

RANCANGAN PENGOLAHAN DATA



Kelompok 3

Travel
POSTCARD

To: _____

NAMA KELOMPOK



PLANE CRASH
DETECTIVES

1. Haarisah Yustika Putri Al-Jufri
> 1941720221
2. Ony Novianti
> 2041720029
3. Reynaldi Fakhri Pratama
> 2041720209



STUDI KASUS



Mencari faktor-faktor yang
mempengaruhi kecelakaan
pesawat di seluruh dunia



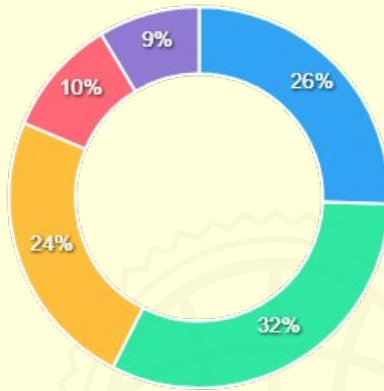
01

FLOWCHART

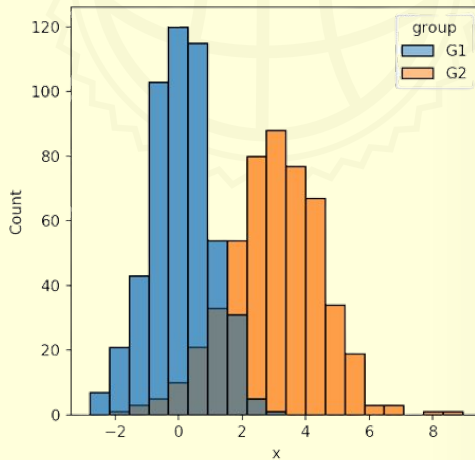
Flowchart untuk melakukan
preprocessing data dan
ETL (*Extract-Transform-Load*)







- series-1
- series-2
- series-3
- series-4
- series-5



02

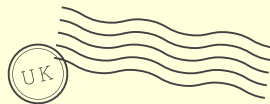
VISUALISASI

Rencana akhir visualisasi data yang akan dibuat



03

PROSES DATA



*Skenario proses data akan
dilakukan dengan SparkSQL,
DataFrames atau Datasets*

6 Tahapan Utama

1. EXTRACT DATA

Melakukan ekstraksi data dari file CSV kemudian di load ke dalam sebuah DataFrame

2. DATA CLEANSING

Melakukan proses pembersihan data :

- Menghapus kolom yang tidak perlu
- Menghapus duplikat
- Mengisi kolom yang hilang

3. DATA TRANSFORMATION

Melakukan proses transformasi data :

- Mengubah format tanggal sesuai standar
- Mengubah lokasi menjadi terpisah, seperti negara, dan lain sebagainya

4. DATA PREPARATION

Menyiapkan variabel-variabel apa saja dari data untuk dianalisis

5. DATA ANALYSIS

Melakukan analisis, misal melakukan clustering untuk menemukan pola tertentu

6. DATA MODELING

Membuat model prediksi untuk menentukan apa saja faktor yang menyebabkan kecelakaan pesawat



PLANE CRASH
DETECTIVES

PASSPORT

THANK
YOU

