

# **MILESTONE TUGAS BESAR I**

## **IF4061 VISUALISASI DATA**

### **Tujuan, Persiapan Data, dan Formulasi Pertanyaan**

Diajukan sebagai tugas Mata Kuliah IF4061 Visualisasi Data



Disusun Oleh:

Jason Rivalino	13521008
Muhammad Equilibrie Fajria	13521047
M. Farrel Danendra Rachim	13521048

**SEKOLAH TEKNIK ELEKTRO DAN INFORMATIKA**  
**INSTITUT TEKNOLOGI BANDUNG**  
**BANDUNG**  
**2025**

## DAFTAR ISI

<b>DAFTAR ISI</b>	<b>2</b>
<b>BAB I</b>	<b>3</b>
<b>DEFINISI TUJUAN DAN PARAMETER</b>	<b>3</b>
A. Topik Utama	3
B. Tujuan	3
C. Target Pengguna	3
D. Intensi	4
<b>BAB II</b>	<b>5</b>
<b>EKSPLORASI DAN PERSIAPAN DATA</b>	<b>5</b>
A. Data Acquisition	5
B. Data Examination	5
C. Penjelasan Tipe Data	6
D. Data Transforming	8
E. Consolidating	11
<b>BAB III</b>	<b>12</b>
<b>RUMUSAN PERTANYAAN</b>	<b>12</b>
A. Editorial Fokus dan Pendekatan Reasoning	12
B. Pembawaan Visualisasi Naratif	13

# **BAB I**

## **DEFINISI TUJUAN DAN PARAMETER**

### **A. Topik Utama**

Topik utama dari data yang akan divisualisasikan dalam pengerjaan Tugas Besar ini yaitu seputar analisis data terkait dengan kondisi kedatangan (*arrival*) dan penundaan (*delay*) yang terjadi di Amerika Serikat dalam kurun waktu 10 tahun mulai dari Agustus 2013 hingga Agustus 2023.

### **B. Tujuan**

Tujuan yang ingin dicapai dari proses visualisasi data dengan topik yang dipilih antara lain sebagai berikut.

- a. Memberikan informasi wawasan kepada pembaca mengenai kinerja dari berbagai maskapai penerbangan Amerika Serikat di berbagai bandara dengan berfokus pada pola kedatangan dan keterlambatan yang terjadi dalam kurun waktu satu dekade terakhir (Agustus 2013 - Agustus 2023).
- b. Menyampaikan analisis terkait faktor-faktor yang berpengaruh atas terjadinya penundaan dalam penerbangan seperti jenis maskapai penerbangan, cuaca, sistem kontrol lalu lintas udara (NAS), keamanan, dan keterlambatan yang terjadi pada pesawat sebelumnya
- c. Memberikan kesempatan kepada pembaca untuk dapat menentukan pemahaman dan pemberian saran yang lebih baik mengenai tantangan operasional dalam industri penerbangan.

### **C. Target Pengguna**

Terdapat tiga target pengguna yang ingin dituju dari proses visualisasi data yang akan dilakukan yaitu sebagai berikut.

- a. Penumpang pesawat secara umum

Dengan adanya visualisasi data terkait topik yang dibahas, informasi yang ada diharapkan dapat membantu penumpang pesawat secara umum untuk melakukan pemilihan terhadap maskapai atau bandara dengan tingkat keterlambatan yang lebih rendah.

b. Agen perencanaan perjalanan

Dengan adanya visualisasi data terkait topik yang dibahas, informasi yang ada diharapkan dapat memberikan wawasan kepada agen perencanaan perjalanan untuk memberikan rekomendasi saran terkait waktu penerbangan secara lebih akurat.

c. Ahli atau penggemar penerbangan

Dengan adanya visualisasi data terkait topik yang dibahas, informasi yang ada diharapkan dapat membantu para ahli ataupun penggemar penerbangan untuk menentukan solusi dari permasalahan penundaan serta membantu pengembangan dalam bidang industri penerbangan di Amerika Serikat yang lebih baik untuk kedepannya.

#### **D. Intensi**

Bagian intensi akan membahas mengenai fungsi yang ingin dipenuhi dan suasana yang ingin digambarkan dari proses visualisasi data yang akan dilakukan. Berikut merupakan rincian penjelasannya.

a. Fungsi

Fungsi yang akan dipenuhi dalam proses visualisasi data yaitu akan menekankan pada fungsi *Explanatory* yang lebih menekankan pada penjelasan informasi kepada pembaca melalui pemberian narasi yang bersifat spesifik dan terfokus secara jelas. Setelah selesai membaca visualisasi ini, pembaca diharapkan agar mengambil konklusi yang sama dengan konklusi yang diambil oleh pembuat visualisasi data tersebut.

b. Suasana (*Tone*)

Suasana yang ingin dibawakan dalam proses visualisasi data yaitu berupa suasana *pragmatic* yang membawa pembaca untuk dapat lebih berpikir secara analitis dan kritis terkait dengan visualisasi data yang ditampilkan.

## BAB II

### EKSPLORASI DAN PERSIAPAN DATA

#### A. *Data Acquisition*

Proses *Data Acquisition* dilakukan dengan melakukan pengambilan data awal yang berasal dari *dataset* yang bersifat *Open Source* sehingga telah tersedia secara langsung dan dapat diakses melalui platform Kaggle dengan link berikut.

Kaggle link: <https://www.kaggle.com/datasets/sriharshaedala/airline-delay>

Adapun *dataset* yang terdapat pada *Kaggle Link* ini juga mengambil referensi sumber data yang berasal dari pemerintah Amerika (*U.S. Government Works*), tepatnya pada instansi pemerintah bagian Bureau of Transportation Statistics, Airline Service Quality Performance 234 dengan referensi link pemerintah adalah berikut.

US Government link: <https://www.usa.gov/government-copyright>

#### B. *Data Examination*

Proses *Data Examination* dilakukan untuk menentukan seberapa tingkat kepercayaan pada *dataset* yang akan dianalisis dan divisualisasikan. Proses *Data Examination* akan meninjau dua hal utama yaitu aspek *Completeness* dan *Quality* dengan rincian penjelasan sebagai berikut.

##### a. *Completeness*

Pada aspek *Completeness*, beberapa hasil yang didapat dari pemeriksaan yang dilakukan terkait dengan kelengkapan kebutuhan data adalah sebagai berikut.

- i. Data yang ada sudah memiliki kelengkapan yang sesuai dengan setiap kategori atribut data yang diperlukan telah mencukupi untuk kebutuhan proses visualisasi data.
- ii. Data yang ada sudah mencakup periode waktu secara jelas dan lengkap dengan rentang waktu yang terdapat dalam data berada diantara Agustus 2013 hingga Agustus 2023.
- iii. Data yang ada sudah mencakup keseluruhan aspek variabel dengan total keseluruhan kolom pada data yaitu sebanyak 21 kolom.

- iv. Data yang ada sudah menyediakan total jumlah *records* yang sudah cukup untuk dianalisis secara menyeluruh dengan total sebanyak 171.666 baris data.

b. *Quality*

Pada aspek *Quality*, beberapa hasil yang didapat dari pemeriksaan yang dilakukan terkait dengan kelengkapan kebutuhan data adalah sebagai berikut.

- i. Data yang ada sudah cukup bagus dan tidak mengandung *error*.
- ii. Data yang ada sudah memberikan informasi yang jelas terkait dengan klasifikasi atribut yang dimiliki masing-masing.
- iii. Data yang ada secara umum sudah memiliki format penulisan yang terstruktur untuk semua *records* dan tidak ada permasalahan terkait dengan karakter data.
- iv. Masih terdapat *records* data yang mengandung *missing value* dengan beberapa atribut tertentu memiliki nilai yang bersifat NULL. Untuk rincian atribut dengan *missing value* yaitu terdapat 6 atribut dengan seluruh *records* terisi lengkap, 14 atribut dengan total 240 *records* yang mengalami *missing value*, dan 1 atribut dengan total 443 *records* yang mengalami *missing value*.
- v. Tidak ada *records* yang mengalami duplikat pada keseluruhan data yang ada.
- vi. Akurasi data sudah terukur dengan baik, meskipun ada beberapa data yang tergolong dalam *extreme outlier*.
- vii. Terdapat sekitar 19.000-22.000 nilai data yang termasuk *outlier* untuk setiap kolom jika dilakukan pengecekan dengan menggunakan metode IQR dan sekitar 1.000-3.000 nilai data *outlier* jika dilakukan pengecekan dengan menggunakan metode Z-Score. Walaupun begitu, hanya sekitar 100-200 data yang dapat dikatakan tergolong dalam *extreme outlier*, yaitu *outlier* yang memiliki nilai yang sangat tidak realistis dibandingkan *outlier* lainnya.

### C. Penjelasan Tipe Data

Berbagai atribut yang ada dalam data yang akan divisualisasikan memiliki jenis tipe dengan kelompok yang berbeda-beda dengan rincian pembagiannya adalah sebagai berikut.

a. *Categorical Nominal*

Beberapa atribut yang termasuk ke dalam tipe data ini adalah berikut.

- *Carrier*: kode maskapai penerbangan
- *Carrier\_name*: nama maskapai penerbangan

- *Airport*: kode bandara
- *Airport\_name*: nama bandara

b. *Categorical Ordinal*

Tidak ada atribut yang termasuk ke dalam tipe data ini.

c. *Quantitative (interval-scale)*

Beberapa atribut yang termasuk ke dalam tipe data ini adalah berikut.

- *Year*: data tahun penerbangan
- *Month*: data bulan penerbangan

d. *Quantitative (ratio-scale)*

Beberapa atribut yang termasuk ke dalam tipe data ini adalah berikut.

- *Arr\_flights*: banyaknya penerbangan kedatangan
- *Arr\_del15*: banyaknya penerbangan kedatangan yang mengalami keterlambatan (*delayed*) selama 15 menit atau lebih
- *Carrier\_ct*: jumlah keterlambatan selama 15 menit atau lebih akibat faktor maskapai penerbangan
- *Weather\_ct*: jumlah keterlambatan selama 15 menit atau lebih akibat faktor cuaca
- *Nas\_ct*: jumlah keterlambatan selama 15 menit atau lebih akibat faktor *National Airspace System* (NAS)
- *Security\_ct*: jumlah keterlambatan selama 15 menit atau lebih akibat faktor keamanan
- *Late\_aircraft\_ct*: jumlah keterlambatan selama 15 menit atau lebih akibat kedatangan pesawat yang terlambat
- *Arr\_cancelled*: jumlah penerbangan yang dibatalkan
- *Arr\_diverted*: jumlah penerbangan yang dialihkan
- *Arr\_delay*: banyaknya total menit dari penundaan penerbangan
- *Carrier\_delay*: total menit penundaan penerbangan akibat faktor maskapai penerbangan
- *Weather\_delay*: total menit penundaan penerbangan akibat faktor cuaca
- *Nas\_delay*: total menit penundaan penerbangan akibat faktor *National Airspace System* (NAS)
- *Security\_delay*: total menit penundaan penerbangan akibat faktor keamanan

- *Late\_aircraft\_delay*: total menit penundaan penerbangan akibat kedatangan pesawat yang terlambat

#### **D. Data Transforming**

##### **a. Transforming for Quality**

Proses *Transforming for Quality* dilakukan dengan tujuan untuk melakukan penyesuaian dan pembersihan terhadap data yang mengalami kesalahan. Pada data yang digunakan ini, terdapat sejumlah proses penyesuaian dan pembersihan yang dilakukan antara lain sebagai berikut.

- i. Melakukan penghapusan pada data yang mengalami duplikat. Dikarenakan tidak ada duplikat pada data yang digunakan, maka proses ini tidak perlu dilakukan.
- ii. Memproses beberapa data yang mengalami *missing value*. Pada data yang ada, dapat dilakukan penghapusan terhadap baris data yang mengalami *missing value* dikarenakan proporsi data yang memiliki *missing value* hanya sedikit yaitu sebanyak 443 *records* dibanding total keseluruhan data yaitu 171.666 *records* (~0.26%).
- iii. Data-data yang memiliki nilai *outlier* dibiarkan karena asumsi data yang diperoleh bernilai realistis dan proporsi data yang termasuk "*outlier*" sangat banyak sehingga tidak mungkin disingkirkan atau diganti nilainya.
- iv. Melakukan *handling* berupa perubahan terhadap kesalahan format penulisan karakter. Dikarenakan tidak ada kesalahan format penulisan karakter, maka proses ini tidak perlu dilakukan.

##### **b. Transforming for Analysis**

Proses *Transforming for Analysis* dilakukan dengan tujuan untuk mempersiapkan dan menyempurnakan data yang didapat untuk analisis dan visualisasi. Pada data yang digunakan ini, terdapat sejumlah proses penyempurnaan yang dilakukan antara lain sebagai berikut.

- i. Melakukan *parsing* variabel yang terdapat pada atribut kolom "*airport\_name*" dengan mengekstraksi nama kota dan nama *state* menjadi dua kolom baru yang terpisah dari nama bandara karena data pada kolom tersebut memiliki tiga variabel yang berbeda dengan format yang *parseable* yaitu {kota tempat lokasi bandara}, {*state* tempat lokasi bandara}: {nama bandara}. Setelah dilakukan



proses *parsing*, kolom "*airport\_name*" yang awalnya hanya satu saja akan dibagi menjadi tiga kolom yaitu "*airport\_city*", "*airport\_state*", dan "*airport\_name*". Berikut merupakan contoh implementasi terhadap datanya.

Contoh data: "Albany, GA: Southwest Georgia Regional".

Hasil parsing:

<i>airport_city</i>	<i>airport_state</i>	<i>airport_name</i>
Albany	GA	Southwest Georgia Regional

- ii. Melakukan penggabungan terhadap beberapa variabel untuk membentuk variabel yang baru. Berdasarkan data yang ada, proses ini tidak diperlukan dikarenakan variabel-variabel yang tersedia sudah cukup sesuai dalam memberikan informasi yang dibutuhkan.
- iii. Mengubah data yang bersifat kualitatif menjadi kode atau kata kunci. Berdasarkan data yang ada, beberapa atribut sudah memiliki isi data dalam bentuk kode atau kata kunci seperti "*carrier\_code*" ataupun "*airport\_code*" sehingga proses ini tidak diperlukan.
- iv. Menentukan nilai baru lainnya. Berdasarkan data yang ada, proses ini tidak diperlukan dikarenakan variabel-variabel yang tersedia sudah cukup sesuai dalam memberikan informasi yang dibutuhkan.
- v. Melakukan perubahan terhadap beberapa kolom data menjadi proporsi persentase. Kalkulasi dilakukan karena jika dilakukan agregasi terhadap nilai keseluruhan dari kolom-kolom tersebut, maka hasilnya akan sama dengan sebuah kolom data yang menampung nilai-nilai data tersebut. Dengan kata lain, ada beberapa atribut data yang menjadi sub kategori dari atribut data tertentu (kategori induk). Nilai kolom sebaiknya diwakilkan menjadi persentase (desimal) untuk kemudahan analisis dan visualisasi proporsi data yang nantinya akan membandingkan kuantitas setiap sub kategori tersebut (misal perbandingan penyebab penundaan penerbangan). Proses ini menjadi penting karena untuk nilai data yang termasuk ke dalam "kategori induk" memiliki nilai yang berbeda-beda untuk setiap baris data. Contoh kondisi perbaikan yang dilakukan adalah berikut.

- Atribut "*arr\_flights*" yang menandakan banyak kedatangan penerbangan suatu pesawat, memiliki subkategori "*arr\_del15*" (banyak kedatangan penerbangan yang tertunda selama 15 menit atau lebih), "*arr\_diverted*" (banyak penerbangan yang dialihkan) dan "*arr\_cancelled*" (banyak penerbangan yang dibatalkan). Oleh karena itu, perlu dibuat kolom baru yang berisi persentase data penerbangan dari total penerbangan "*arr\_flights*", yakni "*arr\_del15\_percentage*", "*arr\_diverted\_percentage*", dan "*arr\_cancelled\_percentage*".
- Atribut "*arr\_del15*" memiliki subkategori banyaknya penerbangan yang ditunda berdasarkan penyebab penundaan penerbangan, yakni "*carrier\_ct*", "*weather\_ct*", "*nas\_ct*", "*security\_ct*", dan "*late\_aircraft\_ct*". Oleh karena itu, perlu dibuat kolom baru yang berisi persentase dari total penerbangan yang ditunda "*arr\_del15*", yakni "*carrier\_ct\_percentage*", "*weather\_ct\_percentage*", "*nas\_ct\_percentage*", "*security\_ct\_percentage*", dan "*late\_aircraft\_ct\_percentage*".
- Atribut "*arr\_delay*" yang menandakan total menit penundaan dari sebuah pesawat memiliki subkategori banyaknya menit penerbangan yang ditunda berdasarkan penyebab penundaan penerbangan, yakni "*carrier\_delay*", "*weather\_delay*", "*nas\_delay*", "*security\_delay*", dan "*late\_aircraft\_delay*". Oleh karena itu, perlu dibuat kolom baru yang berisi persentase dari total menit penundaan "*arr\_delay*", yakni "*carrier\_delay\_percentage*", "*weather\_delay\_percentage*", "*nas\_delay\_percentage*", "*security\_delay\_percentage*", dan "*late\_aircraft\_delay\_percentage*".

Contoh implementasi transformasi:

arr_delay	carrier_delay	weather_delay	nas_delay	security_delay	late_aircraft_delay
1604.00	207.00	0.00	1049.00	0.00	348.00

Agregasi nilai subkategori sama dengan nilai kategori induk "*arr\_delay*":

$$207 + 0 + 1049 + 0 + 348 = 1604.$$

Berdasarkan hasil ini, maka nilai kolom-kolom baru yang telah dikalkulasi proporsinya adalah:

- carrier\_delay\_percentage:  $\frac{207}{1604} * 100\% = 12.90\%$
- weather\_delay\_percentage:  $\frac{0}{1604} * 100\% = 0.00\%$
- nas\_delay\_percentage:  $\frac{1049}{1604} * 100\% = 65.40\%$
- security\_delay\_percentage:  $\frac{0}{1604} * 100\% = 0.00\%$
- late\_aircraft\_delay\_percentage:  $\frac{348}{1604} * 100\% = 21.70\%$

Dari hasil yang didapatkan, untuk kondisi dari keseluruhan tabel baru yang terbentuk adalah sebagai berikut.

arr_delay	carrier_delay	carrier_delay_percentage	weather_delay	weather_delay_percentage
1604.00	207.00	12.90	0.00	0.00

nas_delay	nas_delay_percentage	security_delay	security_delay_percentage	late_aircraft_delay	late_aircraft_delay_percentage
1049.00	65.40	0.00	0.00	348.00	21.70

- vi. Melakukan penghapusan pada data *redundant* yang tidak diperlukan. Berdasarkan data yang ada, proses ini tidak diperlukan dikarenakan variabel-variabel yang tersedia sudah cukup sesuai dalam memberikan informasi yang dibutuhkan.
- vii. Resolusi yang diterapkan pada data ini yaitu *Full Resolution* yang menampilkan data individual secara lengkap.

### E. Consolidating

Berdasarkan data yang digunakan, data hanya diambil dari satu sumber saja dan tidak mengambil referensi dari sumber lainnya sehingga proses *Consolidating* tidak diperlukan dalam visualisasi data yang akan dilakukan.

## BAB III

### RUMUSAN PERTANYAAN

#### A. Editorial Fokus dan Pendekatan *Reasoning*

Editorial fokus yang ingin dituju dari proses visualisasi data yang akan dilakukan yaitu untuk mengidentifikasi pola keterlambatan penerbangan di berbagai maskapai dan bandara di AS selama satu dekade terakhir (2013-2023) dengan fokus utama meliputi.

- a. Tren Keterlambatan Penerbangan
  - i. Bagaimana tren keterlambatan berubah dari tahun ke tahun?
  - ii. Apakah ada pola musiman dalam keterlambatan penerbangan?
- b. Penyebab Utama Keterlambatan
  - i. Seberapa besar masing-masing faktor (maskapai, cuaca, sistem lalu lintas udara (NAS), keamanan, atau keterlambatan pesawat sebelumnya) berkontribusi terhadap frekuensi dan durasi dari terjadinya keterlambatan?
- c. Performa Maskapai dan Bandara
  - i. Maskapai mana yang memiliki tingkat keterlambatan tertinggi dan terendah?
  - ii. Bandara mana yang paling sering mengalami keterlambatan dan apa penyebabnya?

Adapun pendekatan yang akan digunakan dalam proses visualisasi data yaitu dengan pendekatan yang bersifat induktif. Hal ini disebabkan karena analisis ini dimulai dari pengamatan data historis penerbangan (2013-2023) untuk menemukan pola keterlambatan dan penyebabnya. Teknik analisis dan visualisasi dipilih sedemikian rupa untuk menggali penemuan-penemuan menarik dengan mencari pertanyaan data yang terus menarik.

Pendekatan induktif ini memungkinkan pengambilan wawasan berbasis data tanpa asumsi awal yang kuat, sehingga hasil analisis lebih objektif dan akurat. Berikut merupakan rincian kondisi pendekatan yang akan dilakukan.

- a. Inferensi data ke pola: Dengan menganalisis ribuan catatan keterlambatan penerbangan, kita dapat mengidentifikasi tren, musim puncak keterlambatan, dan faktor dominan yang menyebabkannya.

- b. Inferensi pola ke kesimpulan: Setelah pola terbentuk, kita bisa menyimpulkan faktor utama yang menyebabkan keterlambatan dan bagaimana dampaknya terhadap industri penerbangan serta keputusan penumpang.

## **B. Pembawaan Visualisasi Naratif**

### **a. Genre**

Naratif genre yang akan dibawakan pada visualisasi data yang akan dilakukan adalah sebagai berikut.

- i. *Annotated Chart*: Membantu menerangkan tren-tren yang spesifik, seperti alasan-alasan utama penundaan dan maskapai apa saja yang memiliki penundaan yang paling banyak. Dengan *Annotated Chart*, pembaca mampu memahami beberapa insight penting yang menjadi pokok visualisasi data ini dengan memprioritaskan pembacaan aspek data yang superlatif atau signifikan (contoh, pada tahun berapa peningkatan penundaan penerbangan mencapai puncaknya).
- ii. *Partitioned Poster*: Melakukan segmentasi informasi secara visual menyerupai poster. Dalam konteks ini, informasi dipersembahkan dalam bentuk gambar bandara atau pesawat. Hal ini berguna agar pembaca dapat memahami secara intuitif konteks data apa yang dipresentasikan pada visualisasi ini, dalam hal ini mengenai penerbangan. Setelah mengetahui konteks, pembaca dapat mencerna *insight* kunci yang ada dengan mudah.

### **b. Visual Narrative Tactics**

Pola *Visual Narrative Tactics* yang akan dibawakan pada visualisasi data yang akan dilakukan adalah sebagai berikut.

- i. *Visual Structure*: Visualisasi ini menggunakan platform visual yang konsisten berupa pemberian subkategori penjelasan yang terstruktur dan perubahan informasi dari keadaan penerbangan yang jelas setiap periode tertentu. Platform ini berguna bagi pembaca untuk dapat menyerap informasi yang disediakan oleh infografis dengan mudah tanpa distraksi dari informasi redundan serta untuk memberikan wawasan yang lebih jelas dalam penjelajahan visualisasi.
- ii. *Highlighting*: pola ini diperlukan untuk memberikan penekanan pada aspek-aspek informasi yang berbeda pada infografis ini (seperti penyebab, bandara, maskapai, dan lain-lain). Dengan pemberian warna yang mencolok pada komponen yang

dianggap penting seperti penyebab penundaan, komponen penting tersebut akan terlihat secara lebih jelas dan mudah dibedakan dengan komponen lainnya. Selain itu, *framing* yang baik juga diperlukan untuk membuat *insight* kunci yang ditampilkan terlihat rapi. seperti "*Top 5 Airlines with the Most Delays*" dengan penekanan pada ukuran tulisan yang lebih besar.

- iii. *Messaging*: Visualisasi ini memanfaatkan *continuity editing* yang memastikan bahwa pembaca memahami bagaimana aspek-aspek penundaan yang ditampilkan berkaitan secara erat tanpa menyebabkan kebingungan. Hal ini dapat dilakukan dengan menghubungkan informasi antar aspek menggunakan visual *partitioned poster* yang menyatukan *insight* setiap aspek.

c. *Narrative Structure Tactics*

Pola *Narrative Structure Tactics* yang akan dibawakan pada visualisasi data yang akan dilakukan adalah sebagai berikut.

- i. *Ordering*: Jalur linear terstruktur yang telah ditentukan oleh pembuat visualisasi dapat memandu pembaca untuk mencerna informasi yang ditampilkan secara logikal. Urutan linear dapat dimulai dari kiri ke kanan pada visualisasi *partitioned poster*.
- ii. *Messaging: Annotations* dan *Headlines* digunakan untuk mengkomunikasikan observasi kunci dari visualisasi ini, termasuk anotasi dari *annotated chart* dan *headline* untuk setiap aspek data yang disampaikan.

d. Pola Naratif

Pola naratif yang tepat digunakan dalam contoh ini adalah *Martini Glass Structure*, yaitu struktur yang mempersembahkan pembaca pada jalur narasi yang sempit sebelum diikuti oleh eksplorasi secara bebas. Narasi yang mengawali visualisasi ini berisi pertanyaan dan *summary* singkat mengenai data yang divisualisasikan. Setelah memahami narasi tersebut, pembaca dapat mengeksplorasi visualisasi data secara lebih lanjut untuk mengetahui jawaban dari narasi tersebut. Hal ini penting karena pembaca perlu mengetahui informasi penting apa yang ingin disampaikan oleh pembuat visualisasi data tersebut, dan bersesuaian dengan fungsi visualisasi data yang bersifat *explanatory* (menjelaskan *insight* kunci yang perlu diketahui pembaca). Setelah selesai membaca

visualisasi ini, pembaca diharapkan agar mengambil konklusi yang sama dengan konklusi yang diambil oleh pembuat visualisasi data tersebut.