



NARRATING FINDING AND PROBLEM INTO TEXTUAL DESCRIPTION

Presentation by **Harish Muhammad**
Siti Romlah
Bima Fajar Miko

XER△TIC



Content

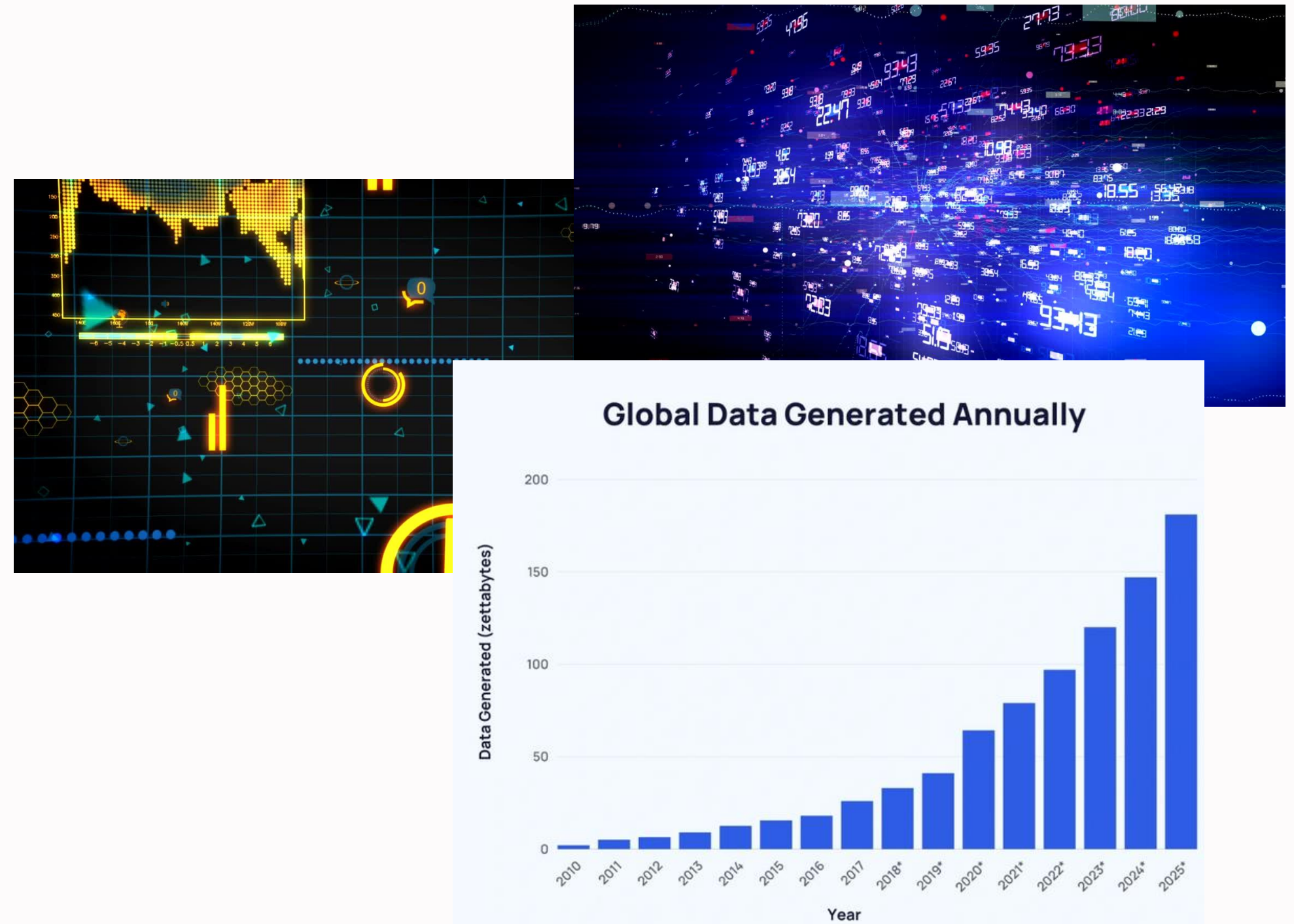
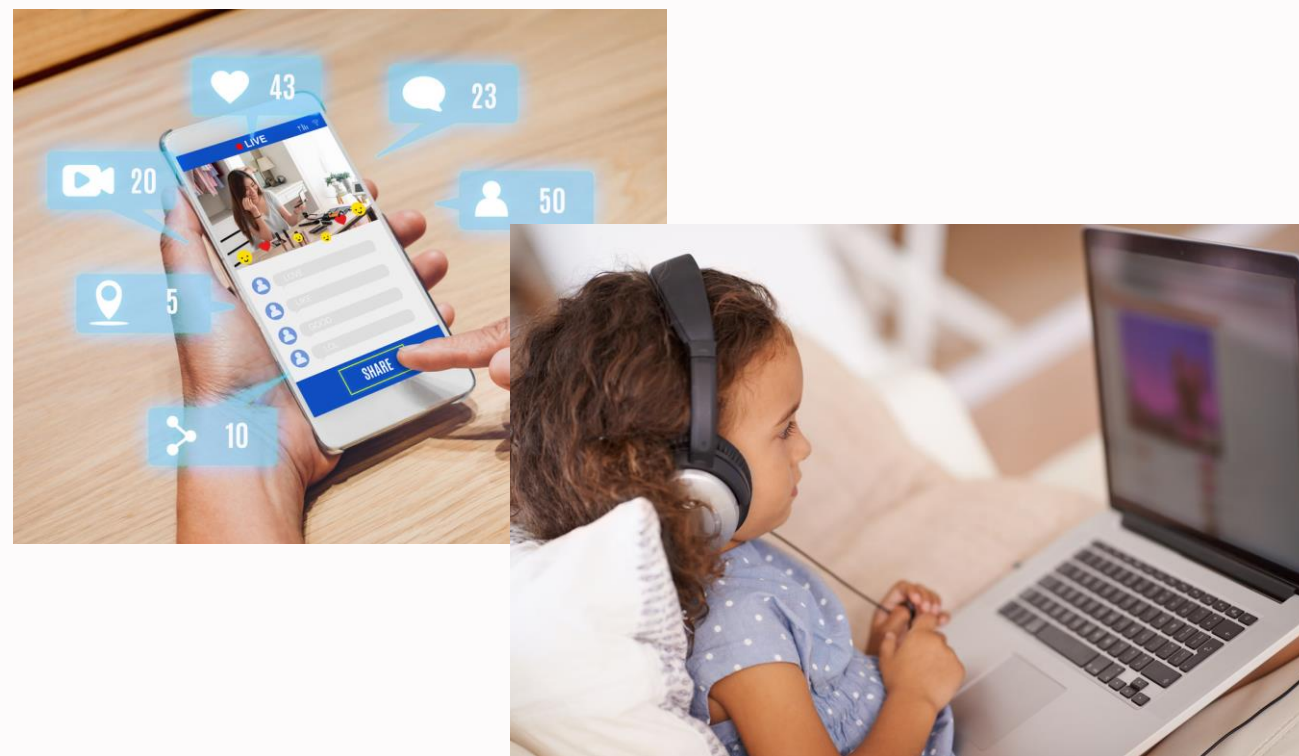
- Context
- Problem Statement
- Goal
- Diagram Flow
- Code implementation
- User Guide
- Mock-up Demonstration
- Progress
- Insights
- Recommendations



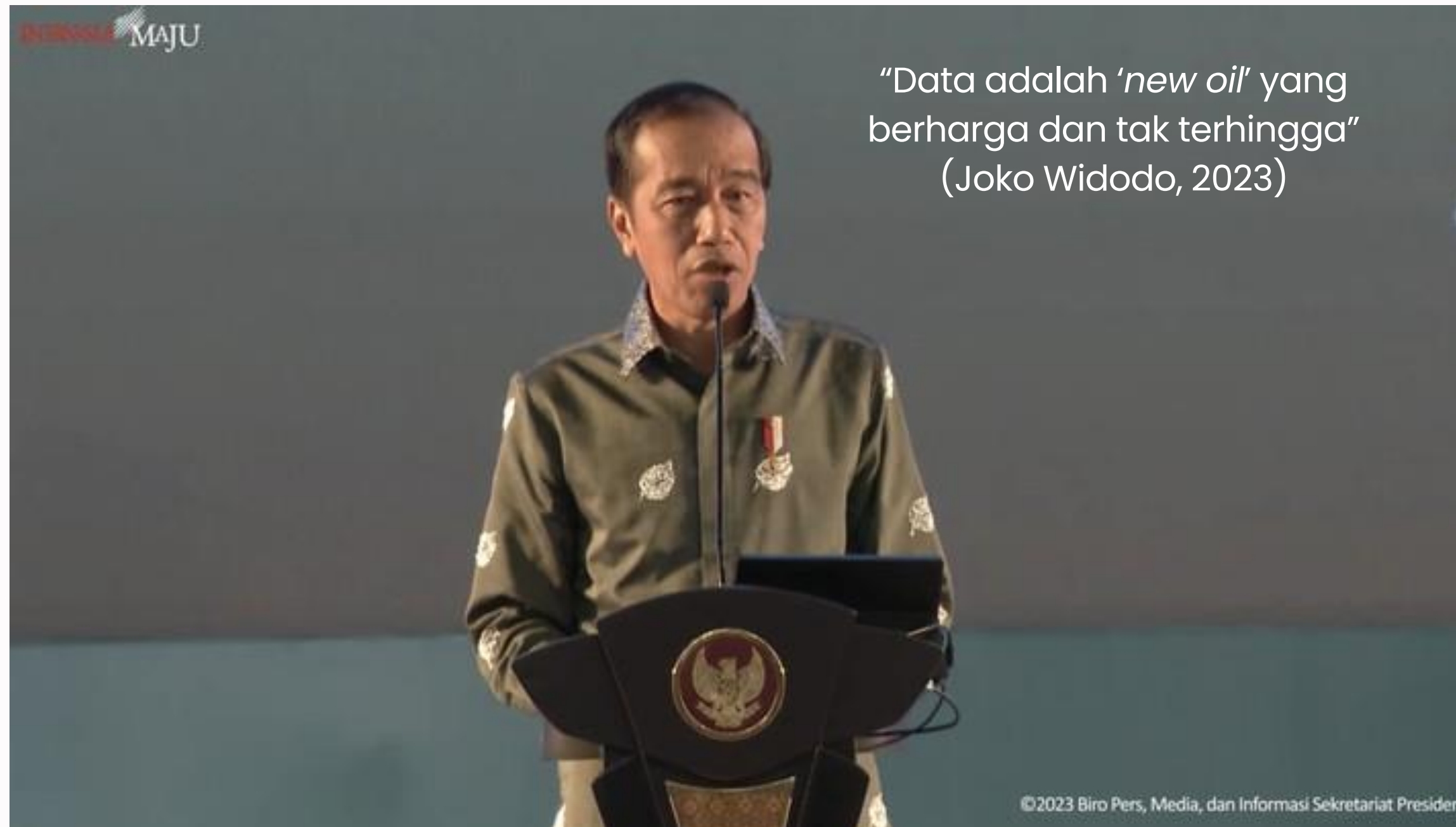
CONTEXT

Digital Era

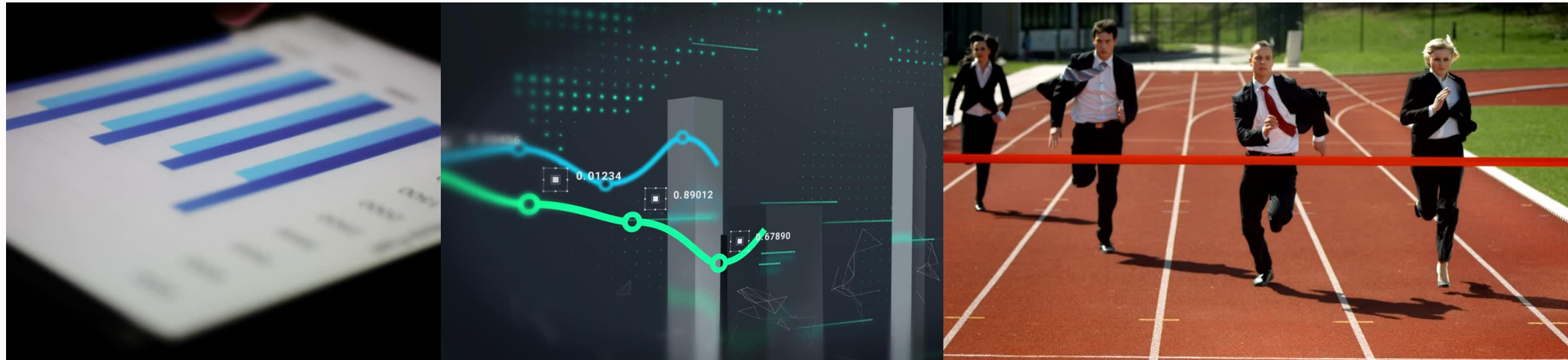
Each day 0,33 ZB or 328, 7 M TB
data generated



DATA is a NEW OIL !!!



Winning the competition



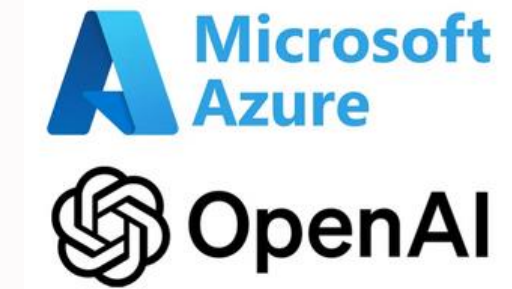
“Data-driven organizations are 23 times more likely to acquire customers, six times as likely to retain customers, and 19 times as likely to be profitable as a result.”

McKinsey
& Company

VISUAL INTO INSIGHT



Grafik --> alat komunikasi efektif
Namun tidak semua bisa menafsirkan grafik



Konversi grafik menjadi narasi dengan
bantuan AI --> semakin produktif

OBJECTIVES



Mengembangkan sebuah skrip yang dapat mengkonversi informasi penting dari chart secara otomatis dalam bentuk narasi (infografis) dengan bantuan large language model (azure openai)

TEKNOLOGI YANG DIGUNAKAN

Large Language Model:

Dalam proyek ini, teknologi Azure OpenAI dengan model gpt 35 turbo digunakan untuk mengotomatiskan proses analisis data grafik menjadi narasi yang informatif.

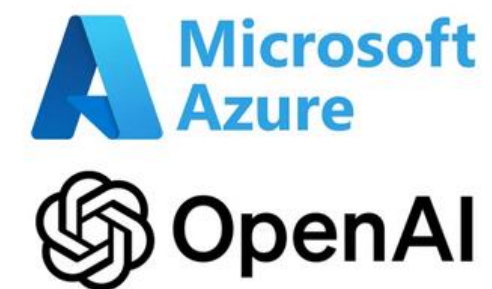
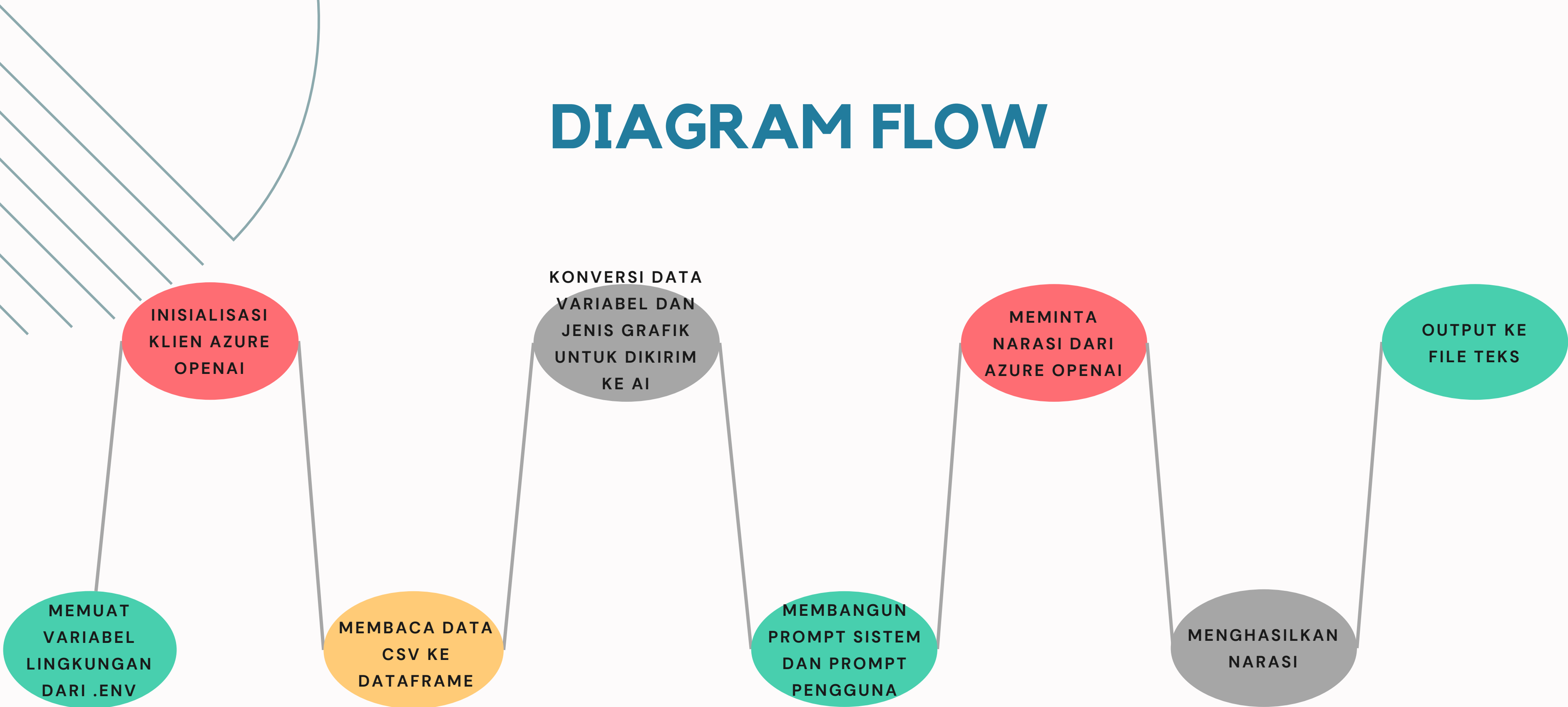


DIAGRAM FLOW



METHODOLOGY

Mengimpor data dari file CSV, mengonversi data tersebut ke dalam format JSON, dan mengirim data ke model Azure OpenAI untuk menghasilkan narasi.

Azure OpenAI API digunakan untuk memanfaatkan model bahasa besar (large language model) dalam menghasilkan narasi dari data grafik.



IMPLEMENTATION

1. Import Library yang diperlukan

```
import pandas as pd
import json
import os
from dotenv import load_dotenv
from openai import AzureOpenAI
```

3. Menginisialisasi klien untuk berkomunikasi dengan Azure OpenAI

```
client = AzureOpenAI(
    azure_endpoint=azure_endpoint,
    api_key=api_key,
    api_version=api_version
)
```

2. Memuat konfigurasi API dari file .env.

```
load_dotenv()
azure_endpoint = os.getenv('AZURE_ENDPOINT')
api_key = os.getenv('API_KEY')
api_version = os.getenv('API_VERSION')
```

4. Fungsi untuk Menghasilkan Narasi dari AI

```
def menghasilkan_narasi(prompt_sistem, prompt_pengguna):
    pesan = [
        {"role": "system", "content": prompt_sistem},
        {"role": "user", "content": prompt_pengguna}
    ]
    respons = client.chat.completions.create(
        model="gpt-35-turbo",
        messages=pesan,
        temperature=0.7,
        max_tokens=800,
        top_p=0.95,
        frequency_penalty=0,
        presence_penalty=0,
        stop=None
    )
    return respons.choices[0].message.content
```


IMPLEMENTATION

5. Fungsi untuk Menghasilkan Insight dari Data Chart

```
def menghasilkan_insight_dari_chart(file_path, kolom_1, kolom_2, kolom_3, judul_chart):  
    data = pd.read_csv(file_path)  
    df = pd.DataFrame(data)  
    data_json = df.to_json(orient="records", lines=True)
```

6. Membuat Deskripsi Chart dan Prompt untuk AI

```
deskripsi_chart = f"""  
Berikut adalah kode Python untuk membuat grafik Stacked bar chart...  
"""  
  
prompt_sistem = f"""  
Anda adalah seorang analis data AI...  
{data_json}  
"""  
  
prompt_pengguna = f"""  
Buatlah narasi yang berisi analisis...  
"""
```

7. Menghasilkan Narasi dari AI

```
narasi = menghasilkan_narasi(prompt_sistem, prompt_pengguna)
```

8. Menyimpan Output ke File Teks

```
output_file = os.path.join(os.path.dirname(file_data), "insight_output_barchart.txt")  
with open(output_file, "w") as file:  
    file.write("Insight Bar Chart:\n")  
    file.write(narasi)
```

User guide:



Instalasi dan setting environment

install library yang diperlukan

Setting konfigurasi API key

Buat file .env berisi Endpoint API

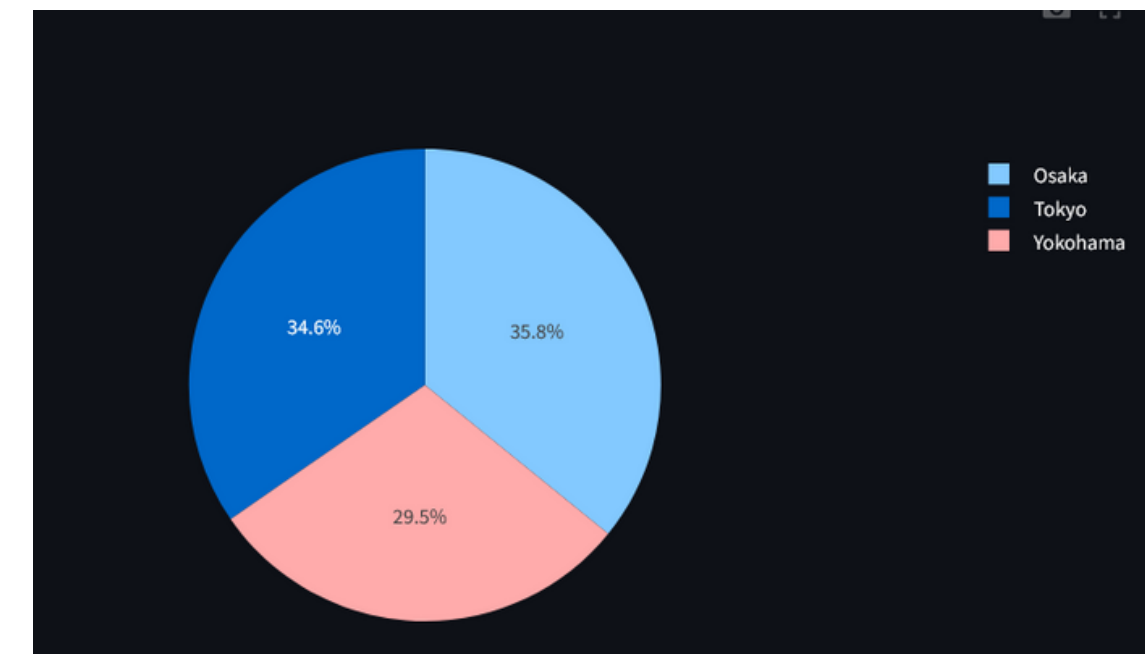
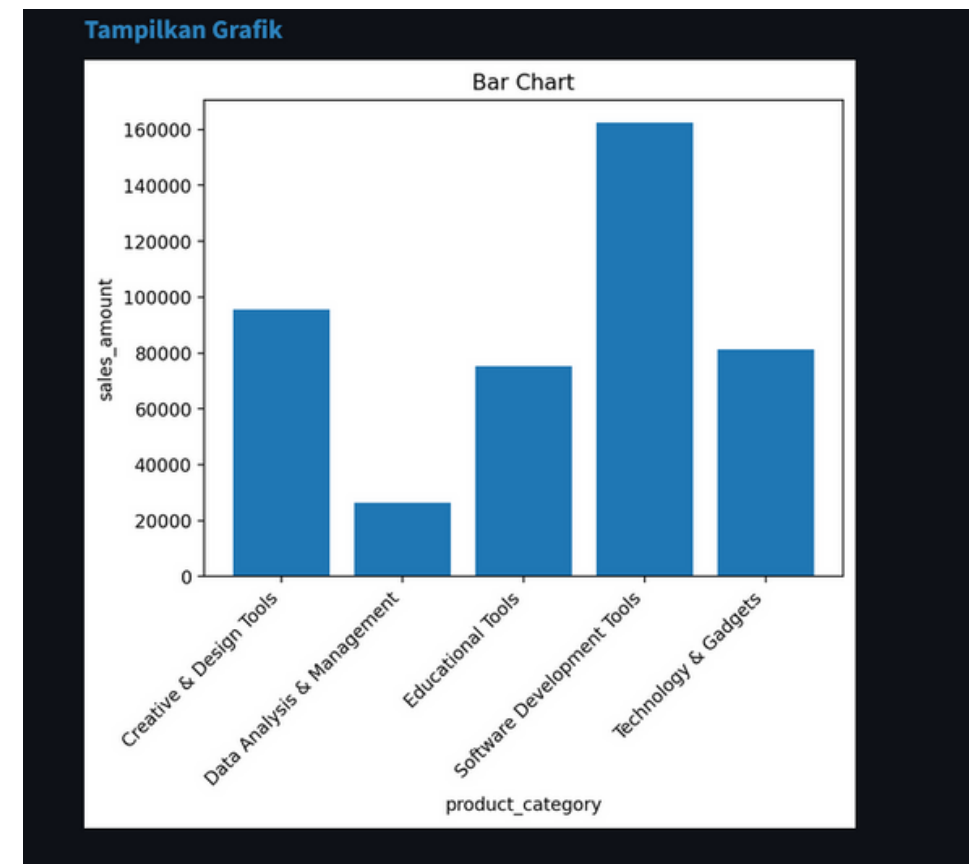
Penggunaan

- upload dataset
- ubah variabel dan jenis grafik
- isi prompt pengguna
- jalankan script di terminal

PROGRES

Menyusun kode dan prompt untuk membuat grafik serta

1. Bar Chart
2. line chart
3. pie chart
4. scatter plot
5. stacked bar chart
6. sankey diagram
7. double line chart
8. barometer/gauge chart
9. waterfall chart
10. funnel chart



PROGRES

Mengubah default sistem prompt :

- menambahkan elemen perhitungan 'akurat'
- Menambahkan elemen 'storytelling'
- menambahkan elemen sasaran audiens

Contoh prompt:

```
Pastikan analisis Anda didasarkan pada perhitungan matematis yang akurat untuk memastikan tidak ada kesalahan atau kekeliruan dalam interpretasi data. Tambahkan elemen storytelling yang membuat analisis ini lebih menarik dan mudah dipahami oleh audiens yang lebih luas yang tidak memiliki latar belakang teknis.
```

PROGRES

sebelum menggunakan sistem prompt baru

Generate Insight

Masukkan pertanyaan Anda di sini...

sales mana yang paling tinggi dari sub kategori

Hasilkan Narasi

Narasi yang Dihasilkan AI

Dari data yang diberikan, kita dapat melihat bahwa sub-kategori dengan penjualan tertinggi adalah "Phones" dengan total penjualan sebesar 330,007.054 dollar. Ini menunjukkan bahwa produk-produk telepon merupakan kontributor utama terhadap pendapatan keseluruhan perusahaan.

Selain itu, kita juga dapat melihat bahwa sub-kategori "Chairs" dan "Storage" juga memiliki penjualan yang signifikan, masing-masing sebesar 328,167.731 dollar dan 223,843.608 dollar. Hal ini menunjukkan bahwa perabotan kantor, seperti kursi dan perabotan penyimpanan, juga memiliki dampak yang signifikan terhadap pendapatan perusahaan.

Namun, di sisi lain, sub-kategori "Fasteners" memiliki penjualan yang jauh lebih rendah, hanya sebesar 3,024.28 dollar. Hal ini menunjukkan bahwa produk-produk fasteners mungkin tidak begitu diminati oleh pelanggan atau mungkin perlu strategi pemasaran yang lebih baik.

Dari insight ini, dapat disimpulkan bahwa perusahaan memiliki potensi untuk meningkatkan penjualan dengan fokus pada produk-produk yang memiliki kontribusi terbesar terhadap pendapatan, seperti telepon dan perabotan kantor, sementara juga perlu memperhatikan produk-produk dengan penjualan rendah untuk meningkatkan strategi pemasaran dan penjualan.

setelah menggunakan sistem prompt baru

Generate Insight

Masukkan pertanyaan Anda di sini...

sales mana yang paling tinggi dari sub kategori

Hasilkan Narasi

Narasi yang Dihasilkan AI

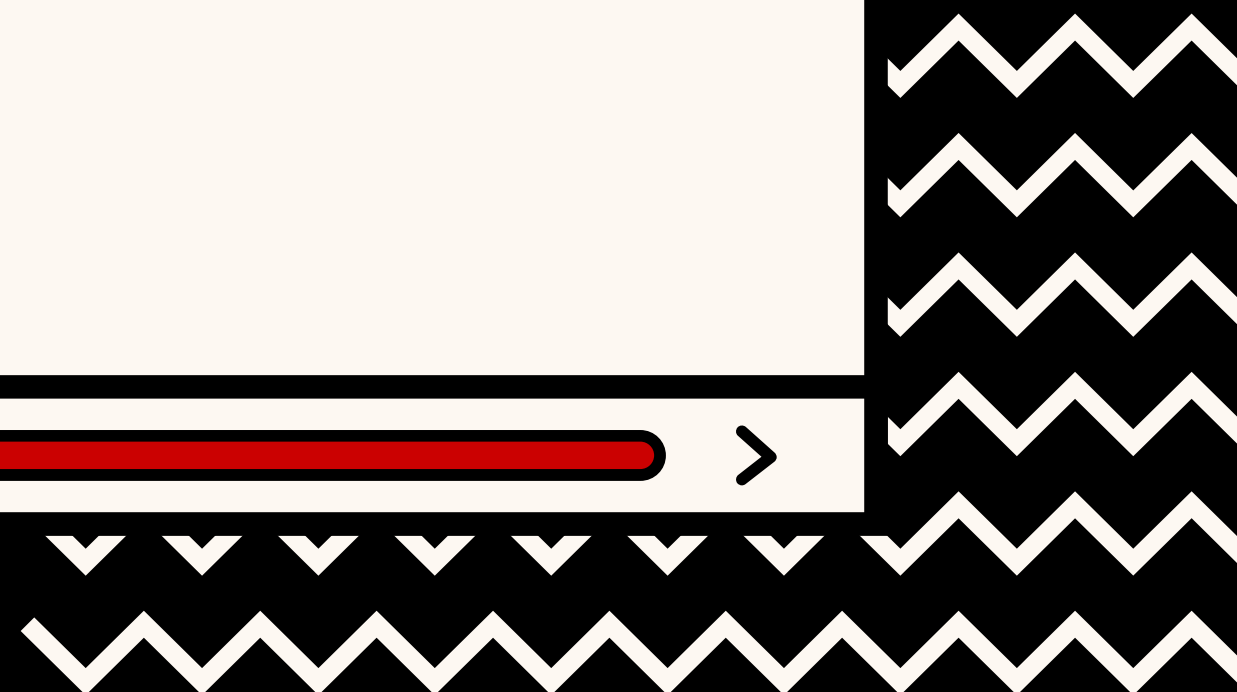
Dari data penjualan yang diberikan, kita dapat melihat bahwa ada berbagai sub-kategori produk yang memiliki penjualan yang berbeda-beda. Dari semua sub-kategori, penjualan tertinggi terjadi pada sub-kategori "Phones" dengan total penjualan sebesar 330,007.054. Hal ini menunjukkan bahwa produk-produk telepon atau smartphone sangat diminati oleh pelanggan, dan mungkin menjadi salah satu produk yang paling banyak dibeli di toko tersebut.

Penjualan terendah terjadi pada sub-kategori "Fasteners" dengan total penjualan hanya sebesar 3,024.28. Ini menunjukkan bahwa produk fasteners atau penjepit mungkin kurang diminati oleh pelanggan, atau mungkin perlu strategi pemasaran yang lebih baik untuk meningkatkan penjualannya.

Dari insight ini, dapat disimpulkan bahwa toko perlu fokus pada peningkatan penjualan produk-produk telepon atau smartphone yang memiliki permintaan tinggi, sementara juga perlu memperhatikan strategi pemasaran untuk meningkatkan penjualan produk fasteners yang lebih rendah. Dengan memahami pola penjualan seperti ini, toko dapat mengatur strategi bisnis yang lebih efektif untuk meningkatkan pendapatan dan keuntungan.

CONCLUSION

Script kode untuk mengkonversi insight dari chart secara otomatis dengan bantuan large language model berhasil dikembangkan



RECOMMENDATION



MODEL DAN TOKEN AI

penggunaan token
disesuaikan dengan
kebutuhan



DATASET

penggunaan dataset
disesuaikan dengan
variabel grafik



JENIS GRAFIK

Menambahkan grafik
agar script bisa
digunakan untuk
aplikasi lebih luas



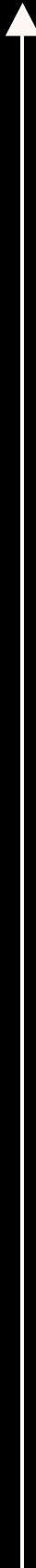
USER INTERFACE

meningkatkan
tampilan mock up



PROMPT

eksperimentasi
menggunakan prompt
baru



DQLab | TETRIS 2024

DEMONSTRATION



Mengubah Insight Dalam Grafik Menjadi Sebuah Narasi



Tampilan Data

	product_category	order_id	product_id	store_id	quantity_sold	sales_amount
0	Creative & Design Tools	310,021	15,418	741	1,876	95,666.52
1	Data Analysis & Management	85,620	5,207	195	526	26,460.29
2	Educational Tools	262,750	15,911	592	1,527	75,221.97
3	Software Development Tools	574,808	32,554	1,266	3,283	162,381.48
4	Technology & Gadgets	264,116	17,315	603	1,495	81,394.17

Upload File Anda Disini

Pilih File

Drag and drop file here

Limit 200MB per file • CSV, XLS, XLSX, JSON

Browse files

df_barchart.csv
350.0B

Pengaturan Grafik

Pilih Tipe Chart:

Bar Chart

Masukkan Kolom (Sumbu X)

product_category

Masukkan Baris (Sumbu Y)

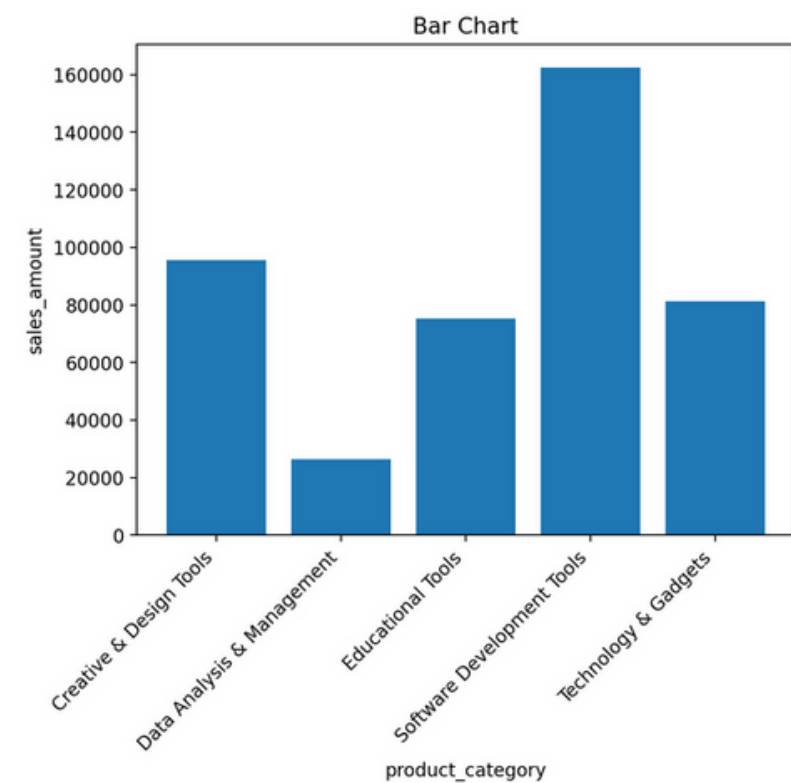
sales_amount

Hasilkan Chart

Tampilan Data

	product_category	order_id	product_id	store_id	quantity_sold	sales_amount
0	Creative & Design Tools	310,021	15,418	741	1,876	95,666.52
1	Data Analysis & Management	85,620	5,207	195	526	26,460.29
2	Educational Tools	262,750	15,911	592	1,527	75,221.97
3	Software Development Tools	574,808	32,554	1,266	3,283	162,381.48
4	Technology & Gadgets	264,116	17,315	603	1,495	81,394.17

Tampilkan Grafik





APPENDIX

Chart and Insight Generator

Pilih file CSV

Drag and drop file here

Limit 200MB per file

Browse files

modified_data.csv

2.3MB

X

Data berhasil dimuat.

	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment
0	CA-2016-152156	2016-11-08	11/11/2016	Second Class	CG-12520	Claire Gute	Consumer
1	CA-2016-152156	2016-11-08	11/11/2016	Second Class	CG-12520	Claire Gute	Consumer
2	CA-2016-138688	2016-06-12	6/16/2016	Second Class	DV-13045	Darrin Van Huff	Corporate
3	US-2015-108966	2015-10-11	10/18/2015	Standard Class	SO-20335	Sean O'Donnell	Consumer
4	US-2015-108966	2015-10-11	10/18/2015	Standard Class	SO-20335	Sean O'Donnell	Consumer

Pilih jenis grafik

Pie Chart

Pilih kolom sumbu X

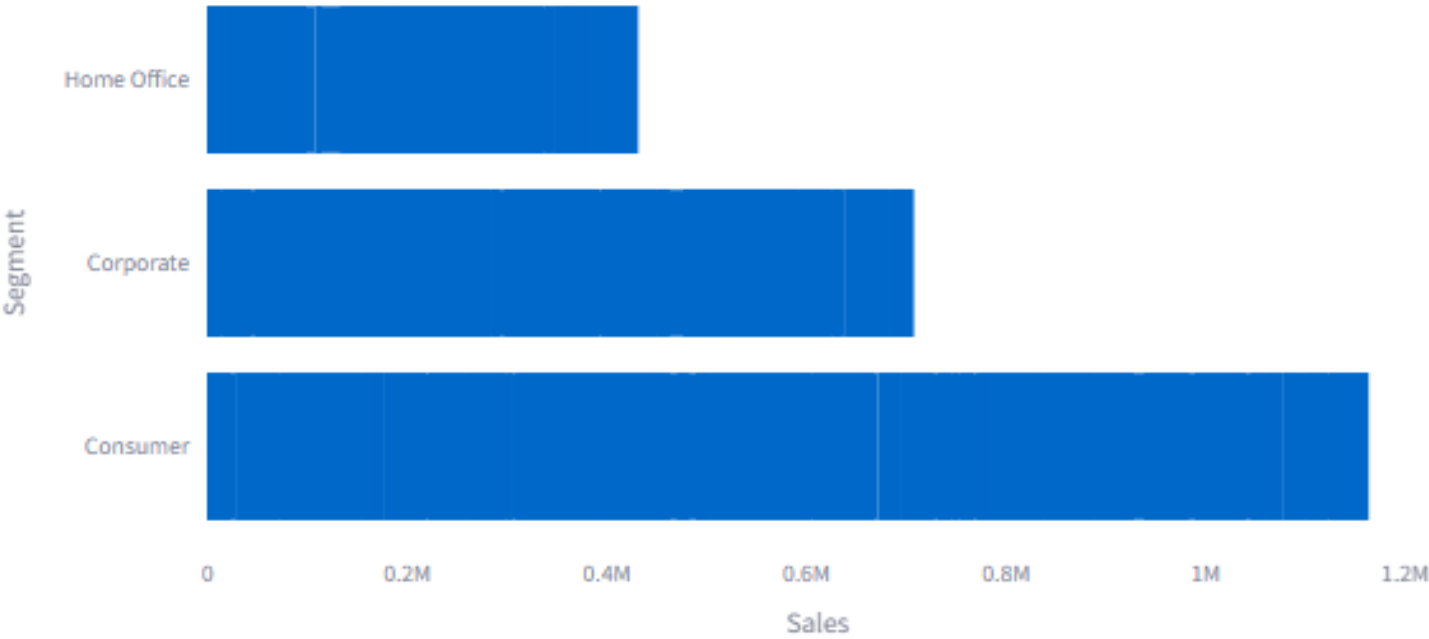
Ship Mode

Pilih kolom sumbu Y

Sales

Buat Grafik

Bar Chart



Tampilkan Narasi

Masukkan prompt untuk AI:

Bisakah anda menjelaskan grafik diatas?

Hasilkan Insight

Insight dari AI:

Tentu! Grafik di atas adalah Bar Chart yang menunjukkan penjualan berdasarkan segmen produk. Sumbu x menunjukkan nilai penjualan dalam satuan mata uang, sedangkan sumbu y menunjukkan segmen produk yang terdiri dari tiga kategori: Konsumen (Consumer), Korporat (Corporate), dan Kebutuhan Proyek (Home Office). Dari grafik ini, kita dapat melihat perbandingan penjualan antara ketiga segmen produk tersebut. Dengan melihat grafik ini, kita dapat menarik kesimpulan tentang kinerja penjualan untuk masing-masing segmen produk.

APPENDIX

