DECODE: JURNAL PENDIDIKAN TEKNOLOGI INFORMASI
P-ISSN: 2775-2984 | E-ISSN: 2775-1813
Vol. 5, No 3, November 2025, Hal: 993-1008
Doi: http://dx.doi.org/10.51454/decode.v5i3.1343
https://journal.umkendari.ac.id/index.php/decode
This work is licensed under a CC BY-SA license.

Permodelan Data Warehouse Penerimaan Negara untuk Mendukung Treasury Open Data Dengan Metode Kimball

Muhammad Alfian Perdana[1*], Ahmad Rivai[1], Indra[1]

[1]Program Studi Magister Ilmu Komputer, Universitas Budi Luhur, Indonesia

Artikel Info

Kata Kunci:

Data warehouse;
Metode kimball;
OLAP;
Penerimaan negara;
Treasury open data.

Keywords:

Data warehouse;
Kimball method;
OLAP;
State revenue;
Treasury open data.

Riwayat Artikel:

Submitted: 22 Agustus 2025
Accepted: 27 Oktober 2025
Published: 30 Oktober 2025

Abstrak: Pengembangan data warehouse dalam rangka pengelolaan penerimaan negara merupakan langkah strategis untuk mendukung transparansi serta efisiensi tata kelola keuangan negara. Penelitian ini akan merancang dan mengimplementasikan model data warehouse berbasis *four step Kimball Method*, dengan fokus pada integrasi data transaksi penerimaan negara ke dalam satu platform terpadu. Sumber data berasal dari berbagai jenis transaksi penerimaan negara. Model data dirancang menggunakan *star schema* yang mencakup satu tabel fakta serta beberapa tabel dimensi. Sistem ini dirancang untuk memenuhi kebutuhan tiga kelompok pemangku kepentingan: (1) pimpinan instansi yang membutuhkan dashboard realisasi yang ringan dan cepat; (2) pengelola keuangan yang memerlukan integrasi data secara akurat; serta (3) pihak-pihak lain yang membutuhkan data terkait penerimaan negara. Desain yang fleksibel memungkinkan eksplorasi data berdasarkan tahun anggaran, jenis penerimaan, serta penerimaan perwilayah. Hasil implementasi menunjukkan bahwa arsitektur data warehouse ini mampu menyediakan data yang konsisten, terstandar, dan mudah diolah untuk keperluan pelaporan dan analitik. Selain meningkatkan efisiensi pelaporan internal, sistem ini juga membuka potensi integrasi dengan inisiatif *Treasury Open Data*, guna memperluas akses publik terhadap data penerimaan negara. Dengan demikian, model ini mendukung penguatan prinsip *good governance* dalam pengelolaan pendapatan nasional.

Abstract: *The development of a data warehouse for managing state revenue represents a strategic initiative to enhance transparency and efficiency in public financial governance. This study designs and implements a data warehouse model based on the Kimball Four-Step Method, focusing on the integration of state revenue transaction data into a unified platform. The data sources consist of various types of revenue transactions. The model is structured using a star schema, comprising one fact table and several dimension tables. The system is designed to serve three groups of stakeholders: (1) institutional leaders requiring lightweight and fast-access dashboards; (2) financial managers demanding accurate data integration; and (3) other stakeholders in need of state revenue information. Its flexible design enables data exploration by fiscal year, revenue type, and regional distribution. The implementation results demonstrate that this data warehouse architecture can provide consistent, standardized, and easily processed data for reporting and analytics. In addition to improving internal reporting efficiency, the system also creates opportunities for integration with the Treasury Open Data initiative, thereby expanding public access to state revenue information. Accordingly, the model supports the strengthening of good governance principles in managing national revenue.*

Corresponding Author:
Muhammad Alfian Perdana
Email: 2311601542@student.budiluhur.ac.id

PENDAHULUAN

Di saat era digital ini, terutama di bidang teknologi yang berkembang sangat cepat. Manusia memasuki era digitalisasi karena banyaknya peran penting teknologi (Tamma et al., 2025). Salah satu bidang yang berkembang di bidang teknologi yaitu produksi data digital. Produksi data yang berkembang secara masif dan luas saat ini menjadikan pemanfaatan data sebagai aset organisasi menjadi hal yang sangat penting, karena dapat menghasilkan informasi yang faktual untuk mendukung arah kebijakan dan kebutuhan organisasi yang dinamis, *agile*, dan responsif terhadap perubahan (Ratu et al., 2023). Pemanfaatan data ini sangat krusial dalam mendukung lembaga pemerintah agar mampu berkompetisi dan memberikan layanan yang berkesinambungan kepada para pemangku kepentingan. Namun disisi lain, Hambatan yang sering ditemui adalah kenyataan bahwa sistem informasi yang ada belum terintegrasi secara sempurna dan mengakibatkan informasi yang tersaji kurang lengkap dan akurat (Bakhri & Nuryamin, 2021). Untuk itulah diperlukan suatu pusat data yang bisa mengakomodir kebutuhan para pemangku kepentingan.

Direktorat Jenderal Perbendaharaan (DJPb) sebagai salah satu unit eselon I di bawah Kementerian Keuangan memiliki kekayaan data yang sangat besar, termasuk di dalamnya data penerimaan pendapatan negara yang sangat detail. Data ini berasal dari berbagai satuan kerja di seluruh Indonesia, mencakup berbagai jenis akun penerimaan seperti pajak, PNBP, hibah, dan lainnya. Jaringan koordinasi DJPb yang menjangkau pusat hingga daerah menjadikan lembaga ini sangat potensial dalam membangun kultur *data-driven organization*, melalui adopsi teknologi terkini yang memungkinkan transformasi data menjadi pengetahuan (*knowledge*) melalui pendekatan data analitik.

Sebagai unit IT di lingkungan DJPb, Direktorat Sistem Informasi dan Teknologi Perbendaharaan (SITP) telah mengembangkan sistem *Treasury Big Data* (TBD) sebagai *platform* terpusat yang menghimpun data dari berbagai sistem aplikasi pengelolaan keuangan negara. Salah satu inisiatif penting dalam TBD adalah pengembangan fitur *Open Data*, yaitu antarmuka aplikasi yang memungkinkan pengguna internal maupun eksternal untuk mengakses dan mengunduh data set yang telah disiapkan untuk keperluan pelaporan, pengawasan, dan analisis kebijakan.

Salah satu data set strategis yang direncanakan untuk dibuka dalam fitur *Open Data* adalah data penerimaan negara. Data ini sangat dibutuhkan, baik oleh instansi vertikal DJPb untuk penyusunan laporan realisasi, maupun oleh pihak eksternal seperti akademisi, pengawas internal, dan perencana kebijakan fiskal. Namun, Sistem informasi yang ada saat ini sebagian besar berbasis *Online Transaction Processing* (OLTP), yang dirancang untuk memproses transaksi harian secara cepat dan aman. Meskipun demikian, OLTP tidak dirancang untuk analisis data yang kompleks karena struktur tabelnya sangat terfragmentasi (Rafique et al., 2024). Kondisi ini menyebabkan kesulitan dalam mengintegrasikan data penerimaan negara yang tersebar di berbagai tabel dan aplikasi, sehingga menghambat kebutuhan analisis manajerial maupun pelaporan publik.

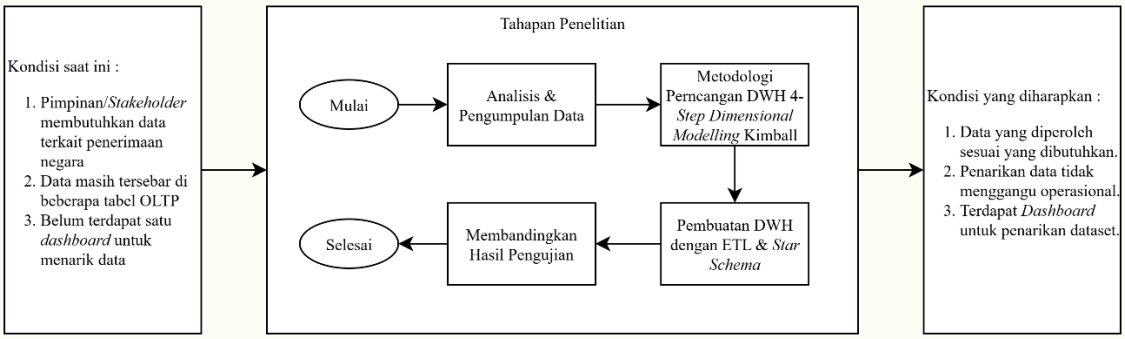
Solusi yang banyak digunakan untuk mengatasi masalah tersebut adalah penerapan *data warehouse* (DW). *Data warehouse* memungkinkan integrasi data dari berbagai sumber, menyediakan model multidimensional yang lebih mudah dianalisis, dan mendukung penerapan *business intelligence* (Triaji et al., 2024). Lebih jauh, penelitian terbaru menunjukkan bahwa DW di sektor publik mampu meningkatkan transparansi fiskal dan efisiensi pelaporan. Misalnya, (Bakhri & Nuryamin, 2021) membangun Eksekutif Pada Yayasan Ummu'l Quro Di Depok dan berhasil mempercepat proses analisis serta mendukung pengambilan keputusan di tingkat eksekutif yayasan. (Asmita et al., 2023) juga menegaskan bahwa metode Kimball efektif dalam membangun DW untuk pengelolaan keamanan informasi di lembaga pemerintahan.

Metode Kimball menjadi salah satu pendekatan yang populer dan banyak digunakan karena berfokus pada kebutuhan bisnis (*bottom-up approach*) dan menghasilkan desain *data mart* yang lebih sederhana dan efisien (R. Kimball and M. Ross, 2013). Dibandingkan metode lain seperti Inmon atau *Data Vault*, Kimball lebih sesuai untuk organisasi yang membutuhkan pelaporan cepat dan dashboard interaktif (Smith & Elshnoudy, n.d.). Dengan menggunakan star schema, model data dapat diakses secara intuitif oleh pengguna, serta meningkatkan performa query pada saat analisis.

Sejalan dengan agenda keterbukaan data (*open data*), pengembangan DW penerimaan negara diharapkan dapat mendukung inisiatif *Treasury Open Data* yang bertujuan meningkatkan transparansi, partisipasi publik, dan tata kelola keuangan yang lebih akuntabel. Oleh karena itu, penelitian ini bertujuan untuk merancang dan mengimplementasikan DW penerimaan negara berbasis Kimball *Method*, dengan harapan dapat menyediakan data yang konsisten, terstandar, serta mudah diakses baik untuk kebutuhan internal maupun publik.

METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan rancangan studi prediktif yang bertujuan membangun model klasifikasi berbasis *data warehouse* penerimaan negara. Tujuan penelitian adalah untuk memberikan gambaran yang jelas tentang fenomena yang diteliti dengan mengkomunikasikan tujuan yang jelas, merencanakan pendekatan pelaksanaan, dan mengumpulkan berbagai jenis data yang akan digunakan untuk membuat laporan penelitian (Hartono et al., 2024). Data diperoleh dari basis data transaksi penerimaan yang dikelola oleh Direktorat Jenderal Perbendaharaan, kemudian diolah melalui proses *Extract, Transform, Load* (ETL) dengan menggunakan aplikasi pentaho sebelum dimuat ke dalam *data warehouse*. Seluruh proses dirancang menggunakan metodologi Kimball *Four-Step* yang meliputi: pemilihan proses bisnis, penentuan granularitas data, identifikasi dimensi, serta identifikasi fakta.



Gambar 1. Bagan Tahapan Penelitian

Berdasarkan alur tahapan penelitian seperti yang tertera pada gambar 1, penelitian ini akan dilakukan melalui beberapa tahapan yang menggambarkan alur kegiatan mulai dari kondisi eksisting sistem, perancangan model *data warehouse*, hingga pengujian hasil implementasi. Setiap tahap disusun selaras dengan pendekatan Kimball *Four-Step Dimensional Modelling* yang terdiri atas pemilihan proses bisnis, penentuan granularitas, identifikasi dimensi, dan identifikasi fakta. Adapun tahapan penelitian secara rinci dijelaskan sebagai berikut:

Analisis dan Pengumpulan Data

Tahapan awal penelitian dimulai dengan melakukan analisis terhadap kondisi eksisting pengelolaan data penerimaan negara. Proses ini mencakup identifikasi kebutuhan informasi yang dibutuhkan oleh pimpinan dan pemangku kepentingan, serta pemetaan sumber data utama dari sistem *Online Transaction Processing* (OLTP) yang dikelola oleh Direktorat Jenderal Perbendaharaan. Data yang dikumpulkan meliputi tanggal transaksi, jenis penerimaan, kode akun, kode satuan kerja, dan nilai transaksi. Proses pengambilan data dilakukan menggunakan query SQL kemudian disimpan ke area *data staging* untuk memudahkan proses integrasi pada tahap berikutnya.

Pre-Processing Data

Setelah data terkumpul, dilakukan proses *pre-processing* untuk meningkatkan kualitas dan konsistensi data sebelum dimuat ke dalam data warehouse. Proses ini meliputi tiga tahapan utama:

1. *Cleaning* Data, yaitu menghapus data duplikat, memperbaiki nilai yang hilang, serta menyeragamkan format penulisan kode akun dan satuan kerja agar sesuai dengan standar sistem keuangan negara.
2. *Transformation* dan *Normalization*, yaitu mengubah struktur dan format data agar sesuai dengan kebutuhan analisis multidimensi, termasuk pembentukan *date hierarchy* dan *surrogate key*.
3. *Integration*, yaitu menggabungkan data dari berbagai sumber ke area *staging* dan memetakan hubungan antar atribut agar konsisten.

Ketiga tahapan ini menjadi pondasi yang penting untuk memastikan integritas dataset sebelum dimuat ke model *star schema* pada data *warehouse.*

Perancangan Data *Warehouse* dengan Metode Kimball *Four-Step*

Setelah data siap, dilakukan perancangan struktur data *warehouse* menggunakan metodologi Kimball *Four-Step Dimensional Mo*delling, yang terdiri dari:

1. Pemilihan Proses Bisnis (*Select the Business Process*) – proses yang dipilih adalah pelaporan realisasi penerimaan negara.
2. Penentuan Granularitas Data (*Declare the Grain*) – ditetapkan pada level transaksi penerimaan per akun, satuan kerja, dan tanggal.
3. Identifikasi Dimensi (*Identify the Dimensions*) – meliputi dimensi waktu, akun penerimaan, satuan kerja, dan kurs.
4. Identifikasi Fakta (*Identify the Facts*) – membentuk tabel fakta utama (fact_table_new) yang menyimpan metrik utama seperti total nilai penerimaan (TOTAL_IDR dan TOTAL_USD) serta *foreign key* ke tabel-tabel dimensi.

Desain ini kemudian diwujudkan dalam bentuk Star Schema yang menjadi dasar penyimpanan data pada *data warehouse*.

Implementasi Proses *Extract, Transform, Load* (ETL)

Tahap ini merupakan implementasi teknis dari rancangan data *warehouse* yang telah disusun. Proses dilakukan menggunakan aplikasi Pentaho *Spoon* dengan tiga langkah utama:

1. *Extract*, merupakan proses penarikan data dari sistem OLTP ke area *staging* setelah melalui proses verifikasi struktur data;
2. *Transform*, merupakan proses untuk membersihkan, menggabungkan, dan menyesuaikan format data dengan model dimensional;
3. Load, merupakan proses untuk memindahkan data hasil transformasi ke area data *mart* dalam bentuk *star schema*.

Proses ETL dijalankan secara terjadwal (*scheduler job*) agar sinkron dengan pembaruan data di sistem operasional.

Visualisasi dan Pengujian Sistem

Setelah data berhasil dimuat ke dalam data *warehouse*, dilakukan tahap visualisasi data melalui *dashboard* interaktif menggunakan Pentaho *Server*. *Dashboard* ini menampilkan indikator utama penerimaan negara seperti total realisasi per periode, akun, satuan kerja, dan mata uang. Pengujian sistem dilakukan dengan dua metode, yaitu:

1. *Blackbox Testing*, yaitu proses untuk memastikan setiap fungsi menampilkan data yang benar sesuai kebutuhan.
2. *User Acceptance Test* (UAT), yang melibatkan delapan pengguna untuk menilai kemudahan dan keakuratan sistem.

Pengujian ini memastikan bahwa sistem telah memenuhi kebutuhan pengguna dan siap untuk integrasi dengan portal *Treasury Open Data*.

Kondisi Akhir yang Diharapkan

Setelah seluruh proses dilalui, diharapkan data *warehouse* yang dibangun dapat menyediakan data penerimaan negara yang konsisten, terstandar, dan mudah diakses tanpa mengganggu sistem operasional utama. Selain mendukung efisiensi pelaporan internal, sistem ini juga memungkinkan

pemanfaatan data publik secara terbuka melalui portal *Treasury Open Data*, sejalan dengan prinsip transparansi dan *good governance*.

Studi Pustaka

Literatur yang digunakan sebagai dasar penelitian diperoleh dari jurnal-jurnal ilmiah bereputasi yang relevan dengan topik *data warehouse*, ETL, dan metode Kimball. Metode Kimball dipilih karena penelitian ini akan berfokus secara khusus pada proses bisnis untuk merancang basis data dalam pembentukan *dashboard* dan pelaporan analitis (Ygnacio et al., 2023). Sumber utama berasal dari publikasi 5 tahun terakhir yang terindeks Scopus maupun Sinta, serta buku acuan klasik mengenai *data warehouse* dan *dimensional modeling*. Selain itu, literatur tambahan yang digunakan meliputi artikel terkait implementasi *business intelligence* di sektor publik dan studi kasus penerapan *star schema* untuk pengelolaan data keuangan. Seluruh referensi dipilih untuk memperkuat landasan konseptual dan metodologi penelitian.

Pengumpulan Data

Tahap pengumpulan data dilakukan sebagai langkah awal dalam penelitian ini. Data yang digunakan bersumber dari sistem *Online Transaction Processing* (OLTP) yang dikelola oleh Direktorat Jenderal Perbendaharaan, yang menyimpan transaksi penerimaan negara secara harian. Data tersebut mencakup informasi pokok seperti tanggal transaksi, jenis penerimaan, kode akun, kode satuan kerja (satker), serta nilai transaksi. Data diperoleh melalui mekanisme *query* langsung terhadap basis data OLTP dengan menggunakan *Structured Query Language* (SQL).

Karena penelitian ini difokuskan pada perancangan data warehouse, data transaksi tidak diambil secara menyeluruh, melainkan dipilih sesuai kebutuhan analisis. Proses seleksi data mengacu pada metodologi Kimball yang menekankan pemilihan proses bisnis relevan sebagai titik awal pembangunan *data mart* (R. Kimball and M. Ross, 2013). Dalam penelitian ini, proses bisnis yang dipilih adalah penerimaan negara, yang selanjutnya dijadikan sebagai tabel fakta utama. Jumlah data yang berhasil dikumpulkan mencakup lebih dari 100.000 baris transaksi dalam periode tahun anggaran 2021–2023. Data tersebut kemudian disimpan dalam format *comma separated values* (CSV) sebelum diproses lebih lanjut dalam tahap preprocessing. Format CSV dipilih karena mendukung portabilitas dan dapat dengan mudah diintegrasikan dengan perangkat ETL (*Extract, Transform, Load*).

Selain pengumpulan data primer dari sistem OLTP, penelitian ini juga mengacu pada literatur terkini untuk memperkuat rancangan skema dan validasi metodologi. Studi (Bahaa et al., 2021a) menegaskan bahwa penerapan *staging area* yang baik dalam proses ETL berperan penting dalam menjaga integritas dan kualitas data sebelum dimuat ke *data warehouse*. Sementara itu, penelitian (Koppula, 2021) menunjukkan bahwa optimasi alur ETL yang terstruktur dapat meningkatkan efisiensi proses pemuatan dan meminimalkan risiko inkonsistensi data. Dengan demikian, tahap pengumpulan data tidak hanya bertujuan untuk mengumpulkan dataset mentah, tetapi juga memastikan bahwa data yang diambil telah sesuai dengan kebutuhan analisis penerimaan negara serta siap untuk ditransformasikan ke dalam model *star schema*.

HASIL DAN PEMBAHASAN

***Pre-Processing* Data**

Pada tahap ini, data transaksi penerimaan negara yang diperoleh dari sistem OLTP perlu melalui proses *pre-processing* sebelum dapat dimasukkan ke dalam *data warehouse*. Proses ini dilakukan untuk mengatasi masalah kualitas data seperti duplikasi, nilai yang hilang, maupun ketidaksesuaian format. *Pre-processing* menjadi langkah krusial karena kualitas data sangat menentukan keandalan model *star schema* dan *performa query* dalam *data warehouse* (Balázs et al., 2021). Adapun tahapan pre-processing data dalam penelitian ini meliputi:

1. *Cleaning Data,* Proses *cleaning data* dilakukan untuk menghapus data duplikat, memperbaiki data yang hilang (*missing values*), serta menstandarkan format atribut sebelum data dimasukkan ke dalam *star schema*. Sebagai contoh, penulisan kode akun dan satuan kerja yang tidak konsisten

diseragamkan agar sesuai dengan standar sistem keuangan negara. Hasil pembersihan menunjukkan bahwa sekitar 3,2% dari data awal dilakukan proses *cleaning data*. Studi yang lebih mutakhir menunjukkan bahwa "*data cleansing is a basic piece of the transformation stage in data warehousing*" (Bahaa et al., 2021b) dan merupakan upaya penting untuk memastikan bahwa hanya data bermutu tinggi yang dimuat ke dalam *data warehouse* (Gupta & Jolly, 2021). Dengan demikian, *cleaning data* menjadi langkah krusial untuk menjaga integritas dan keandalan sistem analitik.

2. *Transformation* dan *Normalization,* Tahap transformasi dilakukan untuk mengubah struktur data agar sesuai dengan kebutuhan analisis multidimensi—meliputi konversi format tanggal ke dalam *date hierarchy* (hari, bulan, tahun), pembentukan *surrogate key* untuk tabel dimensi, serta normalisasi nilai transaksi agar konsisten dan mudah digabungkan. Teknik normalisasi dan standardisasi juga diperlukan agar data dari berbagai sumber memiliki format yang seragam. Penelitian menunjukkan bahwa "*data normalization is a necessary initial stage of data processing … especially if the variables are measured in scales that differ significantly in values*" (Kalinina et al., 2025) dan bahwa transformasi yang tepat dapat meningkatkan kecepatan *query* dan akurasi analisis pada sistem *data warehouse* (Amorim et al., 2023). Dengan demikian, tahap transformasi dan normalisasi tidak hanya meningkatkan konsistensi data, tetapi juga mendukung performa analitik dan kualitas output laporan.
3. *Integration,* Data penerimaan negara yang semula tersebar di berbagai tabel OLTP digabungkan dalam *staging area*. Proses integrasi ini menghubungkan data dari beberapa sumber sehingga dapat dimuat secara utuh ke dalam data warehouse. Integrasi juga melibatkan pemetaan antar atribut (misalnya antara kode akun dan jenis penerimaan) agar hubungan antar data tetap konsisten. Menurut (Asmita et al., 2023), integrasi data yang baik berperan penting dalam mendukung analisis yang komprehensif dan meningkatkan nilai guna data bagi pengambil keputusan.

***Dimensional Modeling* Menggunakan Metode Kimball (*Four-Step Method*)**

1. Menentukan Proses Bisnis (*Select the Business* ***Process*****),** Langkah pertama dalam metode Kimball adalah memilih proses bisnis utama yang akan dianalisis. Proses bisnis ini harus relevan dengan tujuan analisis. Pemilihan proses ini menjadi dasar untuk menentukan lingkup data yang akan dimodelkan. Dalam praktiknya, proses bisnis harus dapat diukur dan memiliki aktivitas yang menghasilkan fakta numerik Al-Rahman et al., 2023). Proses bisnis yang dipilih dalam penelitian ini adalah proses pelaporan realisasi penerimaan negara. Tujuan utamanya adalah menyediakan *data mart* yang mampu menampilkan informasi penerimaan berdasarkan:
   a. Jenis akun penerimaan (pajak, PNBP, hibah, dll.)
   b. Satuan kerja (satker) dan kementerian/lembaga (K/L)
   c. Wilayah geografis satker penyetor
   d. Waktu transaksi (bulanan, triwulan, tahunan)
   e. Jenis mata uang (Rupiah dan valuta asing)
2. Menentukan Granularitas (*Declare the Grain*)

   Granularitas menentukan tingkat detail data dalam tabel fakta. Ini adalah salah satu aspek paling penting karena Pada tahap ini akan diputuskan apa saja yang dapat mewakili atau dipresentasikan oleh sebuah tabel fakta. Menentukan granularitas secara tepat memastikan data dapat digunakan untuk analisis yang akurat dan konsisten (Asmita et al., 2023). *Grain* atau tingkat detail dari data ditetapkan pada level transaksi penerimaan per satker per akun per tanggal. Setiap baris data dalam tabel fakta (fact_table_new) merepresentasikan total penerimaan (dalam IDR dan USD) yang terjadi pada kombinasi:
   a. Tanggal transaksi (GL_DATE)
   b. Kode akun pendapatan (KD_AKUN)
   c. Kode satuan kerja (KD_SATKER)
   d. Mata uang transaksi (CURRENCY_CODE)

   Granularitas ini memungkinkan agregasi dan analisis penerimaan pada berbagai level, seperti per wilayah, akun, K/L, maupun periode waktu tertentu.

3. Mengidentifikasi Dimensi (*Identify the Dimensions*)

Setelah granularitas data ditetapkan, langkah selanjutnya dalam perancangan model *data warehouse* adalah mengidentifikasi dimensi yang akan memberikan konteks pada fakta. Dimensi biasanya mencakup atribut-atribut seperti waktu, lokasi, produk, pelanggan, atau unit kerja, sehingga analisis dapat dilakukan dengan hierarki yang relevan (misalnya: hari → bulan → kuartal → tahun). Dalam penelitian terkini dijelaskan bahwa "*dimensions are descriptive attributes that categorize and describe the facts in a data warehouse*" (Fattah, 2023) dan bahwa desain dimensi yang baik memungkinkan pengguna untuk "*slice and dice*" data dengan berbagai perspektif bisnis. Berdasarkan granularitas yang telah ditetapkan, beberapa dimensi yang diidentifikasi dalam penelitian ini antara lain:

a. Dimensi Waktu (dim_waktu): Menyediakan struktur waktu berdasarkan tahun anggaran (TA), periode (Periode), dan triwulan (Periode_Triwulan).
b. Dimensi Akun/Kelompok Akun (dim_kelompok_akun): Menyediakan klasifikasi jenis akun, kelompok akun, dan nama jenis pendapatan.
c. Dimensi Satuan Kerja (dim_kdsatker): Informasi lengkap tentang satker, Kementerian/Lembaga, Lokasi, KPPN, dan Provinsi.
d. Dimensi Kurs (dim_kurs): Menyimpan nilai tukar mata uang (Kurs) untuk setiap tanggal (GL_Date) dan kodem ata uang (Currency_Code).
e. Dimensi Mata Uang (Diintegrasikan dalam fakta): Walaupun tidak terpisah menjadi tabel sendiri, kode mata uang (Currency_Code) disimpan langsung di tabel fakta dan direferensikan ke dim_kurs.

4. Mengidentifikasi Fakta (*Identify the Facts*)

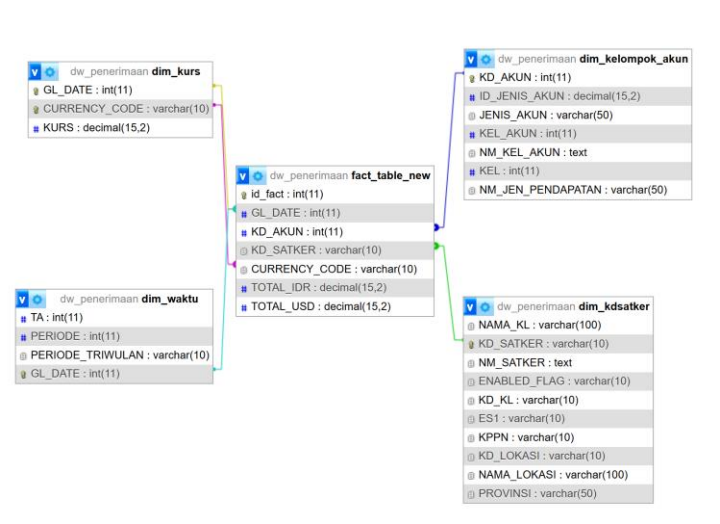
Langkah terakhir adalah mengidentifikasi fakta numerik yang dapat diukur dan berkaitan dengan proses bisnis yang telah dipilih. Fakta ini biasanya berupa jumlah, nilai, atau kuantitas yang dapat dijumlahkan, rata-rata, atau dihitung (seperti total penjualan, jumlah unit, atau biaya). Fakta harus konsisten dengan granularitas yang telah ditetapkan dan harus dapat dikaitkan dengan dimensi yang telah diidentifikasi (Al-Rahman et al., 2023). Tabel fakta utama yang dibentuk adalah fact_table_new, yang berisi metrik-metrik penerimaan negara sebagai berikut:

a. TOTAL_IDR: Nilai penerimaan dalam rupiah
b. TOTAL_USD: Nilai penerimaan dalam dolar (jika dikonversi)
c. KD_AKUN: Foreign key ke dimensi akun
d. KD_SATKER: Foreign key ke dimensi satker
e. GL_DATE: Foreign key ke dimensi waktu
f. CURRENCY_CODE: Foreign key ke dimensi kurs

Struktur ini memungkinkan pengguna melakukan agregasi penerimaan berdasarkan waktu, organisasi, wilayah, jenis akun, dan mata uang.

Membuat Desain *Data Mart* dengan Metode *Star Schema*

Proses pembentukan data mart pada penelitian ini menggunakan metode *star schema*. *Star schema* merupakan metode pemodelan data yang efektif dan sering digunakan dalam pembentukan *data warehouse* untuk mengatur data secara terstruktur dan mudah dipahami. Dalam model *star schema data warehouse*, data disimpan dalam sebuah tabel fakta sentral yang berisi berbagai ukuran yang relevan, dikelilingi oleh tabel-tabel dimensi yang menjelaskan detail lebih lanjut dari ukuran-ukuran tersebut. Adapun bentuk *star schema* untuk penelitian ini dapat dilihat pada gambar 2 berikut:

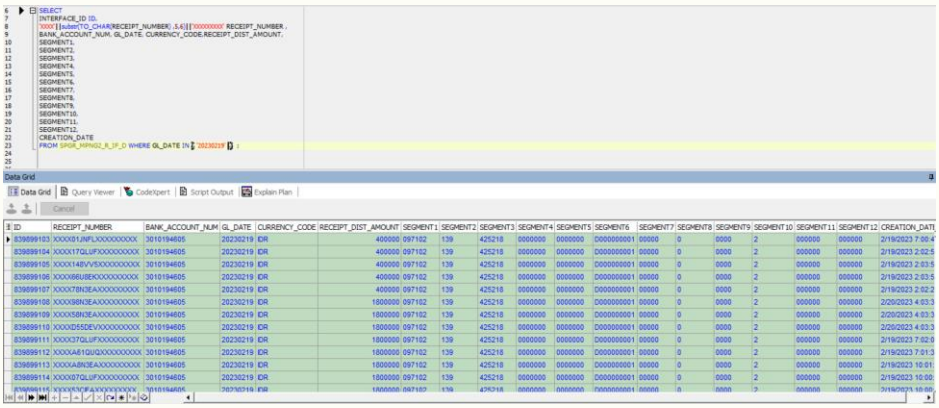


Gambar 2. *Star Schema*

Proses *Extract, Transform, Load* (ETL)

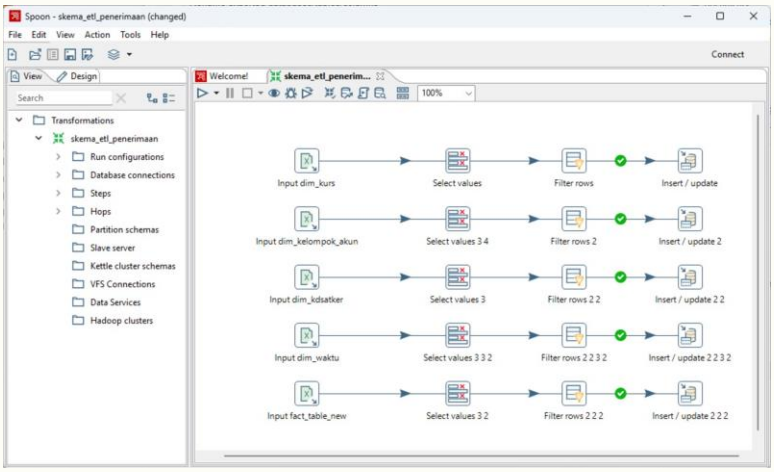
Proses ETL mengacu pada rangkaian operasi berurutan yang mencakup tahap extract, transform, dan load, yang merupakan inti dari arsitektur sistem data warehouse (Sethi, 2021). Proses ETL merupakan inti dan tulang punggung arsitektur data warehouse (Dhaouadi et al., 2022). Tujuannya adalah menstandarkan pemrosesan data dari berbagai sumber sehingga data tersedia dan sesuai untuk digunakan dalam analisis dan pelaporan (Mandala, 2021). Proses ini meliputi ekstraksi data dari beragam sumber—seperti basis data relasional, berkas flat file, *platform* CRM/ERP, atau tipe data semi-terstruktur—kemudian pembersihan dan penyatuan data agar dapat dibebaskan dari duplikasi dan anomali seperti karakter khusus atau missing values, sebelum muat ke target *data warehouse (Khan et al., 2024)*. Selanjutnya, transformasi bisnis diterapkan: termasuk perhitungan agregat, penggabungan tabel, normalisasi nilai, dan pembentukan *surrogate key* agar data siap digunakan dalam model multidimensi. Setelah itu, data yang telah dinormalisasi dan terintegrasi dimuat (load) ke dalam skema *star-schema* atau area *data mart*, dimana kemudian dapat dimanfaatkan untuk kegiatan *business intelligence* dan pengambilan keputusan strategis (Koppichetti, 2022). Berikut proses ETL yang dilakukan pada penelitian ini:

1. *Extract*, Proses ekstraksi adalah tahap pertama dalam proses ETL di mana data diambil dari berbagai sumber untuk kemudian dipindahkan ke area staging (penyimpanan sementara). Tahapan ini dimulai dengan proses pengecekan data pada tabel induk penerimaan negara yang ada pada database OLTP seperti pada gambar 3 untuk selanjutnya dimasukkan ke dalam satu tabel ke area staging.



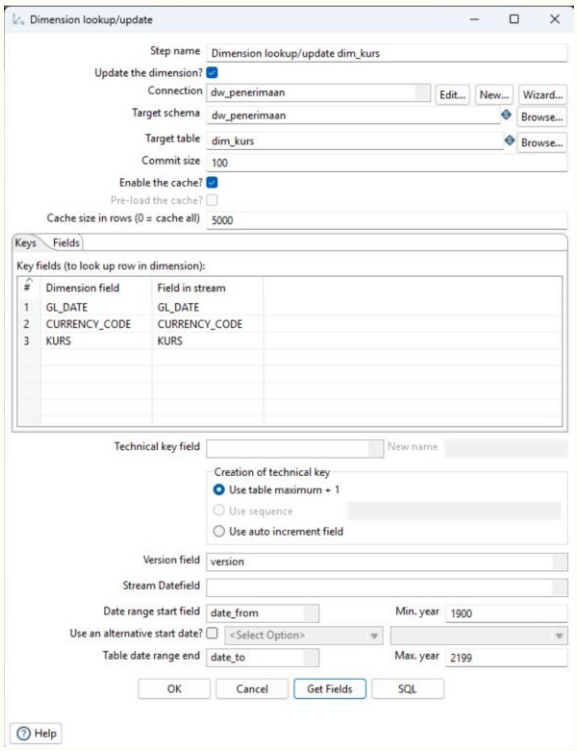
Gambar 3. Proses pengecekan data sumber

Apabila data telah sesuai, tahap selanjutnya melakukan penarikan data tersebut untuk dimasukkan ke dalam tabel di area staging untuk dilakukan proses selanjutnya. Proses penarikan data memakai aplikasi Pentaho Spoon dengan mekanisme seperti yang terlihat pada gambar 4.



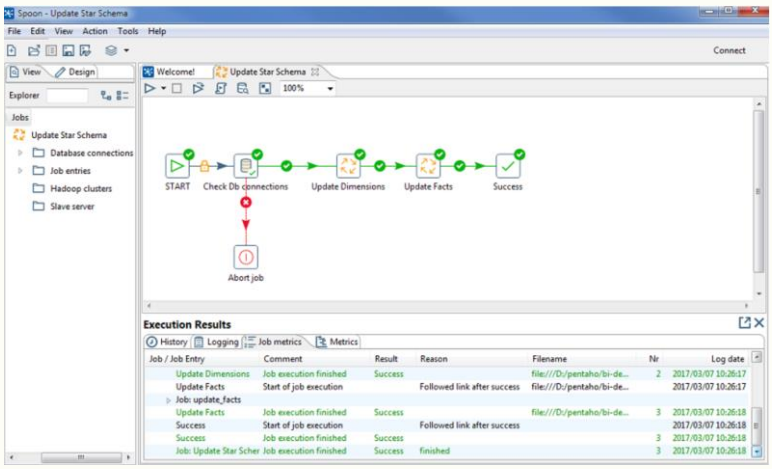
Gambar 4. Proses penarikan data ke area *staging*

2. *Transform*, Setelah data masuk ke dalam area staging, selanjutnya data yang telah diekstraksi diubah dan dikonsolidasikan agar sesuai dengan kebutuhan analisis dan format yang diperlukan oleh data warehouse. Pada tahap ini dilakukan proses pemfilteran dan pembersihan data (*cleansing data*) yaitu menghapus data yang tidak valid atau tidak diperlukan, menghapus duplikasi, dan memastikan kualitas data, penggabungan dan pengagregasian data (*data integration*) yaitu menggabungkan data dari beberapa sumber dan meringkas data untuk analisis yang lebih mudah, validasi dan autentikasi yaitu memastikan data memenuhi aturan bisnis dan persyaratan kepatuhan yang telah ditentukan sebelumnya, dan terakhir melakukan transformasi format yaitu mengubah format data agar sesuai dengan skema data warehouse yang dituju. Gambar 5 di bawah ini menunjukkan salah satu proses pembuatan tabel view dimensi yang nantinya akan dilakukan load ke data mart.



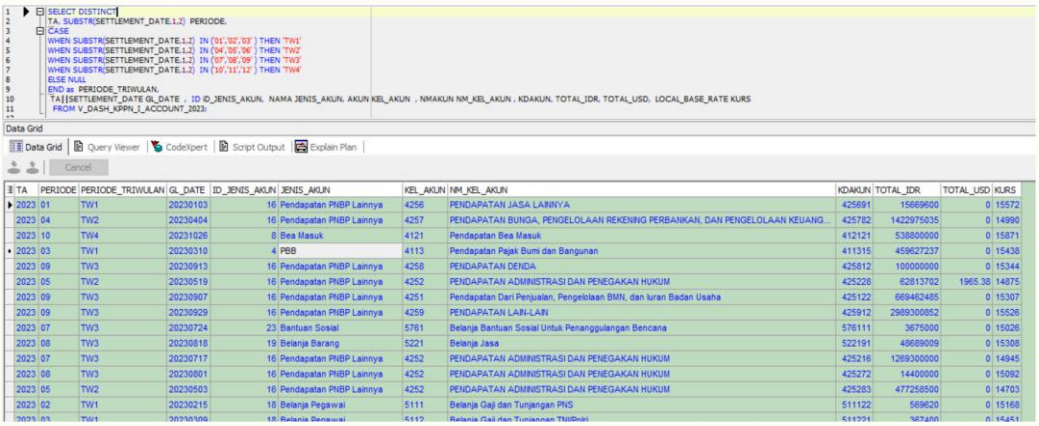
Gambar 5. Pembuatan *view* dimensi kurs

3. *Loading*, Setelah proses transform selesai dilakukan, selanjutnya data yang telah ditransform dimasukkan ke dalam data warehouse (area data mart) untuk dimanfaatkan dalam bentuk dashboard yang akan dibuat, seperti yang terlihat pada gambar 6 ini.



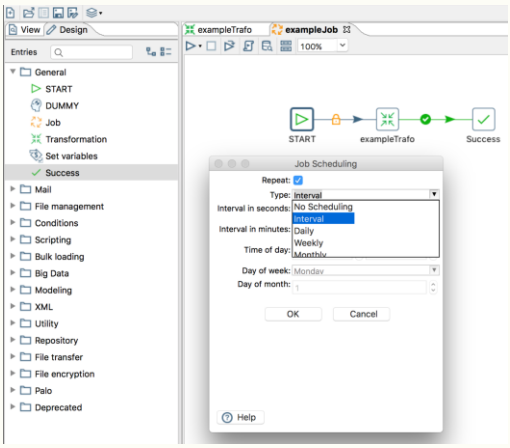
Gambar 6. Proses *load* data

Setelah proses load data berhasil, maka data sudah berhasil mengalir ke tabel fact_table_new. Gambar 7 berikut ini menampilkan data yang berhasil mengalir ke tabel fakta.



Gambar 7. Data Pada Tabel Fakta

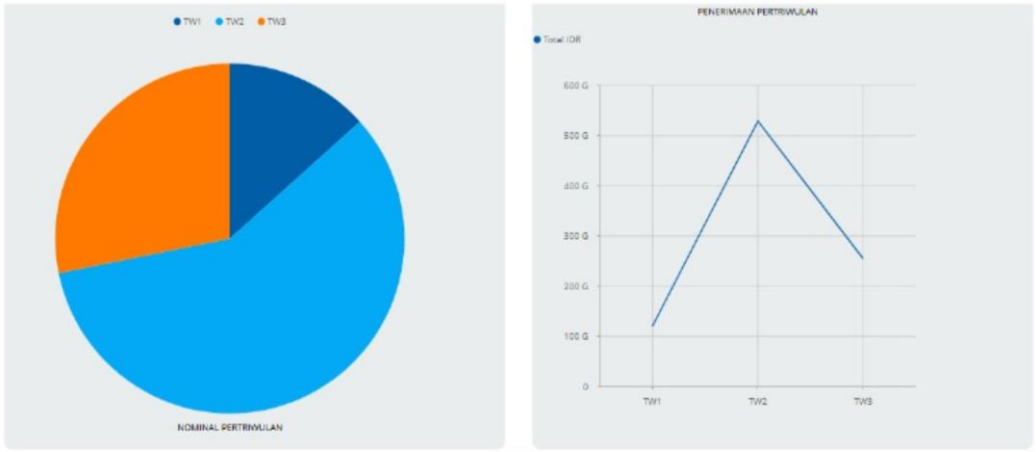
Agar data yang telah dimasukkan ke dalam *data mart* dapat sinkron dengan data real yang ada di database production, maka proses ETL di atas dijalankan secara periodik dengan menggunakan *scheduler*. Pengaturan *job scheduler* dapat dilihat pada gambar 8 di bawah ini.



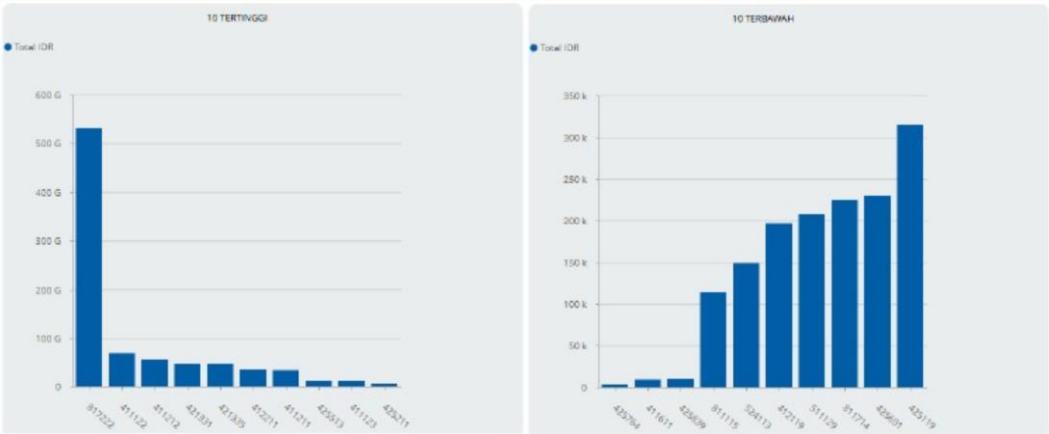
Gambar 8. Pengaturan *job scheduler* harian

Visualisasi Data Pada *Dashboard*

Data yang telah ada pada data warehouse selanjutnya divisualisasikan ke dalam *dashboard* dengan menggunakan aplikasi Pentaho Server untuk dilakukan pengujian keakuratan data yang akan dilakukan oleh pengguna data. Berikut disajikan pada gambar 9 dan 10 beberapa tampilan *dashboard* terkait penggunaan data untuk keperluan monitoring terkait dengan realisasi penerimaan negara.



Gambar 9. Data Rekapitulasi Penerimaan Negara Per Triwulan



Gambar 10. Data 10 akun penerimaan tertinggi dan terendah

Pengujian

Dalam rangka menilai apakah proses pembuatan data warehouse terkait pemanfaatan data penerimaan negara sesuai dengan keinginan pengguna, maka perlu dilakukan pengujian kehandalan data yang telah dibuat tersebut. Pengujian dilakukan dengan metode kuantitatif deskriptif, menggunakan *User Acceptance Test* (UAT) untuk menilai tingkat penerimaan sistem oleh pengguna. Pengukuran dilakukan melalui kuesioner berbasis skala Likert (1-5), yang menghasilkan nilai persentase penerimaan sistem sebesar 97%. Metode ini digunakan untuk memperoleh data numerik terkait tingkat kesesuaian sistem dengan kebutuhan pengguna. Pengujian ini melibatkan 8 (delapan) orang yang terdiri dari tiga orang pejabat struktural setingkat eselon IV dan lima orang staff calon pengguna data ini. Jenis pengujian yang digunakan yaitu *blackbox testing. Blackbox testing* merupakan metode pengujian perangkat lunak yang dilakukan berdasarkan spesifikasi fungsional tanpa memperhatikan struktur internal atau kode sumber dari perangkat lunak tersebut (Rizal & Nuryasin, 2025). adapun hasil pengujiannya sebagai berikut:

1. Dalam pengujian *blackbox* terhadap informasi yang akan ditampilkan pada *dashboard,* diajukan beberapa persyaratan untuk dilakukan pengujian. Hasil pengujian tersebut tersaji pada tabel 1 di bawah ini.

Tabel 1. Hasil pengujian *blackbox*

No	Persyaratan	Hasil
1	Menampilkan total realisasi penerimaan negara yang terdiri atas tanggal penyetoran, jenis penerimaan, penyetor, dan jumlah penerimaan yang disetorkan	Berhasil
2	Menampilkan total realisasi penerimaan per rentang waktu tertentu	Berhasil
3	Menampilkan total realisasi penerimaan per kode akun penerimaan	Berhasil
4	Menampilkan total realisasi penerimaan per penyetor (satuan kerja)	Berhasil
5	Menampilkan total realisasi penerimaan per jenis mata uang (*currency*)	Berhasil

2. Selanjutnya dilakukan pengujian *dashboard* oleh 8 (delapan) orang pengguna dalam bentuk *User Acceptance Test* (UAT) terhadap visualisasi *data warehouse* menggunakan kuesioner yang terdiri dari 5 (lima) pertanyaan mengenai tampilan dan informasi yang disajikan pada *dashboard* yang telah dibuat. Penilaian dilakukan menggunakan skala: sangat setuju nilai 5, setuju nilai 4, kurang setuju nilai 3, tidak setuju nilai 2, dan sangat tidak setuju nilai 1. Hasil dari kuesioner dapat dilihat pada tabel 2 dan persentase penerimaan sebesar 97% dapat dilihat pada tabel 3.

Tabel 2. Hasil Kuesioner UAT

No	Pertanyaan	SS	S	KS	TS	STS
1	Data telah sesuai dengan kebutuhan	8				
2	Data telah valid dan tidak ada error	4	3	1		
3	Informasi grafik sesuai kebutuhan	7	1			
4	Grafik dapat menampilkan informasi	8				
5	Informasi yang tampil mudah dipahami	8				
	Total	**35**	**4**	**1**		

Tabel 3. Skor Hasil Pengujian UAT

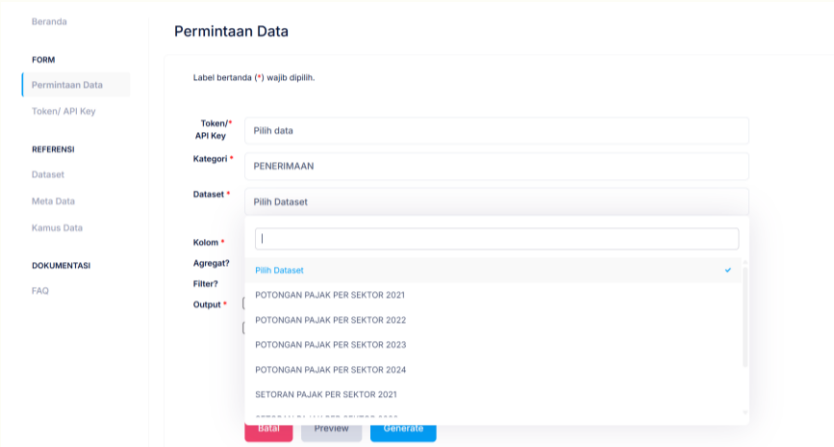
Skor Tertinggi UAT = Jumlah pertanyaan * jumlah responden * skor tertinggi tiap pertanyaan	**5 * 8 * 5 = 200**
Sangat Setuju (SS)	35 * 5 = 175
Setuju (S)	4 * 4 = 16
Kurang Setuju (KS)	1 * 3 = 3
Tidak Setuju (TS)	0 * 2 = 0
Sangat Tidak Setuju (STS)	0 * 1 = 0
Total skor	**194**
Persentase	**(194/200) * 100% = 97%**

Pemanfaatan Data pada Portal Open Data

Setelah proses pengujian keakuratan data yang dilakukan oleh pengguna dinyatakan lulus, selanjutnya data-data tersebut telah dapat dipergunakan oleh para pengguna maupun pihak-pihak lain yang berkepentingan. Proses penarikan data realisasi penerimaan negara yang ada pada data warehouse yang telah dibentuk sebelumnya dapat dilakukan melalui portal *open data treasury* yang telah ada. Pengguna terlebih dahulu login pada portal open data, selanjutnya akan muncul tampilan seperti gambar 11 dibawah ini.

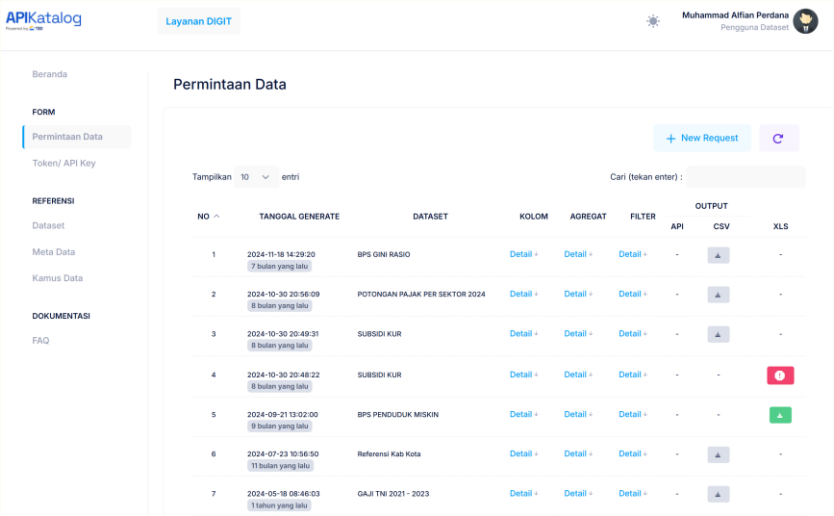
Gambar 11. Tampilan Awal Portal Open Data

Untuk dapat menarik data terkait penerimaan negara, pengguna masuk ke menu Form – Permintaan Data. Tampilan pada menu tersebut dapat dilihat pada gambar 12.



Gambar 12. Tampilan *Form* Permintaan Data

Selanjutnya pengguna tinggal memilih jenis kategori, dataset, dan kolom apa saja yang diperlukan lalu melakukan request ke server data warehouse untuk proses permintaan data. Apabila data telah berhasil terbentuk, maka pengguna dapat mendownload data yang diminta dalam format API/CSV/XLS sesuai dengan permintaannya. Tampilan formnya dapat dilihat pada gambar 13.



Gambar 13. Tampilan *Form* Untuk *download* dataset

Diskusi

Penerapan model data warehouse dalam penelitian ini menunjukkan bahwa pendekatan Kimball Method efektif digunakan untuk membangun sistem pelaporan realisasi penerimaan negara yang bersifat fleksibel, cepat, dan mudah dipahami oleh berbagai pemangku kepentingan. Dimensi-dimensi yang digunakan seperti waktu, akun pendapatan, satuan kerja, wilayah, dan kurs, memungkinkan analisis data yang mendalam dan komprehensif, baik secara vertikal antar organisasi maupun horizontal antar waktu dan wilayah. Dibandingkan dengan kondisi sebelumnya di mana laporan realisasi penerimaan hanya dapat diakses melalui sistem OLTP atau pengolahan manual berbasis Excel, keberadaan sistem data warehouse ini memberikan peningkatan signifikan dalam hal kecepatan akses, akurasi data, dan kemudahan eksplorasi informasi. Proses pelaporan yang sebelumnya membutuhkan waktu berjam-jam untuk agregasi data kini dapat dilakukan secara instan melalui dashboard OLAP yang terhubung langsung ke data mart.

Hasil implementasi juga menunjukkan bahwa sistem mampu mendukung kebutuhan pelaporan yang bersifat periodik (bulanan, triwulanan, dan tahunan), serta mendukung pelacakan data lintas satuan kerja atau berdasarkan jenis akun pendapatan tertentu. Selain itu, pengguna dapat melakukan analisis terhadap dampak fluktuasi nilai tukar terhadap nilai penerimaan dalam mata uang asing, yang sebelumnya sulit dilakukan tanpa proses transformasi data tambahan. Keunggulan lain dari sistem ini adalah kemampuannya untuk berintegrasi secara langsung dengan inisiatif *Treasury Open Data*. Dengan struktur data yang telah disiapkan dalam bentuk dimensional dan terstandarisasi, proses ekspor dan publikasi dataset menjadi lebih mudah dilakukan. Ini mendukung prinsip transparansi fiskal, akuntabilitas publik, dan governance yang berbasis data. Namun demikian, masih terdapat beberapa keterbatasan. Salah satunya adalah ketergantungan pada proses ETL yang perlu dijalankan secara berkala untuk menjaga sinkronisasi data dengan sistem operasional. Selain itu, integrasi data antar sistem sumber masih memerlukan penyesuaian struktural agar lebih optimal dalam proses transformasi dan pemetaan data. Secara keseluruhan, pembangunan data warehouse penerimaan negara dengan pendekatan Kimball tidak hanya mampu meningkatkan efisiensi pelaporan internal, tetapi juga membuka ruang partisipasi publik yang lebih luas melalui data terbuka. Model ini berpotensi untuk direplikasi ke area pengelolaan keuangan negara lainnya, seperti belanja negara, pembiayaan, maupun aset.

KESIMPULAN

Penelitian ini menunjukkan bahwa perancangan data warehouse penerimaan negara dengan pendekatan Kimball *Four-Step Method* mampu menghasilkan sistem pelaporan yang efisien, terstruktur, dan mudah digunakan oleh berbagai pemangku kepentingan. Model dimensional yang dibangun menggunakan skema *star schema* terdiri atas satu tabel fakta dan empat tabel dimensi (waktu, akun, satuan kerja, dan kurs), yang telah diuji melalui proses ETL menggunakan aplikasi Pentaho. Proses ETL berhasil memastikan data dari lebih dari 100.000 baris transaksi penerimaan negara dapat diintegrasikan dan dimuat ke dalam data warehouse tanpa mengganggu sistem operasional utama.

Hasil uji *User Acceptance Test* (UAT) terhadap delapan pengguna menunjukkan tingkat penerimaan sistem sebesar 97%, yang menandakan bahwa data dan visualisasi yang dihasilkan telah sesuai dengan kebutuhan pengguna. Selain itu, dashboard OLAP yang dibangun mampu menampilkan informasi realisasi penerimaan negara berdasarkan waktu, akun, satuan kerja, wilayah, dan mata uang dengan waktu respon yang cepat. Temuan ini memperlihatkan bahwa sistem *data warehouse* yang dikembangkan tidak hanya meningkatkan efisiensi pelaporan internal, tetapi juga membuka peluang integrasi dengan inisiatif *Treasury Open Data*, sebagai bentuk transparansi dan keterbukaan data penerimaan negara kepada publik. Dengan hasil integrasi data yang konsisten dan terstandar, sistem ini dapat menjadi pondasi awal penerapan konsep *data-driven governance* di lingkungan pengelolaan keuangan negara. Namun, untuk pengembangan selanjutnya, disarankan agar sistem ini dioptimalkan melalui otomatisasi proses ETL, penambahan fitur analisis geospasial, integrasi dengan data belanja dan pembiayaan, serta penyediaan API publik agar pengguna eksternal dapat mengakses dataset secara

langsung. Dengan demikian, hasil penelitian ini diharapkan menjadi fondasi bagi pengembangan sistem pelaporan keuangan negara yang lebih modern, transparan, dan berbasis data terintegrasi.

DAFTAR PUSTAKA

Al-Rahman, S. Q. A., Hasan, E. H., & Sagheer, A. M. (2023). Design And Implementation Of The Web (Extract, Transform, Load) Process In Data Warehouse Application. *IAES International Journal of Artificial Intelligence*, *12*(2), 765–775. https://doi.org/10.11591/ijai.v12.i2.pp765-775

Amorim, L. B. V. D, Cavalcanti, G. D. C., & Cruz, R. M. O. (2023). The choice of scaling technique matters for classification performance. *Applied Soft Computing*, *133*. https://doi.org/10.1016/j.asoc.2022.109924

Asmita, M., Henny, H., & Samidi, S. (2023). Data Warehouse Modelling Information Security Log Management in Building a Security Operation Center in Central Government Agencies With Kimball Method. *Jurnal Teknik Informatika (Jutif)*, *4*(4), 987–994. https://doi.org/10.52436/1.jutif.2023.4.4.649

Bahaa, A., Eldemerdash, S. R., & Fahmy, H. (2021). A Systematic Literature Review For Implementing Data Ops In The Data Warehouse Lifecycle During The ETL Phase. *Journal of Computer Science*, *17*(11), 1011–1030. https://doi.org/10.3844/JCSSP.2021.1011.1030

Bakhri, S., & Nuryamin, Y. (2018). Rancangan Data Warehouse Untuk Penunjang Sistem Informasi Eksekutif Pada Yayasan Ummu'l Quro Di Depok. *Jurnal Khatulistiwa Informatika, 4*(1), 149-155.

Balázs, B., Mooney, P., Nováková, E., Bastin, L., & Arsanjani, J. J. (2021). Data quality in citizen science. *The Science of Citizen Science*, 139–157. https://doi.org/10.1007/978-3-030-58278-4_8

Dhaouadi, A., Bousselmi, K., Gammoudi, M. M., Monnet, S., & Hammoudi, S. (2022). Data Warehousing Process Modeling from Classical Approaches to New Trends: Main Features and Comparisons. *Data*, *7*(8), 113. https://doi.org/10.3390/data7080113

Fattah, A. M. M., Ridwan. T. T. Sulistiyowati, N.(2023). Dimensional Data Design for Event Feedback Data Warehouse. *Jurnal Informatika dan Sains*, 6(1), 69-73. https://doi.org/10.31326/jisa.v6i1.1648

Gupta, N., & Jolly, S. (2021). Enhancing Data Quality At ETL Stage Of Data Warehousing. *International Journal of Data Warehousing and Mining*, *17*(1), 74–91. https://doi.org/10.4018/IJDWM.2021010105

Hartono, R., Saputra, H. A., & Triyono, G. (2024). Pemodelan Prediksi Alokasi Pagu Belanja Pegawai dengan Metode Neural Network dan Linear Regression. *Decode: Jurnal Pendidikan Teknologi Informasi*, *4*(3), 865–880. https://doi.org/10.51454/decode.v4i3.708

Kalinina, I., Gozhyj, A., Bidyuk, P., Gozhyi, V., Korobchynskyi, M., & Nadraga, V. (2025). A Systematic Approach to Data Normalization and Standardization in Machine Learning Problems. *Lecture Notes on Data Engineering and Communications Technologies*, *244*, 206–219. https://doi.org/10.1007/978-3-031-88483-2_11

Khan, B., Khan, W., Jan, S., & Chughtai, M. I. (2024). An Overview of ETL Techniques, Tools, Processes and Evaluations in Data Warehousing. *Journal on Big Data*, *6*(1), 1–20. https://doi.org/10.32604/JBD.2023.046223

Kimball, R., & Ross, M. (2013). *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling*. Wiley Computer Publishing.

Koppichetti, R. K. (2022). ETL Strategies for Large-Scale Retail Data Warehouses. *International Journal Research of Leading Publication (IJLRP)*, *3*(8), 1-12. https://doi.org/10.5281/zenodo.15026506

Koppula, R. S. (2021). Optimizing Etl Workflows For Big Data Processing. *International Journal of Core Engineering & Management*, *6(12)*, 118-127.

Mandala, N. R. (2021). ETL in Data Lakes vs. Data Warehouses. *ESP Journal of Engineering & Technology Advancements*, *1*(2), 224–230. https://doi.org/10.56472/25832646/JETA-V1I2P123

Rafique, S., Mushtaq, R., Anum, L., Hamid, K., Iqbal, M. W., & Ruk, S. A. (2024). Analytical Study of OLTP Workload Management in Database Management System. *Journal of Computing & Biomedical Informatics*, *6*(2). https://doi.org/https://doi.org/10.56979/602/2024

Ratu, A., Kusneti, L., & Wijaya, A. (2023). Prototype Data Warehouse Kantor Penilai Publik XYZ Dengan Metode Nine-Step Kimball. *Journal Of Informatics And Busisnes*, *1*(3), 89–96.

Rizal, D. F., & Nuryasin, I. (2025). Implementasi Blackbox Testing Pada Sistem Informasi Sirkulasi Perpustakaan Berbasis Website dengan Teknik Equivalence Partitioning. *Decode: Jurnal Pendidikan Teknologi Informasi*, *5*(1), 65–78. https://doi.org/10.51454/decode.v5i1.1052

Sethi, S. (2021). Designing ETL Pipelines for Scalable Data Processing. *IJIRMPS*, *9*(6), 1-10. https://doi.org/10.5281/ZENODO.14945154

Smith, J., & Elshnoudy, I. A. (2023). A Comparative Analysis of Data Warehouse Design Methodologies for Enterprise Big Data and Analytics. *Emerging Trends in Machine Intelligence and Big Data*, 15(10).

Tamma, R. O., Syahidin, Y., & Yunengsih, Y. (2025). Perancangan Sistem Informasi Pelaporan 10 Besar Tindakan Operasi Rawat Inap Menggunakan Metode Agile. *Decode: Jurnal Pendidikan Teknologi Informasi*, *5*(1), 106–120. https://doi.org/10.51454/decode.v5i1.1070

Triaji, B., Subagyo, A. A., & Rifai, M. A. (2024). Development of a Higher Education Data Warehouse Using the Data Vault 2.0 Method. *Sinkron*, *8*(4), 2591–2602. https://doi.org/10.33395/sinkron.v8i4.14215

Ygnacio, L. A. V., Retuerto, M. G., & Andrade-Arenas, L. (2023). Mobile Application With Business Intelligence To Optimize The Control Process Of Tourist Agencies. *Indonesian Journal of Electrical Engineering and Computer Science*, *29*(3), 1708–1718. https://doi.org/10.11591/ijeecs.v29.i3.pp1708-1718