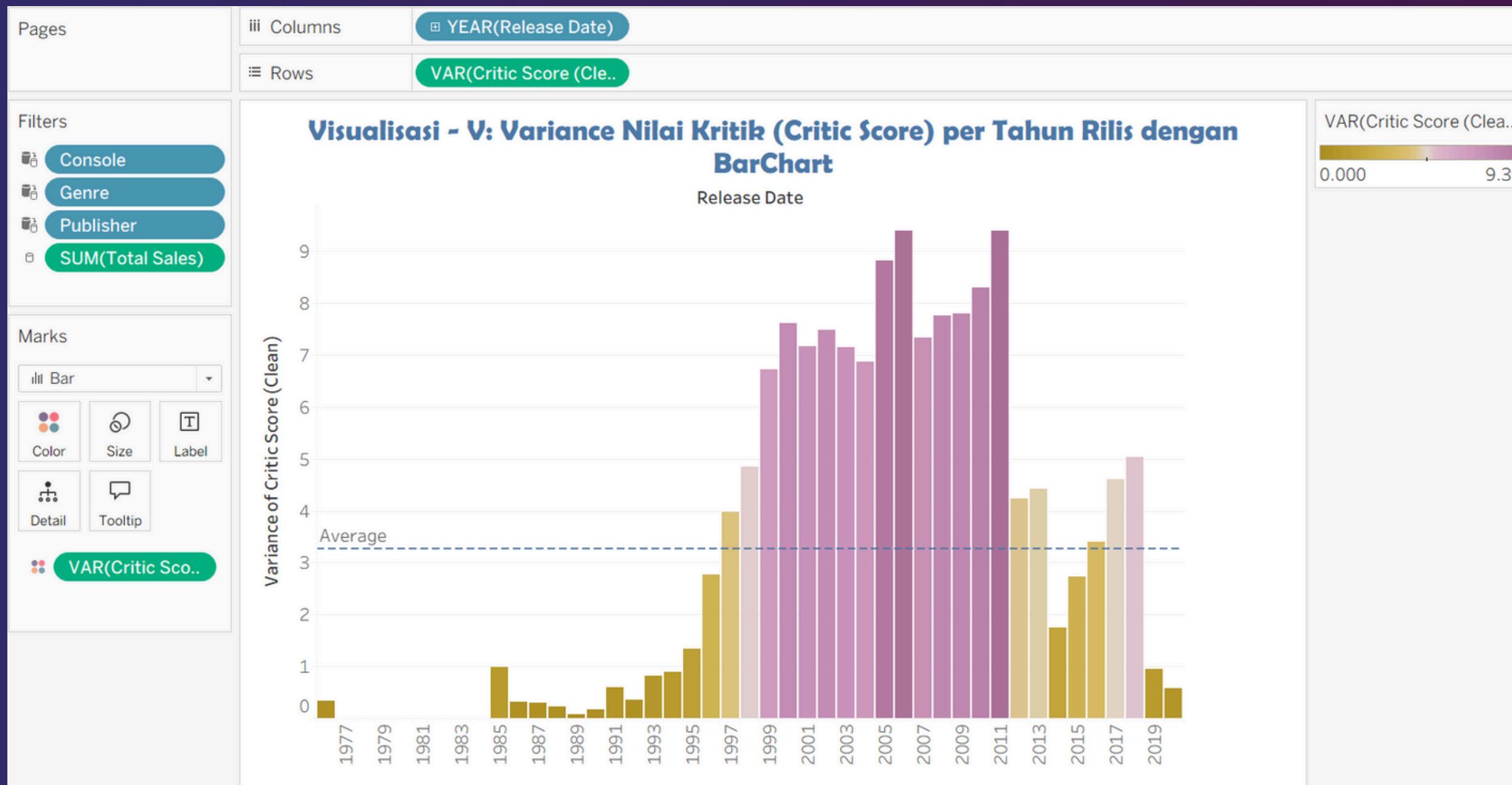
Penerapan Descriptive Analytics:

- Analisis statistik sederhana: mean, median, mode, variance, correlation
 - **Visualisasi III-Rata-Rata skor kritikus (AVG) berdasarkan tanggal rilis game (LineChart)**
 - **Hasil Analisis Deskriptif**
 - Setelah data dibersihkan dan diseragamkan, dilakukan analisis Mode untuk melihat kategori (genre atau console) yang paling sering muncul menggunakan TreeMap.
 - Ukuran kotak merepresentasikan jumlah kemunculan game per genre, sedangkan warna menunjukkan publisher.
 - Hasil menunjukkan bahwa genre Action, Role-Playing, dan Shooter merupakan kategori dengan jumlah game terbanyak.
 - **TreeMap dipilih karena mampu menampilkan banyak kategori sekaligus secara ringkas dan memudahkan identifikasi kategori dominan.**

DATA MODELING

Penerapan Descriptive Analytics:

- Analisis statistik sederhana: mean, median, mode, variance, correlation
 - Visualisasi - V: Variance Nilai Kritik (Critic Score) per Tahun Rilis dengan BarChart



DATA MODELING

Penerapan Descriptive Analytics:

- Analisis statistik sederhana: mean, median, mode, variance, correlation
 - **Visualisasi III-Rata-Rata skor kritikus (AVG) berdasarkan tanggal rilis game (LineChart)**
 - **Hasil Analisis Deskriptif**
 - Tahun dengan variansi tinggi menunjukkan bahwa penilaian kritikus terhadap game yang dirilis di tahun tersebut sangat beragam—ada game yang sangat dipuji dan ada yang dikritik keras.
 - Sebaliknya, tahun dengan variansi rendah menunjukkan penilaian yang lebih konsisten antar game.
 - Analisis ini dapat membantu publisher atau developer memahami tahun-tahun dengan standar penilaian yang lebih stabil atau lebih fluktuatif.

Alasan Pemilihan BarChart:

BarChart dipilih karena mampu menampilkan perbandingan antar tahun secara jelas dan langsung. Dengan tambahan garis rata-rata dan pewarnaan kategori, visualisasi ini memudahkan identifikasi tahun-tahun dengan penyebaran skor yang ekstrem dan membantu dalam analisis tren kualitas game dari perspektif kritikus.



EVALUASI DATA ANALYSIS

Evaluasi Hasil Analisis Keseluruhan :

Dari hasil evaluasi keseluruhan, kami menyimpulkan bahwa semua visualisasi yang kami buat sudah sesuai dengan tujuan proyek.

Kami menemukan bahwa game dengan skor tinggi dari kritikus cenderung memiliki penjualan yang lebih besar, terutama di genre populer seperti Action dan Sports.

Tren skor kritikus memang berfluktuasi, tetapi cenderung menurun karena standar penilaian yang makin ketat.

Selain itu, genre Action, Role-Playing, dan Shooter menjadi yang paling sering muncul di pasar, menandakan fokus industri pada genre dengan daya tarik luas.

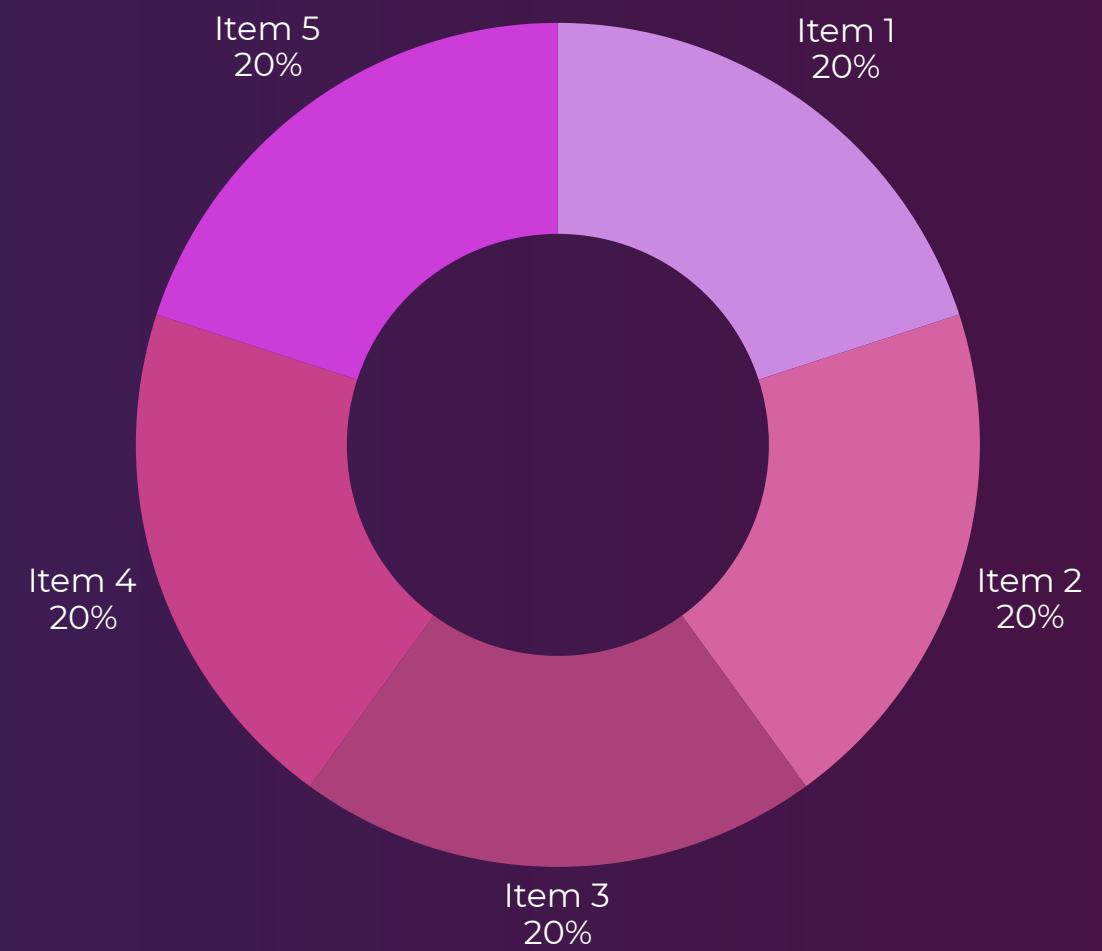
Secara umum, hasil analisis kami menunjukkan bahwa kualitas, genre, dan waktu rilis memiliki pengaruh besar terhadap performa game, dan dashboard ini mampu menggambarkan tren tersebut dengan jelas.



DATA ANALITIK

DEPLOYMENT

Dashboard Tableau



KELOMPOK 2





DATA ANALITIK

THANK YOU!

DATA ANALYSIS IS KEY TO
BUSINESS GROWTH AND
SUCCESS!



KELOMPOK 01

