

LAPORAN PROYEK MATA KULIAH
12S3202 – GUDANG DATA DAN KECERDASAN BISNIS

Data Warehouse-based Dashboard for Kickstarter Dataset

<<https://github.com/walkervalestinuss/proyek-dwbi-g2-2324>>



Disusun oleh:

- 1. 12S21009 - Mikhael Janugrah Pakpahan**
- 2. 12S21011 - Aldi Jeremy Simamora**
- 3. 12S21012 - Walker Valentinus Simanjuntak**
- 4. 12S21014 - Fritz Kevin Manurung**
- 5. 12S21015 - Sitogab Antonio Octavianus Girsang**

FAKULTAS INFORMATIKA DAN TEKNIK ELEKTRO
INSTITUT TEKNOLOGI DEL

2024

DAFTAR ISI

| | |
|--|----|
| DAFTAR ISI..... | 1 |
| 1. Pendahuluan..... | 1 |
| 1.1. Latar Belakang..... | 1 |
| 1.2. Tujuan Pengerjaan Proyek..... | 1 |
| 1.3. Ruang Lingkup..... | 2 |
| 1.4. Tim Pengembang..... | 2 |
| 2. Analisis dan Desain..... | 4 |
| 2.1. Pertanyaan Analisis..... | 4 |
| 2.2. Arsitektur Sistem..... | 4 |
| 2.3. Sumber Data..... | 4 |
| 2.4. Model Dimensional..... | 8 |
| 2.4.1. High-Level Dimensional Model..... | 8 |
| 2.4.2. Dimensional Model Schema..... | 9 |
| 2.4.3. Detailed Dimensional Model..... | 10 |
| 2.5. Extract, Transform, and Load..... | 15 |
| 2.5.1. High-Level Source to Target Map..... | 15 |
| 2.5.2. Detailed ETL Flow for Each Source to Target..... | 17 |
| 2.6. Business Intelligence Application..... | 20 |
| 2.6.1. Business Intelligence Application Specification..... | 20 |
| 2.6.2. Detailed Business Intelligence Application Specification..... | 20 |
| 2.6.3. Business Intelligence Application Mockup..... | 22 |
| 3. Implementasi..... | 23 |
| 3.1. ROLAP Schema..... | 23 |
| 3.2. ETL..... | 24 |
| 3.2.1. Concat Data Source Kickstarter..... | 24 |
| 3.2.2. ETL dari Dataset ke Table Staging Dim Category..... | 25 |
| 3.2.3. ETL dari Dataset ke Table Staging Dim Location..... | 25 |
| 3.2.4. ETL dari dataset ke Table Staging Dim Project..... | 26 |
| 3.2.5. ETL dari dataset ke Table Staging Dim Date..... | 26 |
| 3.2.6. ETL dari dataset ke Table Staging Backer Fact..... | 27 |
| 3.2.7. Job ETL..... | 27 |
| 3.3. MOLAP Schema..... | 28 |
| 3.4. Dashboard..... | 29 |
| 4. Evaluasi..... | 30 |
| 5. Penutup..... | 31 |
| 5.1. Kesimpulan..... | 31 |
| 5.2. Saran..... | 32 |
| LAMPIRAN..... | 33 |
| Lampiran A. Spesifikasi Aplikasi Kecerdasan Bisnis..... | 33 |
| Lampiran B. Spesifikasi Rinci Aplikasi Kecerdasan Bisnis..... | 34 |

1. Pendahuluan

Bagian ini berisi latar belakang, tujuan pengerjaan proyek, ruang lingkup, dan tim pengembang.

1.1. Latar Belakang

Kickstarter adalah platform pendanaan terbesar untuk proyek kreatif di dunia saat ini. Kickstarter juga menyediakan wadah bagi para kreator yang ingin membangun proyek kreatif mereka. Dengan demikian para kreator dapat mempromosikan dan mengajukan pendapat maupun ide terkait proyek yang dimiliki ke seluruh dunia dengan tujuan nantinya akan mendapatkan pendanaan oleh para pendukung yang tertarik dengan proyek tersebut.

Dalam menggunakan Kickstarter, pihak pembuat *campaign* tentu mengharapkan hasil yang memuaskan dan menguntungkan kedua belah pihak. Untuk mendapatkan keberhasilan dalam setiap *campaign* yang akan diadakan, pemilik membutuhkan analisis pemberi dana dari proyek-proyek yang sudah ada sebelumnya seperti informasi terkait lokasi geografis para pemberi dana, rata-rata jumlah dana yang terkumpul dari pemberi dana untuk setiap *campaign*, dan kategori apa yang paling menarik fokus para pemberi dana di tahun 2022, untuk dijadikan pertimbangan oleh pembuat *campaign* di tahun-tahun berikutnya. Sistem Gudang Data dan Kecerdasan Bisnis sesuai untuk menjadi solusi atas persoalan tersebut karena kita akan dapat mengumpulkan dan menganalisis data yang berhubungan dengan informasi yang dibutuhkan oleh setiap pembuat *campaign*.

Berdasarkan kebutuhan yang ada, *Data Warehouse* akan menjadi pilihan kami untuk diimplementasikan. Dengan pengimplementasian *Data Warehouse* kami dapat mengintegrasikan semua data ke dalam satu platform dan melakukan analisis yang mendalam untuk mendukung keputusan para pembuat *campaign*.

Kelompok kami menggunakan Tableau Desktop sebagai *Business Intelligence Application*. Pemilihan ini didasarkan pada beberapa hal yaitu antarmuka aplikasi yang ramah pengguna, Tableau juga menyediakan berbagai jenis visualisasi data yang dapat disesuaikan dengan kebutuhan kelompok kami.

1.2. Tujuan Pengerjaan Proyek

Tujuan dari pengerjaan proyek ini adalah:

1. Mengembangkan sistem *Data Warehouse* dan *Dashboard* sederhana berdasarkan studi kasus dunia nyata, yaitu kampanye penggalangan dana di Kickstarter tahun 2022.

2. Menerapkan konsep dan teknik pengolahan data yang telah dipelajari dalam mata kuliah *Data Warehouse and Business Intelligence (DWBI)*.
3. Dapat membuat setidaknya satu model dimensional dengan satu tabel fakta dan tiga tabel dimensi untuk mengorganisir data dalam bentuk yang mudah dipahami dan dapat digunakan untuk analisis.

1.3. Ruang Lingkup

Ruang lingkup dalam pengerjaan proyek ini adalah:

1. Pendekatan yang kami gunakan adalah *Kimball Life Cycle* yaitu dengan menggunakan model *Star Schema*. Dimana terdapat satu tabel fakta yang terhubung dengan beberapa tabel dimensi. Pendekatan ini juga digunakan karena strukturnya yang lebih sederhana, cocok untuk proyek dengan skala kecil sesuai dengan proyek yang kami kerjakan.
2. Dataset yang digunakan didapat dari Kickstarter Dataset pada laman Web Robots. Data yang kami pakai adalah data tahun 2022.

1.4. Tim Pengembang

Pada Tabel 1 disajikan susunan tim pengembang pada proyek ini.

Tabel 1. Susunan Tim Pengembang

| No. | Nama | Peran | Tanggung Jawab |
|-----|---------------------------------------|---------|--|
| 1 | Sitogab Antonio Octavianus Grisang | Ketua | <ul style="list-style-type: none"> ● High Level Source to Target Map ● Detailed ETL flow for Each Source to Target |
| 2 | Mikhael Janugrah Pakpahan | Anggota | <ul style="list-style-type: none"> ● Detailed ETL flow for Each Source to Target ● Evaluasi |
| 3 | Aldi Jeremy Simamora | Anggota | <ul style="list-style-type: none"> ● High Level Dimensional Model ● Detailed Dimensional Model |

| No. | Nama | Peran | Tanggung Jawab |
|-----|----------------------------------|---------|--|
| 4 | Walker Valentinus Simanjuntak | Anggota | <ul style="list-style-type: none"> ● Extract, Transform, Load ● ROLAP & MOLAP ● Dashboard |
| 5 | Fritz Kevin Manurung | Anggota | <ul style="list-style-type: none"> ● Dimensional Model Schema ● Detailed Dimensional Model |

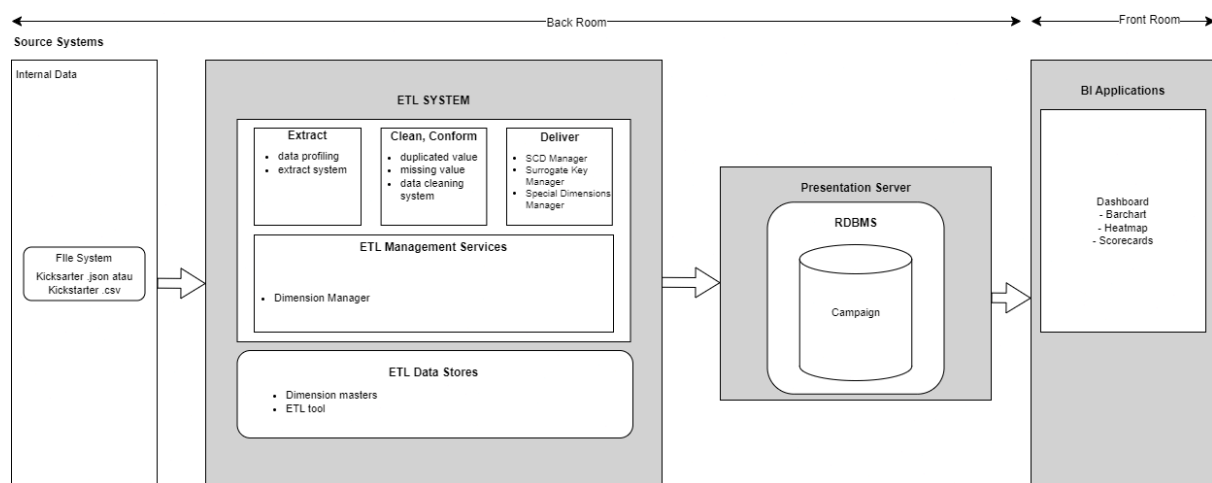
2. Analisis dan Desain

Bagian ini berisi pengamatan dan pendeskripsian komponen-komponen yang akan dikaji. Bagian ini juga akan berisi hasil pengamatan dan pendeskripsian seperti pertanyaan bisnis, arsitektur sistem, sumber data, model dimensional, proses ETL, dan *business intelligence application*.

2.1. Pertanyaan Analisis

1. Apa saja lokasi geografis backer pada umumnya? Menargetkan upaya pemasaran ke area dengan konsentrasi backer potensial yang tinggi)
2. Berapa rata rata amount pledged yang diberikan oleh setiap backer pada setiap campaign? (Memperkirakan potensi pendanaan berdasarkan proyeksi jumlah backer)
3. Apakah backer cenderung fokus pada kategori campaign tertentu? (Memahami preferensi pendukung dan membantu menyesuaikan campaign agar sesuai dengan audience yang tepat)

2.2. Arsitektur Sistem



Gambar 1. Application Architecture Model Sample

2.3. Sumber Data

Sumber data yang digunakan pada proyek berikut adalah file CSV dari dataset Kickstarter Crowdfunding 2022. Bentuk data CSV (*Comma Separated Value*) adalah sebuah format data yang digunakan untuk pertukaran data antar aplikasi. Data dalam format CSV data disusun dalam bentuk tabel yang terdiri dari baris dan kolom. Setiap baris dalam file CSV mewakili satu rekaman data, dan setiap kolom dalam baris tersebut dipisahkan oleh koma.

Tabel 2. Atribut Data CSV

| No. | Nama Atribut | Keterangan | Tipe Data | Panjang Atribut |
|-----|------------------------|--|-----------|-----------------|
| 1. | Backers_count | Menunjukkan jumlah pendukung (backers) proyek | Integer | - |
| 2. | Blurb | Deskripsi singkat dari proyek | Varchar | 255 |
| 3. | Category | Menunjukkan kategori proyek | Varchar | 255 |
| 4. | Converted_pledge | Jumlah dana yang terkumpul dari backers | Integer | - |
| 5. | Country | Singkatan negara tempat proyek berada | Varchar | 2 |
| 6. | Country_display | Keterangan nama negara tempat proyek berada | Varchar | 255 |
| 7. | Created_at | Waktu penunjuk proyek diposting di platform | Varchar | 255 |
| 8. | Creator | Informasi tentang pembuat proyek | Varchar | 255 |
| 9. | Currency | Mata uang yang digunakan untuk target dana dan dana yang terkumpul | Varchar | 3 |
| 10. | Currency_symbol | Simbol mata uang yang digunakan untuk menandai jumlah uang | Varchar | 1 |
| 11. | Currency_trailing_code | Menunjukkan apakah simbol mata uang diikuti oleh kode | Boolean | - |
| 12. | Current_currency | Mata uang saat ini yang digunakan dalam kampanye | Varchar | 3 |
| 13. | Deadline | Tanggal batas akhir untuk mencapai target | Integer | - |

| | | | | |
|-----|-----------------------|--|-----------|-------|
| | | pendanaan | | |
| 14. | Disable_communication | Menunjukkan apakah komunikasi dengan pencipta proyek telah dinonaktifkan | Boolean | - |
| 15. | Friends | Daftar teman yang terhubung dengan proyek | Varchar | 255 |
| 16. | Fx_rate | Tingkat konversi mata uang untuk mata uang yang berbeda | Float | - |
| 17. | Goal | Jumlah dana yang ditargetkan untuk dikumpulkan dalam proyek | Integer | - |
| 18. | Id | ID unik untuk proyek | Integer | - |
| 19. | Is_backing | Menunjukkan apakah proyek ini dapat ditandai sebagai favorit | Boolean | - |
| 20. | Is_starrable | Menunjukkan apakah proyek ini dapat ditandai sebagai favorit | Boolean | - |
| 21. | Is_starred | Menunjukkan apakah pengguna telah menandai proyek tersebut sebagai favorit | Boolean | - |
| 22. | Launched_at | Waktu ketika proyek diluncurkan atau dimulai | Integer | - |
| 23. | Location | Lokasi fisik tempat proyek tersebut berada | Varchar | 255 |
| 24. | Name | Nama atau judul proyek | Varchar | 255 |
| 25. | Permissions | Izin yang diberikan oleh pengguna kepada proyek atau pencipta | Varchar | 255 |
| 26. | Photo | Foto yang terkait dengan proyek | Media JPG | in mb |

| | | | | |
|-----|-------------------|---|----------------|-----|
| 27. | Pledged | Jumlah dana yang telah terkumpul dari pendukung, dalam mata uang asli | Float | - |
| 28. | Profile | Profil proyek atau pencipta | Varchar | 255 |
| 29. | Slug | Bagian dari URL yang mengidentifikasi proyek secara unik | Varchar | 255 |
| 30. | Source_url | URL sumber yang mengarah ke proyek | Varchar (link) | 255 |
| 31. | Spotlight | Menunjukkan apakah proyek tersebut telah dipilih untuk ditampilkan di halaman utama | Boolean | - |
| 32. | Staff_pick | Menunjukkan apakah proyek tersebut telah dipilih oleh staf Kickstarter | Boolean | - |
| 33. | State | Status proyek, misalnya "live" atau "successful" | Varchar | 255 |
| 34. | State_changed_at | Waktu ketika status proyek berubah | Integer | - |
| 35. | Static_usd_rate | Tingkat konversi mata uang untuk USD (Dolar Amerika) yang tetap | Float | - |
| 36. | Urls | URL yang terkait dengan proyek | Varchar | 255 |
| 37. | Usd_exchange_rate | Tingkat konversi mata uang untuk USD yang digunakan dalam kampanye | Float | - |
| 38. | Usd_pledged | Jumlah dana yang terkumpul dari pendukung, dalam USD (Dolar Amerika) | Float | - |

| | | | | |
|-----|----------|--|---------|-----|
| 39. | Usd_type | Tipe mata uang yang digunakan dalam kampanye, misalnya "international" atau "domestic" | Varchar | 255 |
|-----|----------|--|---------|-----|

2.4. Model Dimensional

2.4.1. High-Level Dimensional Model

Tabel 3. Detailed Bus Matrix

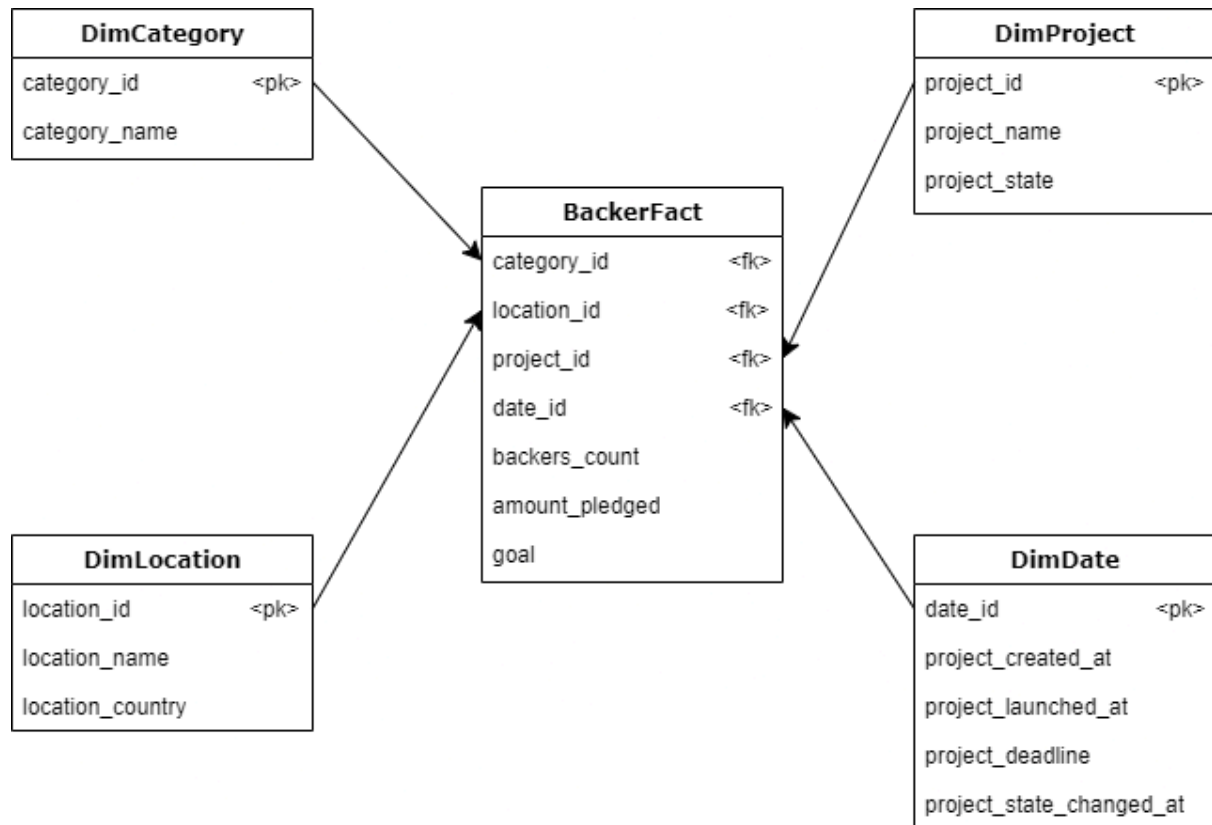
| Business Process | Fact Table | Fact Grain | Grain | Facts | Dim Project | Dim Location | Dim Category | Dim Date |
|-----------------------|-------------|-------------|---|-------------------------------|-------------|--------------|--------------|----------|
| Backers Support Count | Backer Fact | Transaction | Each row represents a single funding transaction by a backer for a project. | backers_count, amount_pledged | X | X | X | X |

Tabel 4. Attributes and Matrices

| Dimension/Fact Name | Attribute/Fact Name | Description | Sample Value |
|---------------------|---------------------|------------------|--------------|
| Campaign Fact | project_id | ID unik proyek | 4465571 |
| | location_id | ID unik lokasi | 1017865 |
| | category_id | ID unik kategori | 292 |
| | date_id | ID unik tanggal | 1 |

| | | | |
|--------------|--------------------------|-------------------------------|--|
| | backers_count | jumlah pendana | 40 |
| | pledged | jumlah dana terkumpul | 15295670 |
| | goal | target dana terkumpul | 1234567 |
| Dim_Project | project_id | ID Unik proyek | 4465571 |
| | project_state | Status proyek | active |
| | project_name | Nama proyek | Things I Have Loved “ A collection (sort of) |
| Dim_Location | location_id | ID unik lokasi | 2363796 |
| | location_name | nama lokasi | Beverly Hills |
| | location_country | Nama lengkap negara | United States |
| Dim_Category | category_id | ID unik kategori | 292 |
| | category_name | nama kategori | Comedy |
| Dim_Date | date_id | ID Unik tanggal | 1 |
| | deadline | tanggal | 01/01/2000 |
| | project_created_at | tanggal proyek dibuat | 01/01/2000 |
| | project_launched_at | tanggal proyek diluncurkan | 01/01/2000 |
| | project_state_changed_at | tanggal status proyek berubah | 1/1/2000 |

2.4.2. Dimensional Model Schema



Gambar 2. Dimensional Model Star Schema

2.4.3. Detailed Dimensional Model

Tabel 5. Backer_Fact

| | |
|--------------------------|---|
| Table Name | Backer_Fact |
| Table Type | Fact |
| Display Name | Campaign |
| Database Schema | CampaignDW |
| Table Description | Backer Fact |
| Comment | You can put technical information here so it doesn't get copied into the database in any way. |
| Biz Filter Logic | |
| Size | one row per execution of a package |

| Generate Script? | | | | | |
|------------------|----------------|-----------------------|----------------|----------------|----------|
| | | | | | |
| Column Name | Display Name | Description | Unknown Member | Example Values | SCD Type |
| project_id | project_id | ID Unik proyek | N/A | 4465571 | key |
| location_id | location_id | ID unik lokasi | N/A | 1017865 | key |
| category_id | category_id | ID unik kategori | N/A | 292 | key |
| date_id | date_id | ID Unik tanggal | N/A | 1 | key |
| backers_count | backers_count | jumlah pendana | N/A | 40 | key |
| pledged | amount_pledged | jumlah dana terkumpul | N/A | 15295670 | key |
| goal | goal | target dana | N/A | 1234567 | key |

Tabel 6. DimProject

| | |
|--------------------------|-------------------|
| Table Name | Dim_Project |
| Table Type | Dimension |
| Display Name | DimProject |
| Database Schema | CampaignDW |
| Table Description | Project Dimension |

| | | | | | |
|-------------------------|---|--------------------|-----------------------|---|-----------------|
| Comment | You can put technical information here so it doesn't get copied into the database in any way. | | | | |
| Biz Filter Logic | | | | | |
| Size | one row per execution of a package | | | | |
| Generate Script? | | | | | |
| | | | | | |
| Column Name | Display Name | Description | Unknown Member | Example Values | SCD Type |
| project_id | project_id | ID Unik proyek | N/A | 4465571 | key |
| project_state | project_state | Status proyek | N/A | active | punch through |
| project_name | project_name | Nama proyek | N/A | Things I Have Loved “ A collection (sort of) | key |

Tabel 7. DimLocation

| | |
|--------------------------|--------------------|
| Table Name | Dim_Location |
| Table Type | Dimension |
| Display Name | DimLocation |
| Database Schema | CampaignDW |
| Table Description | Location Dimension |

| | | | | | |
|-------------------------|---|--------------------|-----------------------|-----------------------|-----------------|
| Comment | You can put technical information here so it doesn't get copied into the database in any way. | | | | |
| Biz Filter Logic | | | | | |
| Size | one row per execution of a package | | | | |
| Generate Script? | | | | | |
| | | | | | |
| Column Name | Display Name | Description | Unknown Member | Example Values | SCD Type |
| location_id | location_id | ID unik lokasi | N/A | 1017865 | key |
| location_name | name | nama lokasi | N/A | Beverly Hills | key |
| location_country | country | negara | N/A | United States | key |

Tabel 8. DimCategory

| | |
|--------------------------|---|
| Table Name | Dim_Category |
| Table Type | Dimension |
| Display Name | Category |
| Database Schema | CampaignDW |
| Table Description | Category Dimension |
| Comment | You can put technical information here so it doesn't get copied into the database in any way. |

| | | | | | |
|-------------------------|------------------------------------|--------------------|-----------------------|-----------------------|-----------------|
| Biz Filter Logic | | | | | |
| Size | one row per execution of a package | | | | |
| Generate Script? | | | | | |
| | | | | | |
| Column Name | Display Name | Description | Unknown Member | Example Values | SCD Type |
| category_id | category_id | ID unik kategori | N/A | 292 | key |
| category_name | category_name | nama kategori | N/A | Comedy | key |

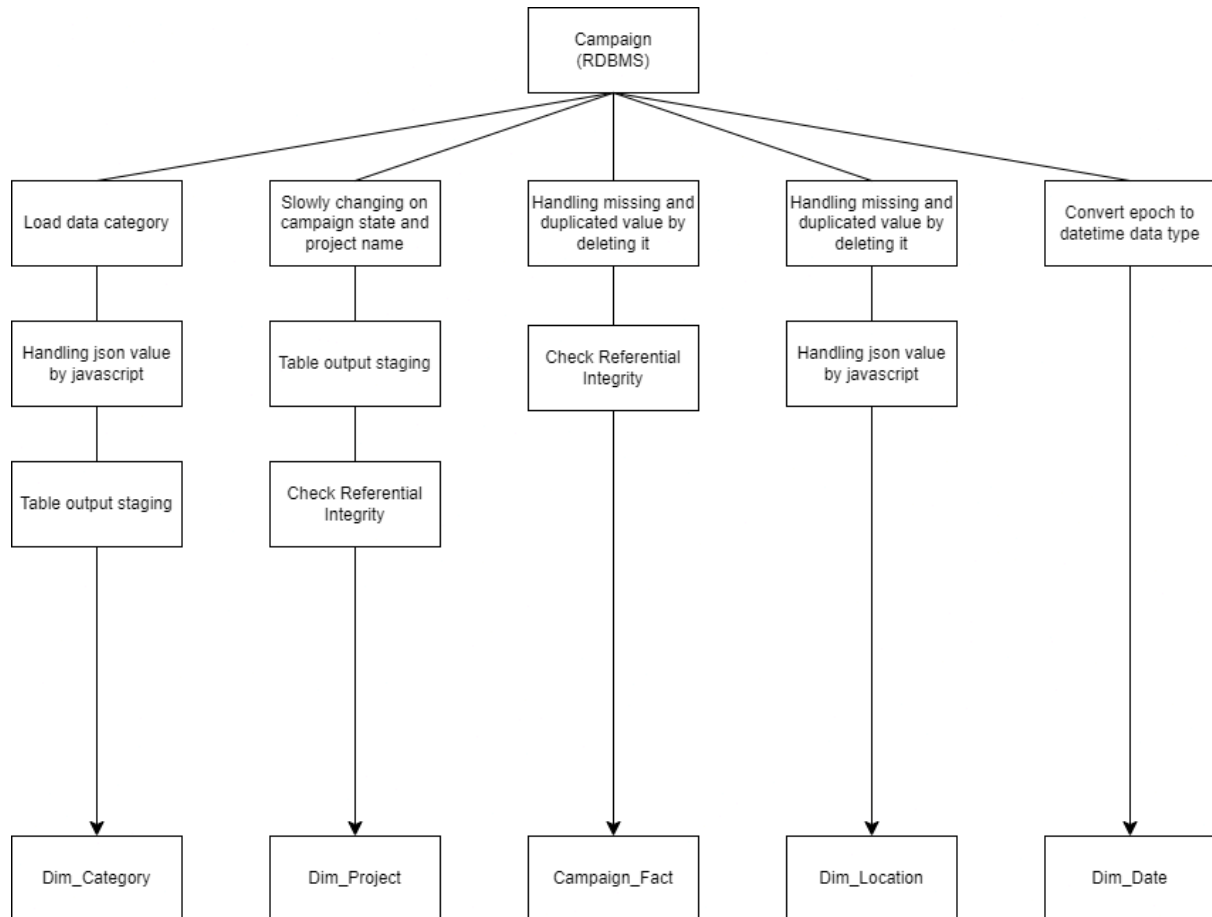
Tabel 9. DimDate

| | |
|--------------------------|---|
| Table Name | DimDate |
| Table Type | Dimension |
| Display Name | Date |
| Database Schema | CampaignDW |
| Table Description | Date Dimension |
| Comment | You can put technical information here so it doesn't get copied into the database in any way. |
| Biz Filter Logic | |
| Size | one row per execution of a package |
| Generate Script? | |

| Column Name | Display Name | Description | Unknown Member | Example Values | SCD Type |
|--------------------------|--------------------------|--|----------------|----------------|----------|
| date_id | date_id | ID unik tanggal | N/A | 1 | key |
| Project_created_at | project_created_at | Tanggal proyek dibuat | N/A | 01/01/2009 | key |
| deadline | deadline | Tanggal batas pengumpulan dana | N/A | 01/01/2009 | |
| Project_launched_at | project_launched_at | Deskripsi panjang tanggal dalam format | N/A | 01/01/2009 | key |
| Project_state_changed_at | project_state_changed_at | Tanggal status proyek berubah | N/A | 1/1/2009 | key |

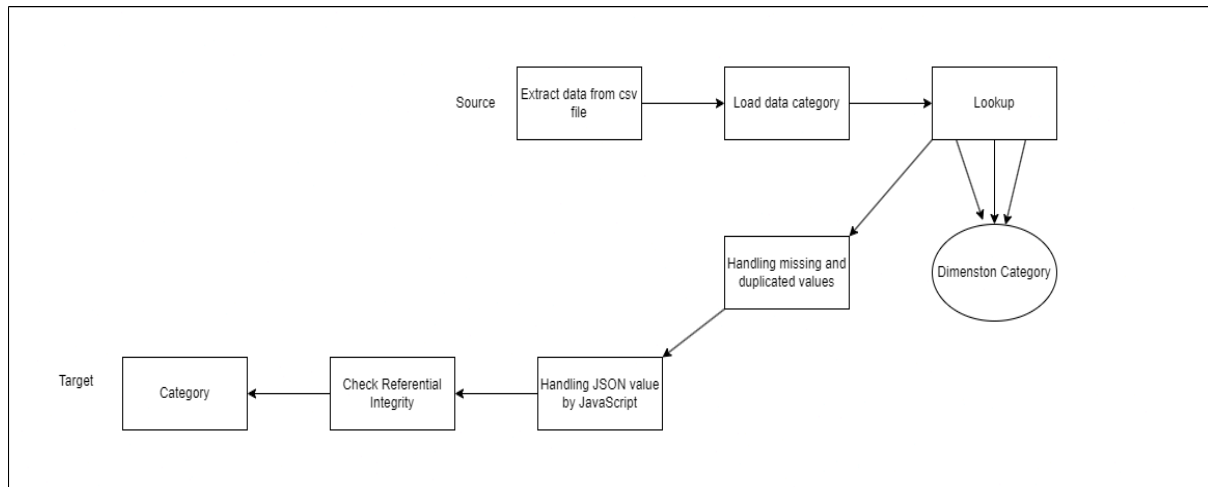
2.5. Extract, Transform, and Load

2.5.1. High-Level Source to Target Map

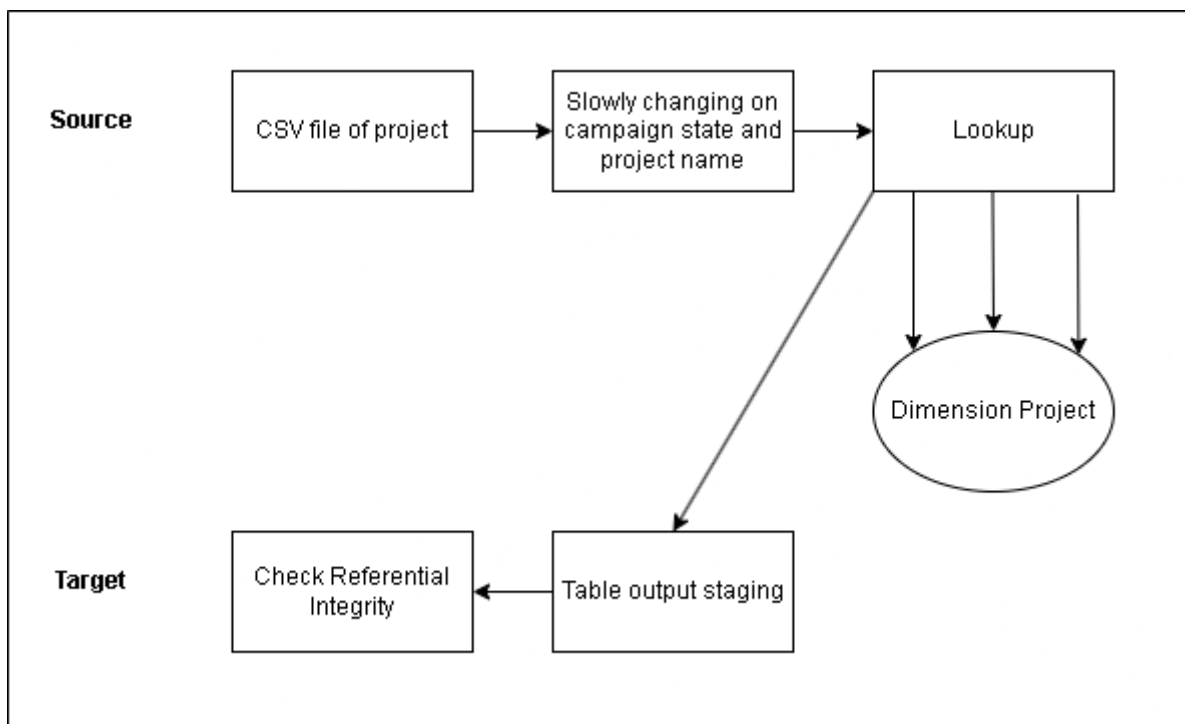


Gambar 3. Diagram Tingkat Tinggi Sumber ke Target

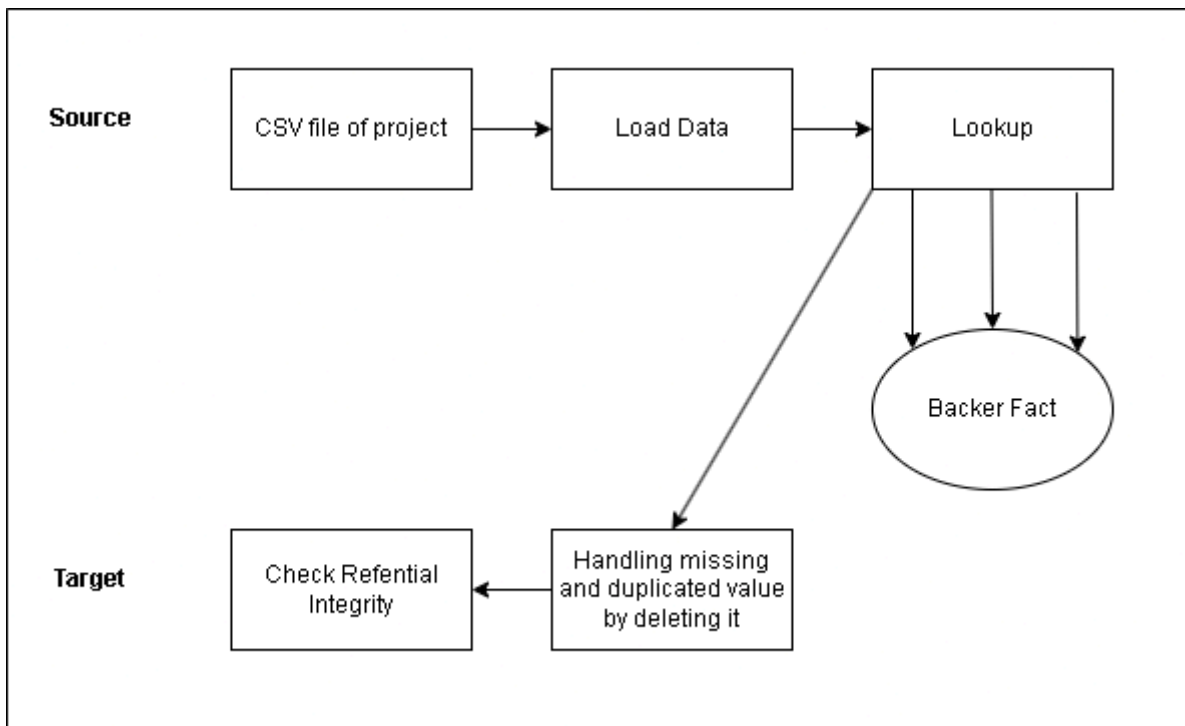
2.5.2. Detailed ETL Flow for Each Source to Target



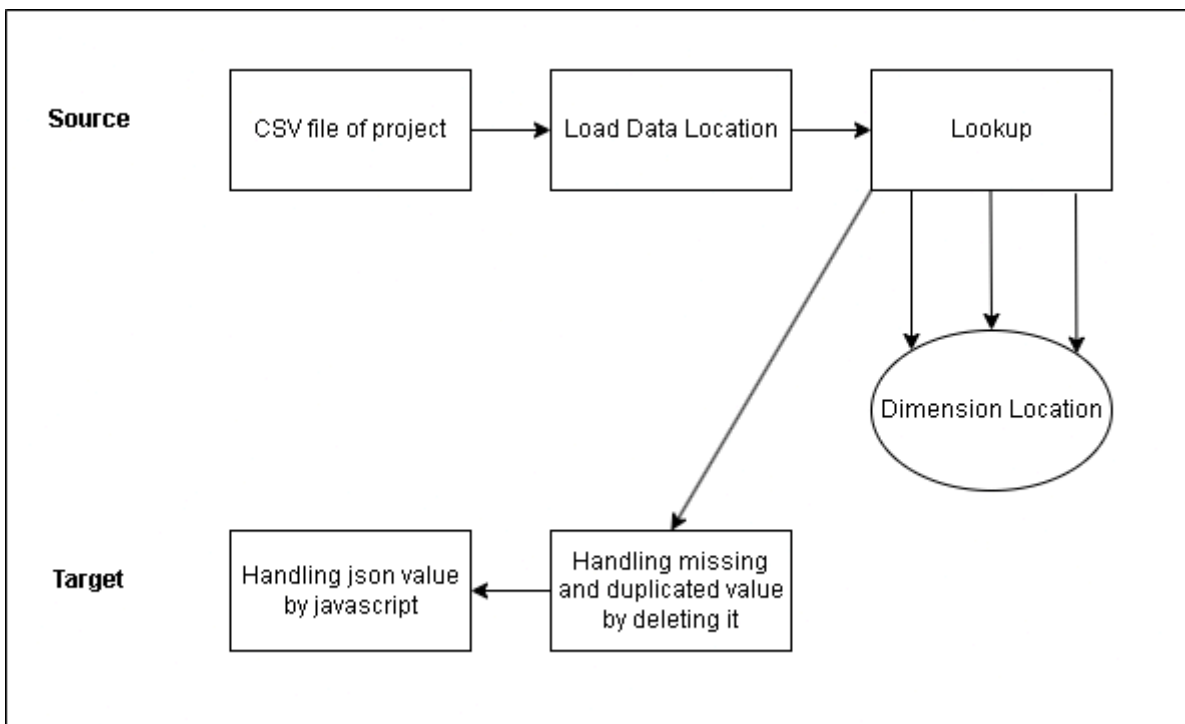
Gambar 4. Diagram Rinci ETL Campaign ke Tabel Category



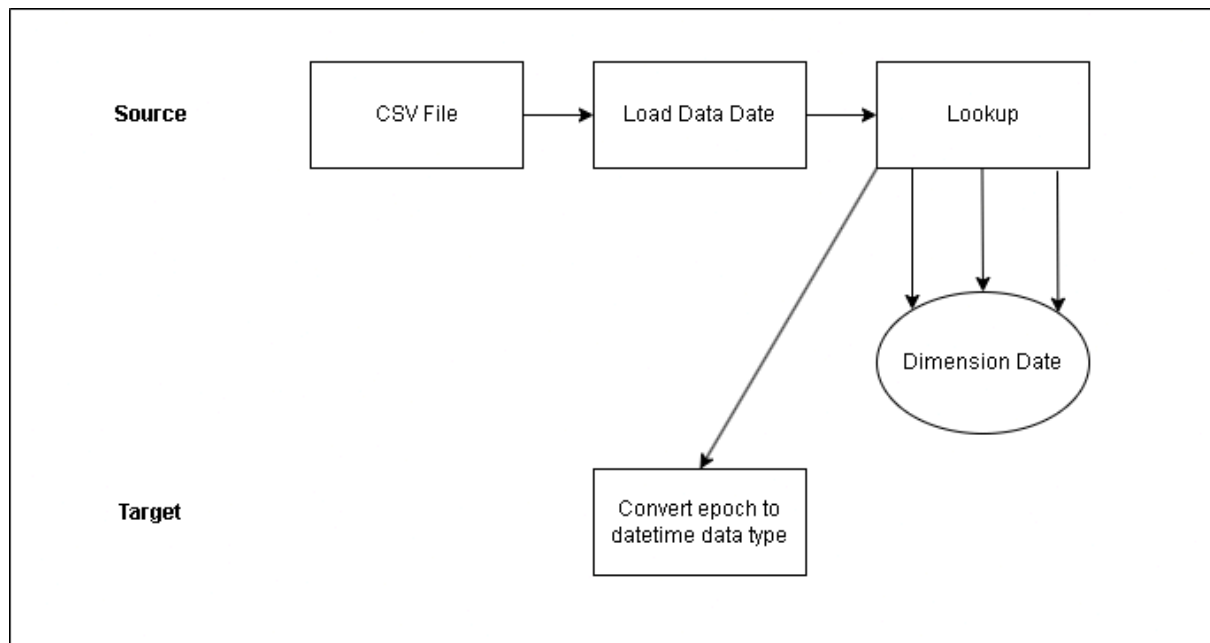
Gambar 5. Diagram Rinci ETL Campaign ke Tabel Project



Gambar 6. Diagram Rinci ETL Campaign ke Tabel Backer_Fact



Gambar 7. Diagram Rinci ETL Campaign ke Tabel Location



Gambar 8. Diagram Rinci ETL Campaign ke Tabel Date

2.6. Business Intelligence Application

2.6.1. Business Intelligence Application Specification

1. Pentaho

Pentaho adalah sebuah platform open-source untuk integrasi data, pelaporan, analisis data, dan manajemen data yang digunakan dalam Business Intelligence (BI). Dikembangkan oleh Hitachi Vantara, Pentaho menyediakan solusi lengkap untuk kebutuhan analisis data bisnis dari berbagai sumber data, mengintegrasikan, dan menyajikan data dalam bentuk yang mudah dipahami dan digunakan untuk pengambilan keputusan.

2. Tableau

Tableau adalah perangkat lunak visualisasi data yang digunakan untuk mengubah data mentah menjadi format yang mudah dimengerti. Ini sangat berguna dalam Business Intelligence (BI) untuk membuat dashboard interaktif dan laporan yang membantu pengguna dalam analisis data. Tableau memungkinkan pengguna untuk menghubungkan, memvisualisasikan, dan berbagi data dengan cepat dan mudah.

3. Microsoft SQL Server

Microsoft SQL Server adalah perangkat lunak yang digunakan untuk mengelola database relasional. Ini mendukung berbagai operasi seperti penyimpanan, pengambilan, dan manipulasi data, serta mendukung transaksi yang aman dan efisien. SQL Server menggunakan Structured Query Language (SQL) untuk berinteraksi dengan databasenya.

2.6.2. Detailed Business Intelligence Application Specification

1. Pentaho

1.1. Transformation

Elemen kunci yang digunakan untuk mengelola dan memproses data, seperti Mengambil data dari berbagai sumber data, seperti database, file flat, file CSV, Excel, sistem ERP, layanan web, dan banyak lagi. Kemudian untuk Transformasi data mencakup pembersihan data (cleaning), penggabungan data (merging), pengelompokan data (grouping), penyaringan data (filtering), dan agregasi data. Yang terakhir adalah Memuat data yang telah diproses dan ditransformasikan ke dalam tujuan akhir, seperti data warehouse, database, atau sistem lain yang digunakan untuk analisis dan pelaporan.

1.2. Connection

Koneksi dalam Pentaho Data Integration (PDI) adalah langkah penting untuk menghubungkan dan mengakses berbagai sumber data. Dalam konteks ETL (Extract,

Transform, Load), koneksi memungkinkan Pentaho untuk mengekstraksi data dari sumber, mentransformasikan, dan kemudian memuatnya ke dalam tujuan akhir. Contohnya adalah menghubungkan ke berbagai jenis database seperti MySQL, PostgreSQL, Oracle, Microsoft SQL Server, SQLite, dan lain-lain.

1.3. Job

Job di Pentaho adalah kumpulan langkah (steps) yang dijalankan secara berurutan atau berdasarkan logika tertentu. Setiap langkah dalam job dapat berupa transformasi, script, kontrol alur (flow control), atau operasi lainnya yang diperlukan dalam proses ETL.

2. Tableau

2.1. Worksheet

Worksheet di Tableau adalah kanvas utama tempat kita membuat dan mengatur visualisasi data. Pada worksheet, Anda dapat memilih dan menampilkan data dari sumber data yang terhubung, membuat berbagai jenis chart, dan mengkonfigurasi visualisasi agar sesuai dengan kebutuhan analisis kita. Setiap worksheet dapat berisi satu atau beberapa visualisasi yang menampilkan data secara interaktif.

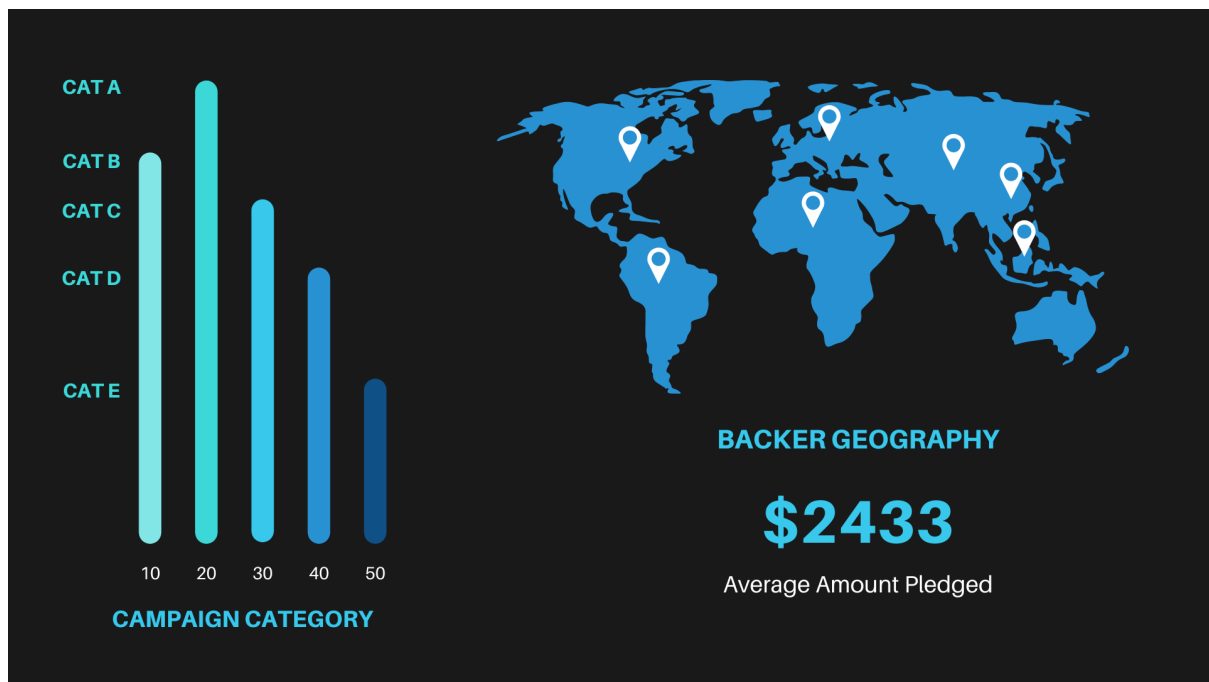
2.2. Filter

Filter pada Tableau digunakan untuk membatasi dan mengontrol data yang ditampilkan dalam visualisasi. Dengan menggunakan filter, kita dapat menyoroti informasi yang relevan, menghilangkan data yang tidak diperlukan, dan membuat analisis lebih fokus dan mudah dipahami.

2.3. Dashboard

Dashboard adalah antarmuka visualisasi data yang kita buat untuk menyajikan informasi penting secara ringkas dan mudah dipahami dari setiap sheet yang ada, sering kali menggunakan grafik, bagan, atau tabel. Dashboard dirancang untuk memberikan gambaran umum tentang kinerja, status, atau kondisi dari berbagai metrik dan data secara real-time atau dalam periode waktu tertentu.

2.6.3. Business Intelligence Application Mockup



Gambar 9. Mockup

Elemen mockup :

Campaign Category : Category A - Category E

Backer Geography : Geografi dari backer yang memiliki konsentrasi potensial tinggi

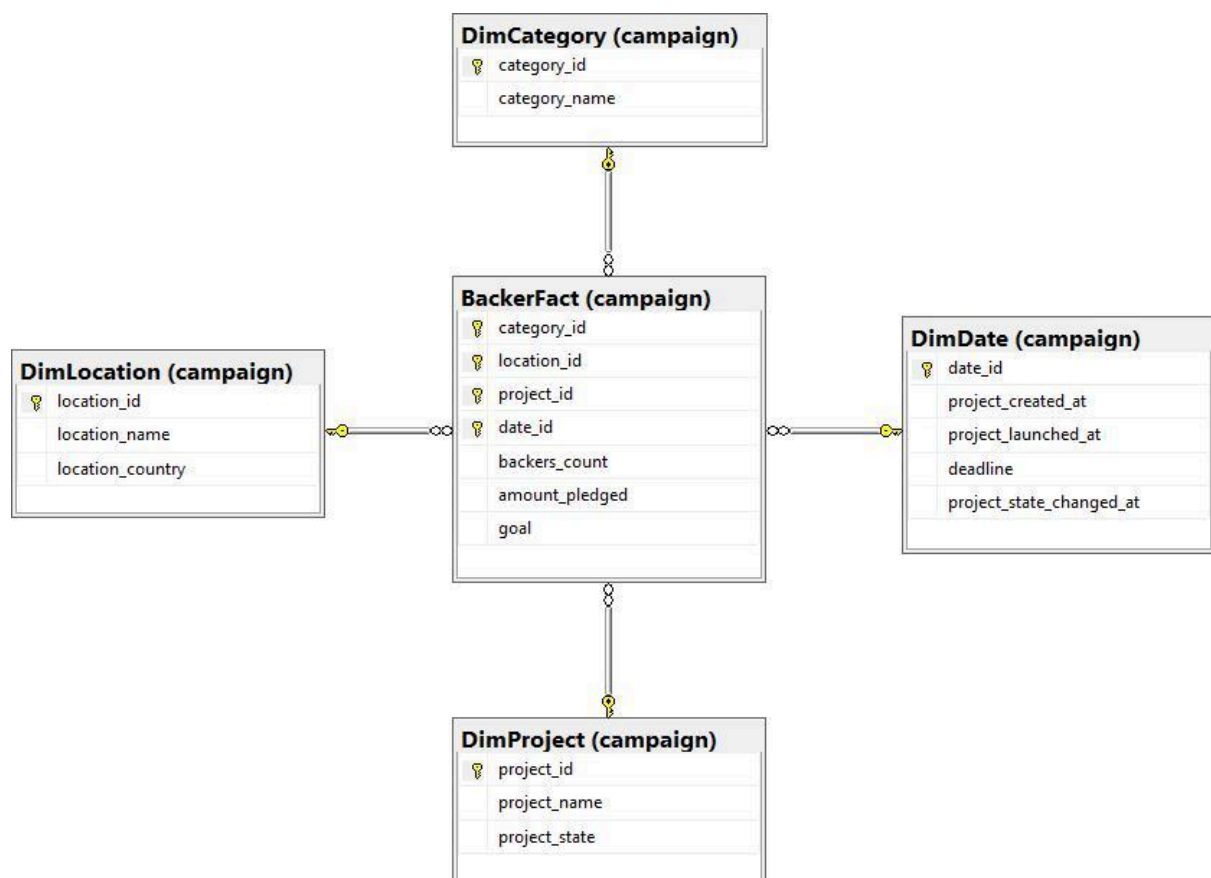
Average Amount : Rata-rata amount pledged

3. Implementasi

Bagian ini berisi ROLAP *schema*, kode program pada ETL, basis data MOLAP, dan *Business Intelligence Front End*.

3.1. ROLAP Schema

ROLAP (Relational Online Analytical Processing) adalah salah satu metode OLAP (Online Analytical Processing) yang digunakan dalam data warehousing untuk memungkinkan analisis multidimensional data yang disimpan dalam sistem manajemen basis data relasional (RDBMS) menggunakan tabel fakta dan dimensi. ROLAP memetakan operasi OLAP ke dalam query SQL yang dijalankan di atas data relasional. Pada tugas kali ini kami mengimplementasikan *star-schema* di mana tabel fakta berada di pusat, dan tabel dimensi mengelilinginya, menyerupai bentuk bintang. Tabel fakta adalah kolom yang menyimpan nilai numerik yang dapat dianalisis. Sedangkan tabel dimensi adalah tabel yang menyediakan konteks untuk data dalam tabel fakta dan berisi deskripsi atau atribut yang terkait dengan data transaksi.

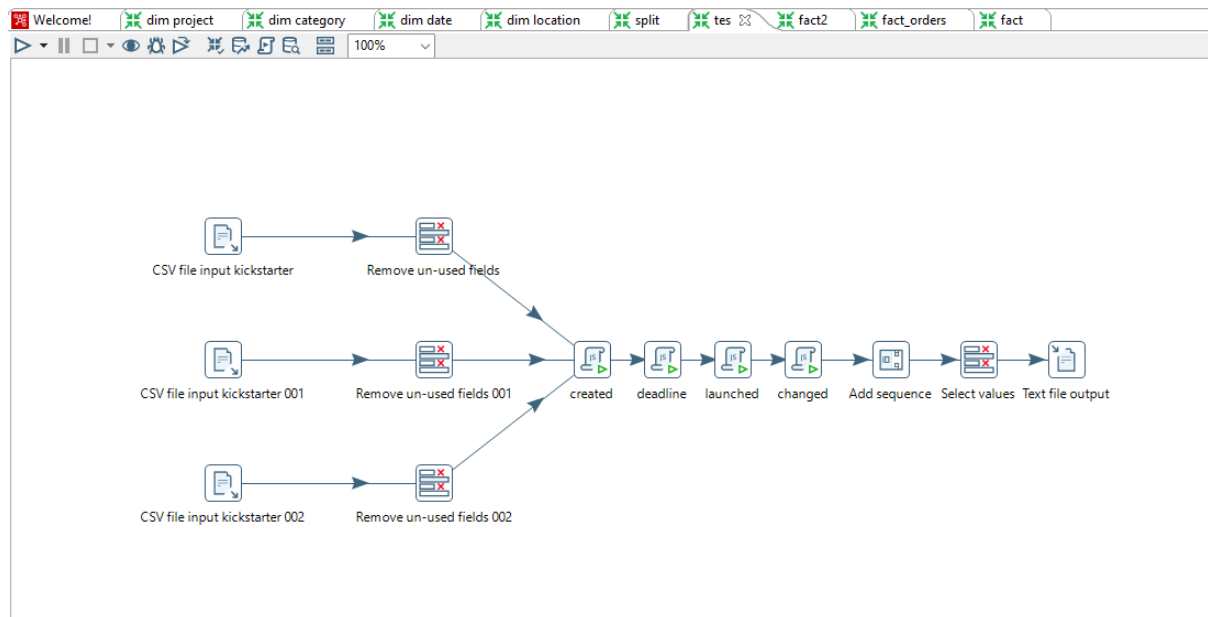


Gambar 10. Rolap Schema

3.2. ETL

ETL (Extract, Transform, Load) adalah proses yang digunakan dalam manajemen data untuk mentransfer data dari berbagai sumber ke dalam gudang data atau sistem penyimpanan data lainnya. Ini melibatkan serangkaian langkah yang diterapkan secara berurutan untuk memastikan data yang relevan, bersih, dan siap digunakan untuk analisis atau pelaporan.

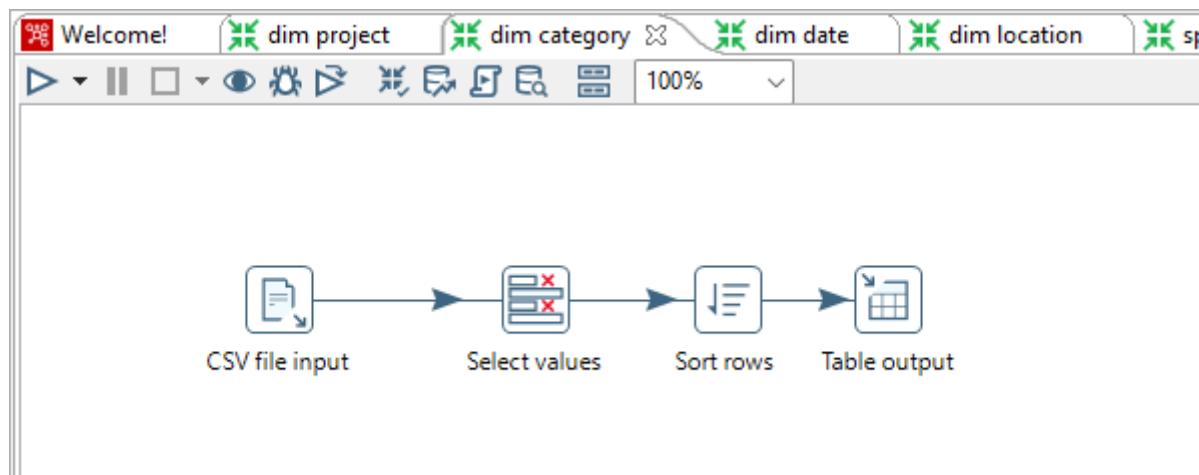
3.2.1 Concat Data Source Kickstarter



Gambar 11. Concat Data Source

Langkah pertama yang kami lakukan adalah menggabungkan dataset Kickstarter yang diunduh dalam bentuk file zip, yang berisi tiga file CSV yaitu, Kickstarter001, Kickstarter002, dan Kickstarter. Proses ini dimulai dengan memeriksa setiap file CSV untuk memahami struktur dan isinya, kemudian memilih kolom-kolom yang relevan untuk digunakan dan ditransformasi. Setelah kolom-kolom yang diperlukan dipilih, kami menggunakan JavaScript untuk memisahkan dan mengubah nilai-nilai JSON dalam beberapa kolom menjadi format yang lebih mudah dikelola. Kemudian kami melakukan perubahan terhadap semua nilai epoch dari Kickstarter menjadi Datetime. Jika semua langkah ini berhasil dijalankan tanpa masalah, maka proses penggabungan file dianggap selesai. Hasil dari langkah ini adalah satu file CSV yang merupakan gabungan dari ketiga file tersebut.

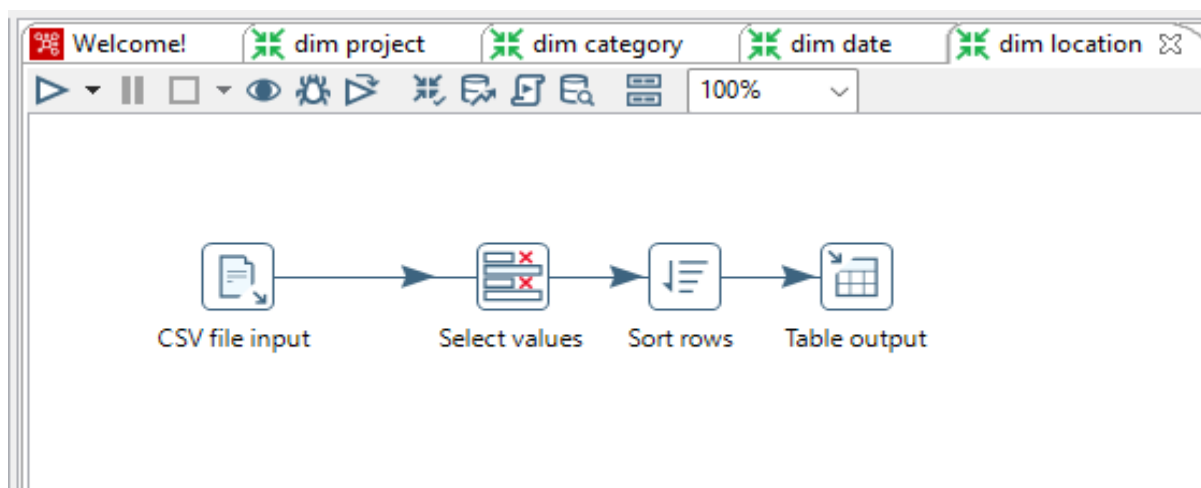
3.2.2 ETL dari Dataset ke Table Staging Dim Category



Gambar 12. Table Staging Category

Pada langkah kedua, kami melakukan transformasi tambahan untuk staging dimensi category ke dalam database. Proses dimulai dengan memasukkan file CSV Kickstarter yang sudah digabungkan ke dalam proses transformasi. Kemudian, kami melakukan pemilihan kolom yang relevan yang akan digunakan untuk dimensi category. Setelah kolom yang dibutuhkan dipilih, kami mengurutkan nama kategori secara alfabetis. Tahap terakhir adalah memasukkan hasil transformasi ke dalam database menggunakan tabel output. Penting untuk dicatat bahwa kami harus mengunduh driver yang sesuai agar dapat menghubungkan Spoon (perangkat lunak ETL) dengan database yang akan digunakan. Dalam praktikum kali ini, kami menggunakan Microsoft SQL Server dan driver JDBC untuk menghubungkan Spoon dengan database tersebut.

3.2.3 ETL dari Dataset ke Table Staging Dim Location

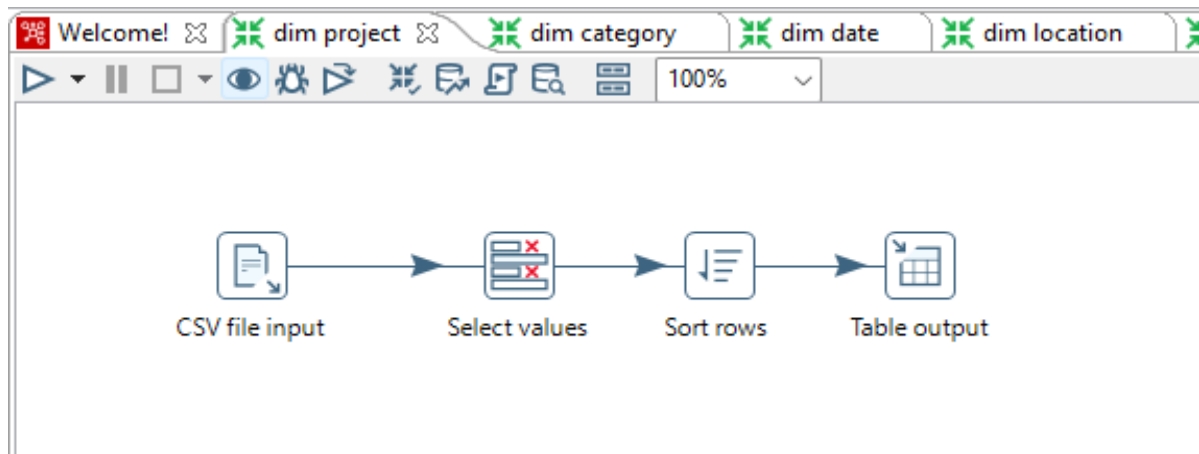


Gambar 13. Table Staging Location

Pada langkah ketiga, kami melakukan transformasi tambahan untuk staging dimensi location ke dalam database. Proses dimulai dengan memasukkan file CSV Kickstarter yang sudah

digabungkan ke dalam proses transformasi. Kemudian, kami melakukan pemilihan kolom yang relevan yang akan digunakan untuk dimensi location. Setelah kolom yang dibutuhkan dipilih, kami mengurutkan nama kategori secara alfabetis. Tahap terakhir adalah memasukkan hasil transformasi ke dalam database menggunakan tabel output.

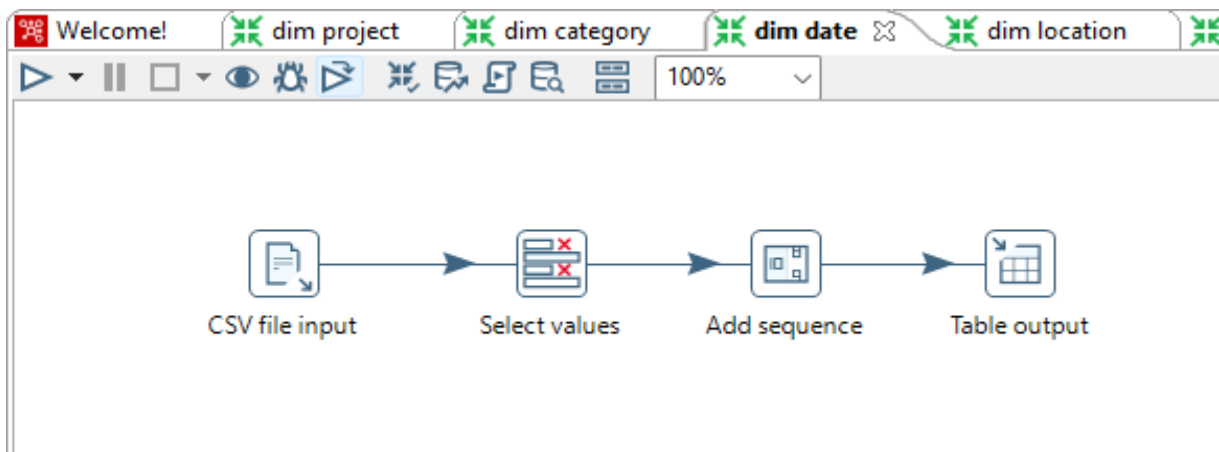
3.2.4 ETL dari dataset ke Table Staging Dim Project



Gambar 14. Table Staging Project

Pada langkah keempat, kami melakukan transformasi tambahan untuk staging dimensi project ke dalam database. Proses dimulai dengan memasukkan file CSV Kickstarter yang sudah digabungkan ke dalam proses transformasi. Kemudian, kami melakukan pemilihan kolom yang relevan yang akan digunakan untuk dimensi project. Setelah kolom yang dibutuhkan dipilih, kami mengurutkan nama project secara alfabetis. Tahap terakhir adalah memasukkan hasil transformasi ke dalam database menggunakan tabel output.

3.2.5 ETL dari dataset ke Table Staging Dim Date

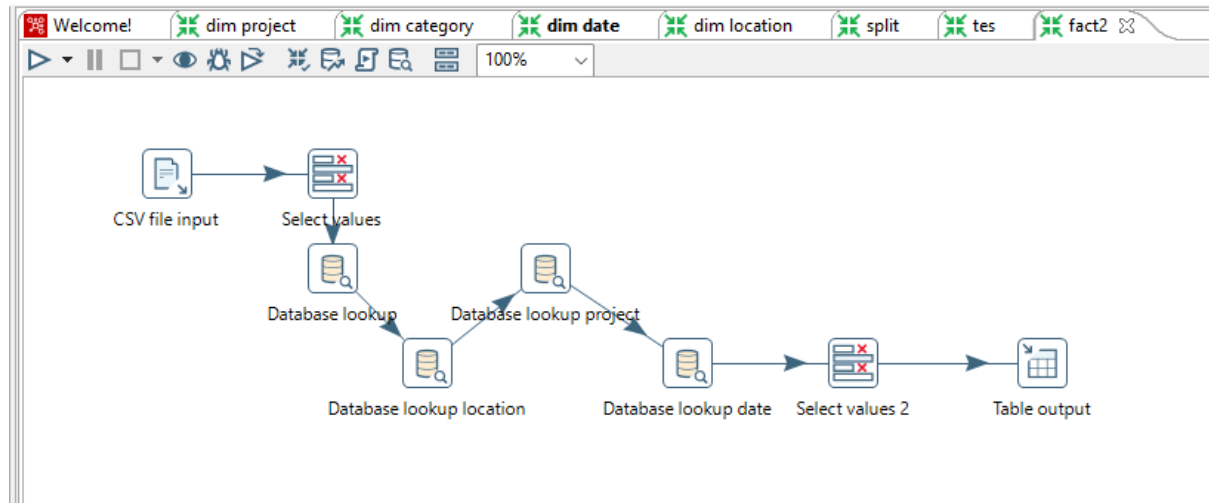


Gambar 15. Table Staging Date

Pada langkah kelima, kami melakukan transformasi tambahan untuk staging dimensi date ke dalam database. Proses dimulai dengan memasukkan file CSV Kickstarter yang sudah digabungkan ke dalam proses transformasi. Kemudian, kami melakukan pemilihan kolom

yang relevan yang akan digunakan untuk dimensi project. Setelah kolom yang dibutuhkan dipilih, kami mengurutkan nama project secara alfabetis. Tahap terakhir adalah memasukkan hasil transformasi ke dalam database menggunakan tabel output.

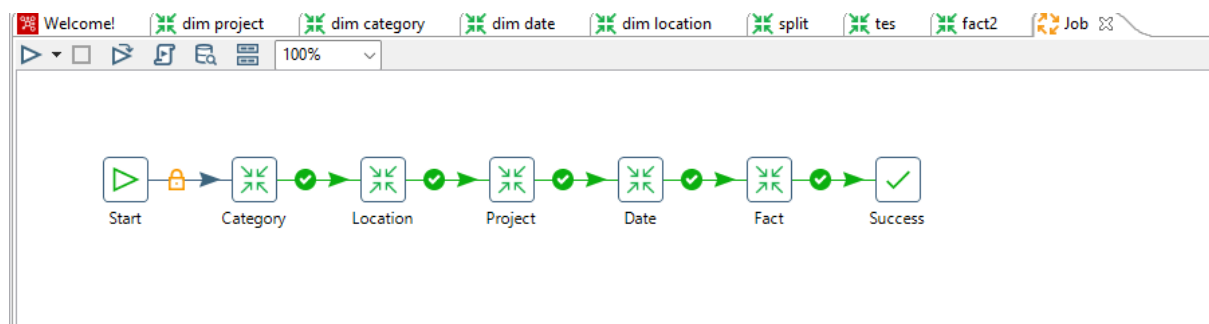
3.2.6 ETL dari dataset ke Table Staging Backer Fact



Gambar 16. Table Staging Backer Fact

Pada langkah keenam, kami melakukan penggabungan primary key dari setiap dimensi untuk dimasukkan ke dalam fact table. Proses ini melibatkan penggunaan database lookup untuk mengambil data yang diperlukan dari setiap dimensi yang telah dibuat sebelumnya di database. Kami juga menambahkan tiga kolom kuantitatif, yaitu *backers_count*, *pledged*, dan *goal*, yang akan digunakan dalam melakukan analisis data. Kolom-kolom ini penting karena akan digunakan untuk menjawab pertanyaan analisis yang telah kami susun sebelumnya. Hasil dari langkah ini adalah pembentukan sebuah fact table yang disebut *Backer_Fact*, yang berisi data yang telah digabungkan dari berbagai dimensi dan kolom-kolom kuantitatif yang telah ditambahkan.

3.2.7 Job ETL

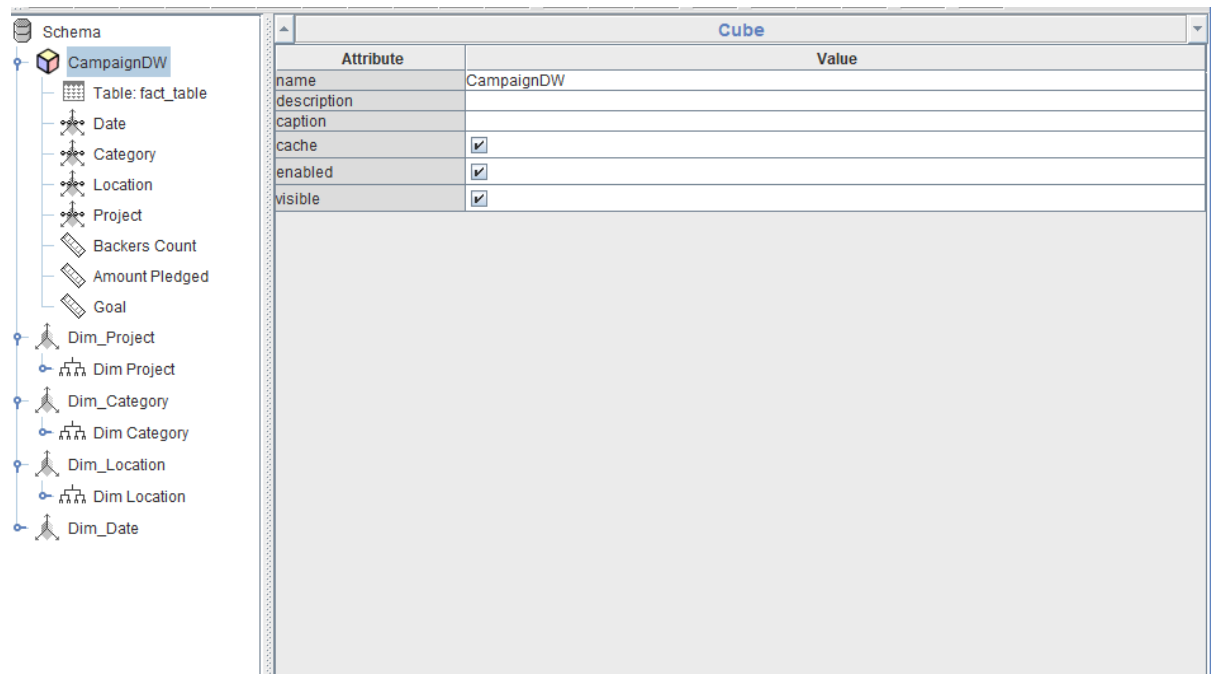


Gambar 17. Table Staging Job ETL

Job pada Pentaho digunakan untuk mengatur dan menjalankan serangkaian langkah atau transformasi secara bersamaan atau secara terjadwal. Pada job ini kami mengelola alur kerja yang lebih kompleks, yang melibatkan beberapa transformasi atau langkah yang harus

dijalankan secara berurutan atau dalam kondisi tertentu. Dengan menggunakan job, kami dapat mengatur bagaimana dan kapan setiap langkah dalam alur kerja akan dieksekusi.

3.3. MOLAP Schema

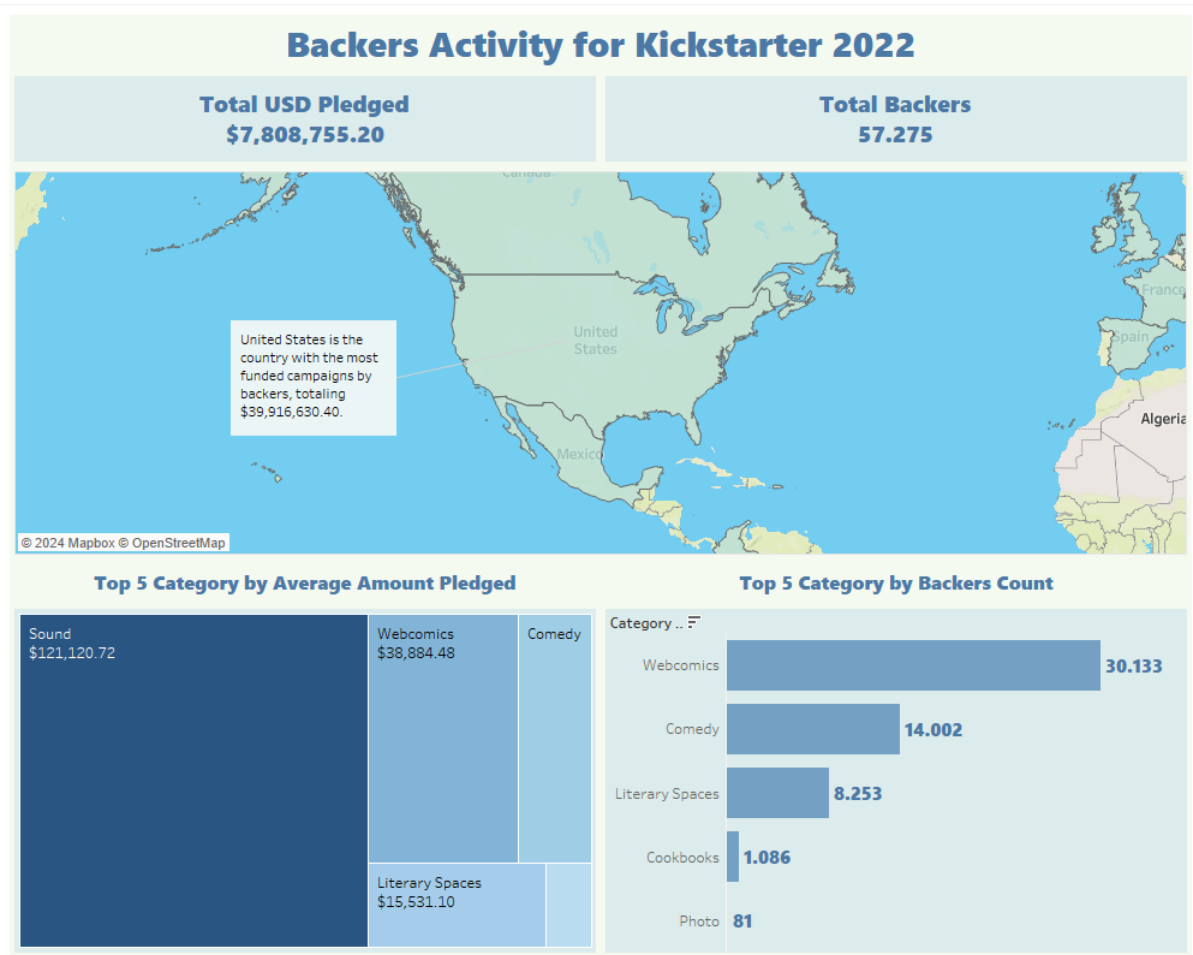


| Attribute | Value |
|-------------|-------------------------------------|
| name | CampaignDW |
| description | |
| caption | |
| cache | <input checked="" type="checkbox"/> |
| enabled | <input checked="" type="checkbox"/> |
| visible | <input checked="" type="checkbox"/> |

Gambar 18. Molap Schema

3.4. Dashboard

https://public.tableau.com/views/KickstarterAnalysis2022/Dashboard1?:language=en-US&publish=yes&:sid=&:display_count=n&:origin=viz_share_link

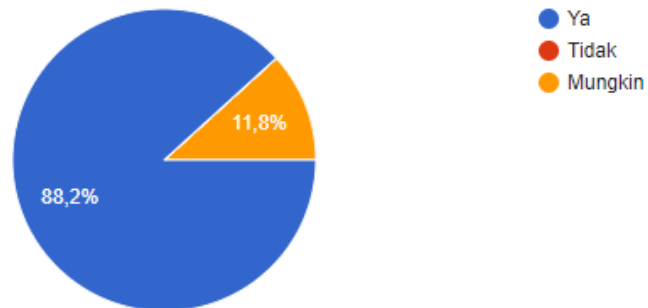


Gambar 19. Dashboard

4. Evaluasi

Apakah Dashboard di atas menjawab pertanyaan analisis nomor 1?

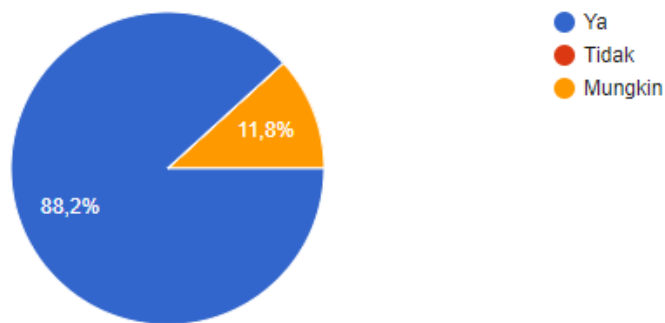
17 jawaban



Gambar 20. Hasil Evaluasi 1

Apakah Dashboard di atas menjawab pertanyaan analisis nomor 2?

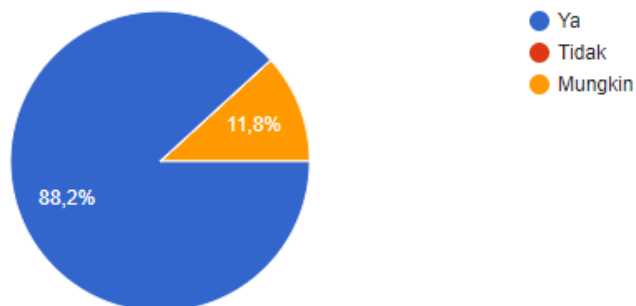
17 jawaban



Gambar 21. Hasil Evaluasi 2

Apakah Dashboard di atas menjawab pertanyaan analisis nomor 3?

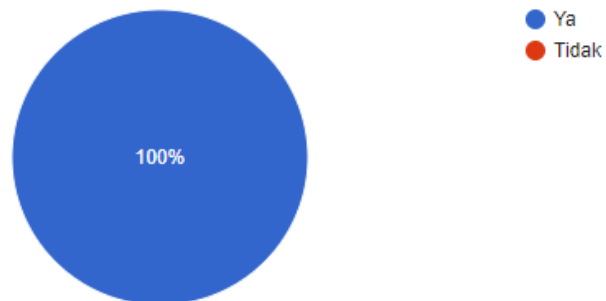
17 jawaban



Gambar 22. Hasil Evaluasi 3

Apakah warna yang digunakan nyaman untuk dilihat?

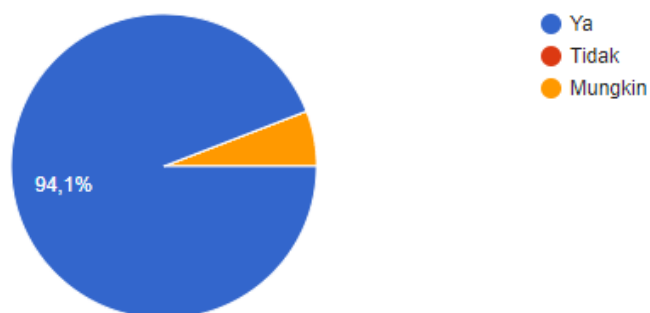
17 jawaban



Gambar 23. Hasil Evaluasi 4

Apakah informasi dari dashboard mudah dipahami?

17 jawaban



Gambar 24. Hasil Evaluasi 5

5. Penutup

5.1. Kesimpulan

Dari dashboard "Backers Activity for Kickstarter 2022", kita dapat menyimpulkan beberapa poin penting terkait aktivitas pendukung di Kickstarter pada tahun tersebut. Total dana yang terkumpul melalui platform ini mencapai \$7,808,755.20, dengan jumlah pendukung atau backers sebanyak 57,275 orang. Amerika Serikat merupakan negara dengan kontribusi pendanaan terbesar, yaitu \$39,916,630.40. Kategori "Sound" menjadi kategori dengan jumlah dana rata-rata tertinggi yang dikumpulkan, yakni \$121,120.72, diikuti oleh "Webcomics" dengan \$38,884.48, dan "Comedy" dengan \$33,882.44. Di sisi lain, kategori "Webcomics" menarik jumlah pendukung terbanyak dengan 30,133 orang, diikuti oleh "Comedy" dengan 14,002 pendukung dan "Literary Spaces" dengan 8,253 pendukung. Data ini menunjukkan bahwa Amerika Serikat memainkan peran dominan dalam pendanaan Kickstarter, dan kategori tertentu seperti "Sound" dan "Webcomics" berhasil menarik perhatian lebih banyak pendukung dan dana dibandingkan kategori lainnya. Dashboard ini memberikan gambaran yang jelas tentang bagaimana dana dan keterlibatan pendukung terdistribusi di berbagai kategori proyek di Kickstarter sepanjang tahun 2022.

5.2. Saran

Berikut beberapa saran untuk pengerjaan proyek analisis Kickstarter:

1. Pastikan Anda memahami tujuan utama dari proyek ini. Apakah Anda ingin menganalisis tren pendanaan, kategori proyek yang paling populer, atau demografi pendukung? Tujuan yang jelas akan memandu setiap langkah proses.
2. Pastikan data Kickstarter yang digunakan telah divalidasi dan dibersihkan. Tangani missing values dan data duplikat.
3. Lakukan transformasi data yang diperlukan untuk memastikan data siap digunakan dalam analisis. Ini termasuk normalisasi, agregasi, dan pembentukan dimensi dan fakta yang relevan.

LAMPIRAN

Lampiran A. Spesifikasi Aplikasi Kecerdasan Bisnis

| No | Nama Aplikasi Kecerdasan Bisnis | Deskripsi Singkat | Kategori Aplikasi | Nama Grup Pengguna | Skor Kebutuhan | Level Usaha dalam Pengembangannya | Tipe Aplikasi | Elemen | Komentar |
|----|---------------------------------|--|--------------------------------|--------------------|----------------|-----------------------------------|-----------------|----------------------|----------|
| 1 | Pentaho | Pentaho adalah sebuah platform intelijen bisnis (Business Intelligence) dan analitik data yang komprehensif, yang menyediakan berbagai alat untuk pengumpulan, integrasi, analisis, dan visualisasi data. | Business Intelligence Software | Kelompok 02_DWBI | 9/10 | Sangat Tinggi | Perangkat Lunak | Transformation, Job | |
| 2 | Tableau | Tableau adalah aplikasi kecerdasan bisnis yang digunakan untuk melakukan visualisasi dan analisa terhadap data dari sumber data. | Business Intelligence Software | Kelompok 02_DWBI | 9/10 | Sangat Tinggi | Perangkat Lunak | Worksheet, Dashboard | |
| 3 | Microsoft SQL Server | Microsoft SQL Server adalah perangkat lunak yang digunakan untuk mengelola database relasional. Ini mendukung berbagai operasi seperti penyimpanan, pengambilan, dan manipulasi data, serta mendukung transaksi yang aman dan efisien. SQL Server menggunakan Structured Query Language (SQL) untuk berinteraksi dengan databasenya. | Database Design | Kelompok 02_DWBI | 9/10 | Sangat Tinggi | Perangkat Lunak | Schema, Database | |

Lampiran B. Spesifikasi Rinci Aplikasi Kecerdasan Bisnis

Aplikasi Tableau

| No | Elemen/Atribut | Lokasi | Tipe Fungsi | Nilai Default | Sumber | Dibuat di | Kueri | Komentar |
|----|----------------|----------|---|---|---|--|-------|--|
| 1 | Worksheet | Aplikasi | Measures, Dimension, Aggregation, Visualization | Nilai default dalam konteks worksheet Tableau dapat merujuk pada pengaturan awal atau nilai yang diambil secara otomatis oleh Tableau saat membuat visualisasi, seperti format tampilan, warna, ukuran, atau filter default yang diterapkan pada visualisasi. | Sumber data yang digunakan berasal dari database MSSQL. | Worksheet dapat diakses dari halaman awal membuka aplikasi | - | Worksheet adalah fitur yang digunakan sebagai area kerja utama untuk membuat visualisasi data. |
| 2 | Filter | Aplikasi | Digunakan untuk mengolah data sesuai dengan kebutuhan. | Nilai yang ada berbeda-beda tergantung pada konfigurasi dan pengaturan yang digunakan sesuai kebutuhan. | Sumber data yang digunakan berasal dari database MSSQL. | Dapat dibuat di sheet maupun dashboard. | - | Filter adalah fitur yang digunakan untuk menyaring data dalam sumber data yang dianalisis. |
| 3 | Dashboard | Aplikasi | Digunakan untuk menyatukan semua worksheet yang telah dibuat. | Nilai yang ada berbeda-beda tergantung pada konfigurasi dan pengaturan yang digunakan sesuai kebutuhan. | Sumber data yang digunakan berasal dari database MSSQL. | Dapat dibuat melalui tab dashboard. | - | Tampilan akhir yang akan diunggah ke Tableau Public untuk dipakai oleh pengguna. |

Aplikasi Pentaho

| No | Elemen/Atribut | Lokasi | Tipe Fungsi | Nilai Default | Sumber | Dibuat di | Kueri | Komentar |
|----|----------------|------------|---|---------------|---|-----------|-------|----------|
| 1 | Transformation | Local File | Workflow Execution | - | Sumber data yang berasal dari file CSV. | PDI | - | - |
| 2 | Connection | Database | Database Connection | - | Sumber data diperoleh dari Pentaho yang merupakan hasil proses ETL dan di staging ke MSSQL. | PDI | - | - |
| 3 | Job | Local File | Menjalankan job dari transformation yang telah dibuat | - | Sumber data diperoleh dari proses Transformation. | PDI | - | - |

Referensi

dataset : <https://webrobots.io/kickstarter-datasets/>