**LAPORAN PROYEK MATA KULIAH**

**12S3202 – GUDANG DATA DAN KECERDASAN BISNIS**

***Data Mart-based Dashboard for Kickstarter Crowdfunding (2022)***



**Disusun oleh:**

1. **12S21046 - Ruth Marelisa Hutagalung**
2. **12S21048 - Nessy Pentasonia Pangaribuan**
3. **12S21050 - Jessica Wasty Sitorus**
4. **12S21052 - Griselda**
5. **12S21054 - Diah Anastasya Sihombing**

**FAKULTAS INFORMATIKA DAN TEKNIK ELEKTRO**

**INSTITUT TEKNOLOGI DEL**

**2024**

# DAFTAR ISI

[DAFTAR ISI 1](#_Toc166959467)

[1. Pendahuluan 1](#_Toc166959468)

[1.1. Latar Belakang 1](#_Toc166959469)

[1.2. Tujuan Pengerjaan Proyek 2](#_Toc166959470)

[1.3. Ruang Lingkup 2](#_Toc166959471)

[1.4. Tim Pengembang 3](#_Toc166959472)

[2. Analisis dan Desain 5](#_Toc166959473)

[2.1. Pertanyaan Analisis 5](#_Toc166959474)

[2.2. Arsitektur Sistem 5](#_Toc166959475)

[2.3. Sumber Data 6](#_Toc166959476)

[2.4. Model Dimensional 10](#_Toc166959477)

[2.4.1. *High-Level Dimensional Model* 10](#_Toc166959478)

[2.4.2. *Dimensional Model Schema* 10](#_Toc166959479)

[2.4.3. *Detailed Dimensional Model* 11](#_Toc166959480)

[2.5. *Extract, Transform, and Load* 13](#_Toc166959481)

[2.5.1. *High-Level Source to Target Map* 14](#_Toc166959482)

[2.5.2. Detailed ETL Flow for Each Source to Target 15](#_Toc166959483)

[2.6. *Business Intelligence Application* 18](#_Toc166959484)

[2.6.1. *Business Intelligence Application Specification* 18](#_Toc166959485)

[2.6.2. *Detailed Business Intelligence Application Specification* 19](#_Toc166959486)

[2.6.3. *Business Intelligence Application Mockup* 20](#_Toc166959487)

[3. Implementasi 21](#_Toc166959488)

[3.1. ROLAP *Schema* 21](#_Toc166959489)

[3.2. ETL 23](#_Toc166959490)

[3.3. MOLAP Schema 29](#_Toc166959491)

[3.4. *Dashboard* 30](#_Toc166959492)

[4. Evaluasi 31](#_Toc166959493)

[5. Penutup 35](#_Toc166959494)

[5.1. Kesimpulan 35](#_Toc166959495)

[5.2. Saran 35](#_Toc166959496)

[LAMPIRAN 36](#_Toc166959497)

[Lampiran A. Spesifikasi Aplikasi Kecerdasan Bisnis 36](#_Toc166959498)

[Lampiran B. Spesifikasi Rinci Aplikasi Kecerdasan Bisnis 37](#_Toc166959499)

**DAFTAR TABEL**

[Tabel 1. 1 Susunan Tim Pengembang 3](#_Toc166959523)

[Tabel 2. 1 Atribut pada Sumber Data 6](#_Toc166959540)

**DAFTAR GAMBAR**

[Gambar 2. 1 Arsitektur Sistem 5](#_Toc166959621)

[Gambar 2. 2 Bus Matrix 10](#_Toc166959622)

[Gambar 2. 3 Dimensional Model Shema 11](#_Toc166959623)

[Gambar 2. 4 Detailed DimLocation 12](#_Toc166959624)

[Gambar 2. 5 Detailed DimCategory 12](#_Toc166959625)

[Gambar 2. 6 Detailed DimDate 12](#_Toc166959626)

[Gambar 2. 7 Detailed DimProject 13](#_Toc166959627)

[Gambar 2. 8 Detailed DimCreator 13](#_Toc166959628)

[Gambar 2. 9 Detailed FactCrowdfunding 13](#_Toc166959629)

[Gambar 2. 10 High-Level to Target Map 15](#_Toc166959630)

[Gambar 2. 11 Detailed ETL Flow DimCategory 16](#_Toc166959631)

[Gambar 2. 12 Detailed ETL Flow DimCreator 16](#_Toc166959632)

[Gambar 2. 13 Detailed ETL Flow DimLocation 17](#_Toc166959633)

[Gambar 2. 14 Detailed ETL Flow DimDate 17](#_Toc166959634)

[Gambar 2. 15 Detailed ETL Flow DimProject 18](#_Toc166959635)

[Gambar 2. 16 Detailed ETL Flow FactCrowdfunding 18](#_Toc166959636)

[Gambar 2. 17 Business Intelligence Application Mockup 20](#_Toc166959637)

[Gambar 3. 1 ROLAP Schema 21](#_Toc166959669)

[Gambar 3. 2 Merge Data 23](#_Toc166959670)

[Gambar 3. 3 Split Data 1 24](#_Toc166959671)

[Gambar 3. 4 Split Data 2 24](#_Toc166959672)

[Gambar 3. 5 Remove 25](#_Toc166959673)

[Gambar 3. 6 DimCategory 25](#_Toc166959674)

[Gambar 3. 7 DimCreator 26](#_Toc166959675)

[Gambar 3. 8 DimDate 26](#_Toc166959676)

[Gambar 3. 9 DimLocation 27](#_Toc166959677)

[Gambar 3. 10 DimProject 27](#_Toc166959678)

[Gambar 3. 11 FactTable 28](#_Toc166959679)

[Gambar 3. 12 MOLAP Shema 1 29](#_Toc166959680)

[Gambar 3. 13 MOLAP Shema 2 30](#_Toc166959681)

[Gambar 3. 14 Dashboard Analysis 30](#_Toc166959682)

[Gambar 4. 1 Penjelasan Pertanyaan Analisis 31](#_Toc166959683)

[Gambar 4. 2 Dashboard Analysis 31](#_Toc166959684)

[Gambar 4. 3 Penjelasan Pengisian 32](#_Toc166959685)

[Gambar 4. 4 Analysis terhadap Pertanyaan Analisis nomor 1 32](#_Toc166959686)

[Gambar 4. 5 Analysis terhadap Pertanyaan Analisis nomor 2 32](#_Toc166959687)

[Gambar 4. 6 Analysis terhadap Pertanyaan Analisis nomor 3 33](#_Toc166959688)

[Gambar 4. 7 Pemahaman Responden terhadap Dashboard 33](#_Toc166959689)

[Gambar 4. 8 Pengaruh Warna terhadap Pemahaman 33](#_Toc166959690)

# Pendahuluan

*Crowdfunding* telah menjadi salah satu cara yang populer bagi individu, *startup*, dan perusahaan untuk menggalang dana untuk proyek-proyek kreatif dan inovatif mereka. Platform *Crowdfunding* seperti Kickstarter telah menjadi tempat utama bagi para pengusaha untuk memperoleh dukungan finansial dari masyarakat luas. Namun, dengan munculnya banyak proyek *Crowdfunding* setiap hari, menjadi semakin penting untuk dapat menganalisis tren dan pola yang muncul di dalamnya. Dengan terus bertambahnya jumlah proyek crowdfunding setiap harinya, kemampuan untuk mengidentifikasi dan menganalisis tren serta pola yang ada menjadi semakin penting. Ini tidak hanya tentang pemahaman pasar yang lebih baik tetapi juga tentang memiliki kemampuan untuk mengarahkan atau mengelola proyek dengan sukses melalui berbagai tantangan dan peluang yang ada di lingkungan crowdfunding. Memahami tren terkini membantu pengusaha membuat kampanye crowdfunding yang menarik dan relevan. Dengan analisis data, mereka bisa mengenali pola sukses dan menghindari kesalahan umum, menyesuaikan rencana mereka dengan permintaan pasar, dan mengarahkan proyek mereka menuju inovasi. Ini memfasilitasi pencarian sumber pendanaan yang strategis, memungkinkan pengusaha bergerak maju dengan sumber daya yang paling mendukung visi mereka. Singkatnya, analisis data memberikan panduan yang jelas dalam mengarahkan atau mengelola proyek dengan sukses melalui berbagai tantangan dan peluang yang ada di lingkungan crowdfunding.

## Latar Belakang

Persoalan yang kami temukan terkait dengan objek kaji yang kami pilih adalah kesulitan dalam menganalisis dan memahami tren serta pola yang muncul dalam proyek *Crowdfunding* di platform Kickstarter. Volume besar data yang dihasilkan oleh ribuan proyek yang aktif setiap saat membuat sulit untuk melihat gambaran keseluruhan, mengidentifikasi tren yang signifikan, dan memprediksi faktor-faktor yang mempengaruhi keberhasilan proyek.

Dengan adanya persoalan ini, kami memiliki solusi yaitu mengimplementasikan *Data Mart* sebagai bagian dari Sistem Gudang Data kami. *Data Mart* merupakan *subset* dari *Data Warehouse* yang terfokus pada kebutuhan spesifik dari satu departemen atau fungsi bisnis. Kami memilih *Data Mart* karena ini memungkinkan kami untuk menyimpan data yang relevan dengan proyek *Crowdfunding* secara terpisah, sehingga memudahkan akses dan analisis data dengan lebih cepat dan efisien.

Tipe *Business Intelligence* *application* yang akan kami implementasikan adalah dashboard interaktif. *Dashboard* ini akan menyajikan informasi secara visual tentang performa proyek *Crowdfunding*, termasuk metrik kunci seperti tingkat pendanaan, jumlah penggalang dana, dan tren penggalangan dana dari waktu ke waktu. Kami memilih *dashboard* interaktif karena ini memungkinkan para pengguna untuk dengan cepat memahami informasi yang kompleks dan membuat keputusan yang tepat berdasarkan wawasan yang diberikan. Selain itu, kemampuan interaktifnya akan memungkinkan para pengguna untuk mengeksplorasi data lebih lanjut sesuai dengan kebutuhan mereka.

## Tujuan Pengerjaan Proyek

Tujuan utama dari proyek ini adalah sebagai berikut.

1. Pengumpulan Data: Mengumpulkan data dari platform Kickstarter secara berkala untuk mencakup sebanyak mungkin proyek *Crowdfunding* yang sedang berlangsung.
2. Pembuatan *Data Mart*: Membangun struktur *Data Mart* yang sesuai untuk menyimpan dan mengelola data proyek *Crowdfunding* dengan efisien.
3. Pengembangan *Dashboard*: Mengembangkan *dashboard* interaktif yang memungkinkan pengguna untuk melihat dan menganalisis data proyek *Crowdfunding* secara intuitif.
4. Analisis Tren dan Pola: Melakukan analisis mendalam terhadap data proyek *Crowdfunding* untuk mengidentifikasi tren, pola, dan faktor-faktor yang mempengaruhi keberhasilan proyek.

## Ruang Lingkup

Ruang Lingkup dari pekerjaan proyek ini adalah sebagai berikut.

1. Sumber Data:

Data akan diperoleh langsung dari platform Kickstarter. Data akan mencakup informasi tentang proyek-proyek *Crowdfunding* yang sedang berlangsung, termasuk detail proyek, pendanaan, tanggal mulai dan berakhir, serta informasi dana.

1. Pendekatan Pengembangan Sistem:

Kami akan menggunakan Kimball *Approach* dalam pengembangan sistem. Pendekatan ini akan memungkinkan kami untuk fokus pada pembangunan *Data Mart* yang terpisah untuk menyimpan data proyek *Crowdfunding* dengan lebih terstruktur dan efisien.

1. Pengembangan Dashboard:

Kami akan mengembangkan dashboard interaktif menggunakan *tools* tableau. *Dashboard* akan menyajikan informasi tentang performa proyek *Crowdfunding* secara visual dengan menggunakan grafik, diagram, dan tabel yang interaktif.

1. Analisis Tren dan Pola

Kami akan melakukan analisis tren dan pola menggunakan teknik analisis data *clustering.* Analisis ini akan dilakukan untuk mengidentifikasi faktor-faktor yang mempengaruhi keberhasilan proyek *Crowdfunding* dan tren-tren yang muncul dari waktu ke waktu.

## Tim Pengembang

Pada Tabel 1 disajikan susunan tim kopengembang pada proyek ini.

Tabel 1. 1 Susunan Tim Pengembang

| No. | Nama | Peran | Tanggung Jawab |
| --- | --- | --- | --- |
|  | Nessy Pentasonia Pangaribuan | Ketua | * Mengerjakan BAB 1 * Membuat *sheets* visualisasi dan dashboard * Mengerjakan bagian Evaluasi |
|  | Griselda | Anggota | * Pertanyaan Analisis * Bux Matrix * DimCategory, DimCreator, DimDate, DimLocation, dan DimProject pada ETL * MOLAP |
|  | Ruth Marelisa Hutagalung | Anggota | * Mengerjakan BAB 1 * Merge data, split data, remove dan fact table pada ETL * ROLAP * Mengerjakan Bab 5 |
|  | Jessica Wasty Sitorus | Anggota | * Mengerjakan BAB 2 * Membuat High-Level Source to Target Map * Membuat Detailed ETL Flow for Each Source to Target |
| 5 | Diah Anastasya Sihombing | Anggota | * Mengerjakan BAB 2 * Pertanyaan Analisi * Arsitektur Sitem * Dimensional Model Shema * Detailed Dimensional Model |

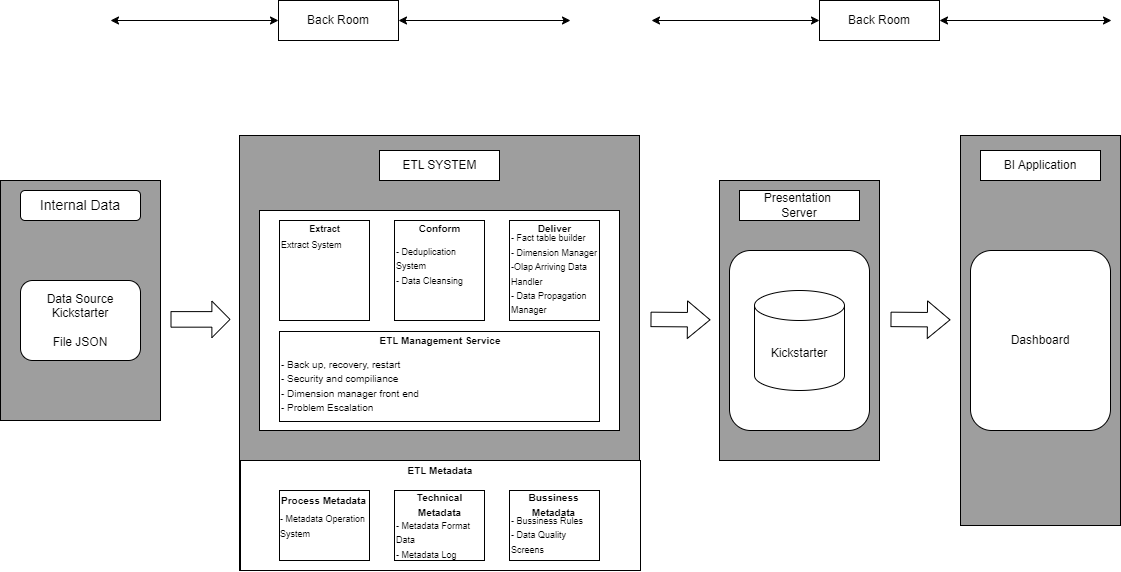
# Analisis dan Desain

Bagian ini berisi pertanyaan analisis yang perlu disertakan dalam pengembangan sistem, arsitektur sistem yang direncanakan, sumber data yang akan digunakan, dan model dimensional yang relevan dengan pertanyaan analisis yang telah dibuat.

## Pertanyaan Analisis

1. Bagaimana tren jumlah dana pendanaan proyek di Kickstarter pada tahun 2022?
2. Apa strategi pendanaan yang paling efektif yang digunakan oleh proyek di setiap kategori di Kickstarter pada tahun 2022?
3. Apakah ada kategori tertentu di Kickstarter yang mendapatkan jumlah dana pendanaan paling tinggi pada tahun 2022? Bagaimana perbandingannya dengan jumlah dana pendanaan pada kategori lainnya?

## Arsitektur Sistem



Gambar 2. 1 Arsitektur Sistem

## Sumber Data

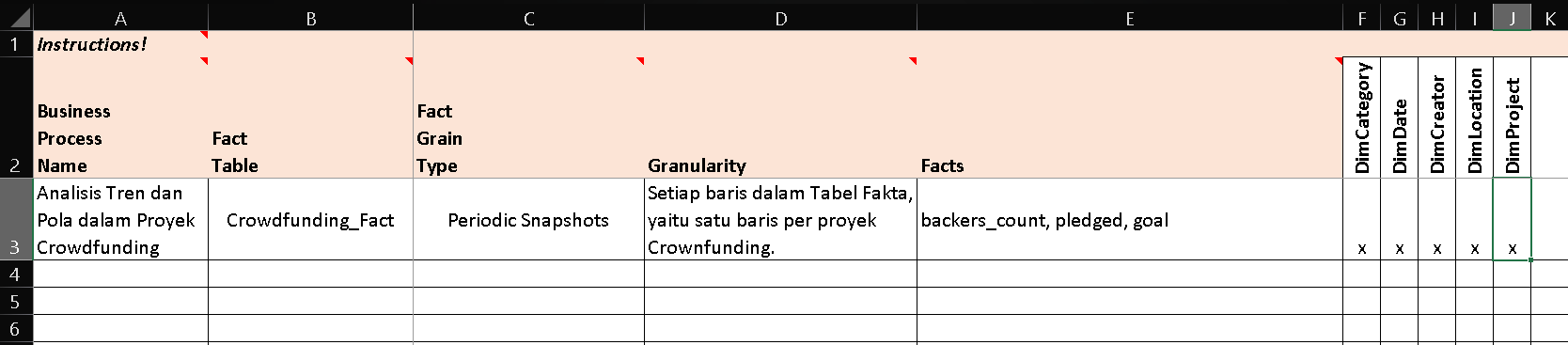
Kickstarter adalah sebuah platform crowdfunding online yang menghasilkan data dalam berbagai format seperti spreadsheet, RDBMS, JSON, atau format data lainnya, tergantung pada cara pengumpulan dan penyimpanan data tersebut. Dalam kasus ini, sumber data yang digunakan adalah file .csv yang berasal dari data Kickstartarter pada tahun 2022.

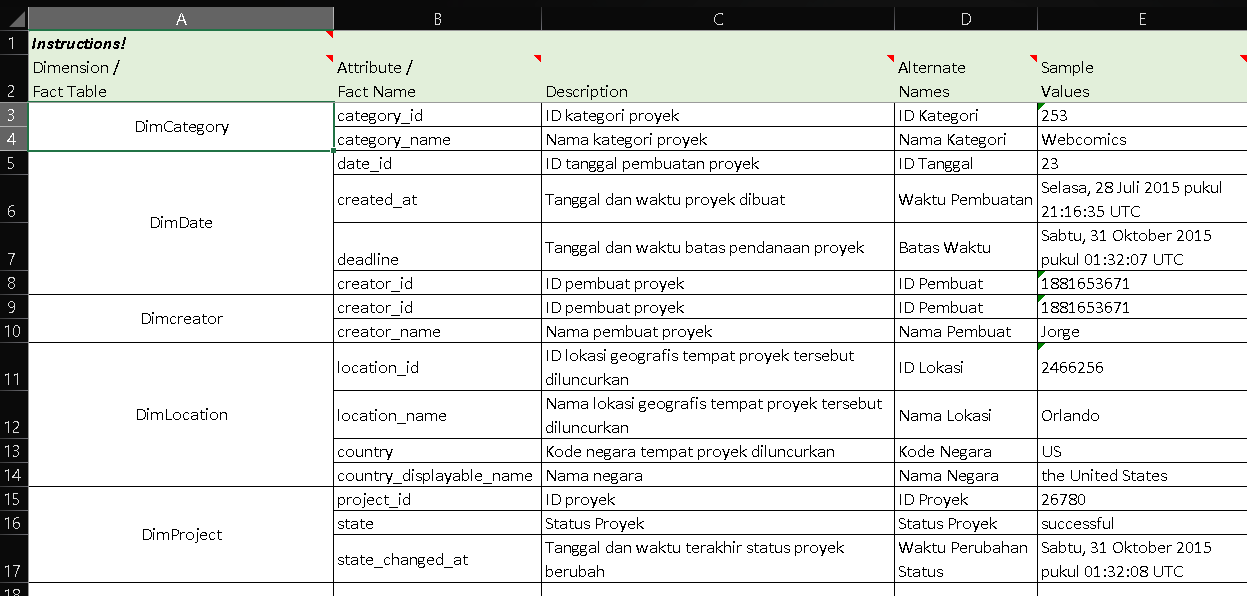
Tabel 2. 1 Atribut pada Sumber Data

| **Nama Atribut** | **Keterangan** | **Tipe Data** | **Panjang Atribut** |
| --- | --- | --- | --- |
| Backers\_count | Jumlah orang yang telah mendukung proyek | Integer | 10 |
| Blurb | Deskripsi singkat proyek yang digunakan untuk mempromosikan proyek | Integer | 255 |
| Category | Kategori proyek | String | 255 |
| Converted\_pledged\_amount | Jumlah dana yang terkumpul untuk proyek | Integer | 10 |
| Country | Singkatan negara asal pembuat proyek | String | 255 |
| Country\_displayable\_name | Nama lengkap negara asal proyek | String | 255 |
| Created\_at | Tanggal dan waktu proyek dibuat | Date |  |
| Creator | Nama pembuat proyek | String | 255 |
| Currency | Mata uang yang digunakan untuk pendanaan proyek | String | 8 |
| Currency\_symbol | Simbol mata uang yang digunakan untuk pendanaan proyek | String | 8 |
| Currency\_trailing\_code | Kode mata uang yang digunakan untuk pendanaan proyek. | Boolean |  |
| Current\_currency | Mata uang yang digunakan untuk menampilkan informasi tentang proyek di Kickstarter. | String | 8 |
| Deadline | Tanggal dan waktu batas pendanaan proyek. | Integer | 10 |
| Friends | Berisi keterangan kolega/teman | String | 255 |
| Disable\_communication | Atribut boolean yang menunjukkan apakah pembuat proyek telah menonaktifkan kemampuan untuk menerima pesan dari pendukung. | Boolean |  |
| Fx\_rate | Nilai tukar antara mata uang asli proyek dan mata uang yang digunakan di Kickstarter | Float | 16 |
| Goal | Target pendanaan yang ingin dicapai proyek | Float | 8 |
| Id | ID unik untuk setiap proyek di Kickstarter | Integer | 10 |
| Is\_backing | Atribut boolean yang menunjukkan apakah pengguna yang melihat data ini sedang mendukung proyek tersebut, | Boolean |  |
| Is\_starrable | Atribut boolean yang menunjukkan apakah pengguna yang melihat data ini dapat menjadikan proyek ini favorit | Boolean |  |
| Is\_starred | Atribut boolean yang menunjukkan apakah pengguna yang melihat data ini telah menjadikan proyek ini favorit. | Boolean |  |
| Launched\_at | Tanggal dan waktu proyek diluncurkan | Date |  |
| Location | Lokasi proyek | String | 255 |
| Name | Nama dari yang melakukan penggalangan dana | String | 255 |
| Photo | Url atau link menuju foto proyek | String | 255 |
| Pledged | Jumlah dana yang telah dikumpulkan proyek dalam mata uang aslinya | Float | 16 |
| Profile | Informasi detail mengenai pembuat proyek | String | 255 |
| Slug | Tautan singkat yang dapat digunakan untuk mengakses proyek | String | 255 |
| Source\_url | URL asal proyek jika berasal dari luar platform | String | 255 |
| Spotlight | Sebuah atribut boolean yang menunjukkan apakah proyek ditampilkan di bagian khusus atau kurasi platform Kickstarter. | Boolean |  |
| Staff\_pick | Sebuah atribut boolean yang menunjukkan apakah proyek dipilih oleh tim Kickstarter sebagai proyek favorit mereka. | Boolean |  |
| State | Status proyek | String | 255 |
| State\_changed\_at | Tanggal dan waktu terakhir status proyek berubah | Date |  |
| Static\_usd\_rate | Nilai tukar dolar Amerika Serikat (USD) yang statis terhadap mata uang proyek. | Float | 16 |
| Urls | URL terkait proyek | String | 255 |
| Usd\_exchange\_rate | Nilai tukar mata uang proyek ke USD pada saat data diambil. | Float | 16 |
| Usd\_pledged | Jumlah dana yang telah dihimpun proyek dikonversi ke USD. | Float | 16 |
| Usd\_type | Jenis mata uang USD yang digunakan. | String | 255 |

## Model Dimensional

### *High-Level Dimensional Model*





Gambar 2. 2 Bus Matrix

Dapat dilihat gambar menampilkan bux matrix yang digunakan untuk menganalisis tren dan pola dalam proyek *crowdfunding.* Tabel fakta yang digunakan adalah “Crowdfunding Fact”, yang mencatat data secara berkala sebagai snapshot periodik.

### *Dimensional Model Schema*

*Dimesional Model Schema* terdiri dari sebuah tabel fakta yang menjadi inti, dikelilingi oleh tabel-tabel dimensi yang berisi data referensi yang umumnya dapat dinormalisasi. Dengan skema ini, kueri dapat direspons dengan cepat dan sumber daya database dapat dimanfaatkan secara efisien. Dapat dilihat bahwa faktanya yaitu Crowdfunding Fact yang dikelilingi oleh tabel tabel dimensi yaitu dimensi lokasi, dimensi category, dimensi date, dimensi creator, dan dimensi project.



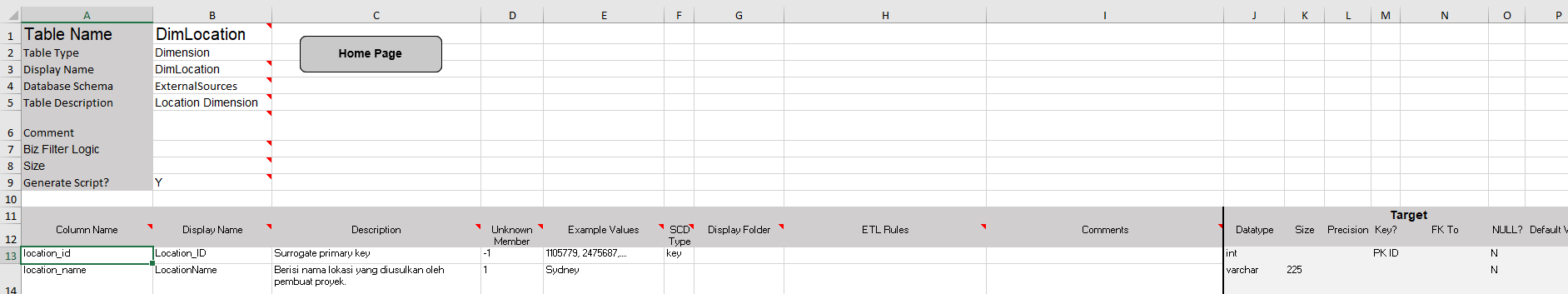
Gambar 2. 3 Dimensional Model Shema

### *Detailed Dimensional Model*

Kickstarter.xlsx adalah dataset yang digunakan untuk proyek crowdfunding dalam Detailed-Dimensional-Modeling-Workbook-KimballU.xlsx. Detailed dimensional modeling adalah teknik untuk merancang struktur data yang sangat rinci untuk analisis bisnis. Dalam konteks ini, tujuannya adalah merancang struktur data untuk menganalisis crowdfunding di platform Kickstarter. Dimensi utama dalam dataset ini mencakup Location, Category, Creator, Date, Projectm Backer, dan Goal dengan setiap dimensi memiliki beberapa atribut yang didefinisikan sebagai kolom dalam tabel.

* **Dim\_Location**

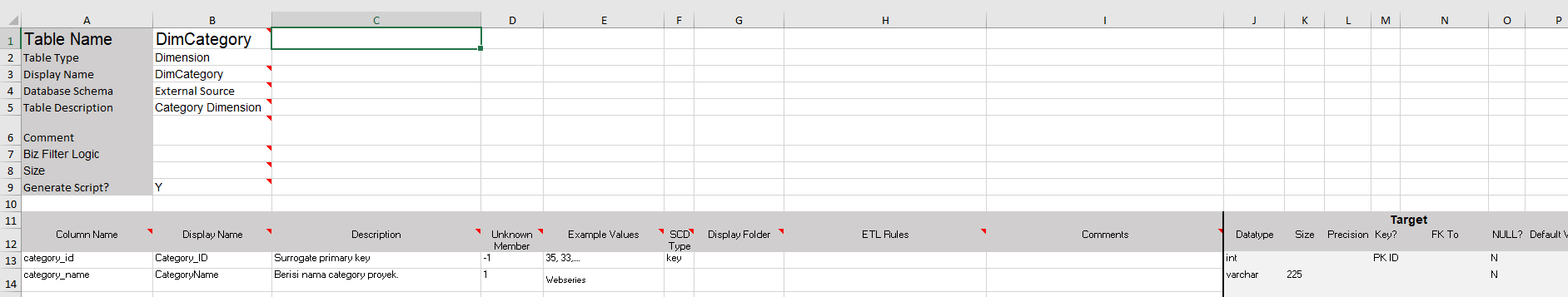
Dim\_*Location* berisi nama lokasi yang diusulkan oleh pembuat proyek. Dim\_*Location* ini berisi atribut *location\_id* dan *location\_name*.



Gambar 2. 4 Detailed DimLocation

* **Dim\_Category**

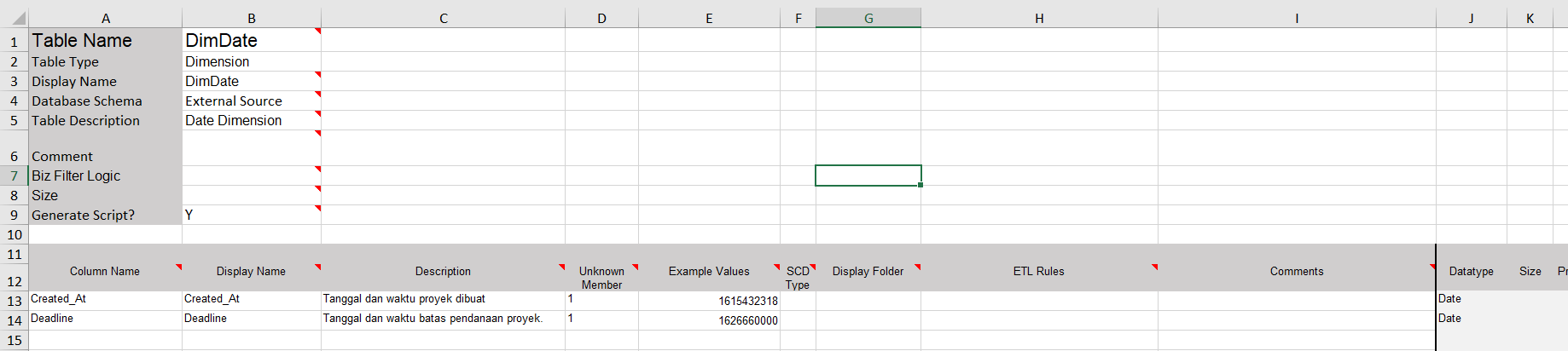
Dim\_*Category* berisi nama kategori proyek. Dim\_*Category* ini berisi atribut *category\_id* dan *category\_name*.



Gambar 2. 5 Detailed DimCategory

* **Dim\_Date**

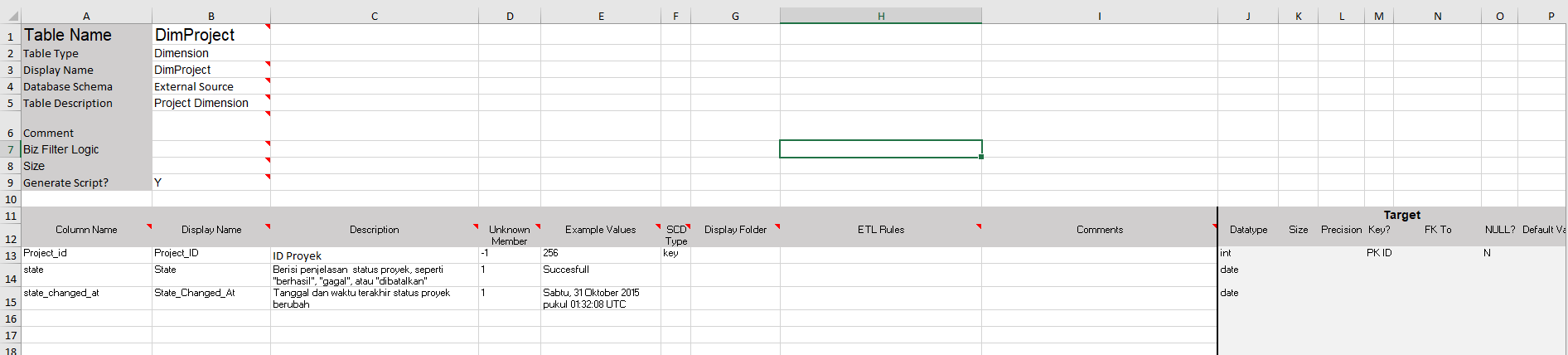
Dim\_*Date* berisi tanggal dan waktu yang diusulkan oleh pembuat proyek. Dim\_*Date* ini berisi atribut *created\_at* dan *deadline*.



Gambar 2. 6 Detailed DimDate

* **Dim\_Project**

Dim\_*Project* berisi penjelasan status proyek, yaitu "berhasil", "gagal", atau "dibatalkan". Dim\_*State* ini berisi atribut *project\_id, state, dan state\_changed\_at.*

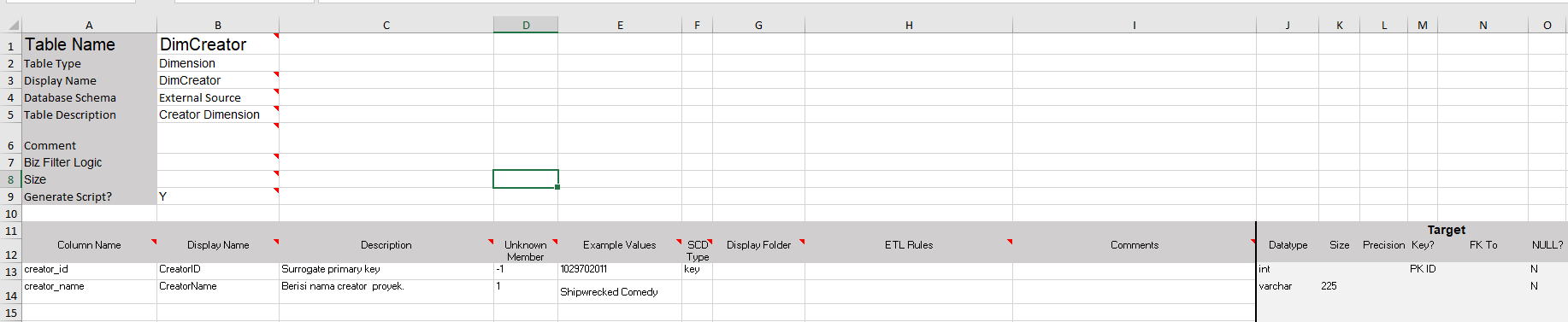


Gambar 2. 7 Detailed DimProject

**Gambar 2.4.3 Detailed Dimensional Model DimProject**

* **Dim\_Creator**

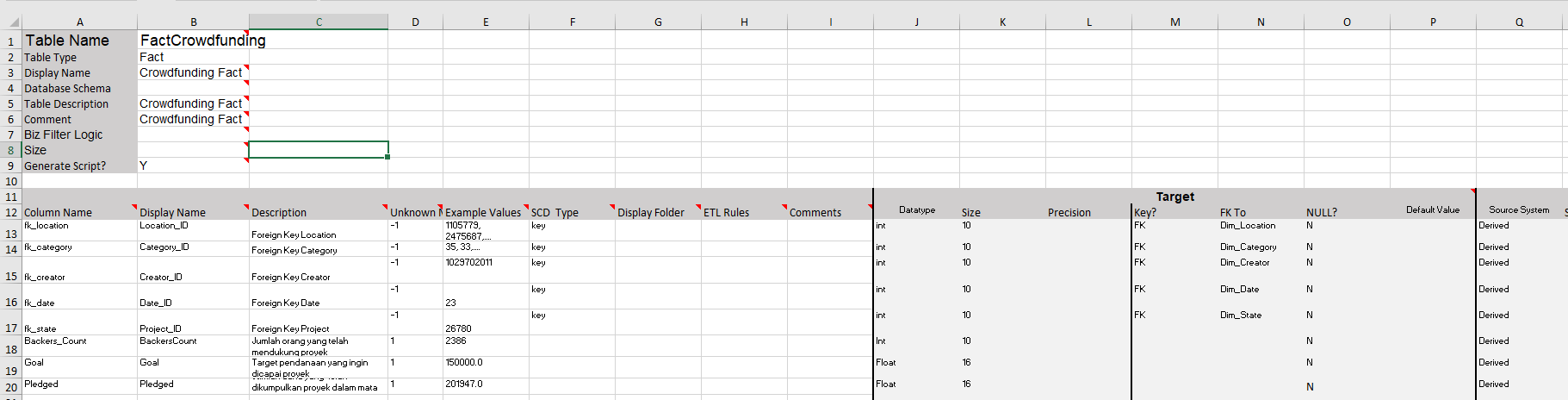
Dim\_*Creator* berisi nama dan id dari pengusul proyek. Dim\_*Creator* ini berisi atribut *creator\_id* dan *creator\_name*.



Gambar 2. 8 Detailed DimCreator

* **Fact\_Crowdfunding**

*Fact\_Crowdfunding* berisi fakta apa saja yang digunakan pada proyek. *Fact\_Crowdfunding* ini berisi atribut fk\_location, fk\_category, fk\_creator, fk\_date, fk\_project, backers\_count, goal, dan pledged.

****

Gambar 2. 9 Detailed FactCrowdfunding

## *Extract, Transform, and Load*

ETL (Extract, Transform, Load) adalah proses integrasi data dari sumber yang berbeda dan memindahkannya ke *database* yang terpusat, seperti *data warehousing*. ETL mencakup ekstraksi data dari sumber eksternal, transformasi data untuk memenuhi kebutuhan database yang terpusat, dan memuat data ke database tersebut. Tujuan dari ETL adalah untuk memastikan bahwa data yang digunakan oleh organisasi adalah akurat, konsisten, dan tersedia untuk analisis.

* 1. Extract

Pada tahap ini, data diambil dari sumber-sumber yang berbeda seperti file, database, dan aplikasi web. Data yang diambil melalui tahap ini dapat berupa data raw atau data yang sudah diproses sebelumnya. Dalam tahap ini, software ETL akan mengkonfigurasi koneksi dengan sumber data dan membuat skrip untuk mengambil data.

* 1. Transform

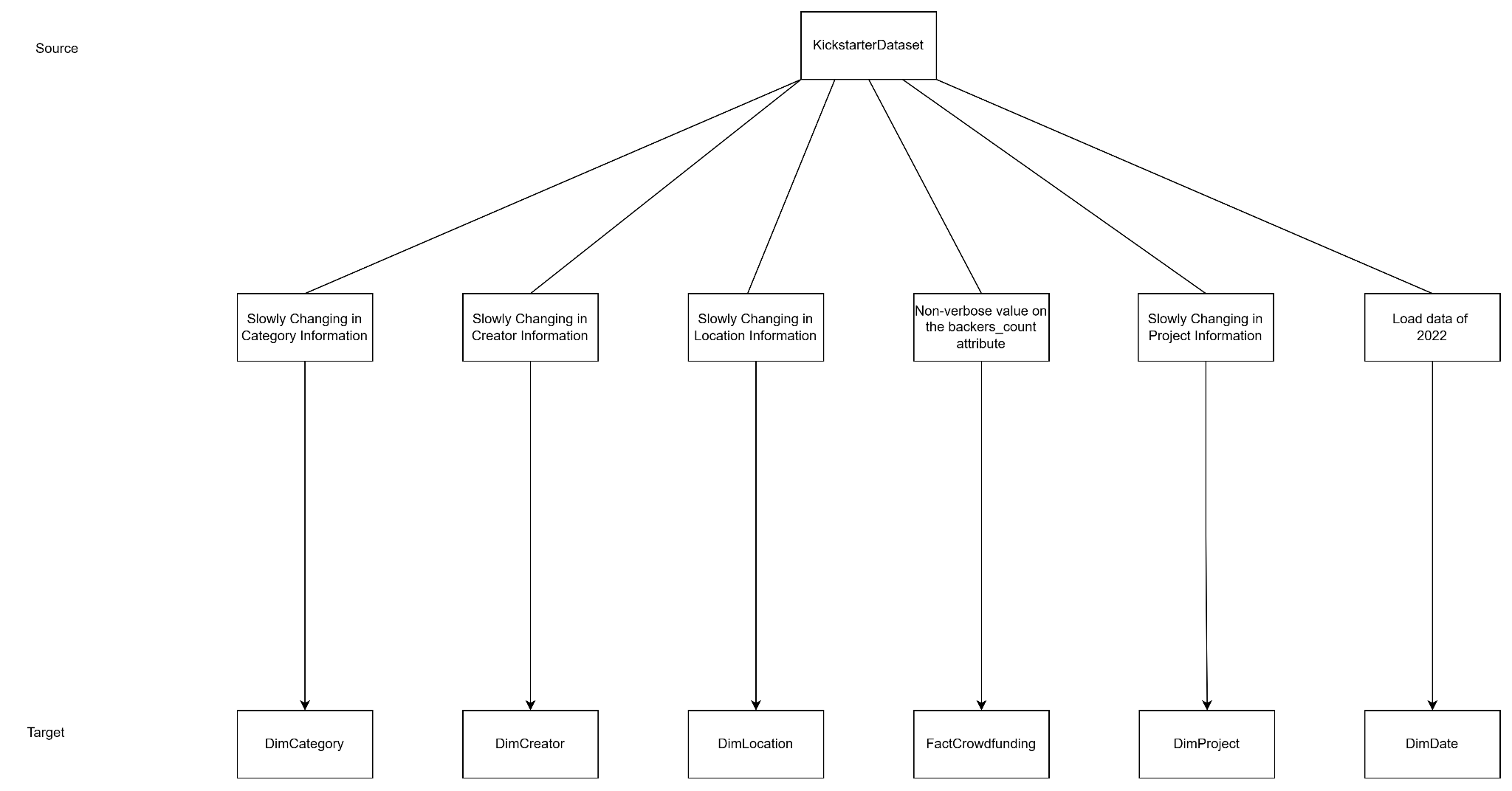
Pada tahap ini, data yang diambil dalam tahap ekstraksi akan diolah dan diformat sesuai dengan kebutuhan. Tahap ini meliputi beberapa proses seperti pembersihan data, normalisasi data atau normalisasi database, dan penambahan kolom. Dalam tahap ini, software ETL akan mengkonfigurasi alur data yang akan diolah dan membuat skrip untuk memproses data.

* 1. Load

Pada tahap ini, data yang sudah diolah dalam tahap transformasi akan dimuat ke dalam database terpusat. Dalam tahap ini, software ETL akan mengkonfigurasi koneksi dengan database terpusat dan membuat skrip untuk memuat data.

### *High-Level Source to Target Map*

High-Level Source to Target Mapping adalah serangkaian instruksi transformasi data yang menentukan bagaimana mengonversi struktur dan konten data dalam sistem sumber ke struktur dan konten yang dibutuhkan dalam sistem target. Ketika data dipindahkan dari satu sistem ke sistem lain, hampir tidak mungkin memiliki situasi di mana sistem sumber dan sistem target memiliki skema yang sama. Oleh karena itu, diperlukan mekanisme yang memungkinkan pengguna untuk memetakan atribut mereka dalam sistem sumber ke atribut dalam sistem target. Proses ini menjadi lebih rumit ketika ada data yang harus dipindahkan ke gudang data pusat dari berbagai sumber data, masing-masing memiliki skema yang berbeda. Sehingga dengan dilakukan analisis terhadap data kickstarter tersebut dapat ditentukan dimensi-dimensi yang relevan dalam terhadap pertanyaan analisis yang telah dilakukan, dimensi -dimensi tersebut yaitu DimCategory, DimCreator, DimLocation, DimState, dan DimDate dan satu tabel fakta (FactCrowdfunding).



Gambar 2. 10 High-Level to Target Map

Gambar ini menunjukkan Pemetaan Tingkat Tinggi Sumber ke Target untuk proyek migrasi data. Sumbernya adalah KickstarterDataset, yang merupakan kumpulan data proyek Kickstarter. Targetnya adalah data warehouse, yang merupakan tempat penyimpanan untuk menyimpan dan menganalisis data. Peta menunjukkan bahwa data dari KickstarterDataset akan dimigrasikan ke tiga dimensi (DimCategory, DimCreator, dan DimLocation) dan satu tabel fakta (FactCrowdfunding). Data juga akan dimigrasikan ke dua tabel dimensi (DimState dan DimDate).

### Detailed ETL Flow for Each Source to Target

* **DimCategory**

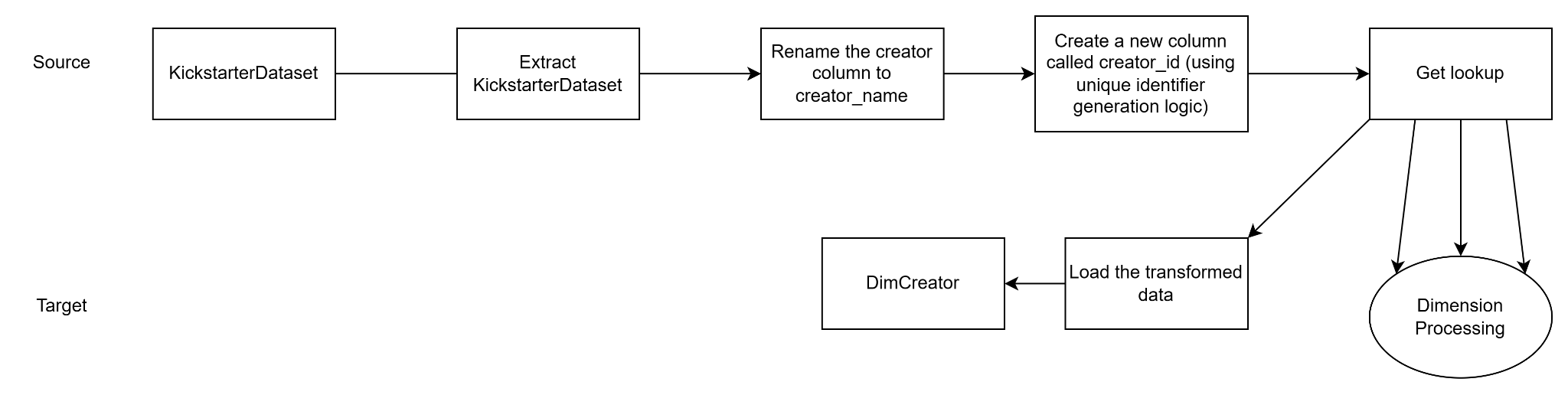
Alur ETL DimCategory dimulai dengan ekstraksi dataset Kickstarter, di mana data mentah diambil untuk disiapkan untuk proses selanjutnya. Selanjutnya, ekstraksi data kategori dilakukan dengan menyaring dan mengambil informasi terkait kategori dari dataset, serta pembuatan kolom baru category\_id untuk memberikan identifier unik pada setiap kategori. Setelah itu, kategori yang tidak relevan disaring untuk menjaga kualitas dan relevansi data. Selanjutnya, data kategori diproses dan dimensi diolah untuk memastikan konsistensi dan kelengkapan data sebelum dimuat ke dalam tabel target DimCategory untuk analisis dan pelaporan lebih lanjut.

****

Gambar 2. 11 Detailed ETL Flow DimCategory

* **DimCreator**

Alur DimCreator dimulai dengan ekstraksi data dari KickstarterDataset. Kolom yang berisi informasi pencipta proyek diubah namanya menjadi creator\_name. Kemudian, sebuah kolom baru bernama creator\_id dibuat dan diisi dengan identifier unik untuk setiap pencipta. Data dimensi diproses untuk memastikan integritas dan kesiapan sebelum digunakan dalam analisis atau pelaporan lebih lanjut. Terakhir, data yang telah ditransformasi ini dimuat ke dalam tabel DimCreator.

****

Gambar 2. 12 Detailed ETL Flow DimCreator

* **DimLocation**

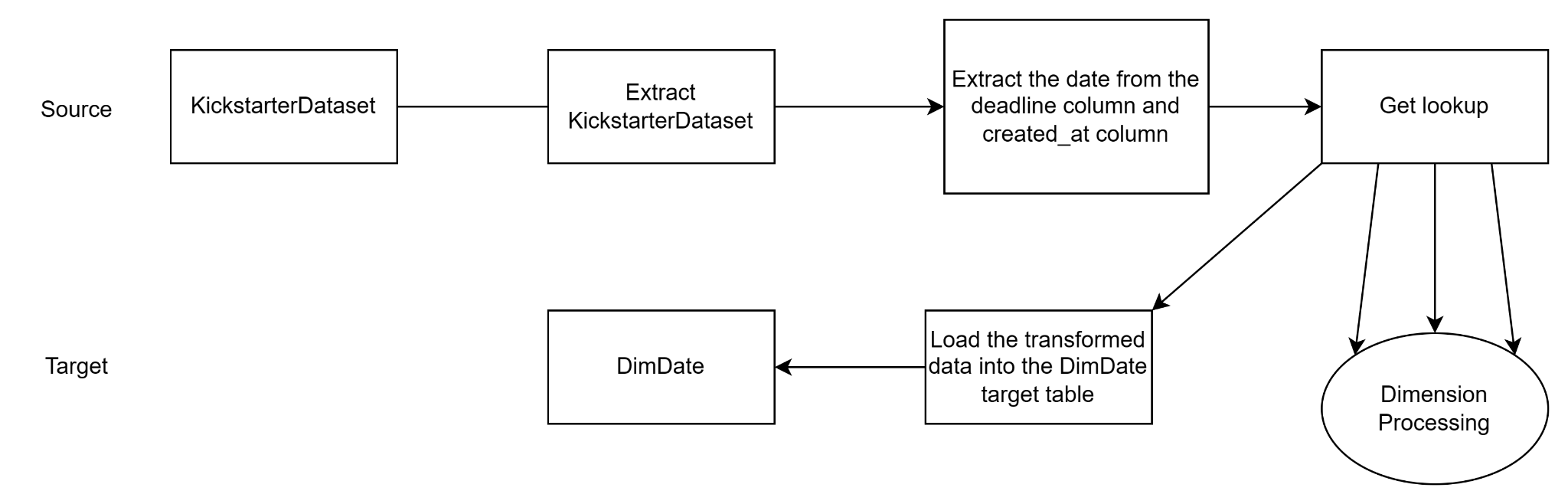
Alur DimLocation dalam proses ETL dimulai dengan ekstraksi data dari KickstarterDataset dan identifikasi serta penyelesaian masalah data. Nama negara yang disingkat kemudian ditangani dan dipetakan ke nama lengkap menggunakan tabel lookup. Data lokasi yang telah difilter dan divalidasi disimpan dalam tabel target DimLocation. Terakhir, dilakukan pemrosesan dimensi lebih lanjut jika diperlukan untuk analisis dan pelaporan.

****

Gambar 2. 13 Detailed ETL Flow DimLocation

* **DimDate**

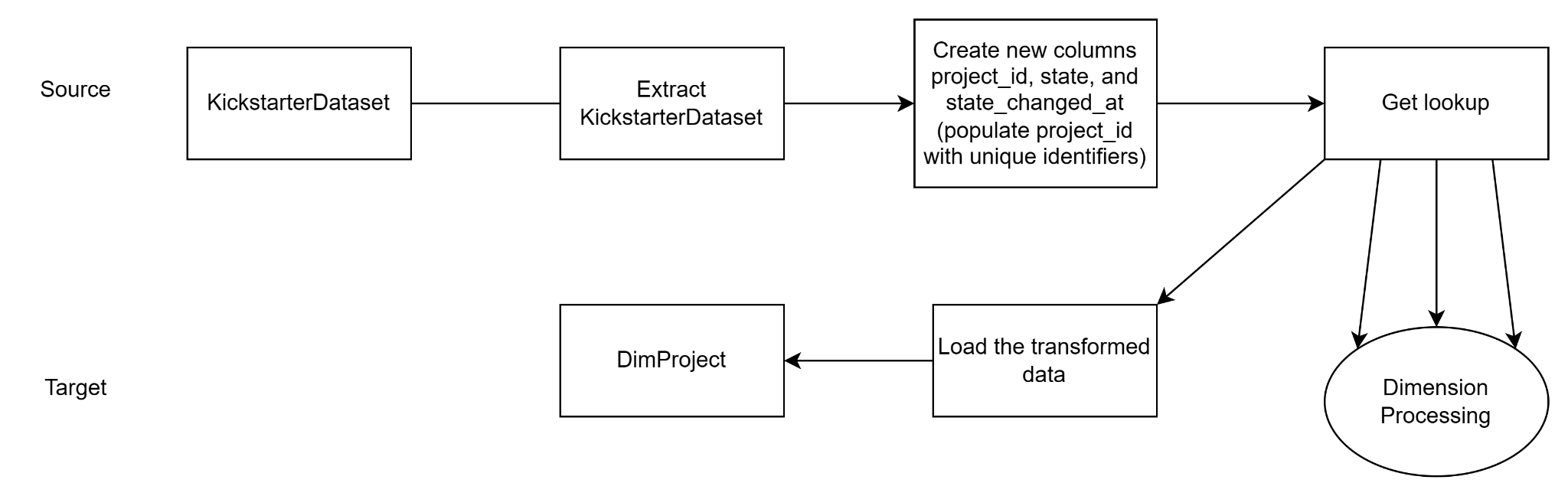
Proses ETL untuk tabel DimDate dimulai dengan ekstraksi data tanggal dari kolom created\_at dan deadline dalam dataset Kickstarter. Setelah data tanggal ini diekstrak, dilakukan lookup untuk memastikan konsistensi dan integritas data. Akhirnya, data yang telah divalidasi ini dimuat ke dalam tabel DimDate, diikuti oleh pemrosesan dimensi untuk menyiapkan data bagi analisis lebih lanjut.

****

Gambar 2. 14 Detailed ETL Flow DimDate

* **DimProject**

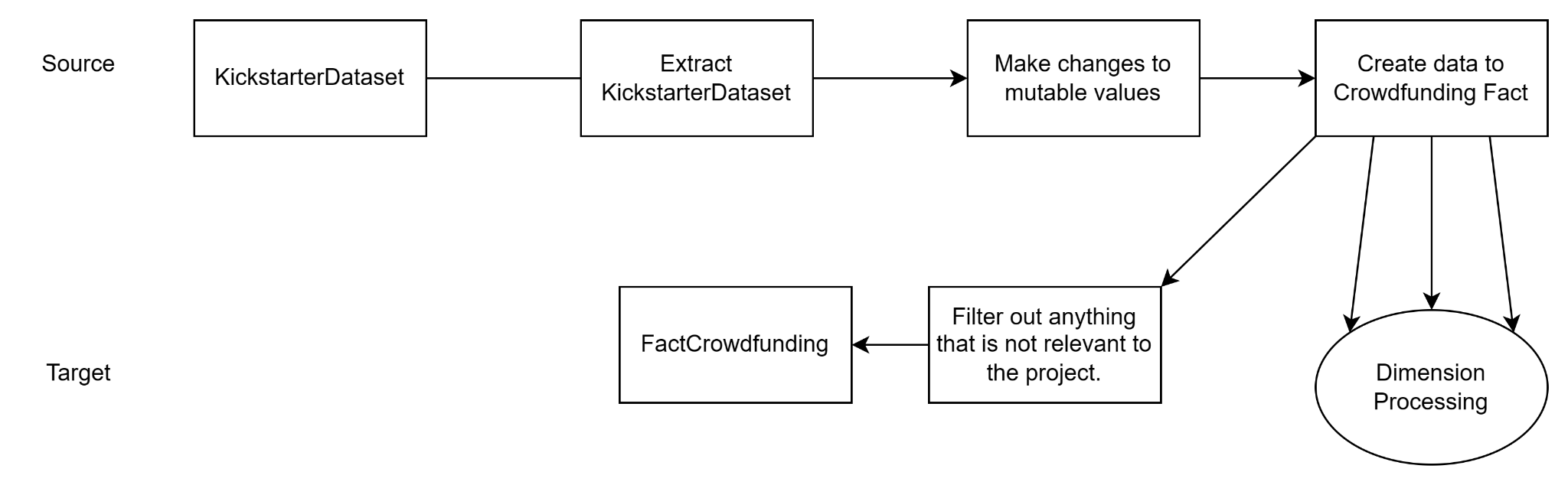
Proses ETL untuk DimProject dimulai dengan ekstraksi data dari dataset Kickstarter dan pembuatan kolom baru seperti project\_id, state, dan state\_changed\_at. Setelah itu, dilakukan lookup untuk memverifikasi data yang telah dibuat. Akhirnya, data yang telah diverifikasi dimuat ke dalam tabel DimProject dan diikuti oleh pemrosesan dimensi untuk validasi dan integritas lebih lanjut.

****

Gambar 2. 15 Detailed ETL Flow DimProject

* **FactCrowdfunding**

Proses ETL untuk FactCrowdfunding dimulai dengan ekstraksi data dari dataset Kickstarter dan melakukan perubahan pada nilai yang diperlukan. Data yang tidak relevan kemudian disaring, dan catatan fakta crowdfunding yang baru dibuat. Akhirnya, data yang telah diproses dimuat ke dalam tabel FactCrowdfunding dan dilakukan pemrosesan dimensi untuk validasi lebih lanjut.

****

Gambar 2. 16 Detailed ETL Flow FactCrowdfunding

## *Business Intelligence Application*

### *Business Intelligence Application Specification*

Terdapat pada lampiran A:

1. Tren Pendanaan Crowdfunding

Crowdfunding menunjukkan pertumbuhan pesat dengan peningkatan jumlah proyek dan dana yang terkumpul. Globalisasi membuat crowdfunding menjadi fenomena internasional, dengan proyek dari berbagai negara. Inovasi dalam metode pendanaan terus berkembang, menarik lebih banyak pendukung.

1. Pola Kategori Proyek Crowdfunding

Setiap kategori memiliki platform yang lebih sesuai, seperti Kickstarter untuk proyek kreatif. Pola ini membantu mengarahkan pendukung ke proyek yang sesuai dengan minat mereka.

1. Faktor-Faktor yang Mempengaruhi Keberhasilan Proyek Crowdfunding

Keberhasilan dipengaruhi oleh kualitas presentasi proyek, promosi yang efektif melalui media sosial dan jaringan pribadi, serta target pendanaan yang realistis. Imbalan yang menarik dan transparansi penggunaan dana sangat penting untuk membangun kepercayaan. Dukungan awal yang kuat juga memberikan momentum penting untuk kampanye.

### *Detailed Business Intelligence Application Specification*

Terdapat pada lampiran B:

1. Pola Kategori Proyek Crowdfunding

Tanggal: Tanggal dan waktu proyek dibuat

Jumlah Proyek: Jumlah proyek yang akan dilakukan

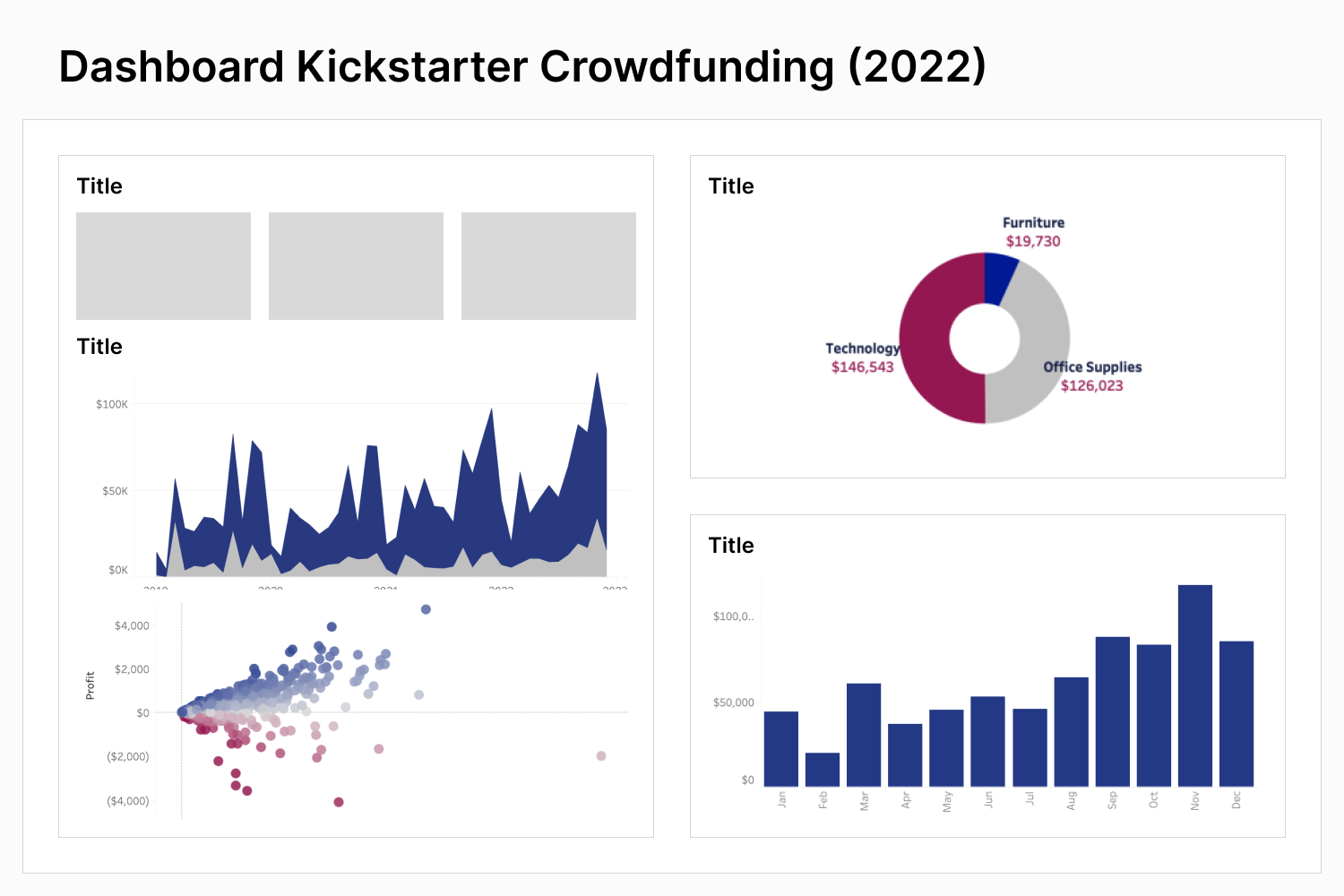
Jumlah Dana yang Dikumpulkan: Jumlah uang yang telah dikumpulkan proyek sejauh ini

1. Tren Pendanaan Crowdfunding

Kategori Proyek: Nama kategori proyek

Jumlah Proyek: Jumlah proyek yang akan dilakukan

### *Business Intelligence Application Mockup*



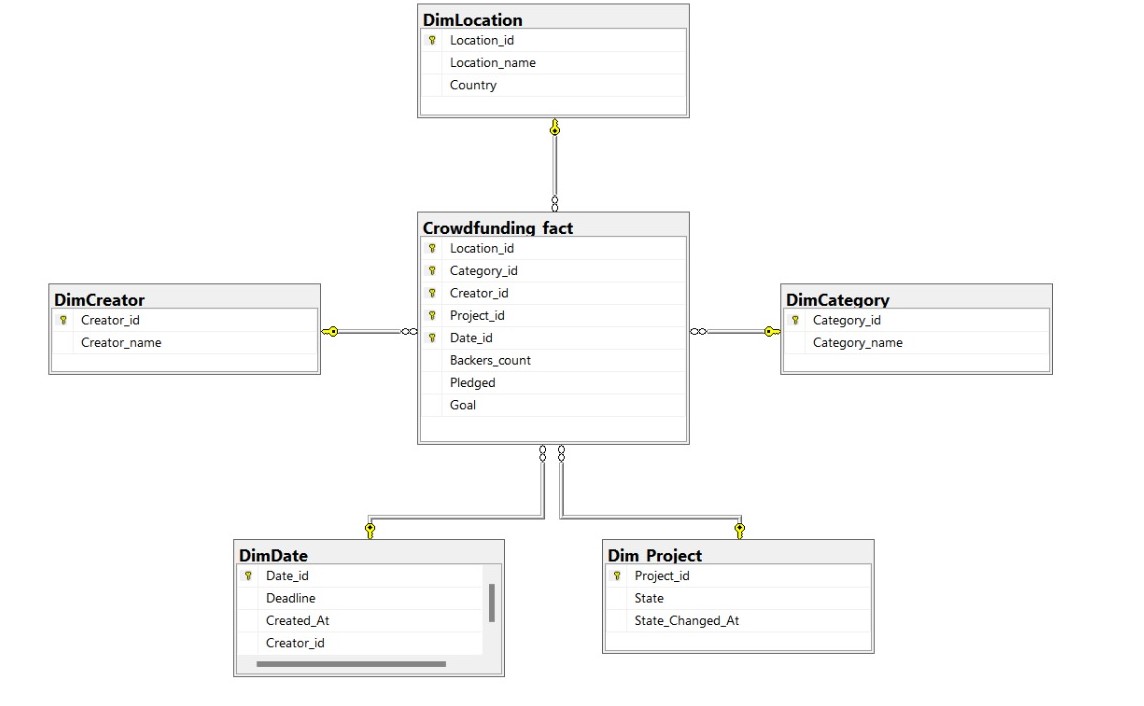
Gambar 2. 17 Business Intelligence Application Mockup

# Implementasi

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

## ROLAP *Schema*

Dalam skema ini, ada sebuah tabel fakta, bernama Crowdfunding fact, dikelilingi oleh beberapa tabel dimensi: DimLocation, DimCreator, DimDate, DimCategory dan DimProject. Setiap tabel dimensi berisi atribut-atribut yang memberikan konteks dan detail tambahan untuk tabel fakta.



Gambar 3. 1 ROLAP Schema

Tabel fakta Crowdfunding menyimpan data transaksi inti yang terkait dengan proyek crowdfunding. Ini mencakup atribut-atribut berikut:

* Location\_id: Unique identifier for the crowdfunding project.
* Category\_id: Pengidentifikasi unik untuk kategori proyek.
* Project\_id: Pengidentifikasi unik untuk proyek urun dana.
* Date\_id: Pengidentifikasi unik untuk tanggal yang terkait dengan transaksi.
* Backers\_count: Jumlah total pendukung yang telah berkontribusi pada proyek.
* Pledged: Jumlah total uang yang dijaminkan ke proyek.
* Goal: Target jumlah uang yang ingin dikumpulkan oleh proyek.

Tabel-tabel ini terhubung ke tabel fakta melalui hubungan kunci asing.

1. DimLocation:

Location\_id: Pengenal unik untuk lokasi.

Location\_name: Nama lokasi.

Country: Negara tempat lokasi berada

1. DimCreator:

Creator\_id: Pengenal unik untuk pembuat proyek.

Creator\_name: Nama pembuat proyek.

1. DimDate:

Date\_id: Pengenal unik untuk tanggal.

Date: Tanggal yang sebenarnya.

1. DimProject:

Project\_id: Unique identifier for the crowdfunding project.

State: Status proyek saat ini (misalnya, active, funded, failed).

State\_changed\_at: Tanggal ketika status proyek terakhir kali diubah.

1. DimCategory:

Category\_id: Pengidentifikasi unik untuk kategori proyek.

Category\_name: Pengidentifikasi nama untuk kategori proyek.

Gambar tersebut menunjukkan bahwa tabel DimLocation, DimCreator, dan DimProject memiliki hubungan satu-ke-banyak dengan tabel fakta Crowdfunding. Ini berarti bahwa setiap record di tabel fakta Crowdfunding dapat dikaitkan dengan satu record di setiap tabel dimensi. Tabel DimDate memiliki hubungan banyak-ke-satu dengan tabel fakta Crowdfunding. Ini berarti bahwa setiap record di tabel fakta Crowdfunding dapat dikaitkan dengan beberapa record di tabel DimDate.

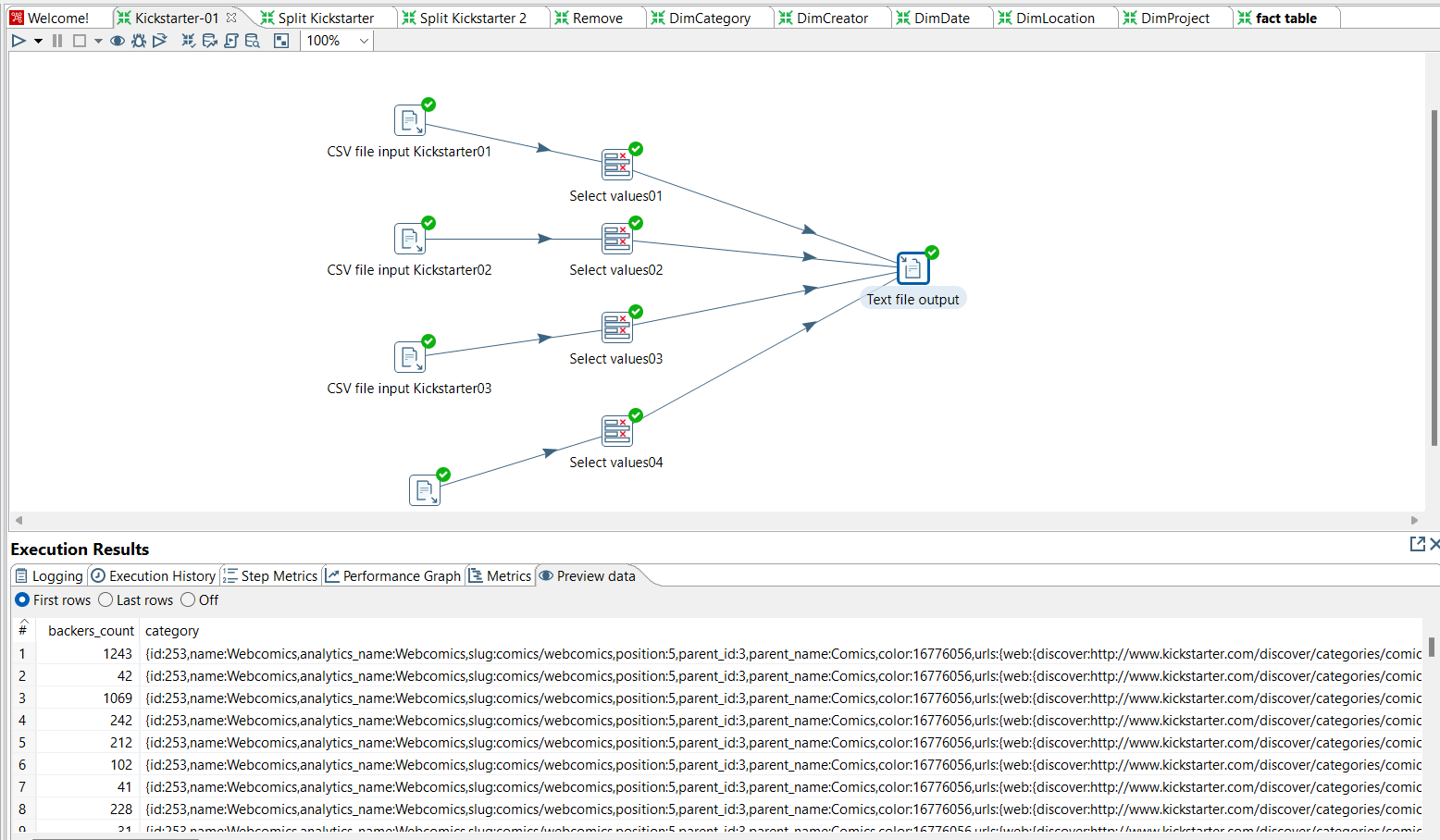
Hubungan kunci asing antara tabel-tabel tersebut memastikan bahwa data konsisten dan catatan-catatan dapat dihubungkan dengan benar.

## 

## ETL

* Merge Data

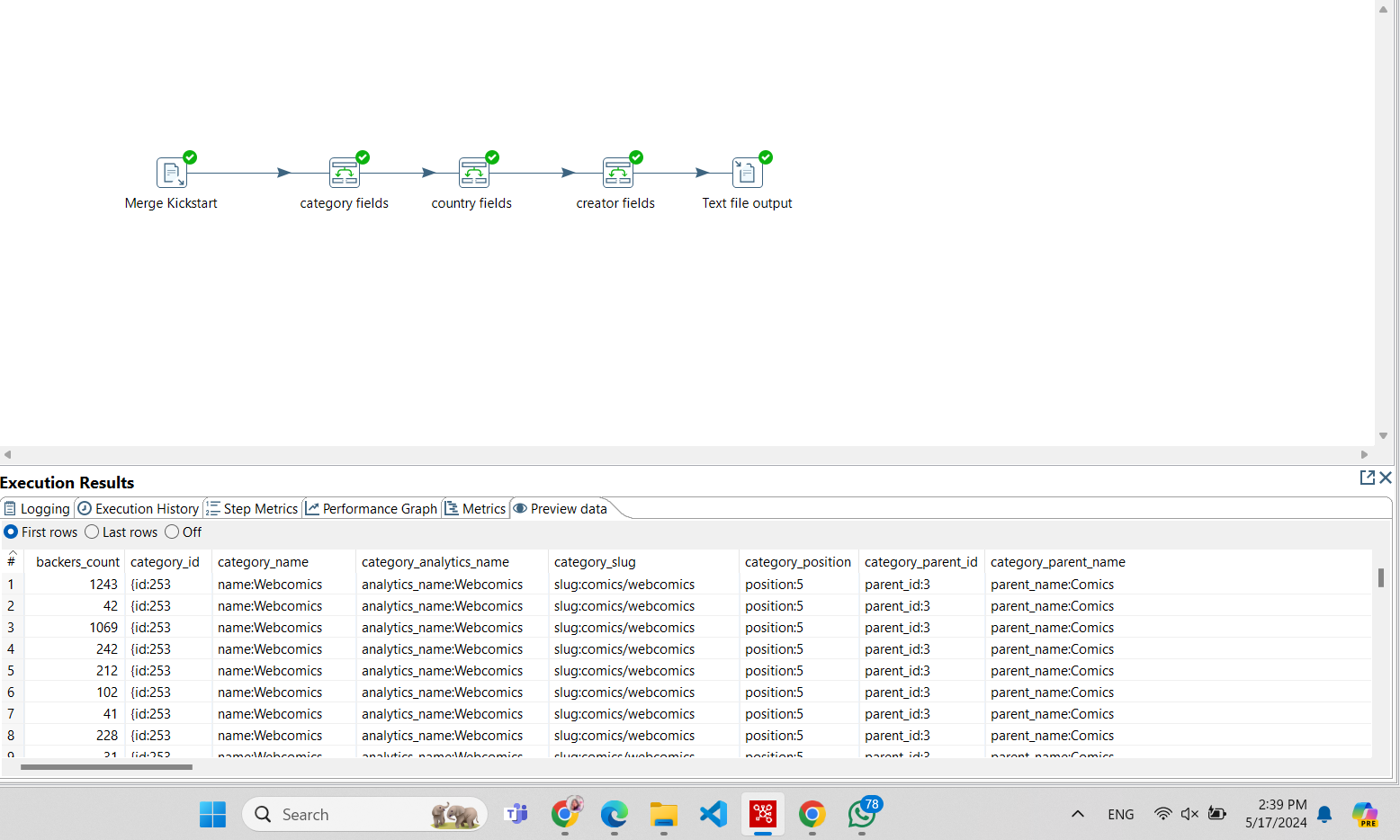
Berikut adalah Cuplikan ETL untuk mengintegrasi dua atau lebih data menjadi satu data yang dapat digunakan pada pengerjaan proyek ini. Gambar 3. 2 menampilkan ETL Merge Data



Gambar 3. 2 Merge Data

* Split Data Pertama

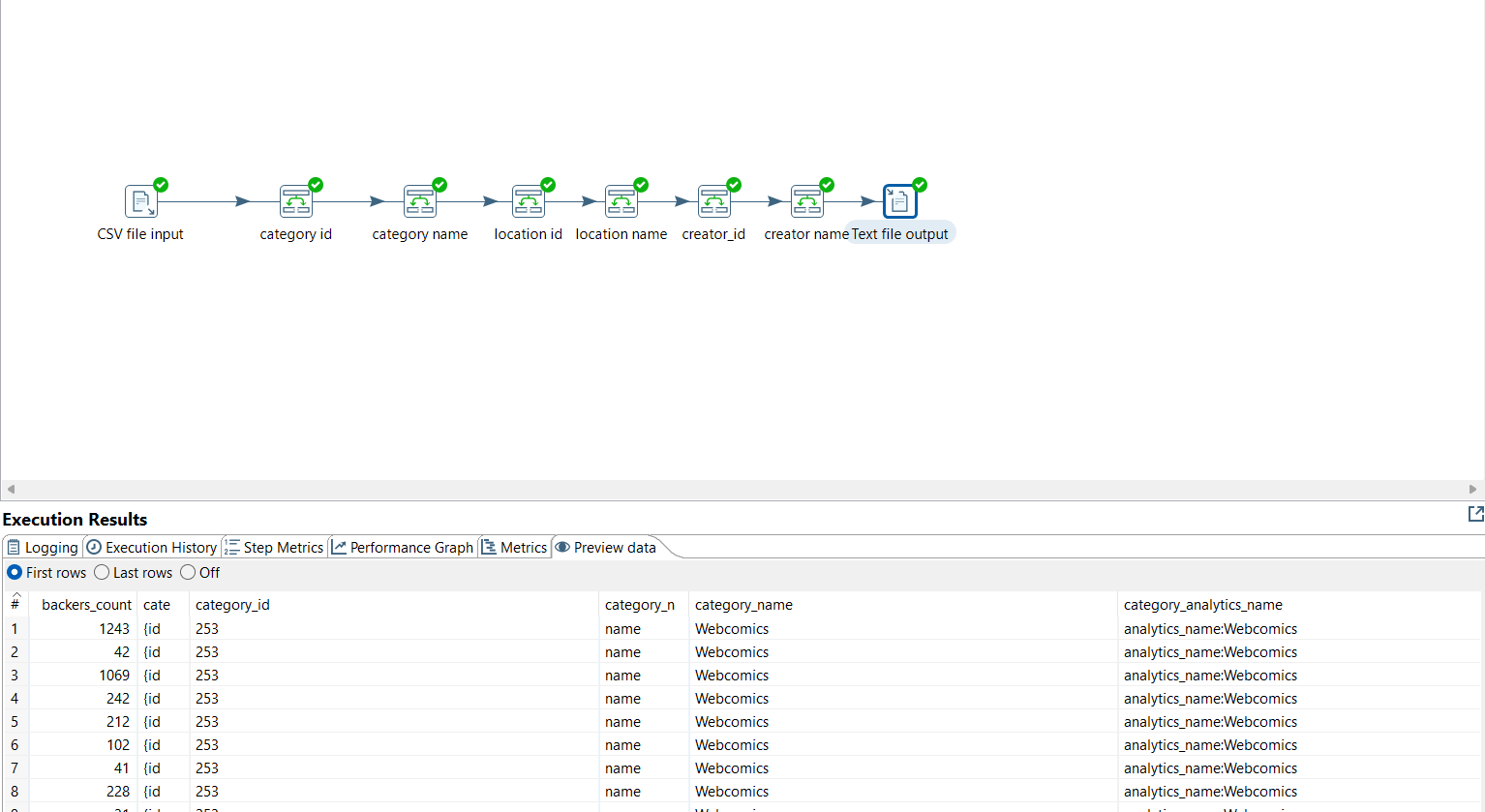
Pada beberapa kolom tertentu seperti *category, country,* dan *creator* memiliki informasi lebih dari satu nilai. Oleh karena itu, kelompok kami melakukan split data untuk memisahkan antar data tersebut agar dapat digunakan. Gambar 3. 3 menampilkan ETL untuk Split Data pertama dan juga *output* dari ETL yang telah dibuat.



Gambar 3. 3 Split Data 1

* Split Data 2

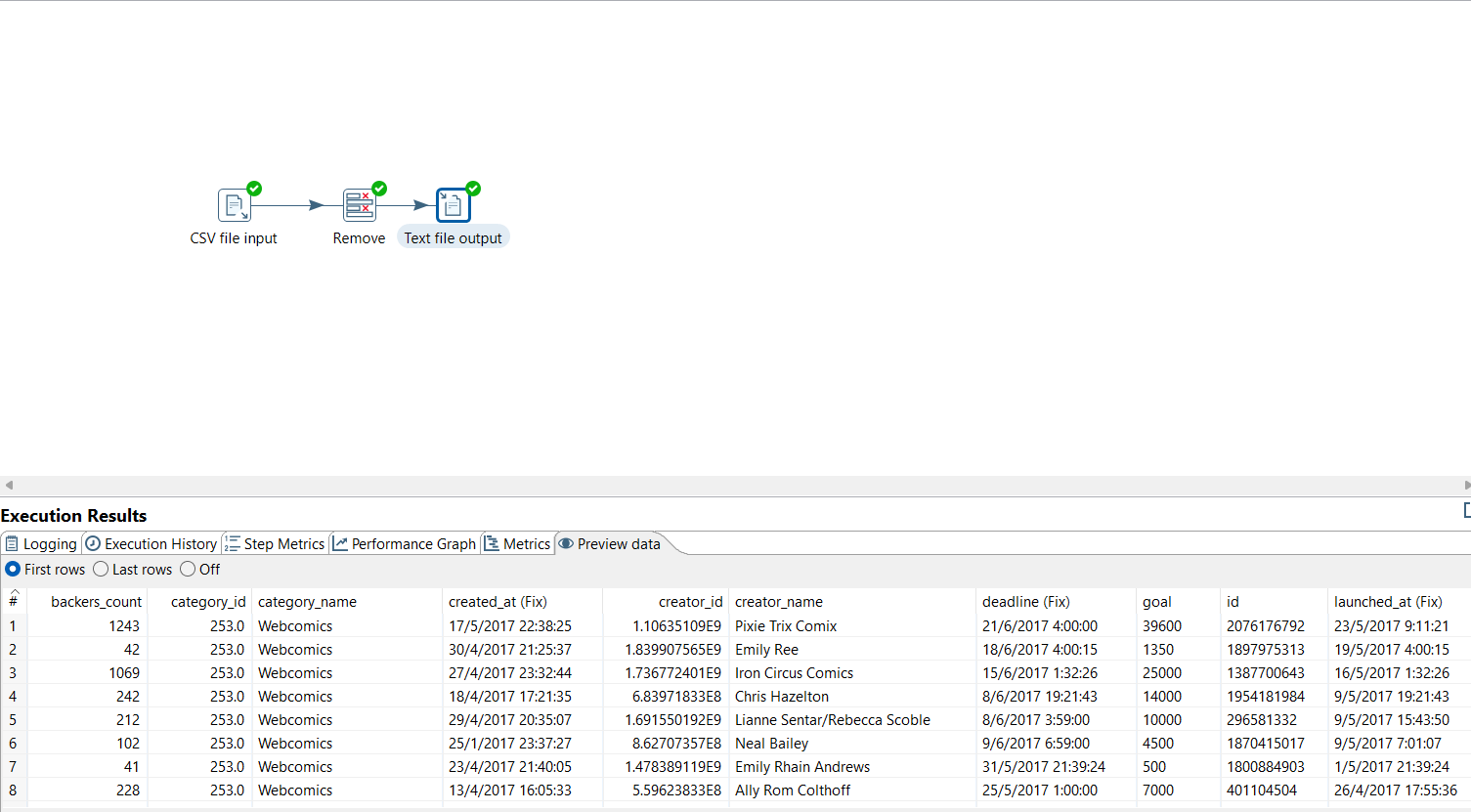
Pada split data kedua kami melakukan pemisahaan dengan nilai yang akan digunakan nantinya, berikut pada gambar 3. 4 adalah tampilan split data kedua yang kami lakukan.



Gambar 3. 4 Split Data 2

* Remove

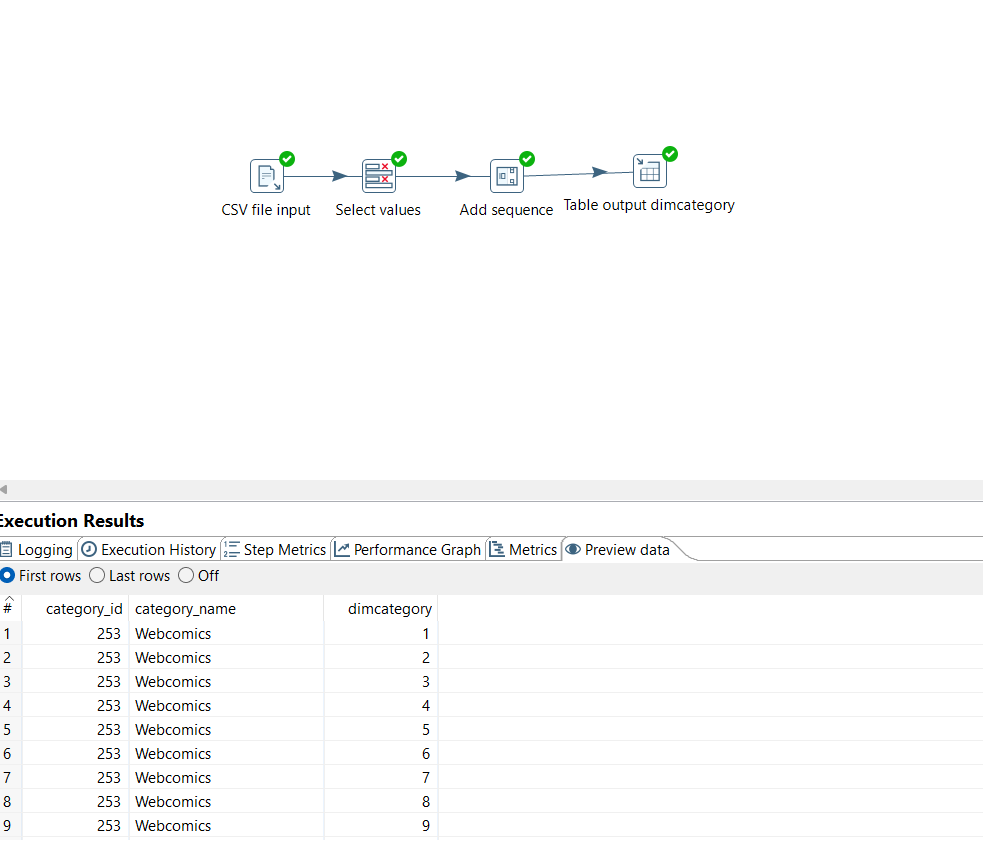
Pada gambar 3. 5 menampilkan ETL untuk menghapus kolom yang nilainya tidak kami gunakan.



Gambar 3. 5 Remove

* DimCategory

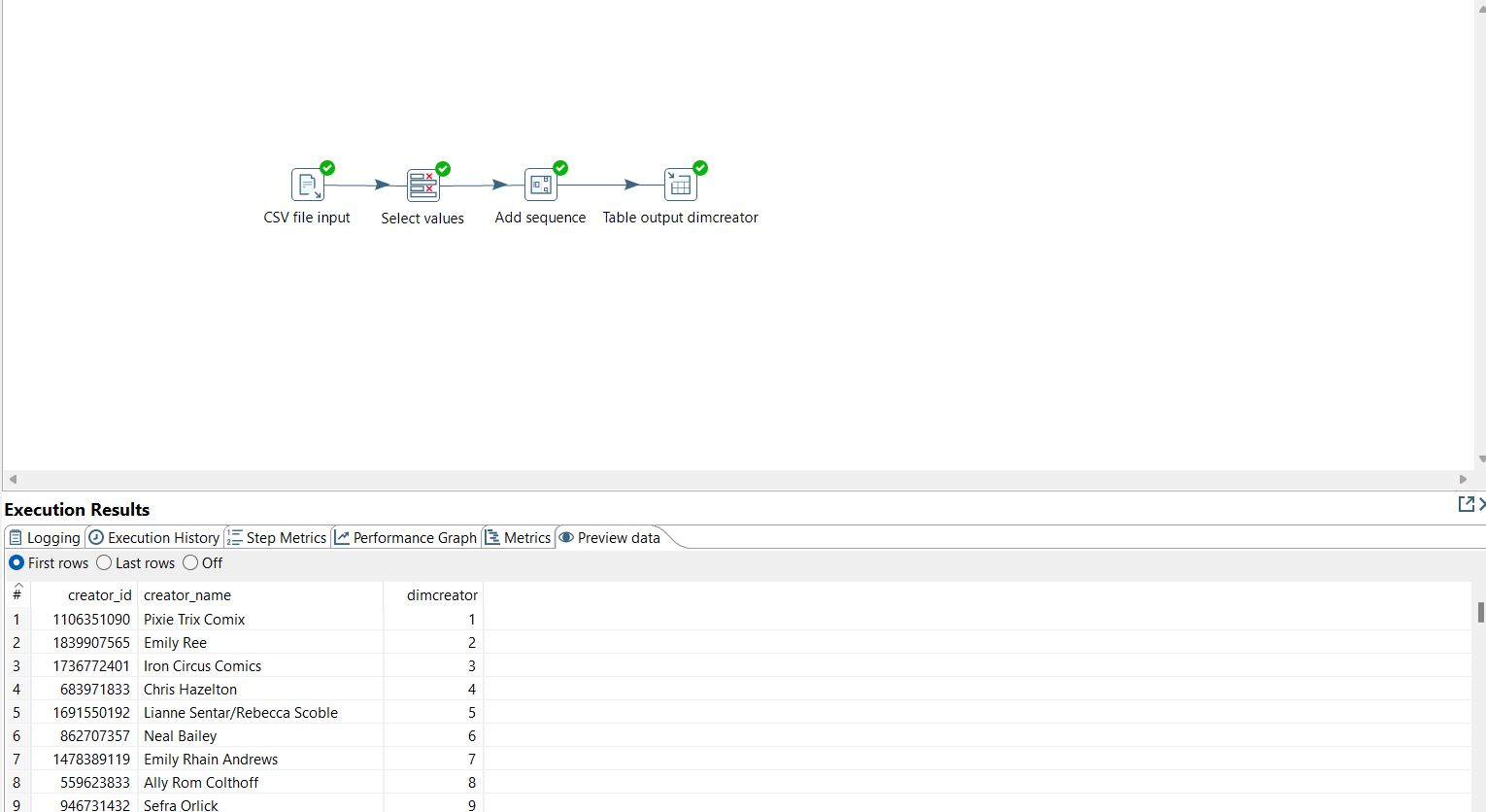
Pada gambar 3.6 ditampilkan proses ETL untuk DimCategory, di mana kami hanya mengambil atribut yang ada pada DimCategory.



Gambar 3. 6 DimCategory

* DimCreator

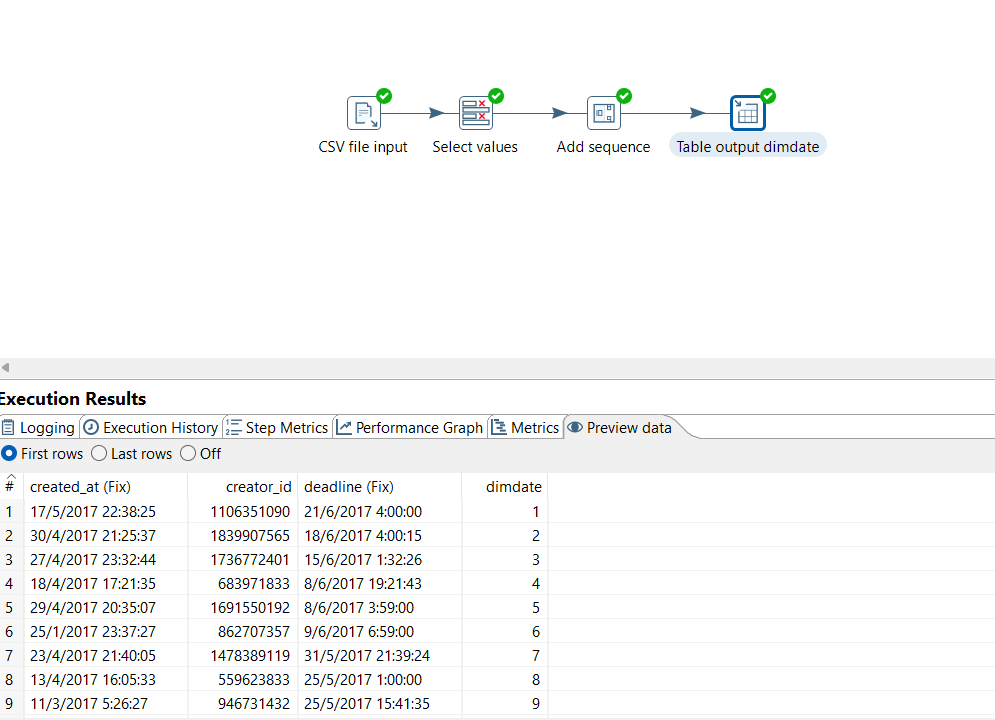
Pada gambar 3.7 ditampilkan proses ETL untuk DimCreator, di mana kami hanya mengambil atribut yang ada pada DimCreator.



Gambar 3. 7 DimCreator

* DimDate

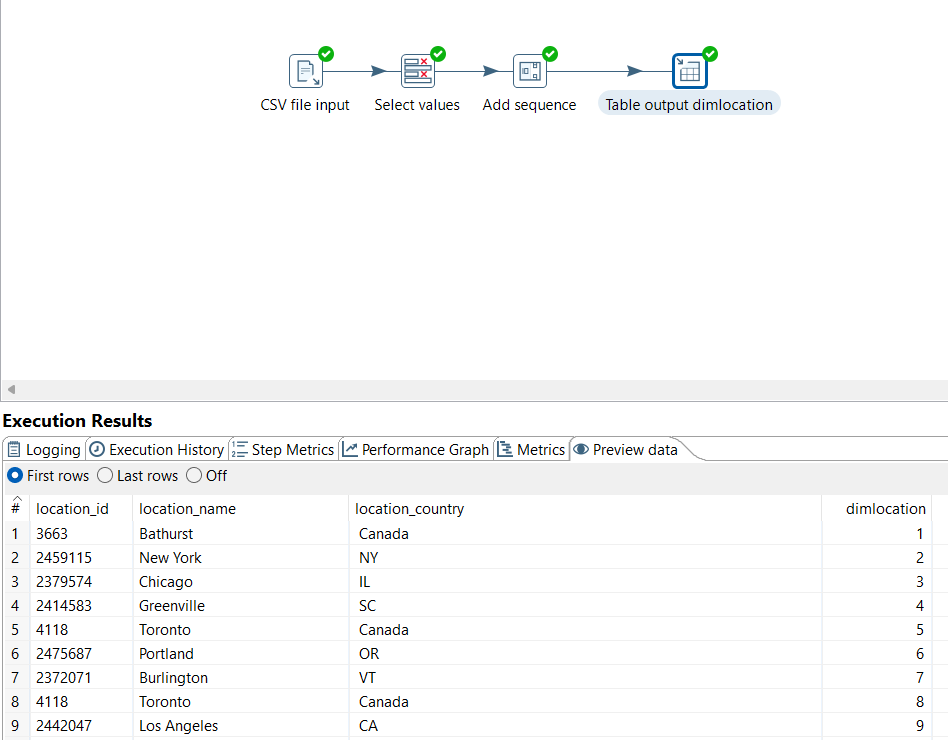
Pada gambar 3.8 ditampilkan proses ETL untuk DimDate, di mana kami hanya mengambil atribut yang ada pada DimDate.



Gambar 3. 8 DimDate

* DimLocation

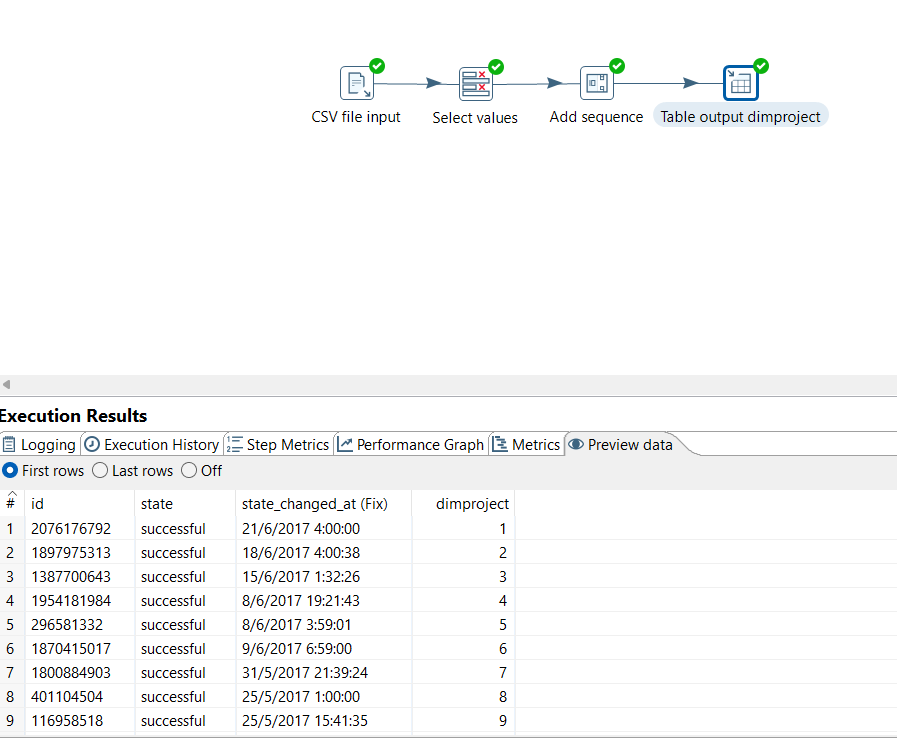
Pada gambar 3.9 ditampilkan proses ETL untuk DimLocation, di mana kami hanya mengambil atribut yang ada pada DimLocation.



Gambar 3. 9 DimLocation

* DimProject

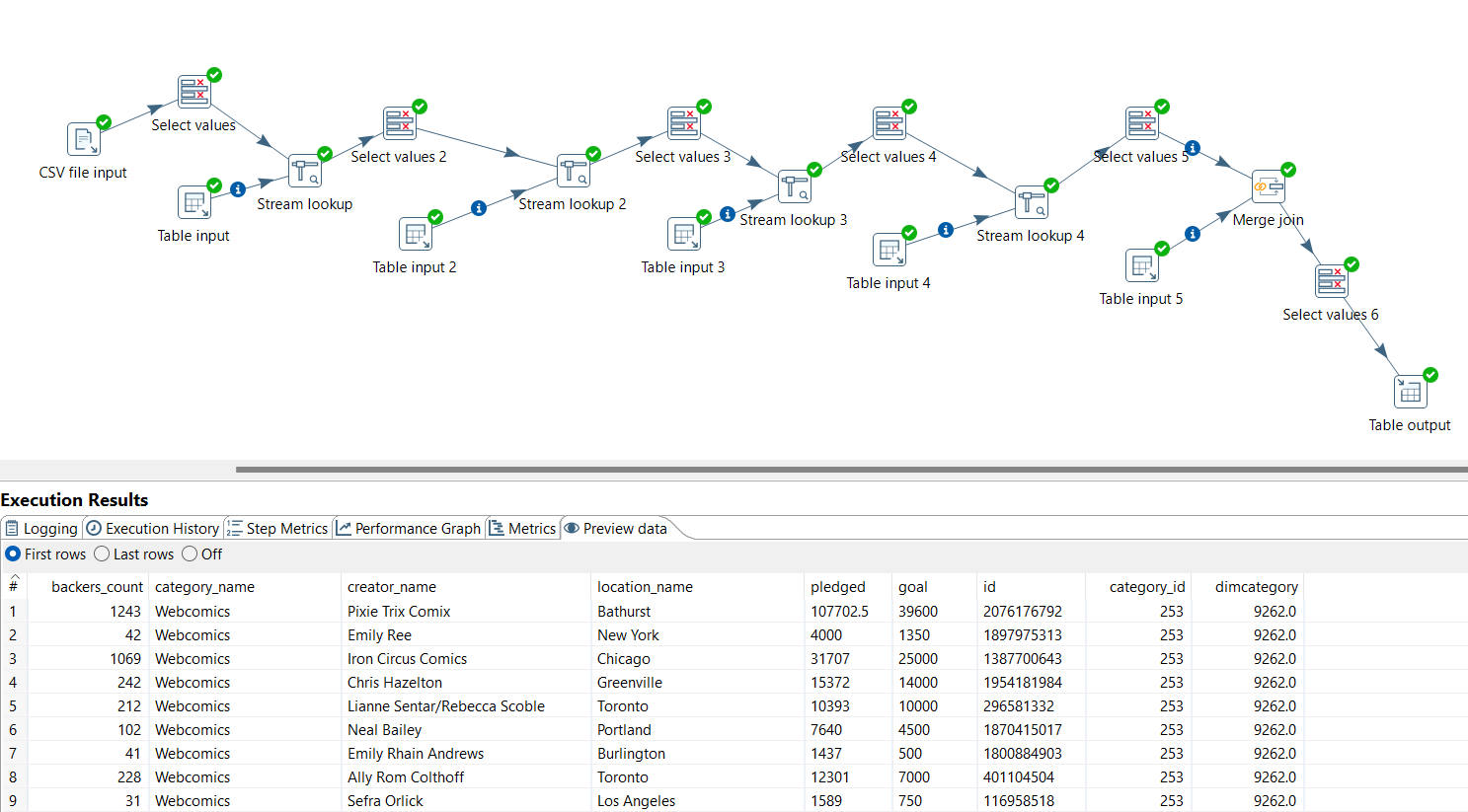
Pada gambar 3.10 ditampilkan proses ETL untuk DimProject, di mana kami hanya mengambil atribut yang ada pada DimProject.



Gambar 3. 10 DimProject

* FactTable

Pada gambar 3.11 menampilkan ETL untuk fact table.



Gambar 3. 11 FactTable

Data yang digunakan untuk mengisi fact table adalah, backers\_count, category\_name, creator\_name, location\_name, pledge, goal, project\_id, category\_id. Kolom-kolom ini dapat berupa metrik dan dimensi yang digunakan dalam tabel fakta untuk data proyek Kickstarter.

1. Metrics:

backers\_count (number of backers)

pledged (total amount pledged)

goal (funding goal)

1. Dimensions:

category\_name (category of the project)

creator\_name (name of the project creator)

location\_name (location of the project creator)

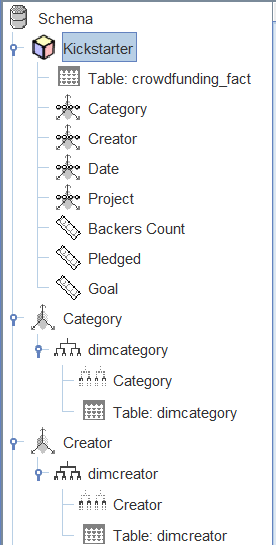
id (project ID)

category\_id (ID of the project category)

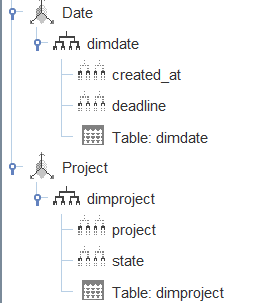
Metrik dan dimensi aktual yang digunakan dalam tabel fakta akan bergantung pada kebutuhan bisnis tertentu dan bagaimana data akan dianalisis.

## MOLAP Schema

Pada gambar 3. 12 dan 3. 13 adalah cuplikan cube shema kickstarter. Pada schema tersebut, terdapat dimensi Category, Creator, Date dan Project. Selain itu, measure yang digunakan adalah backers count, pledged dan juga goal.

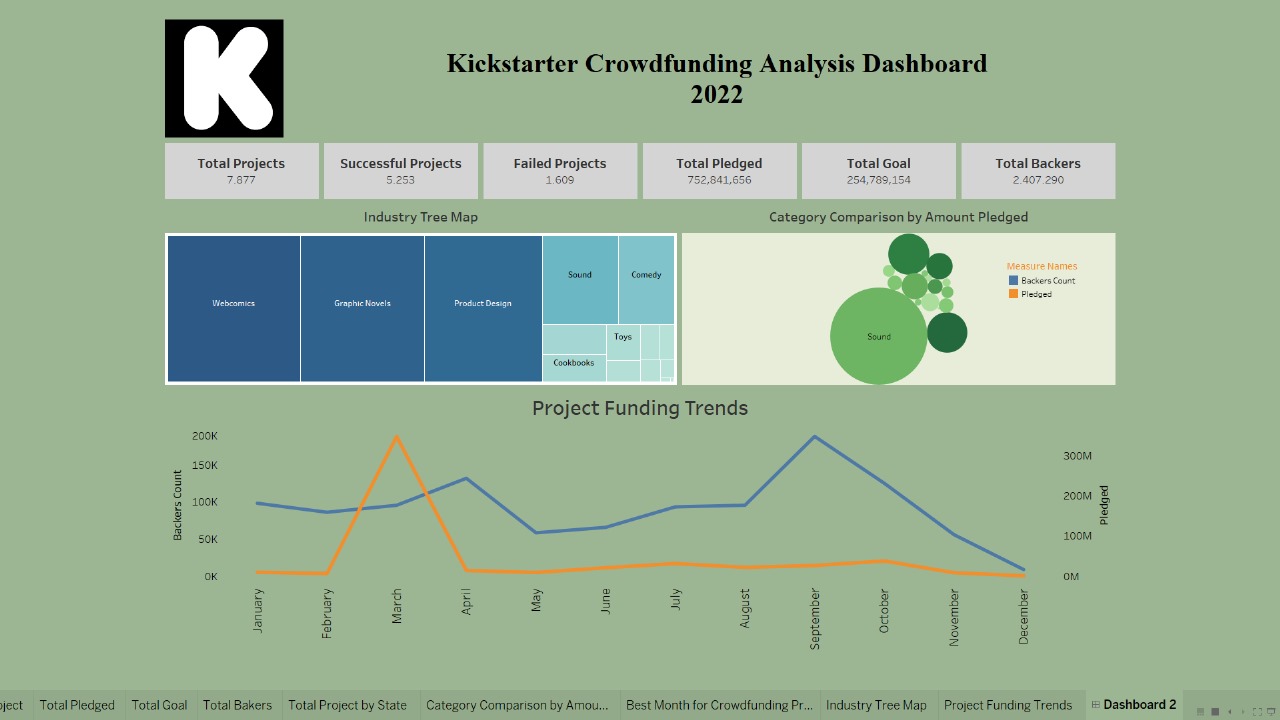


Gambar 3. 12 MOLAP Shema 1



Gambar 3. 13 MOLAP Shema 2

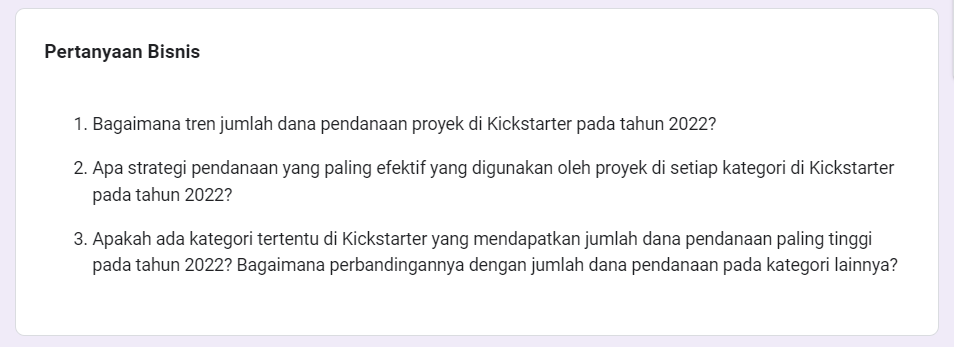
## *Dashboard*



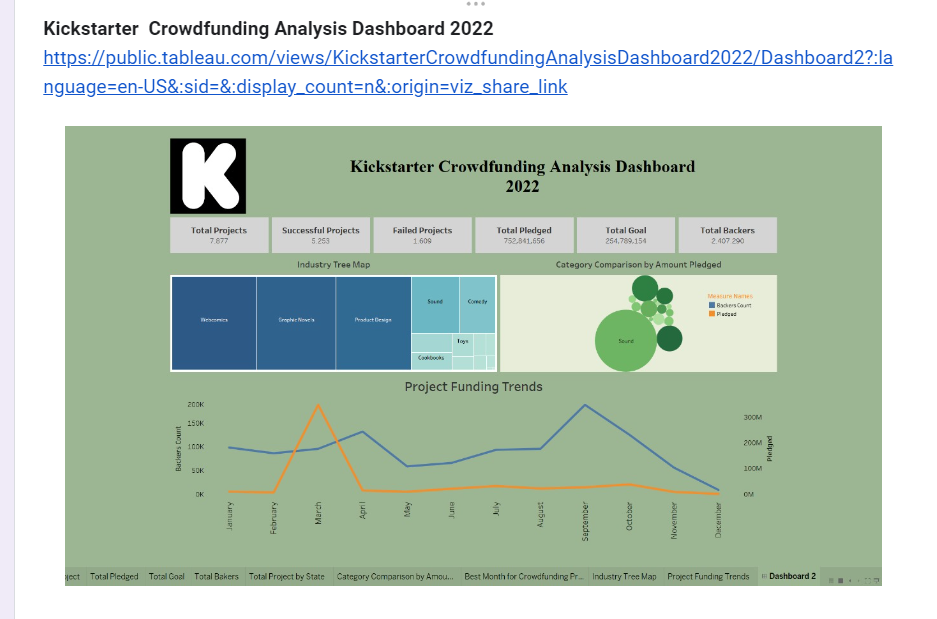
Gambar 3. 14 Dashboard Analysis

# Evaluasi

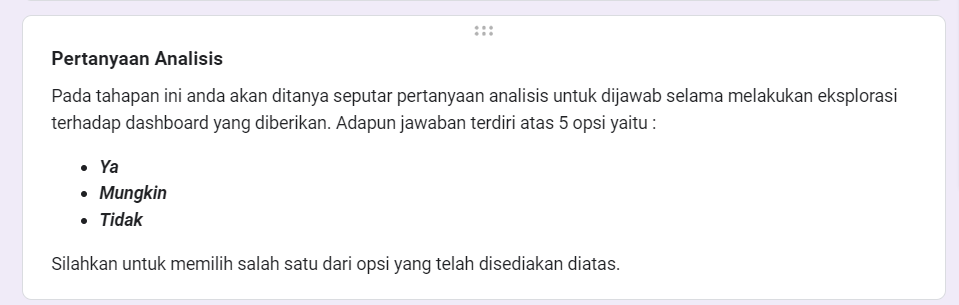
Untuk mengevaluasi dashboard yang telah kami bangun, kami melakukan survei sederhana dengan jumlah responden sebanyak 14 responden. Pada formulir yang kami berikan, kami menyertakan pertanyaan bisnis yang kami analisis, tampilan dashboard yang kami bangun, dan beberapa pertanyaan analisis.



Gambar 4. 1 Penjelasan Pertanyaan Analisis



Gambar 4. 2 Dashboard Analysis



Gambar 4. 3 Penjelasan Pengisian



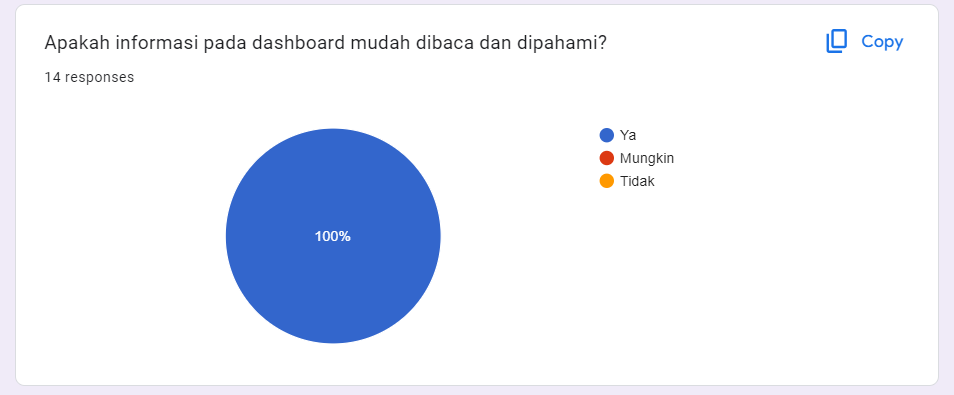
Gambar 4. 4 Analysis terhadap Pertanyaan Analisis nomor 1



Gambar 4. 5 Analysis terhadap Pertanyaan Analisis nomor 2



Gambar 4. 6 Analysis terhadap Pertanyaan Analisis nomor 3



Gambar 4. 7 Pemahaman Responden terhadap Dashboard



Gambar 4. 8 Pengaruh Warna terhadap Pemahaman

Dari survei yang kami lakukan, kami mendapatkan hasil sebagai berikut:

* Sebanyak 14 responden dapat mengetahui tren pendanaan proyek di Kickstarter pada tahun 2022 dengan melihat dashboard yang disajikan
* Sebanyak 14 responden dapat mengetahui strategi pendaan yang paling efektif yang digunakan oleh proyek di setiap kategori di Kickstarter pada tahun 2022 dengan melihat dashboard yang

# Penutup

## Kesimpulan

Dari analisis yang dilakukan pada data dalam studi kasus Kickstarter, beberapa alat dan perangkat lunak digunakan untuk mendukung proses pengerjaan Dashboard for Kickstarter. Alat dan perangkat lunak tersebut meliputi Pentaho, Schema Workbench, MySQL, Tableau, serta alat ETL (Extract, Transform, Load). Alat ETL berperan dalam mengumpulkan data dari berbagai sumber, memproses data, dan mentransfernya ke sistem penyimpanan data seperti data mart. Penggunaan perangkat lunak ini memungkinkan penggabungan data dari berbagai sumber, analisis data, dan pembuatan visualisasi interaktif.

Dengan demikian, Dashboard for Kickstarter yang dihasilkan dapat menjawab pertanyaan analisis sebagai berikut:

1. Bagaimana tren jumlah dana pendanaan proyek di Kickstarter pada tahun 2022?
2. Apa strategi pendanaan yang paling efektif yang digunakan oleh proyek di setiap kategori di Kickstarter pada tahun 2022?
3. Apakah ada kategori tertentu di Kickstarter yang mendapatkan jumlah dana pendanaan paling tinggi pada tahun 2022? Bagaimana perbandingannya dengan jumlah dana pendanaan pada kategori lainnya?

## Saran

Dalam laporan ini, kami memberikan rekomendasi untuk mengembangkan Dashboard Interaktif untuk platform Kickstarter. Pertama, identifikasikan tujuan dashboard dan tentukan metrik-metrik yang relevan seperti tingkat keberhasilan project. Persiapkan data dengan hati-hati dan pilih atribut-atribut yang penting seperti kategori, waktu, dan jumlah pendukung. Rancang tata letak dan visualisasi yang efektif dengan memilih grafik yang tepat dan menggunakan elemen visual yang jelas. Integrasikan dengan sumber data eksternal untuk memperkaya analisis. Tingkatkan performa dashboard dengan memilih indeks yang tepat dan memperbarui infrastruktur jika diperlukan. Dengan mengimplementasikan rekomendasi-rekomendasi ini, akan dapat mengembangkan Dashboard Interaktif yang kuat dan informatif untuk platform Kickstarter.

# 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 | Tren Pendanaan Crowdfunding | Menganalisis tren pendanaan crowdfunding dari waktu ke waktu. | Tren dan Pola | Kelompok\_5\_DWBI | 8 | 4 | Line Graph | Tanggal, Jumlah Proyek, Jumlah Dana yang Dikumpulkan |  |
| 2 | Pola Kategori Proyek Crowdfunding | Menganalisis pola kategori proyek crowdfunding. | Tren dan Pola | Kelompok\_5\_DWBI | 7 | 3 | Bar Chart | Kategori Proyek, Jumlah Proyek |  |
| 3 | Faktor-Faktor yang Mempengaruhi Keberhasilan Proyek Crowdfunding | Menganalisis faktor-faktor yang mempengaruhi keberhasilan proyek crowdfunding. | Prediksi Keberhasilan | Kelompok\_5\_DWBI | 8 | 4 | Decision Tree | Kategori Proyek, Jumlah Dana yang Dikumpulkan, Lokasi Proyek |  |

# 

# Lampiran B. Spesifikasi Rinci Aplikasi Kecerdasan Bisnis

**Aplikasi** Tren Pendanaan Crowdfunding

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| No | Elemen/Atribut | Lokasi | Tipe Fungsi | Nilai *Default* | Sumber | Dibuat di | Kueri | Komentar |
| 1 | Tanggal |  | Tanggal dan waktu proyek dibuat | Nilai default dapat diubah sesuai dengan kebutuhan atau karakteristik data yang ingin ditampilkan | Kickstarter | Line Graph dibuat pada dashboard | Query diperoleh dengan cara mengakses atau mengolah data yang diperoleh dari sumber data |  |
| 2 | Jumlah Proyek |  | Jumlah proyek yang akan dilakukan | Nilai default dapat diubah sesuai dengan kebutuhan atau karakteristik data yang ingin ditampilkan | Kickstarter | Line Graph dibuat pada dashboard | Query diperoleh dengan cara mengakses atau mengolah data yang diperoleh dari sumber data |  |
| 3 | Jumlah Dana yang Dikumpulkan |  | Jumlah uang yang telah dikumpulkan proyek sejauh ini | Nilai default dapat diubah sesuai dengan kebutuhan atau karakteristik data yang ingin ditampilkan | Kickstarter | Line Graph dibuat pada dashboard | Query diperoleh dengan cara mengakses atau mengolah data yang diperoleh dari sumber data |  |

**Aplikasi** Pola Kategori Proyek Crowdfunding

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| No | Elemen/Atribut | Lokasi | Tipe Fungsi | Nilai *Default* | Sumber | Dibuat di | Kueri | Komentar |
| 1 | Kategori proyek |  | Nama kategori proyek | Nilai default dapat diubah sesuai dengan kebutuhan atau karakteristik data yang ingin ditampilkan | Kickstarter | Bar Chart dibuat pada dashboard | Query diperoleh dengan cara mengakses atau mengolah data yang diperoleh dari sumber data |  |
| 2 | Jumlah proyek |  | Jumlah proyek yang akan dilakukan | Nilai default dapat diubah sesuai dengan kebutuhan atau karakteristik data yang ingin ditampilkan | Kickstarter | Bar Chart dibuat pada dashboard | Query diperoleh dengan cara mengakses atau mengolah data yang diperoleh dari sumber data |  |