Data Analytics for Business 2024

Project Based Learning

Study Case

[Paper ID]

CS 04 - 5:

1. Muhammad Fahmi Hutomo KM-CS04151
2. Marsyanda Nur Zahra KM-CS04340
3. Syahirotul Ambar Maulidiyah KM-CS04067

Kata Pengantar

Puji syukur kami panjatkan kepada Tuhan Yang Maha Esa atas limpahan rahmat dan karunia-Nya, sehingga kami dapat menyelesaikan laporan Project-Based Learning Bitlabs Academy yang menggunakan studi kasus dari Paper ID ini dengan baik. Laporan ini merupakan salah satu bentuk implementasi pembelajaran berbasis proyek yang bertujuan untuk mengasah kemampuan analitis dan pemecahan masalah secara terstruktur.

Kami menyampaikan terima kasih sebesar-besarnya kepada pihak Bitlabs Academy yang telah memberikan bimbingan dan fasilitas dalam proses pembelajaran ini. Ucapan terima kasih juga kami sampaikan kepada para mentor dan instruktur yang telah memberikan arahan serta dukungan selama pelaksanaan proyek ini. Tak lupa, kami haturkan rasa terima kasih kepada rekan-rekan peserta yang telah bekerja sama dan berbagi pengalaman, sehingga proses pembelajaran menjadi lebih bermakna.

Melalui proyek ini, kami berharap dapat mengembangkan pemahaman yang lebih mendalam terhadap analisis studi kasus, mengaplikasikan teori-teori yang relevan, serta memperkuat kemampuan kami dalam menghadapi tantangan dunia profesional.

Kami menyadari bahwa laporan ini masih jauh dari sempurna. Oleh karena itu, kami sangat terbuka terhadap saran dan kritik yang membangun untuk pengembangan ke depan. Semoga laporan ini dapat bermanfaat bagi para pembaca dan menjadi referensi yang berguna bagi kegiatan serupa di masa yang akan datang.

Surabaya, 08 Desember 2024

Kelompok 05 - CS 04

Daftar Isi

Profil Perusahaan

Paper.id adalah platform penyedia jasa penerbitan faktur dan pembayaran digital yang dirancang khusus untuk bisnis di Indonesia. Platform ini memungkinkan pengguna untuk membuat faktur tanpa batas dengan mudah, mengotomatiskan pengingat pembayaran, dan mengintegrasikan berbagai metode pembayaran seperti kartu kredit dan kode QR. Dengan fokus pada penyederhanaan transaksi keuangan, Paper.id memberdayakan bisnis untuk mengelola proses faktur mereka dengan lebih efisien dan akurat.

Sejak diluncurkan, Paper.id telah mencapai pertumbuhan yang luar biasa, dengan lebih dari 600.000 pengguna yang memanfaatkan layanannya. Platform ini telah memfasilitasi pemrosesan lebih dari 8 juta faktur, dengan total nilai pembayaran digital melebihi Rp 10 triliun. Dampak signifikan ini menunjukkan bahwa Paper.id berperan dalam mentransformasi pengelolaan keuangan bagi usaha kecil dan menengah di Indonesia.

Dengan antarmuka yang ramah pengguna dan fitur yang tangguh, Paper.id menjadi alat penting bagi bisnis yang ingin mengoptimalkan operasional keuangannya. Perusahaan ini terus berinovasi dan beradaptasi dengan kebutuhan pengguna yang terus berkembang, memastikan posisinya sebagai sumber daya berharga dalam ekonomi digital Indonesia. Untuk informasi lebih lanjut, kunjungi situs web mereka di www.paper.id.

Gambar 1. *Website Paper.id*

Latar Belakang Masalah

Baru-baru ini, perusahaan mengalami peningkatan transaksi fraud yang berdampak negatif pada pendapatan dan kepercayaan pelanggan. Tugas Anda adalah melakukan analisis lanjutan terhadap fraud untuk mengidentifikasi pola perilaku yang mencurigakan serta mendeteksi hubungan yang mencurigakan antara pembeli dan penjual. Anda juga akan menggunakan query SQL tingkat lanjut dan analisis mendalam untuk memberikan wawasan yang dapat ditindaklanjuti serta rekomendasi guna mengurangi risiko fraud.

Data dan Sumber Data

Adapun data pada perusahaan sebagai berikut:

1. **Digital Payment Transaction Data:**
   Columns: dpt_id, buyer_id, seller_id, transaction_amount, payment_method_name, payment_provider_name, transaction_created_datetime, transaction_updated_datetime, dpt_promotion_id
2. **Digital Payment Request Data:**
   Columns: dpt_id, total_fee_amount, document_type_name
3. **Promotion Data:**
   Columns: dpt_promotion_id, promotion_code, promotion_name, transaction_promo_cashback_amount
4. **Company Data:**
   Columns: company_id, company_kyc_status_name, company_kyb_status_name, company_type_group, company_phone_verified_flag, company_email_verified_flag, user_fraud_flag, testing_account_flag, blacklist_account_flag, company_registered_datetime

Alur Pengerjaan Study Case

A. Tools yang digunakan dalam proses menganalisis data

Dalam proses analisis data, berbagai alat digunakan untuk menangani tugas spesifik yang berbeda. BPMN dengan Draw.io digunakan untuk memodelkan proses bisnis secara visual, memudahkan pemangku kepentingan dalam memahami alur dan interaksi antar proses. DBeaver yang terhubung dengan PostgreSQL dipilih karena kemampuannya dalam mengelola dan mengeksekusi query SQL yang optimal untuk database ini, serta menyediakan antarmuka pengguna yang memudahkan manipulasi dan visualisasi data. Python digunakan karena kekuatannya dalam mengolah dan menganalisis data besar melalui pustaka seperti Pandas dan NumPy, serta kemampuannya untuk terintegrasi dengan berbagai alat dan database. Sementara itu, Tableau digunakan untuk membuat visualisasi data yang interaktif dan mudah dipahami, memungkinkan pemangku kepentingan mengeksplorasi data secara real-time. Kombinasi alat-alat ini memungkinkan alur kerja yang efisien dalam memproses, menganalisis, dan menyajikan data secara menyeluruh.

B. Proses Analisis Data

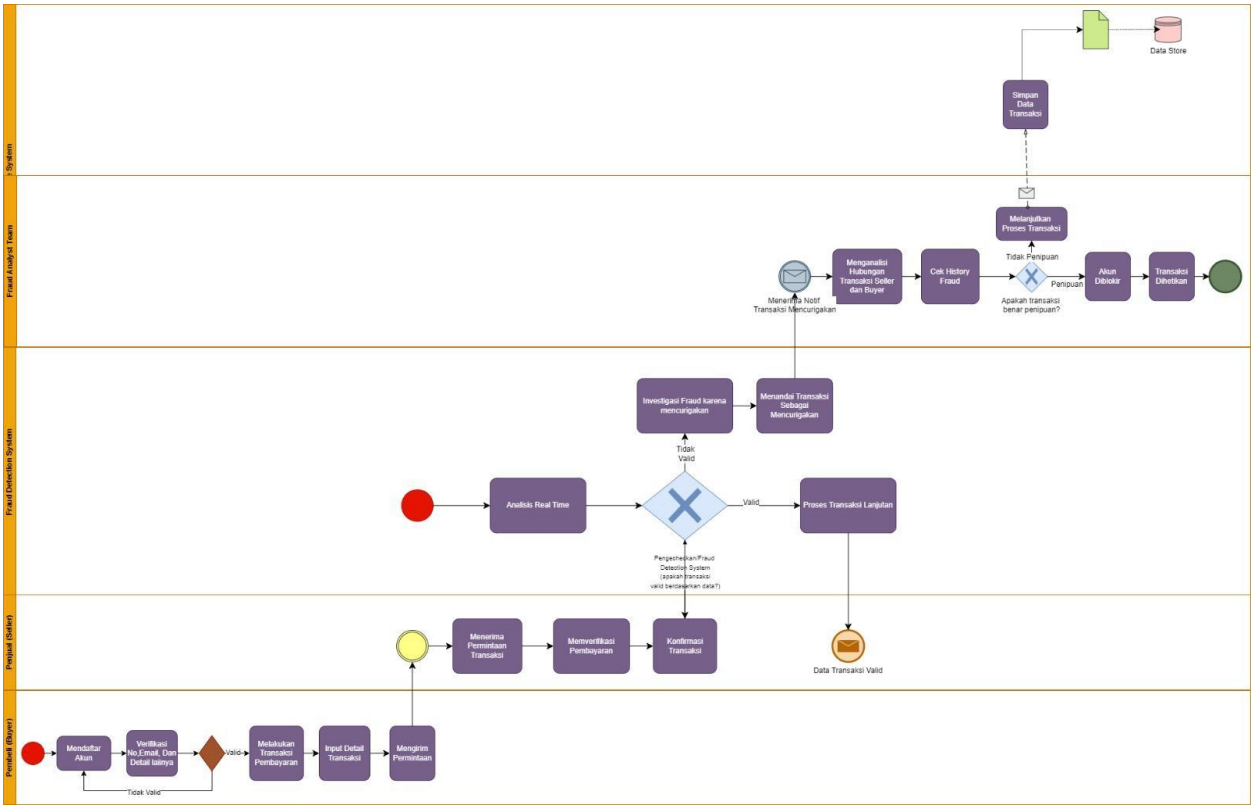
Proses analisis dimulai dengan data cleaning, di mana data yang hilang, outlier, atau inkonsistensi seperti nilai yang tidak sesuai pada kolom terkait fraud (misalnya, user_fraud_flag) diidentifikasi dan diperbaiki. Setelah itu, dilakukan feature engineering untuk membuat fitur turunan seperti Buyer-Seller Relationship Score dan Transaction Frequency Metrics untuk mendeteksi anomali dalam interaksi pembeli-penjual atau pola transaksi yang tidak biasa. Selanjutnya, data yang telah dibersihkan dan ditransformasi melalui scaling dan normalization (misalnya, normalisasi jumlah transaksi) dianalisis lebih lanjut melalui Exploratory Data Analysis (EDA), yang melibatkan analisis distribusi transaksi, frekuensi pasangan pembeli-penjual, dan penggunaan promosi. Visualisasi data seperti grafik transaksi dari waktu ke waktu dan jaringan pembeli-penjual digunakan untuk mengidentifikasi pola aktivitas mencurigakan. Pada tahap berikutnya, analisis lebih lanjut dilakukan menggunakan SQL queries untuk mendeteksi anomali transaksi, kolusi pembeli-

penjual, serta penyalahgunaan promosi, dengan menggunakan query SQL yang kompleks dan stored procedures untuk otomatisasi laporan fraud. Social Network Analysis kemudian digunakan untuk mengidentifikasi hubungan antara pembeli dan penjual yang terlibat dalam fraud melalui visualisasi jaringan dan cohort analysis untuk melacak perilaku transaksi yang berulang. Terakhir, hasil analisis tersebut disajikan dalam bentuk Tableau dashboards untuk memantau tren fraud secara real-time, dengan kemampuan untuk menelusuri dan memfilter data untuk mendalami hubungan yang mencurigakan dan memberikan rekomendasi tindakan pencegahan.

BAB I

[Business] Business Analysis

Proses deteksi penipuan melalui pendekatan Business Process Model Notation (*BPMN*) dirancang untuk memetakan alur kerja secara menyeluruh, mulai dari inisiasi transaksi hingga penyelesaian kasus penipuan. Model ini mencakup langkah-langkah kunci seperti inisiasi transaksi, deteksi penipuan secara real-time, penandaan pengguna mencurigakan, hingga resolusi kasus, dengan tujuan untuk meningkatkan keandalan sistem pembayaran. Dengan visualisasi BPMN, perusahaan dapat lebih mudah mengidentifikasi area yang perlu dioptimalkan, sehingga mendukung upaya menjaga kepercayaan pelanggan dan keberlanjutan operasional di ekosistem digital. Berikut adalah penjelasan rinci dari model BPMN yang mengilustrasikan proses deteksi penipuan (fraud flow analysis) di Paper.id.



Gambar 1.1 *BPMN fraud flow analysis*

Berdasarkan **Gambar 1.1** *BPMN fraud flow analysis* diatas menunjukan bisnis proses sebagai berikut *:*

1. **Proses Pendaftaran**: Proses ini merupakan langkah awal yang harus dilalui oleh pengguna sebelum dapat melakukan transaksi di platform. Pada tahap ini, pengguna (baik pembeli maupun penjual) diwajibkan untuk:
    - Melakukan Pendaftaran Akun: Pengguna mengisi data pribadi seperti nama, alamat email, dan nomor telepon.
    - Verifikasi Email dan Nomor Telepon: Sistem mengirimkan kode OTP (One-Time Password) atau tautan verifikasi ke email dan nomor telepon pengguna. Langkah ini dilakukan untuk memastikan bahwa data yang diberikan valid dan dapat dihubungi.
2. **Proses Inisiasi Transaksi** : Setelah pendaftaran dan verifikasi selesai, pengguna dapat memulai transaksi. Proses ini melibatkan dua pihak utama, yaitu pembeli (Buyer) dan penjual (Seller):
    - Pembeli (Buyer):

      Melakukan Transaksi Pembayaran: Pembeli memulai dengan menginput detail pembayaran seperti jumlah, metode pembayaran, dan informasi terkait lainnya.

      Mengirim Permintaan Transaksi: Setelah detail transaksi lengkap, permintaan dikirimkan kepada penjual untuk diproses lebih lanjut.
    - Penjual (Seller):

      Menerima Permintaan Transaksi: Penjual menerima notifikasi bahwa pembeli telah mengirimkan detail transaksi.

      Memverifikasi Pembayaran: Penjual memastikan bahwa pembayaran sudah dilakukan dan valid.

      Mengonfirmasi Transaksi: Setelah pembayaran terverifikasi, transaksi dikonfirmasi untuk diproses lebih lanjut.
3. **Proses Analisis Real-Time oleh Sistem Deteksi Fraud** : Tahap ini menjadi inti dari sistem keamanan untuk mendeteksi potensi kecurangan dalam transaksi. Prosesnya meliputi:
    - Penerimaan Data Transaksi: Sistem menerima data transaksi dari penjual, termasuk informasi seperti jumlah transaksi, identitas pembeli, dan riwayat transaksi.

- Analisis Real-Time: Sistem menggunakan algoritma dan model analitik untuk memeriksa transaksi dalam waktu nyata. Indikator seperti pola transaksi mencurigakan, hubungan antara penjual dan pembeli. Dimana jika transaksi valid, ditandai sebagai aman dan jika transaksi mencurigakan, sistem memberikan notifikasi kepada tim analisis fraud untuk tindakan lebih lanjut.

4. **Investigasi dan Tindak Lanjut Fraud** : Ketika sebuah transaksi ditandai sebagai mencurigakan, tim analisis fraud akan mengambil alih untuk melakukan investigasi lebih mendalam. Proses ini melibatkan langkah-langkah seperti:
    - Investigasi Manual oleh Tim Analisis Fraud: Tim menganalisis data transaksi yang mencurigakan menggunakan metode analitik lanjutan dan hubungan antara pembeli dan penjual diperiksa untuk mengidentifikasi adanya pola fraud, seperti transaksi fiktif atau kolusi.
    - Penandaan Fraud: Jika ditemukan bukti kecurangan, transaksi akan ditandai sebagai fraud.
    - Keputusan Akhir: Jika transaksi dinyatakan valid, proses dilanjutkan seperti biasa. Jika terbukti fraud, transaksi dihentikan, akun diblokir, dan pelaku dicatat dalam sistem blacklist.

Dari analisis fraud flow yang diterapkan dalam bisnis proses Paper.id, terlihat bahwa sistem telah dirancang untuk mendeteksi dan menangani potensi kecurangan. Proses pendaftaran dengan verifikasi email untuk memastikan hanya pengguna valid yang dapat mengakses platform. Sistem deteksi fraud real-time yang mengandalkan analitik data untuk mengidentifikasi pola mencurigakan. Proses ini diperkuat oleh investigasi manual oleh tim analisis fraud untuk memastikan keakuratan hasil deteksi. Penandaan akun mencurigakan dan pemblokiran otomatis menciptakan lapisan perlindungan tambahan, sementara pencatatan pelaku dalam sistem blacklist memperkuat keamanan ekosistem. Pendekatan ini tidak hanya meminimalkan risiko penipuan tetapi juga menjaga kepercayaan pengguna terhadap layanan Paper.id.

BAB II

Data Analytics using Python Programming

2.1 Data Cleaning

Data cleaning dilakukan menggunakan Python dan library Pandas, dengan seluruh proses dilakukan di Google Colab. Dataset yang digunakan terdiri dari empat tabel, yang masing-masing dimuat dari Google Drive ke dalam lingkungan Google Colab. Proses pembersihan data bertujuan untuk meningkatkan kualitas data untuk menangani nilai duplikat, nilai hilang dan outliers. Sehingga siap digunakan dalam analisis lebih lanjut. Beberapa langkah utama yang dilakukan selama proses ini sebagai berikut:

1. Import data
2. Baca Info data
3. Hapus data duplikat.
4. Handling missing values:
    - dpt_promotion_id kosong di dataframe transaction dianggap transaksi tidak menggunakan promo, diisi dengan ’ no promotion’.
    - baris dengan dpt_promotion_id kosong di dataframe promotion dihapus karena hanya satu baris dengan promotion_code dan promotion_name yang juga kosong disertai cashback 0.
    - promotion_code dan promotion_name yang kosong akan diisi dengan 'no promotion' jika promotion['dpt_promotion_id] == 'no promotion'.
    - promotion_code dan promotion_name yang kosong akan diisi dengan 'unknown' jika promotion['dpt_promotion_id] != 'no promotion'.
    - company_type_group kosong pada dataframe user diisi dengan ‘unknown’.
1. Outliers seperti transaction_amount yang terlampau besar nilainya tetap dipertahankan untuk dianalisis lebih lanjut

   Selain langkah-langkah utama dalam proses pembersihan data, tahap ini juga mencakup identifikasi terhadap data yang hilang atau tidak konsisten, khususnya pada kolom-kolom yang berkaitan dengan transaksi fraud. Langkah ini bertujuan untuk mengungkap pola atau anomali yang dapat memberikan gambaran lebih jelas terkait aktivitas mencurigakan di

dalam transaksi. Berdasarkan analisis yang dilakukan, beberapa temuan penting berhasil diidentifikasi, yaitu:

1. User Tidak Terdaftar Sebagai Seller
   Terdapat 318 user yang tidak terdaftar namun tercatat melakukan transaksi sebagai seller.
2. User Fraud Tidak Di-blacklist
   Sebanyak 660 user yang ditandai fraud tidak dimasukkan ke dalam daftar blacklist.
3. User Fraud dengan Status KYC Tidak Dibekukan
   Terdapat 11 user yang terdeteksi fraud tetapi status KYC (*Know Your Customer*)nya tidak dibekukan.
4. Transaksi dengan Identitas Tidak Tervalidasi
   Lebih dari 50% dari total transaksi melibatkan seller yang memiliki identitas belum tervalidasi, termasuk nomor telepon dan email yang tidak diverifikasi, serta status KYC (*Know Your Customer*) dan KYB (*Know Your Business*) yang belum tervalidasi, dan akun tersebut bukan termasuk akun uji coba (testing account)
5. Self-Transaction (Transaksi Diri Sendiri)
   Sebanyak 46% dari total transaksi yang terjadi adalah self-transaction, di mana buyer dan seller merupakan entitas (company) yang sama.
   - Dari jumlah tersebut, 24 buyer melakukan self-transaction lebih dari 100 kali, yang memperkuat indikasi pola perilaku yang tidak wajar.
   - Namun, self-transaction yang menggunakan promo tercatat hanya sebesar 0,19% dari total transaksi, sehingga dampak penyalahgunaan promo relatif kecil.

2.2 Feature Engineering

Pada tahap ini dibuat beberapa fitur turunan atau kolom baru yang dapat digunakan untuk melakukan analisis lebih lanjut.

1. Pertama, buyer-seller relationship_score. Relationship_score 1(sangat lemah), 2(lemah), 3(sedang), 4(kuat), dan 5(sangat kuat).
2. Kedua, transaction frequency metrics. Ada tiga metrik yang dibuat, yaitu burst_activity, unusual_gap, dan burst_amount. Ketiganya hanya bernilai 1 (ada) atau 0 (tidak ada).
3. Ketiga, promotion exploitation indicator. Metrik is_promotion_exploitation bernilai 1 (ada) atau 0 (tidak ada).

2.3 Scaling and Normalization

Pada transaction_amount dilakukan tranformasi logaritmik sebelum normalisasi karena distribusi datanya memiliki nilai skewness 139. Sedangkan transaction_created_datetime setelah diubah ke detik distribusi datanya mendekati normal dengan skewness hanya -0,24 sehingga tidak

dilakukan transformasi. Tahap terakhir dari proses ini adalah melakukan normalisasi transaction_amount yang sudah ditransformasi dan transaction_created_datetime dalam detik menggunakan MinMaxScaler().

BAB III
Exploratory Data Analysis (EDA)

3.1 Exploratory Data Analysis

Pada tahap ini dilakukan analisis eksplorasi data (Exploratory Data Analysis/EDA) secara mendalam untuk mendeteksi pola yang berkaitan dengan perilaku fraud. Analisis mencakup

beberapa aspek utama, seperti distribusi nilai transaksi, frekuensi pasangan buyer-seller, serta penggunaan promo. Selain itu, investigasi lebih lanjut dilakukan untuk mengidentifikasi pola pada transaksi yang berulang atau nilai transaksi yang secara abnormal tinggi. Hasil dari analisis ini memberikan beberapa informasi penting sebagai berikut:

3.1.1 Analisis Distribusi Transaction_amount

Analisis ini bertujuan untuk memahami distribusi nilai transaksi, termasuk rentang transaksi yang mencakup nilai minimum hingga maksimum. Distribusi yang tidak merata dengan sejumlah kecil transaksi bernilai tinggi menjadi salah satu perhatian utama. Sehingga tahap ini diperoleh beberapa informasi di antaranya:

- Nilai transaksi minimum sebesar 0 dan maksimum sekitar 20 Milyar.
- Nilai transaksi 0 terjadi sebanyak 26 transaksi.
- Nilai transaksi lebih dari 0 hingga 1K terjadi sebanyak 16 transaksi atau 0.032% dari keseluruhan transaksi.
- Nilai transaksi lebih dari 1K hingga 10K terjadi sebanyak 538 transaksi atau 1.076% dari keseluruhan transaksi.
- Nilai transaksi lebih dari 10K hingga 100K terjadi sebanyak 12454 transaksi atau 24.908% dari keseluruhan transaksi.
- Nilai transaksi lebih dari 100K hingga 1M terjadi sebanyak 8726 transaksi atau 17.452% dari keseluruhan transaksi.
- Nilai transaksi lebih dari 1M hingga 10M terjadi sebanyak 13380 transaksi atau 26.760% dari keseluruhan transaksi.
- Lebih dari 10M hingga 100M: 13228 transaksi 26.456% dari keseluruhan transaksi
- Lebih dari 100M hingga 1B: 1628 transaksi 3.256% dari keseluruhan transaksi
- Lebih dari 1B hingga 10B: 3 transaksi 0.006% dari keseluruhan transaksi
- Lebih dari 10B hingga 30B: 1 transaksi 0.002% dari keseluruhan transaksi

3.1.2 Analisis Frekuensi Transaksi Pasangan Buyer-Seller

Analisis ini fokus pada frekuensi transaksi berulang antara pasangan buyer-seller. Frekuensi transaksi yang sangat tinggi dari beberapa pasangan buyer-seller mengindikasikan

adanya potensi pola transaksi yang mencurigakan.Dimana dapat diperoleh beberapa informasi diantaranya:

- Terdapat 10354 pasangan buyer-seller pada data transaction.
- Transaksi berulang lebih dari 0 hingga 5 kali dilakukan oleh 8398 pasangan (81.11% dari total pasangan buyer-seller).
- Transaksi berulang lebih dari 5 hingga 10 kali dilakukan oleh 1017 pasangan (9.82% dari total pasangan buyer-seller).
- Transaksi berulang lebih dari 10 hingga 20 kali dilakukan oleh 587 pasangan (5.67% dari total pasangan buyer-seller).
- Transaksi berulang lebih dari 20 hingga 30 kali dilakukan oleh 158 pasangan (1.53% dari total pasangan buyer-seller)
- Transaksi berulang lebih dari 30 hingga 40 kali dilakukan oleh 60 pasangan (0.58% dari total pasangan buyer-seller)
- Transaksi berulang lebih dari 40 hingga 50 kali dilakukan oleh 39 pasangan (0.38% dari total pasangan buyer-seller)
- Transaksi berulang lebih dari 50 hingga 100 kali dilakukan oleh 63 pasangan (0.61% dari total pasangan buyer-seller)
- Transaksi berulang lebih dari 100 hingga 150 kali dilakukan oleh 15 pasangan (0.14% dari total pasangan buyer-seller)
- Transaksi berulang lebih dari 150 hingga 200 kali dilakukan oleh 7 pasangan (0.07% dari total pasangan buyer-seller)
- Transaksi berulang lebih dari 200 hingga 250 kali dilakukan oleh 5 pasangan (0.05% dari total pasangan buyer-seller)
- Transaksi berulang lebih dari 250 hingga 300 kali dilakukan oleh 3 pasangan (0.03% dari total pasangan buyer-seller)
- Transaksi berulang lebih dari 300 hingga 350 kali dilakukan oleh 1 pasangan (0.01% dari total pasangan buyer-seller)
- Transaksi berulang lebih dari 350 hingga 1250 kali dilakukan oleh 0 pasangan (0.00% dari total pasangan buyer-seller)
- Transaksi berulang lebih dari 1250 hingga 1300 kali dilakukan oleh 1 pasangan (0.01% dari total pasangan buyer-seller)

3.1.3 Analisis Penggunaan Promo

Penggunaan promo dalam transaksi dievaluasi untuk mengidentifikasi pola dan potensi penyalahgunaan. Top 3 kode promo paling sering digunakan juga diidentifikasi dalam analisis ini. Dimana dapat diperoleh beberapa informasi diantaranya:

- Transaksi yang menggunakan promo terjadi sebanyak 937 kali atau hanya 1,87% dari total transaksi.
- Top 3 promo paling banyak digunakan yaitu, promotion-219036467 sebanyak 243 kali, promotion-214984720 sebanyak 88 kali, dan promotion-188676794 sebanyak 69 kali.
- Buyer menggunakan promo sebanyak 8 kali: 1 (0.14%)
- Buyer menggunakan promo sebanyak 6 kali: 2 (0.27%)
- Buyer menggunakan promo sebanyak 5 kali: 2 (0.27%)
- Buyer menggunakan promo sebanyak 4 kali: 7 (0.96%)
- Buyer menggunakan promo sebanyak 3 kali: 26 (3.56%)
- Buyer menggunakan promo sebanyak 2 kali: 109 (14.93%)
- Buyer menggunakan promo sebanyak 1 kali: 583 (79.86%)
- Buyer menggunakan satu kode promo sebanyak 5 kali: 2 (0.23%)
- Buyer menggunakan satu kode promo sebanyak 4 kali: 1 (0.11%)
- Buyer menggunakan satu kode promo sebanyak 3 kali: 4 (0.46%)
- Buyer menggunakan satu kode promo sebanyak 2 kali: 44 (5.03%)
- Buyer menggunakan satu kode promo sebanyak 1 kali: 823 (94.16%)

3.1.4 Analisis Pola Transaksi Berulang, Transaksi dengan Nilai Tinggi, dan Self-transaction

Analisis ini mendalami pola transaksi berulang, termasuk transaksi dengan nilai tinggi dan self-transaction, di mana buyer dan seller merupakan entitas yang sama. Beberapa rentang transaksi menunjukkan proporsi yang signifikan dari self-transaction, mengindikasikan potensi penyalahgunaan sistem. Dimana tahap ini dapat diperoleh beberapa informasi diantaranya:

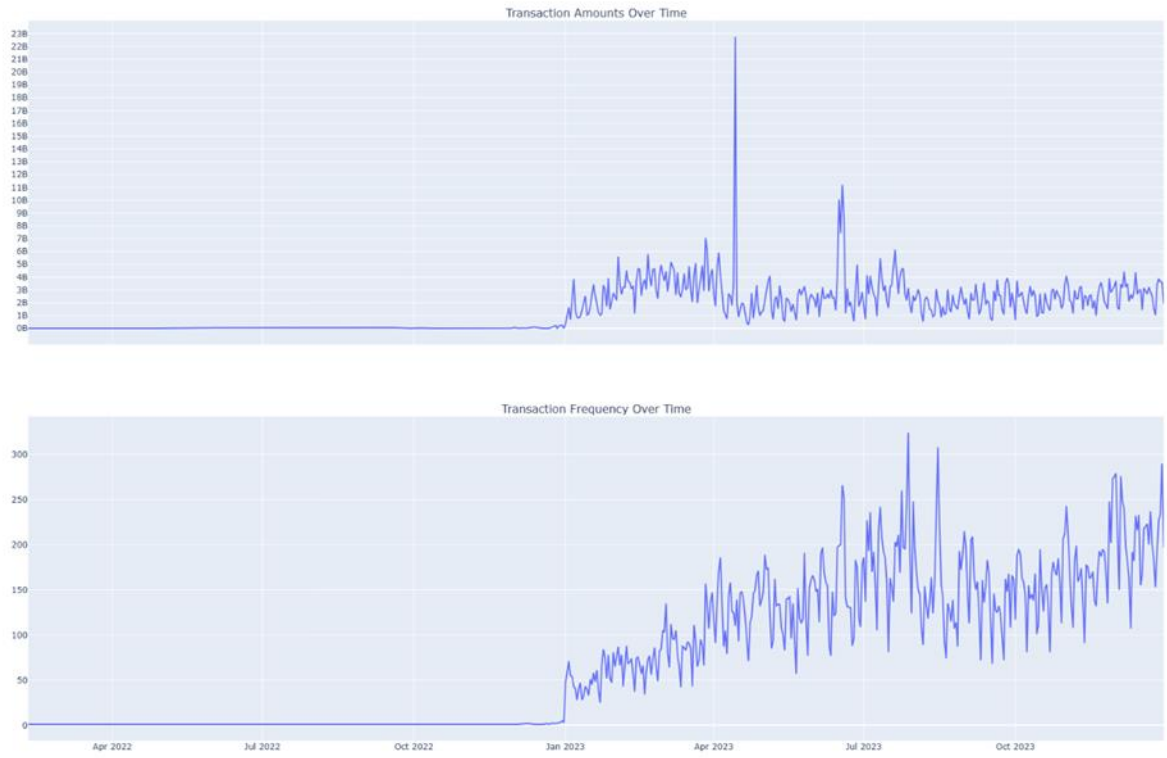
- Terdapat 32 pasangan buyer-seller yang melakukan transaksi berulang lebih dari 100 kali dan 24 di antaranya (75%) nya merupakan self-transaction.
- Nilai transaksi 0 terjadi sebanyak 26 transaksi, 1 transaksi merupakan self-transaction, (3.85%) dari keseluruhan transaksi di rentang ini.

- Nilai transaksi lebih dari 0 hingga 1K terjadi sebanyak 16 transaksi, 8 transaksi merupakan self-transaction, (50.00%) dari keseluruhan transaksi di rentang ini.
- Nilai transaksi lebih dari 1K hingga 10K terjadi sebanyak 538 transaksi, 500 transaksi merupakan self-transaction, (92.94%) dari keseluruhan transaksi di rentang ini.
- Nilai transaksi lebih dari 10K hingga 100K terjadi sebanyak 12454 transaksi, 6643 transaksi merupakan self-transaction, (53.34%) dari keseluruhan transaksi di rentang ini.
- Nilai transaksi lebih dari 100K hingga 1M terjadi sebanyak 8726 transaksi, 5749 transaksi merupakan self-transaction, (65.88%) dari keseluruhan transaksi di rentang ini.
- Nilai transaksi lebih dari 1M hingga 10M terjadi sebanyak 13380 transaksi, 6874 transaksi merupakan self-transaction, (51.38%) dari keseluruhan transaksi di rentang ini.
- Nilai transaksi lebih dari 10M hingga 100M terjadi sebanyak 13228 transaksi, 3364 transaksi merupakan self-transaction, (25.43%) dari keseluruhan transaksi di rentang ini.
- Nilai transaksi lebih dari 100M hingga 1B terjadi sebanyak 1628 transaksi, 85 transaksi merupakan self-transaction, (5.22%) dari keseluruhan transaksi di rentang ini.
- Nilai transaksi lebih dari 1B hingga 10B terjadi sebanyak 3 transaksi, 3 transaksi merupakan self-transaction, (100.00%) dari keseluruhan transaksi di rentang ini.
- Nilai transaksi lebih dari 10B hingga 30B terjadi sebanyak 1 transaksi, 1 transaksi merupakan self-transaction, (100.00%) dari keseluruhan transaksi di rentang ini.

3.2 Visualization

Visualisasi data dilakukan untuk memberikan gambaran yang lebih jelas terkait pola transaksi, hubungan antar pengguna, serta penggunaan promo. Dengan pendekatan ini, potensi anomali atau indikasi fraud dapat lebih mudah diidentifikasi secara intuitif. Beberapa fokus utama dalam visualisasi data dirangkum sebagai berikut

3.2.1 Visualisasi Total Transaksi Harian dari Waktu ke Waktu



Gambar 3.2.1 *Visualisasi Total Transaksi Harian dari Waktu ke Waktu*

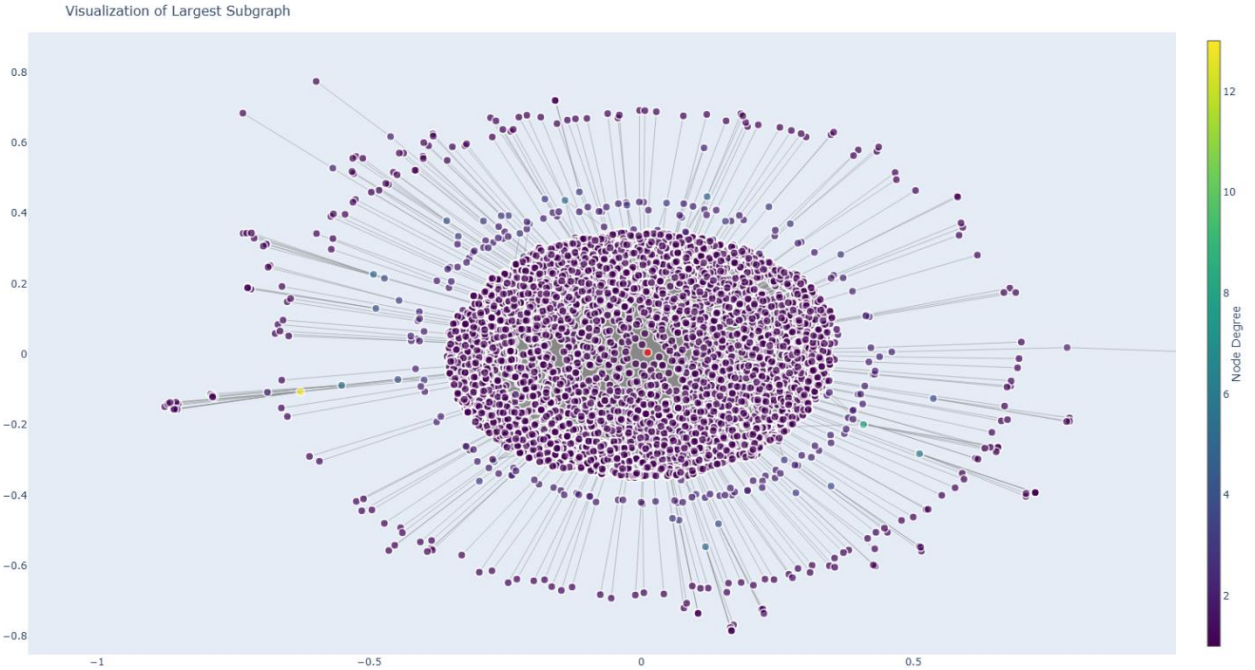
Pada kuarter kedua tahun 2023 terjadi dua kali lonjakan total nilai transaksi harian.

- Lonjakan pertama pada tanggal 14 April terjadi karena ada satu transaksi dengan nilai sangat tinggi, sekitar 20 Milyar. Sebenarnya separuh transaksi pada tanggal tersebut hanya berada di kisaran 5 hingga 6 digit dan total transaksi yang terjadi hanya 110 transaksi.
- Lonjakan kedua pada tanggal 18 Juni tidak sebesar lonjakan pertama dan lebih disebabkan karena separuh transaksi pada hari tersebut bernilai sedikitnya 8 digit dan transaksi yang terjadi pada tanggal tersebut sebanyak 266 kali.

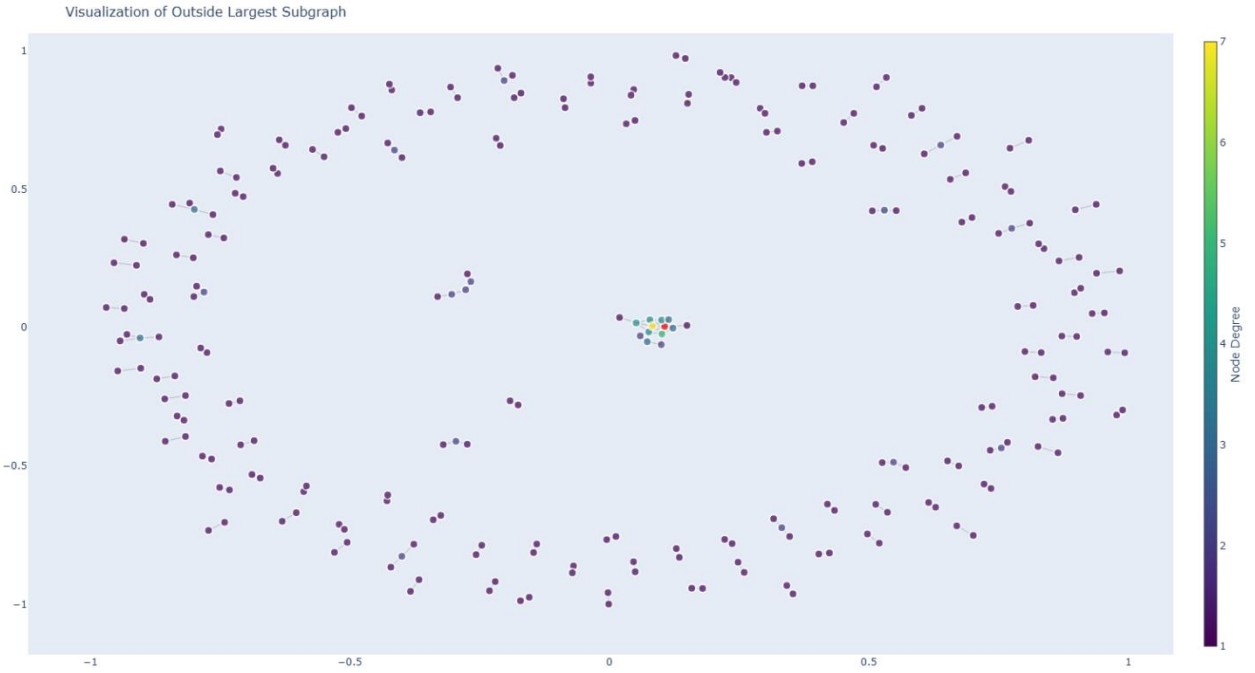
Secara umum, frekuensi transaksi tidak memengaruhi total nilai transaksi harian.

3.2.2 Graf Jaringan Buyer-Seller

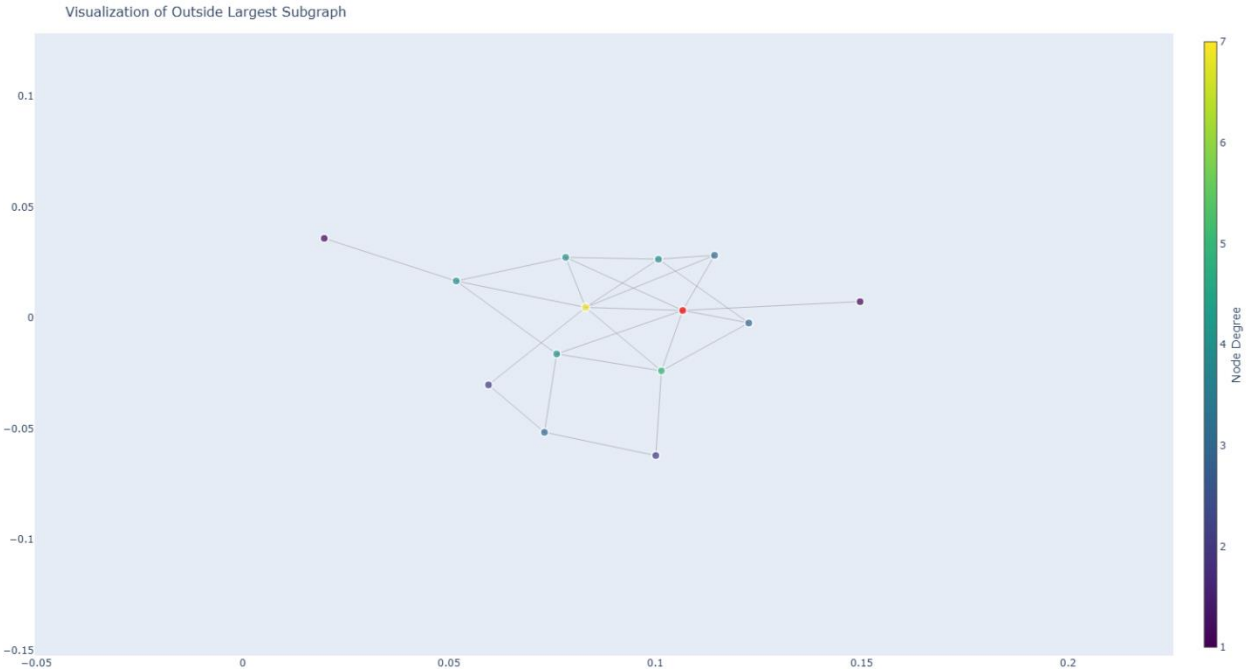
Transaksi buyer-seller pada data transaction membentuk banyak graf kecil yang hanya terdiri dari beberapa node,dengan total node dari semua grafik kecil ini hanya 252 node. Di samping itu, ada satu graf yang sangat besar dengan jumlah node 6591 node, dengan node pusat terhubung ke 6356 node lain. Visualisasinya seperti yang gambar di bawah ini.



Node pusat berwarna merah berada ditengah, terhubung dengan node lainnya yang sebagian besar hanya memiliki satu atau dua edge, tetapi ada juga beberapa node yang memiliki edge lebih dari 5 hingga 13. Pada bagian selanjutnya akan dilakukan analisis jaringan sosial untuk mengetahui perilaku transaksi fraud berdasarkan graf ini.

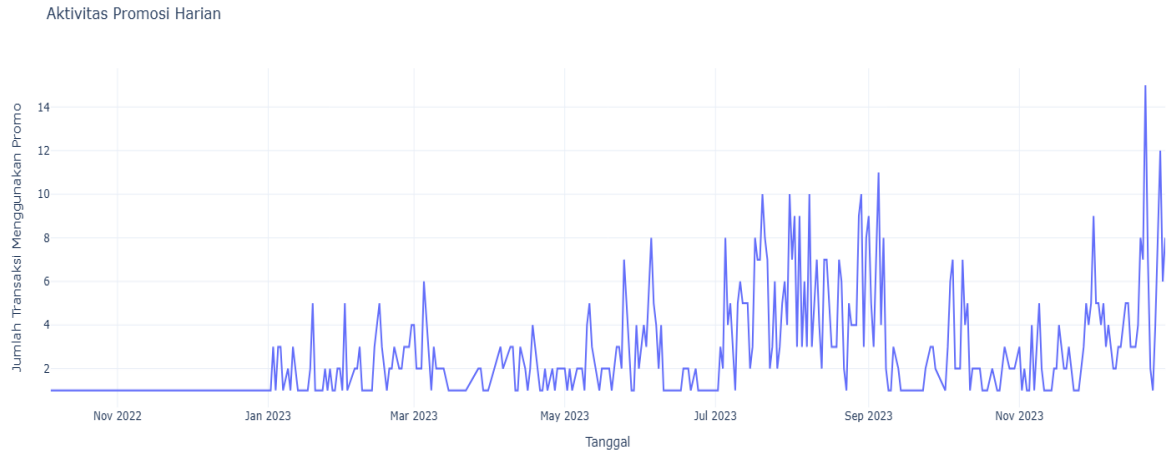


Kebanyakan graf-graf kecil ini hanya terdiri dari 2 sampai 3 node, tetapi ada satu graf yang cukup besar jika dibanding yang lain. Untuk lebih jelasnya seperti gambar di bawah ini



Setelah dilakukan analisis tidak ditemukan hal-hal mencurigakan dalam transaksi user-user dari jaringan ini. Status user pada data user juga tidak masalah, tidak ada yang tergolong fraud atau di blacklist.

3.2.3 Tren Transaksi Harian yang Menggunakan Promo



Secara umum sepanjang tahun 2023 penggunaan promo tampak fluktuatif dari hari ke hari, tetapi pada tiga kuartal pertama, penggunaan promo paling banyak terjadi pada tanggal 5 bulan terakhir di kuartal tersebut. Sedangkan pada kuartal ke-empat, penggunaan promo terbesar masih terjadi di bulan terakhir kuartal, tetapi tidak terjadi pada tanggal 5 seperti biasanya, kali ini terjadi pada tanggal 22 Desember dan menjadi aktivitas penggunaan promo tertinggi sepanjang tahun 2023.

3.2.4 Visualisasi Exploitasi Penggunaan Promo

Dari visualisasi di atas (buyer berwarna biru tua), tampak adanya transaksi berturut-turut yang menggunakan sebanyak 5 kali dalam rentang waktu 23 hari. Hal ini menunjukkan adanya eksploitasi promosi dalam waktu yang relatif singkat.

BAB IV
[SQL] Advanced SQL Queries and Stored Procedures

4.1 Advanced SQL Queries for Fraud Detection

4.1.1 Advanced SQL Queries

Identifikasi transaksi yang mencurigakan atau fraud menjadi salah satu aspek penting dalam menjaga integritas sistem dan keamanan data, terutama dalam sistem transaksi digital yang semakin kompleks. Dengan adanya potensi risiko penyalahgunaan dan kegiatan fraud, sangat penting bagi organisasi untuk memantau setiap transaksi yang terjadi dan mendeteksi pola-pola yang tidak biasa. Salah satu cara yang efektif untuk mendeteksi transaksi yang mencurigakan adalah dengan menggunakan analisis berbasis data, di mana SQL (Structured Query Language. Adapun cara untuk mendeteksinya sebagai berikut:

4.1.1.1 Transaksi Jauh di Luar Rentang Normal

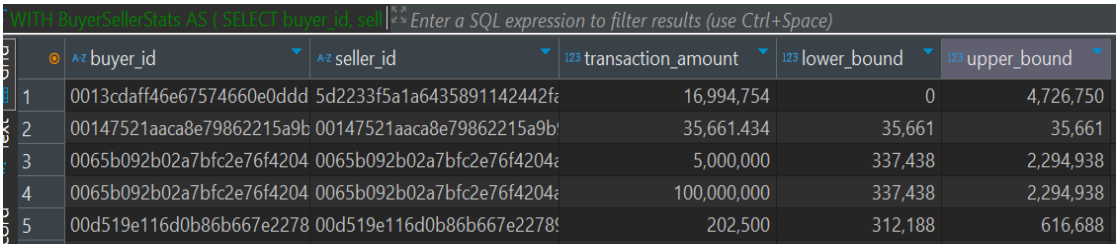
Untuk mendeteksi transaksi yang tidak biasa atau anomali adalah dengan menganalisis perilaku transaksi berdasarkan pasangan pembeli (buyer) dan penjual (seller). Salah satu metode yang dapat digunakan untuk mendeteksi anomali ini adalah dengan menggunakan query SQL sebagai berikut :

Kode Program 4.1.1. *Exploratory Data Analysis* pada Python.

```
-- 4.1.a Transaction Anomalies: Transactions with values far outside the normal range for each buyer/seller pair.
-- Transaksi Jauh di Luar Rentang Normal
WITH BuyerSellerStats AS (
    SELECT
        buyer_id,
        seller_id,
        PERCENTILE_CONT(0.25) WITHIN GROUP (ORDER BY transaction_amount) AS Q1,
        PERCENTILE_CONT(0.75) WITHIN GROUP (ORDER BY transaction_amount) AS Q3
    FROM transaction
    GROUP BY buyer_id, seller_id
),
IQRCalculation AS (
    SELECT
        bss.buyer_id,
        bss.seller_id,
        bss.Q1,
        bss.Q3,
        (bss.Q3 - bss.Q1) AS IQR,
        GREATEST(0, ROUND((bss.Q1 - 1.5 * (bss.Q3 - bss.Q1)))) AS lower_bound,
        ROUND((bss.Q3 + 1.5 * (bss.Q3 - bss.Q1))) AS upper_bound
    FROM BuyerSellerStats bss
),
Anomalies AS (
    SELECT
        t.buyer_id,
        t.seller_id,
        t.transaction_amount,
        ic.lower_bound,
        ic.upper_bound
    FROM transaction t
    JOIN IQRCalculation ic
        ON t.buyer_id = ic.buyer_id AND t.seller_id = ic.seller_id
    WHERE t.transaction_amount < ic.lower_bound
       OR t.transaction_amount > ic.upper_bound
)

SELECT
    buyer_id,
    seller_id,
    transaction_amount,
    lower_bound,
    upper_bound
```

```
FROM Anomalies
ORDER BY buyer_id, seller_id, transaction_amount;

select count(*) from Anomalies;
```

Melalui analisis menggunakan Interquartile Range (IQR), kita dapat menentukan batas-batas bawah (lower_bound) dan atas (upper_bound) dari transaksi yang wajar untuk setiap pasangan pembeli-penjual. Dengan membandingkan nilai transaksi aktual terhadap batas-batas ini, kita dapat mengidentifikasi transaksi yang memiliki nilai jauh di luar rentang normal dan dianggap mencurigakan



	buyer_id	seller_id	transaction_amount	lower_bound	upper_bound
1	0013cdaff46e67574660e0ddd	5d2233f5a1a6435891142442f	16,994,754	0	4,726,750
2	00147521aaca8e79862215a9b	00147521aaca8e79862215a9b	35,661.434	35,661	35,661
3	0065b092b02a7bfc2e76f4204	0065b092b02a7bfc2e76f4204	5,000,000	337,438	2,294,938
4	0065b092b02a7bfc2e76f4204	0065b092b02a7bfc2e76f4204	100,000,000	337,438	2,294,938
5	00d519e116d0b86b667e2278	00d519e116d0b86b667e2278	202,500	312,188	616,688

Dalam konteks ini, hasil query yang dihasilkan akan menampilkan pasangan buyer-seller yang memiliki transaksi dengan nilai "transaction_amount" yang tergolong anomali, di luar batas IQR yang telah ditentukan. Selain itu, hasil query juga akan menampilkan nilai lower_bound dan upper_bound untuk memberikan gambaran lebih jelas mengenai rentang transaksi yang wajar bagi masing-masing pasangan.

4.1.1.2 Analisis Hubungan Pembeli-Penjual

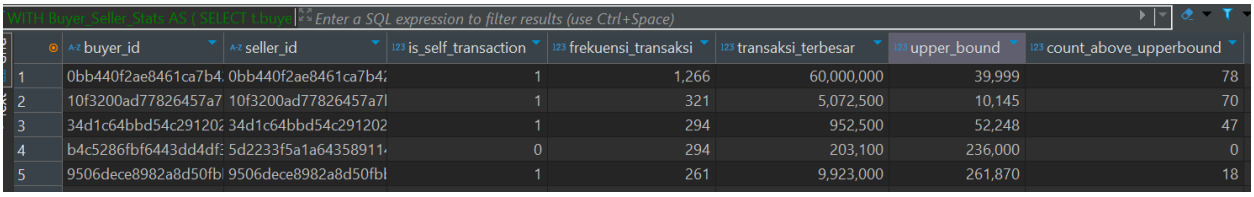
Kode Program 4.1.2. *Exploratory Data Analysis* pada Python.

```
-- 4.1.b Buyer-Seller Relationship Analysis: buyer-seller pairs with unusually high transaction frequencies or amounts
-- Analisis Hubungan Pembeli-Penjual
WITH Buyer_Seller_Stats AS (
    SELECT
        t.buyer_id,
        t.seller_id,
        COUNT(t.dpt_id) AS frekuensi_transaksi,
        MAX(t.transaction_amount) AS transaksi_terbesar,
        PERCENTILE_CONT(0.25) WITHIN GROUP (ORDER BY t.transaction_amount) AS Q1,
        PERCENTILE_CONT(0.75) WITHIN GROUP (ORDER BY t.transaction_amount) AS Q3
    FROM
        transaction t
    GROUP BY
        t.buyer_id, t.seller_id
),
IQR_Calculations AS (
    SELECT
```

```
        bss.buyer_id,
        bss.seller_id,
        bss.frekuensi_transaksi,
        bss.transaksi_terbesar,
        bss.Q1,
        bss.Q3,
        (bss.Q3 - bss.Q1) AS IQR,
        (bss.Q3 + 1.5 * (bss.Q3 - bss.Q1)) AS upper_bound
    FROM
        Buyer_Seller_Stats bss
),
Anomalies AS (
    SELECT
        ic.buyer_id,
        ic.seller_id,
        ic.frekuensi_transaksi,
        ic.transaksi_terbesar,
        ic.upper_bound,
        COUNT(CASE WHEN t.transaction_amount > ic.upper_bound THEN 1 END) AS
count_above_upperbound
    FROM
        transaction t
    JOIN
        IQR_Calculations ic
        ON t.buyer_id = ic.buyer_id AND t.seller_id = ic.seller_id
    GROUP BY
        ic.buyer_id, ic.seller_id, ic.frekuensi_transaksi, ic.transaksi_terbesar, ic.upper_bound
),
Final_Output AS (
    SELECT
        a.*,
        CASE WHEN a.buyer_id = a.seller_id THEN 1 ELSE 0 END AS is_self_transaction,
        bu.user_fraud_flag AS buyer_fraud_flag,
        bu.blacklist_account_flag AS buyer_blacklist_flag,
        su.user_fraud_flag AS seller_fraud_flag,
        su.blacklist_account_flag AS seller_blacklist_flag
    FROM
        Anomalies a
    LEFT JOIN
        users bu ON a.buyer_id = bu.company_id
    LEFT JOIN
        users su ON a.seller_id = su.company_id
)
SELECT
    buyer_id,
    seller_id,
    is_self_transaction,
    frekuensi_transaksi,
    ROUND(transaksi_terbesar) AS transaksi_terbesar,
    ROUND(upper_bound) AS upper_bound,
    count_above_upperbound
--    buyer_fraud_flag,
--    buyer_blacklist_flag,
--    seller_fraud_flag,
```

```
71. --    seller_blacklist_flag
72. FROM
73.     Final_Output
74. ORDER BY
75.     frekuensi_transaksi DESC, transaksi_terbesar DESC;
```

Hasil queri ini menampilkan pasangan buyer-seller dengan frekuensi transaksi tinggi, transaksi terbesar mereka, batas atas transaksi yang dianggap anomali berdasarkan analisis IQR ”upper_bound”, dan total transaksi anomalinya”count_above_upperbound”. Is_self_transaction disertakan juga untuk memperkuat sentimen kecurigaan pada setiap pasangan jika diketahui melakukan self-transaction.



	buyer_id	seller_id	is_self_transaction	frekuensi_transaksi	transaksi_terbesar	upper_bound	count_above_upperbound
1	0bb440f2ae8461ca7b4.	0bb440f2ae8461ca7b4	1	1,266	60,000,000	39,999	78
2	10f3200ad77826457a7	10f3200ad77826457a7	1	321	5,072,500	10,145	70
3	34d1c64bbd54c291202	34d1c64bbd54c291202	1	294	952,500	52,248	47
4	b4c5286fbf6443dd4df:	5d2233f5a1a64358911	0	294	203,100	236,000	0
5	9506dece8982a8d50fb	9506dece8982a8d50fb	1	261	9,923,000	261,870	18

Pasangan buyer-seller nomor satu dianggap mencurigakan dan berpotensi melakukan kolusi karena banyak sekali melakukan self-transaction dan banyak juga nilai transaksi yang terhitung anomali jika dibandingkan dengan nilai-nilai transaksi yang biasanya

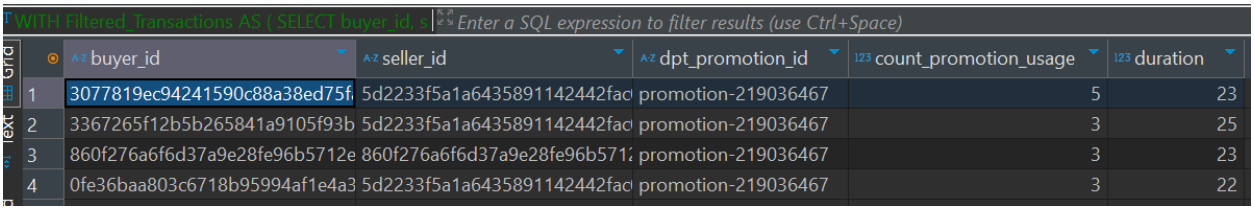
4.1.1.3 Deteksi Penyalahgunaan Promosi

Kode Program 4.1.2. *Exploratory Data Analysis* pada Python.

```
1.  -- Promotion Misuse Detection: Users excessively used promotions within a short period of time
2.  -- Deteksi Penyalahgunaan Promosi
3.  WITH Filtered_Transactions AS (
4.      SELECT
5.          buyer_id,
6.          seller_id,
7.          dpt_promotion_id,
8.          transaction_created_datetime::timestamp AS created_at
9.      FROM
10.         transaction
11.     WHERE
12.         dpt_promotion_id <> 'no promotion'
13. ),
14. Ranked_Transactions AS (
15.     SELECT
16.         buyer_id,
17.         seller_id,
18.         dpt_promotion_id,
19.         created_at,
20.         ROW_NUMBER() OVER (
21.             PARTITION BY buyer_id, seller_id, dpt_promotion_id
22.             ORDER BY created_at
```

```
        ) AS rank,
        COUNT(*) OVER (PARTITION BY buyer_id, seller_id, dpt_promotion_id) AS count_promotion_usage
    FROM
        Filtered_Transactions
),
Duration_Calculation AS (
    SELECT
        buyer_id,
        seller_id,
        dpt_promotion_id,
        COUNT(*) AS count_promotion_usage,
        MIN(created_at) AS first_usage,
        MAX(created_at) AS last_usage,
        EXTRACT(DAY FROM (MAX(created_at) - MIN(created_at))) AS duration
    FROM
        Ranked_Transactions
    GROUP BY
        buyer_id, seller_id, dpt_promotion_id
)
SELECT
    buyer_id,
    seller_id,
    dpt_promotion_id,
    count_promotion_usage,
    duration
FROM
    Duration_Calculation
where
    a. (count_promotion_usage >= 3) and (duration <= 30)
ORDER BY
    count_promotion_usage DESC, duration DESC;
```

Hasil queri ini menampilkan pasangan buyer-seller, promo id unik yang digunakan berkali-kali, banyaknya promo digunakan, dan durasi penggunaan promo berulang.



	buyer_id	seller_id	dpt_promotion_id	count_promotion_usage	duration
1	3077819ec94241590c88a38ed75f	5d2233f5a1a6435891142442fac	promotion-219036467	5	23
2	3367265f12b5b265841a9105f93b	5d2233f5a1a6435891142442fac	promotion-219036467	3	25
3	860f276a6f6d37a9e28fe96b5712e	860f276a6f6d37a9e28fe96b571	promotion-219036467	3	23
4	0fe36baa803c6718b95994af1e4a3	5d2233f5a1a6435891142442fac	promotion-219036467	3	22

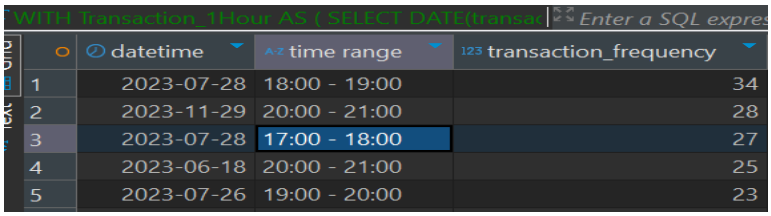
Pasangan buyer-seller nomor 1 dianggap menyalahgunakan atau mengeksploitasi promo karena menggunakan satu jenis promo yang sama sebanyak lima kali dalam waktu relatif singkat, yaitu 23 hari.

4.1.1.4 Waktu yang Mencurigakan

Kode Program 4.1.2. *Exploratory Data Analysis* pada Python.

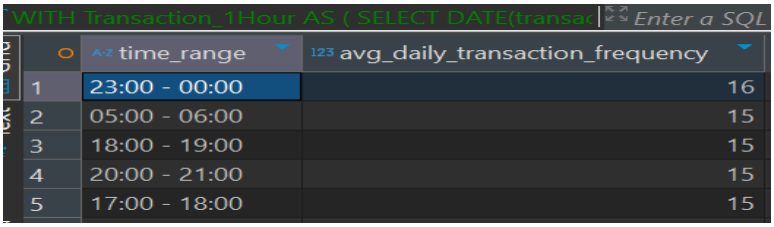
```
--Suspicious Timing: Transactions at irregular hours or intervals (e.g., many transactions in a short time span).
-- waktu yang mencurigakan


WITH Transaction_1Hour AS (
    SELECT
        DATE(transaction_created_datetime::timestamp) AS transaction_date,
        DATE_TRUNC('hour', transaction_created_datetime::timestamp)
        + INTERVAL '1 hour' * FLOOR(EXTRACT(MINUTE FROM transaction_created_datetime::timestamp) / 60) AS transaction_1hour_start,
        COUNT(*) AS transaction_frequency
    FROM
        transaction
    WHERE
        EXTRACT(HOUR FROM transaction_created_datetime::timestamp) NOT BETWEEN 9 AND 16
    GROUP BY
        DATE(transaction_created_datetime::timestamp),
        DATE_TRUNC('hour', transaction_created_datetime::timestamp)
        + INTERVAL '1 hour' * FLOOR(EXTRACT(MINUTE FROM transaction_created_datetime::timestamp) / 60)
),
Formatted_1Hour AS (
    SELECT
        transaction_date,
        TO_CHAR(transaction_1hour_start, 'HH24:MI') || ' - ' ||
        TO_CHAR(transaction_1hour_start + INTERVAL '1 hour', 'HH24:MI') AS time_range,
        transaction_frequency
    FROM
        Transaction_1Hour
    where transaction_frequency >= 12
)
--SELECT
--    transaction_date AS datetime,
--    time_range AS "time range",
--    transaction_frequency
--FROM
--    Formatted_1Hour
--ORDER BY
--    /*datetime, "time range",*/ transaction_frequency DESC;


select
    a. time_range,
    b. round(avg(transaction_frequency)) avg_daily_transaction_frequency
from Formatted_1Hour
group by time_range
order by avg_daily_transaction_frequency desc;
```

Hasil queri ini menampilkan tanggal dan rentang waktu per jam pada waktu ireguler, yang mana pada rentang waktu dan tanggal tersebut terjadi transaksi sedikitnya 12 kali (di luar 09.00-17.00).



	datetime	time range	transaction_frequency
1	2023-07-28	18:00 - 19:00	34
2	2023-11-29	20:00 - 21:00	28
3	2023-07-28	17:00 - 18:00	27
4	2023-06-18	20:00 - 21:00	25
5	2023-07-26	19:00 - 20:00	23

Ternyata tidak jarang transaksi dilakukan pada waktu ireguler.



	time_range	avg_daily_transaction_frequency
1	23:00 - 00:00	16
2	05:00 - 06:00	15
3	18:00 - 19:00	15
4	20:00 - 21:00	15
5	17:00 - 18:00	15

Rentang waktu 23.00-00.00 dianggap paling mencurigakan karena rata-rata transaksi hariannya terbesar.

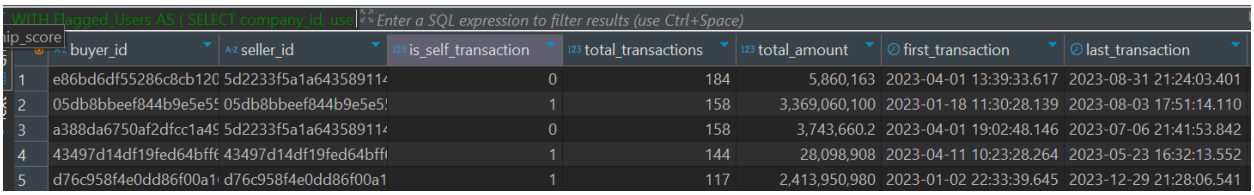
4.1.1.5 Koneksi Pengguna yang Ditandai

Kode Program 4.1.2. *Exploratory Data Analysis* pada Python.

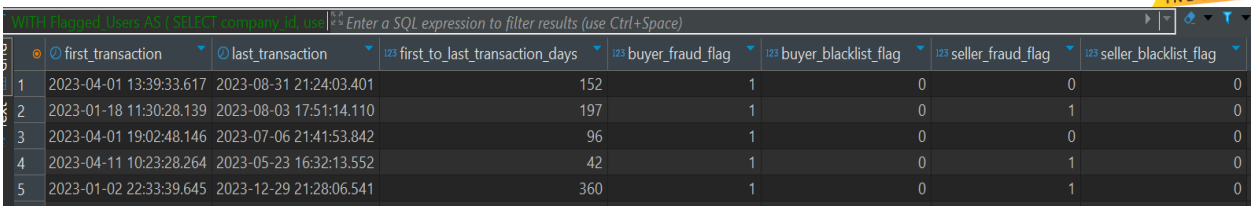
```
---------------------------------- Flagged User Connections 1----------------------------------------
WITH Flagged_Users AS (
    SELECT
        company_id,
        user_fraud_flag,
        blacklist_account_flag
    FROM
        users
    WHERE
        user_fraud_flag = 1 OR blacklist_account_flag = 1
),
User_Connections AS (
    SELECT
        ft.dpt_id,
        ft.buyer_id,
        ft.seller_id,
        fu_buyer.user_fraud_flag AS buyer_fraud_flag,
        fu_buyer.blacklist_account_flag AS buyer_blacklist_flag,
        fu_seller.user_fraud_flag AS seller_fraud_flag,
        fu_seller.blacklist_account_flag AS seller_blacklist_flag,
        ft.transaction_amount,
        ft.transaction_created_datetime::timestamp AS transaction_time
    FROM
        transaction ft
    LEFT JOIN
        users fu_buyer ON ft.buyer_id = fu_buyer.company_id
    LEFT JOIN
        users fu_seller ON ft.seller_id = fu_seller.company_id
    WHERE
```

```
        fu_buyer.user_fraud_flag = 1 OR
        fu_buyer.blacklist_account_flag = 1 OR
        fu_seller.user_fraud_flag = 1 OR
        fu_seller.blacklist_account_flag = 1
),
Flagged_Transactions AS (
    SELECT
        buyer_id,
        seller_id,
        CASE WHEN buyer_id = seller_id THEN 1 ELSE 0 END AS is_self_transaction,
        COUNT(*) AS total_transactions,
        SUM(transaction_amount) AS total_amount,
        MIN(transaction_time) AS first_transaction,
        MAX(transaction_time) AS last_transaction,
        DATE_PART('day', MAX(transaction_time) - MIN(transaction_time)) AS
first_to_last_transaction_days,
        MAX(buyer_fraud_flag) AS buyer_fraud_flag,
        MAX(buyer_blacklist_flag) AS buyer_blacklist_flag,
        MAX(seller_fraud_flag) AS seller_fraud_flag,
        MAX(seller_blacklist_flag) AS seller_blacklist_flag
    FROM
        User_Connections
    GROUP BY
        buyer_id, seller_id
)
SELECT
    buyer_id,
    seller_id,
    is_self_transaction,
    total_transactions,
    total_amount,
    first_transaction,
    last_transaction,
    first_to_last_transaction_days,
    buyer_fraud_flag,
    buyer_blacklist_flag,
    seller_fraud_flag,
    seller_blacklist_flag
FROM
    Flagged_Transactions
ORDER BY
    total_transactions DESC, total_amount DESC;
```

Hasil queri ini menampilkan rekam jejak transaksi user yang fraud atau di-blacklist



	buyer_id	seller_id	is_self_transaction	total_transactions	total_amount	first_transaction	last_transaction
1	e86bd6df55286c8cb120	5d2233f5a1a643589114	0	184	5,860,163	2023-04-01 13:39:33.617	2023-08-31 21:24:03.401
2	05db8bbeef844b9e5e5	05db8bbeef844b9e5e5	1	158	3,369,060,100	2023-01-18 11:30:28.139	2023-08-03 17:51:14.110
3	a388da6750af2dfcc1a4	5d2233f5a1a643589114	0	158	3,743,660.2	2023-04-01 19:02:48.146	2023-07-06 21:41:53.842
4	43497d14df19fed64bff	43497d14df19fed64bff	1	144	28,098,908	2023-04-11 10:23:28.264	2023-05-23 16:32:13.552
5	d76c958f4e0dd86f00a1	d76c958f4e0dd86f00a1	1	117	2,413,950,980	2023-01-02 22:33:39.645	2023-12-29 21:28:06.541



WITH Flagged_Users AS (SELECT company_id, us... | Enter a SQL expression to filter results (use Ctrl+Space)

	first_transaction	last_transaction	first_to_last_transaction_days	buyer_fraud_flag	buyer_blacklist_flag	seller_fraud_flag	seller_blacklist_flag
1	2023-04-01 13:39:33.617	2023-08-31 21:24:03.401	152	1	0	0	0
2	2023-01-18 11:30:28.139	2023-08-03 17:51:14.110	197	1	0	1	0
3	2023-04-01 19:02:48.146	2023-07-06 21:41:53.842	96	1	0	0	0
4	2023-04-11 10:23:28.264	2023-05-23 16:32:13.552	42	1	0	1	0
5	2023-01-02 22:33:39.645	2023-12-29 21:28:06.541	360	1	0	1	0

Pasangan buyer-seller nomor 4 merupakan self-transaction yang mana usernya termasuk user fraud. User ini melakukan 144 transaksi hanya dalam 42 hari tetapi belum diblacklist.

4.2 SQL Joins for User-Company Fraud Insights

Untuk mengidentifikasi dan mencegah tindakan fraud dalam sistem transaksi, sangat penting untuk menganalisis tidak hanya perilaku transaksi pengguna, tetapi juga hubungan antara pengguna dengan perusahaan tempat mereka bertransaksi. Dengan menggabungkan data transaksi dengan informasi terkait status KYC (Know Your Customer), KYB (Know Your Business), dan riwayat flag fraud yang dimiliki oleh perusahaan, kita dapat memperoleh wawasan lebih dalam mengenai pola-pola yang dapat mengindikasikan potensi fraud. Adapun querynya sebagai berikut:

Kode Program 4.2. *User-Company Fraud Insights*

```
SELECT
  u.company_id,
  u.company_kyb_status_name,
  u.company_kyc_status_name,
  COUNT(DISTINCT t.dpt_id) AS total_transactions, -- Menghitung transaksi unik
  SUM(t.transaction_amount) AS total_transaction_amount,
  COUNT(DISTINCT t.buyer_id) AS distinct_buyers,
  COUNT(DISTINCT t.seller_id) AS distinct_sellers,
  SUM(CASE WHEN u.user_fraud_flag = 1 THEN 1 ELSE 0 END) AS fraud_flags_on_users,
  SUM(CASE WHEN u.blacklist_account_flag = 1 THEN 1 ELSE 0 END) AS blacklist_flags_on_users
FROM transaction t
LEFT JOIN
  "user" u ON t.buyer_id = u.company_id
LEFT JOIN
  "user" u_seller ON t.seller_id = u_seller.company_id
WHERE t.transaction_created_datetime BETWEEN '2022-01-01' AND '2023-12-31'
GROUP BY u.company_id, u.company_kyb_status_name, u.company_kyc_status_name
HAVING
  SUM(CASE WHEN u.user_fraud_flag = 1 THEN 1 ELSE 0 END) > 0
  OR SUM(CASE WHEN u.blacklist_account_flag = 1 THEN 1 ELSE 0 END) > 0
ORDER BY total_transaction_amount DESC;
```

Dari hasil query tersebut didapat hasil sebagai berikut:

	company_id	company_kyb_status_name	company_kyc_status_name	total_transactions
1	dd9a9195c84e334de861047a82118fd24ada1030d0688ba6	BELUM_VALIDASI	AKUN_DIBEKUKAN	1
2	f66bd25e3b356f4a49d12a8086bd1985d18e748612a685d79	KYB_DITOLAK	AKUN_DIBEKUKAN	43
3	417924a2a195717000c1352d251149042a326a750799205d62	KYB_DITOLAK	AKUN_DIBEKUKAN	62
4	45c5dbe8c3a70cd95386e2382fdd09312e15e23435c473dc8	BELUM_VALIDASI	AKUN_DIBEKUKAN	55
5	be5de79f79b08a72ba4f2d5eea5190410ba7b2a714d3ec16d	VALIDASI_BERHASIL	AKUN_DIBEKUKAN	48
6	ff297354f3294e46eb09787675d6e2847cd76be14cba7d2361	BELUM_VALIDASI	AKUN_DIBEKUKAN	67
7	0770cab03ae9d39f7cb590d27a2e3c2f3cfddb570654b821c1	DALAM_PENGAJUAN	AKUN_DIBEKUKAN	29
8	3b7a45da383c9a053b40ab63338abc9c6a8b798615d7f70c2	BELUM_VALIDASI	AKUN_DIBEKUKAN	51
9	ef958867c89944a43d26df0ada8ff693ff1b7142cbafe0bfb3e	BELUM_VALIDASI	AKUN_DIBEKUKAN	13
10	d30fc2e358f80fd6c9d64b1836bdfe278d7434e0ca5c75823e	BELUM_VALIDASI	AKUN_DIBEKUKAN	24
11	e631239a0bb5f3adb5d42720d2230e270ce2971616c28377f5	BELUM_VALIDASI	AKUN_DIBEKUKAN	10
12	502180e2fb19b8a5eab58cd3515d8c328d52ac618b3dad29b	BELUM_VALIDASI	AKUN_DIBEKUKAN	43
13	e0722d2eea37431e2391f759e123a42f7fe0a552f6059d41a68	BELUM_VALIDASI	AKUN_DIBEKUKAN	15
14	e3c3494b220dbbe9b2595c7e69c891083a0349739ac47e866	BELUM_VALIDASI	AKUN_DIBEKUKAN	31
15	7707af85b39db2a77fbcb926dd474674ad920042a7e6436fc4	BELUM_VALIDASI	AKUN_DIBEKUKAN	11
16	06aa8c62e26b49cfe99f34d583c7c197756cbe35878bcab654	BELUM_VALIDASI	AKUN_DIBEKUKAN	10
17	05db8bbeef844b9e5e5599dd8caff2d34c7789d5e827c4c01	BELUM_VALIDASI	AKUN_DIBEKUKAN	158
18	5b47a8acf64f61df2bf077f1bb038c000cd6b3c27386a5d1f74	VALIDASI_BERHASIL	AKUN_DIBEKUKAN	15
19	b13aa5ba4ea9bb2443cac2caa92fb8d6b3a86f3d3b35b5da3	BELUM_VALIDASI	AKUN_DIBEKUKAN	42
20	c5109073df78842679dff210146130ac51ca3809917f9959a25	BELUM_VALIDASI	AKUN_DIBEKUKAN	22

	total_transactions	total_transaction_amount	distinct_buyers	distinct_sellers	fraud_flags_on_users	blacklist_flags_on_users
1	1	20,140,103,700	1	1	1	0
2	43	14,218,475,500	1	1	43	0
3	62	12,414,680,100	1	1	62	0
4	55	5,762,689,000	1	1	55	0
5	48	4,953,873,400	1	1	48	0
6	67	4,746,853,400	1	1	67	0
7	29	4,471,152,100	1	1	29	0
8	51	4,324,023,300	1	1	51	0
9	13	4,317,177,900	1	1	13	0
10	24	4,164,875,520	1	1	24	0
11	10	4,100,000,000	1	1	10	0
12	43	4,089,625,090	1	1	43	0
13	15	3,907,510,270	1	2	15	0
14	31	3,600,375,040	1	2	31	0
15	11	3,594,322,940	1	1	11	0
16	10	3,550,500,860	1	1	10	0
17	158	3,369,060,860	1	1	158	0
18	15	3,325,749,250	1	1	15	0
19	42	3,325,296,380	1	3	42	0
20	22	3,199,762,690	1	3	22	0

BAB V

SQL Views and Stored Procedures

5.1 SQL View

Melalui pembuatan SQL Views, kita dapat merancang laporan fraud secara reguler dengan dua jenis tampilan yang sangat berguna

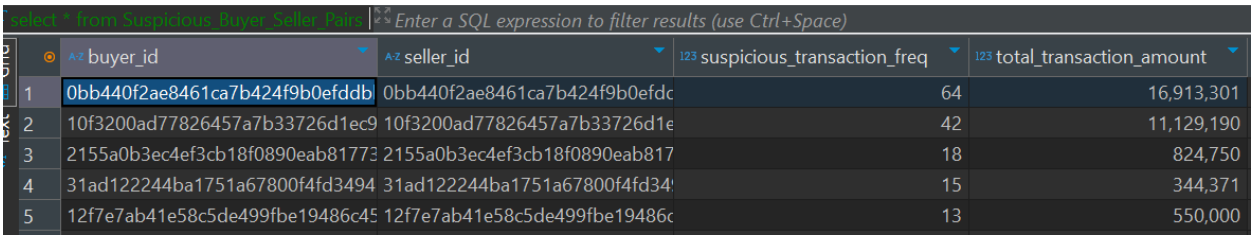
5.1.1 View Pasangan buyer-seller paling mencurigakan

View ini memberikan ringkasan tentang hubungan pembeli-penjual yang paling mencurigakan, berdasarkan frekuensi transaksi dan jumlah transaksi yang dilakukan.. Adapun querynya sebagai berikut :

Kode Program 4.3.1.1.

```
use pbl_paper_id;

select count(*) from transaction_frequency_metrics;

/*----------------------------------  SQL View ----------------------------------------------*/
-- 1.1 Membuat view untuk pasangan pembeli-penjual yang paling mencurigakan berdasarkan fitur baru
-- burst_activity, unusual_gap, burst_amount dibuat pada tahap feature engineering di task 2
-- Pasangan buyer-seller paling mencurigakan
CREATE VIEW Suspicious_Buyer_Seller_Pairs AS
SELECT
    buyer_id,
    seller_id,
    COUNT(dpt_id) AS suspicious_transaction_freq,
    ROUND(SUM(transaction_amount)) AS total_transaction_amount
FROM
    transaction_frequency_metrics
WHERE
    (burst_activity = 1 AND burst_amount = 1)
    OR (unusual_gap = 1 AND burst_amount = 1)
GROUP BY
    buyer_id,
    seller_id
ORDER BY
    suspicious_transaction_freq DESC;

select * from Suspicious_Buyer_Seller_Pairs;
```

Dimana didapat hasil sebagai berikut:



	buyer_id	seller_id	suspicious_transaction_freq	total_transaction_amount
1	0bb440f2ae8461ca7b424f9b0efddb	0bb440f2ae8461ca7b424f9b0efdc	64	16,913,301
2	10f3200ad77826457a7b33726d1ec9	10f3200ad77826457a7b33726d1e	42	11,129,190
3	2155a0b3ec4ef3cb18f0890eab81773	2155a0b3ec4ef3cb18f0890eab817	18	824,750
4	31ad122244ba1751a67800f4fd3494	31ad122244ba1751a67800f4fd34	15	344,371
5	12f7e7ab41e58c5de499fbe19486c45	12f7e7ab41e58c5de499fbe19486c	13	550,000

View ini menampilkan pasangan buyer-seller yang dianggap mencurigakan berdasarkan indikator burs_activity, unusual_gap, dan burst_amount yang dihasilkan dari bagian feature engineering, juga frekuensi transaksi mencurigakan yang mereka lakukan serta total transaksi mencurigakannya.

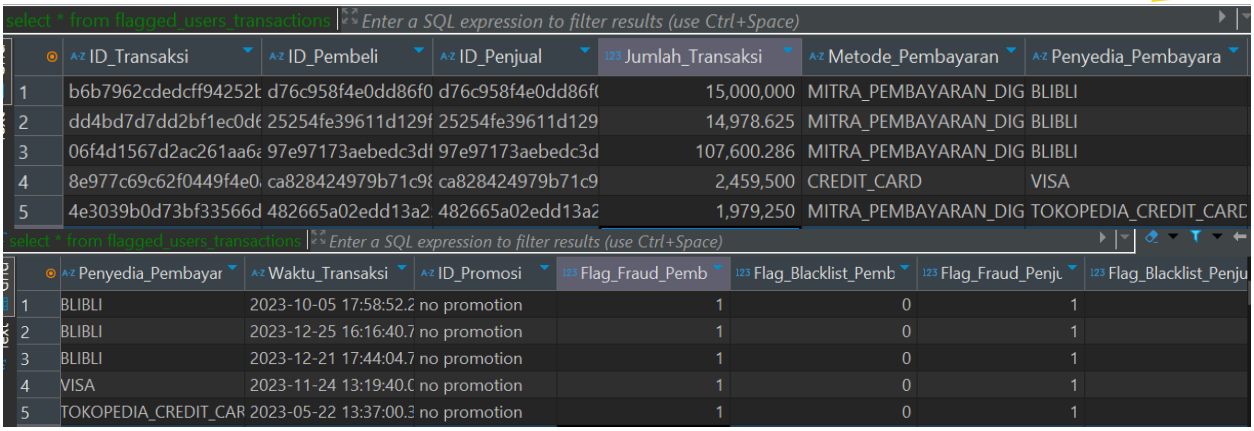
5.1.2 View - Pengguna yang Ditandai dan Transaksi Mereka

View ini berfokus pada pengguna yang telah terflag atau dimasukkan dalam daftar hitam, bersama dengan transaksi yang melibatkan mereka. Adapun querynya sebagi berikut :

Kode Program 4.3.1.2.

```
-- 1.2 Membuat View untuk pengguna yang terindikasi fraud atau di-blacklist dan transaksi mereka
-- Pengguna yang ditandai dan transaksi mereka
CREATE VIEW Flagged_Users_Transactions AS
SELECT
    t.dpt_id AS ID_Transaksi,
    t.buyer_id AS ID_Pembeli,
    t.seller_id AS ID_Penjual,
    t.transaction_amount AS Jumlah_Transaksi,
    t.payment_method_name AS Metode_Pembayaran,
    t.payment_provider_name AS Penyedia_Pembayaran,
    t.transaction_created_datetime AS Waktu_Transaksi,
    t.dpt_promotion_id AS ID_Promosi,
    ub.user_fraud_flag AS Flag_Fraud_Pembeli,
    ub.blacklist_account_flag AS Flag_Blacklist_Pembeli,
    us.user_fraud_flag AS Flag_Fraud_Penjual,
    us.blacklist_account_flag AS Flag_Blacklist_Penjual,
    CASE
        WHEN t.buyer_id = t.seller_id THEN 1
        ELSE 0
    END AS Flag_Self_Transaction

FROM
    transaction t
LEFT JOIN
    user ub
ON
    t.buyer_id = ub.company_id
LEFT JOIN
    user us
ON
    t.seller_id = us.company_id
WHERE
    (ub.user_fraud_flag = 1 OR ub.blacklist_account_flag = 1)
    OR
    (us.user_fraud_flag = 1 OR us.blacklist_account_flag = 1);


select count(*) from flagged_users_transactions;

select * from flagged_users_transactions;
```

Dimana didapat hasil sebagai berikut:



	ID_Transaksi	ID_Pembeli	ID_Penjual	Jumlah_Transaksi	Metode_Pembayaran	Penyedia_Pembayara
1	b6b7962cdedcff94252b	d76c958f4e0dd86f0	d76c958f4e0dd86f0	15,000,000	MITRA_PEMBAYARAN_DIG	BLIBLI
2	dd4bd7d7dd2bf1ec0d0	25254fe39611d129f	25254fe39611d129	14,978.625	MITRA_PEMBAYARAN_DIG	BLIBLI
3	06f4d1567d2ac261aa6a	97e97173aebedc3df	97e97173aebedc3d	107,600.286	MITRA_PEMBAYARAN_DIG	BLIBLI
4	8e977c69c62f0449f4e0	ca828424979b71c9f	ca828424979b71c9	2,459,500	CREDIT_CARD	VISA
5	4e3039b0d73bf33566d	482665a02edd13a2	482665a02edd13a2	1,979,250	MITRA_PEMBAYARAN_DIG	TOKOPEDIA_CREDIT_CARD

	Penyedia_Pembayar	Waktu_Transaksi	ID_Promosi	Flag_Fraud_Pemb	Flag_Blacklist_Pemb	Flag_Fraud_Penju	Flag_Blacklist_Penju
1	BLIBLI	2023-10-05 17:58:52.2	no promotion	1	0	1	
2	BLIBLI	2023-12-25 16:16:40.7	no promotion	1	0	1	
3	BLIBLI	2023-12-21 17:44:04.7	no promotion	1	0	1	
4	VISA	2023-11-24 13:19:40.0	no promotion	1	0	1	
5	TOKOPEDIA_CREDIT_CAR	2023-05-22 13:37:00.3	no promotion	1	0	1	

View ini menampilkan detail transaksi beserta status buyer dan seller sehingga dapat diketahui apakah transaksi tertentu dilakukan oleh user yang fraud atau tidak.

5.2 Stored Procedures

Untuk meningkatkan efisiensi dan akurasi dalam deteksi fraud, stored procedures dapat digunakan untuk mengotomatisasi proses analisis dan identifikasi transaksi mencurigakan. Dengan menggunakan stored procedures, organisasi dapat menjalankan serangkaian query atau logika deteksi fraud secara otomatis, tanpa perlu melibatkan intervensi manual setiap kali analisis dilakukan. Adapaun Store procedures yang dilakukanialah:

5.2.1 Laporan Penipuan Bulanan

Prosedur ini menghasilkan report transaksi bulanan yang tergolong fraud atau mencurigakan. Dari prosedur ini kita dapat melihat total nilai transaksi fraud, berapa fraud user yang terlibat dalam transaksi fraud, berapa berapa pasang buyer-seller ditandai fraud yang bertransaksi, dan berapa pasang buyer-seller yang terindikasi melakukan transaksi mencurigakan (berdasarkan indikator burs_activity, unusual_gap, dan burst_amount) tetapi belum ditandai sebagai fraud user. Harapannya prosedur ini dapat digunakan untuk mendeteksi user-user yang belum ditandai walau sebenarnya mereka melakukan transaksi fraud. Adapun querynya sebagai berikut :

Kode Program 4.1.2.

```
1. use pbl_paper_id;
2.
3. /*----------------------------------------SQL Stored Procedures----------------------------------*/
4.
```

```
-- 2.1 Monthly Fraud Report Procedure
-- laporan penipuan bulanan
DELIMITER $$

CREATE PROCEDURE MonthlyFraudReport(IN report_month VARCHAR(7))
BEGIN
    DECLARE total_fraud_amount DOUBLE;
    DECLARE total_fraud_user_count INT;
    DECLARE total_suspicious_count INT;
    DECLARE total_fraud_transaction_count INT;

    -- Hitung Total Fraud Amount
    SELECT
        SUM(tfm.transaction_amount)
    INTO total_fraud_amount
    FROM transaction_frequency_metrics tfm
    LEFT JOIN user u_buyer ON tfm.buyer_id = u_buyer.company_id
    LEFT JOIN user u_seller ON tfm.seller_id = u_seller.company_id
    WHERE
        (u_buyer.user_fraud_flag = 1 OR u_seller.user_fraud_flag = 1)
        AND DATE_FORMAT(tfm.transaction_created_datetime, '%Y-%m') = report_month;

    -- Hitung Total Fraud Users
    SELECT
        COUNT(DISTINCT CASE WHEN u_buyer.user_fraud_flag = 1 THEN tfm.buyer_id END)
        + COUNT(DISTINCT CASE WHEN u_seller.user_fraud_flag = 1 THEN tfm.seller_id END)
    INTO total_fraud_user_count
    FROM transaction_frequency_metrics tfm
    LEFT JOIN user u_buyer ON tfm.buyer_id = u_buyer.company_id
    LEFT JOIN user u_seller ON tfm.seller_id = u_seller.company_id
    WHERE
        (u_buyer.user_fraud_flag = 1 OR u_seller.user_fraud_flag = 1)
        AND DATE_FORMAT(tfm.transaction_created_datetime, '%Y-%m') = report_month;

    -- Hitung Total Fraud Transactions (Buyer-Seller Pairs)
    SELECT
        COUNT(DISTINCT CONCAT(tfm.buyer_id, '-', tfm.seller_id))
    INTO total_fraud_transaction_count
    FROM transaction_frequency_metrics tfm
    LEFT JOIN user u_buyer ON tfm.buyer_id = u_buyer.company_id
    LEFT JOIN user u_seller ON tfm.seller_id = u_seller.company_id
    WHERE
        (u_buyer.user_fraud_flag = 1 OR u_seller.user_fraud_flag = 1)
        AND DATE_FORMAT(tfm.transaction_created_datetime, '%Y-%m') = report_month;

    -- Hitung Total Suspicious Buyer-Seller Pairs
    SELECT
        COUNT(DISTINCT CONCAT(tfm.buyer_id, '-', tfm.seller_id))
    INTO total_suspicious_count
    FROM transaction_frequency_metrics tfm
    WHERE
        ((tfm.burst_activity = 1 AND tfm.burst_amount = 1)
        OR (tfm.unusual_gap = 1 AND tfm.burst_amount = 1))
        AND DATE_FORMAT(tfm.transaction_created_datetime, '%Y-%m') = report_month
```

```
        AND CONCAT(tfm.buyer_id, '-', tfm.seller_id) NOT IN (
            SELECT DISTINCT CONCAT(f_tfm.buyer_id, '-', f_tfm.seller_id)
            FROM transaction_frequency_metrics f_tfm
            LEFT JOIN user f_u_buyer ON f_tfm.buyer_id = f_u_buyer.company_id
            LEFT JOIN user f_u_seller ON f_tfm.seller_id = f_u_seller.company_id
            WHERE
                (f_u_buyer.user_fraud_flag = 1 OR f_u_seller.user_fraud_flag = 1)
                AND DATE_FORMAT(f_tfm.transaction_created_datetime, '%Y-%m') = report_month
        );

    -- Hasil Akhir
    SELECT
        CAST(total_fraud_amount AS UNSIGNED) AS fraud_transaction_amounts,
        total_fraud_user_count AS fraud_users,
        total_fraud_transaction_count AS fraud_transactions,
        total_suspicious_count AS suspicious_buyer_seller_pairs;
END $$

DELIMITER ;


CALL MonthlyFraudReport('2023-03');
```

```
CALL MonthlyFraudReport('2023-03') | Enter a SQL expression to filter results (use Ctrl+Space)
```

	fraud_transaction_amounts	fraud_users	fraud_transactions	suspicious_buyer_seller_pairs
1	53,421,614,792	176	146	41

Pada bulan Maret 2023 total transaksi fraud (transaksi di mana buyer atau seller ditandai fraud) mencapai sekitar 53 Miliar. Ada 146 pasangan buyer-seller yang terlibat dalam transaksi ini. Beberapa transaksi melibatkan buyer dan seller di mana keduanya merupakan user yang ditandai fraud. Disamping itu, pada periode bulan ini, ditemukan 41 pasangan buyer-seller yang melakukan transaksi mencurigakan, tetapi keduanya belum ada yang ditandai fraud.

5. 2.2 Deteksi Penyalahgunaan Otomatis

Prosedur ini menampilkan report buyer_id yang menggunakan jenis promo yang sama berulang kali sedikitnya 3 kali. Sebagian buyer_id menggunakannya dalam waktu relatif singkat sehingga dapat dianggap melakukan penyalahgunaan.. Adaun querynya sebagai berikut:

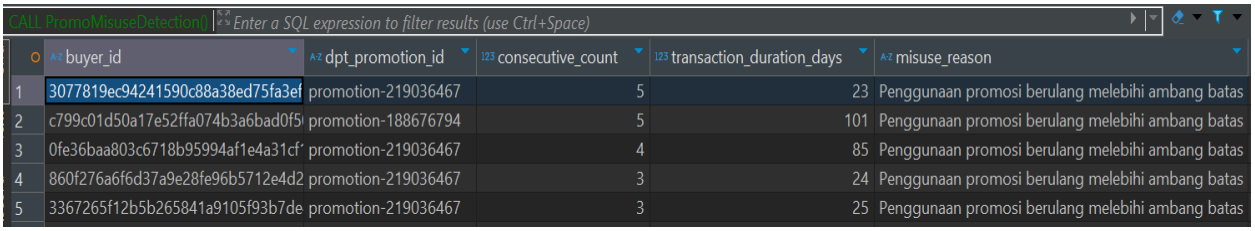
Kode Program 4.3.2.2. *User-Company Fraud Insights*

```
1.  -- 2.2 Automated Promotion Misuse Detection
2.  -- Deteksi penyalahgunaan otomatis
3.  DELIMITER //
4.
5.  CREATE PROCEDURE PromoMisuseDetection()
6.  BEGIN
```

```
    -- Hapus tabel sementara jika sudah ada
    DROP TEMPORARY TABLE IF EXISTS temp_promo_transactions;
    DROP TEMPORARY TABLE IF EXISTS promo_counts;

    -- Tabel sementara untuk menyimpan data transaksi dengan lag calculation
    CREATE TEMPORARY TABLE temp_promo_transactions AS
    SELECT
        buyer_id,
        dpt_promotion_id,
        transaction_created_datetime,
        -- Penanda transaksi sebelumnya
        LAG(transaction_created_datetime) OVER (PARTITION BY buyer_id, dpt_promotion_id
ORDER BY transaction_created_datetime) AS prev_transaction_datetime
    FROM transaction
    WHERE
        dpt_promotion_id IS NOT NULL
        AND dpt_promotion_id <> 'no promotion'; -- Mengabaikan transaksi tanpa promosi

    -- Variabel untuk menghitung consecutive promo count
    SET @row_num := 0;
    SET @prev_buyer := NULL;
    SET @prev_promo := NULL;

    -- Tabel sementara kedua untuk menghitung consecutive promo count
    CREATE TEMPORARY TABLE promo_counts AS
    SELECT
        buyer_id,
        dpt_promotion_id,
        transaction_created_datetime,
        -- Menghitung transaksi berturut-turut
        @row_num := IF(@prev_buyer = buyer_id AND @prev_promo = dpt_promotion_id,
@row_num + 1, 1) AS consecutive_promo_count,
        @prev_buyer := buyer_id,
        @prev_promo := dpt_promotion_id
    FROM temp_promo_transactions
    ORDER BY buyer_id, dpt_promotion_id, transaction_created_datetime;

    -- Tampilkan data misuse ke output jika melebihi threshold
    SELECT
        buyer_id,
        dpt_promotion_id,
        MAX(consecutive_promo_count) AS consecutive_count,
        DATEDIFF(MAX(transaction_created_datetime), MIN(transaction_created_datetime)) AS
transaction_duration_days,
        'Penggunaan promosi berulang melebihi ambang batas' AS misuse_reason
    FROM promo_counts
    GROUP BY buyer_id, dpt_promotion_id
    HAVING MAX(consecutive_promo_count) >= 3  -- Ambang batas deteksi
    ORDER BY consecutive_count DESC, transaction_duration_days ASC;

END //

DELIMITER ;

```

```
-- Menjalankan Stored Procedure
CALL PromoMisuseDetection();

select* from log_promosi_misuse;
```

Dari hasil query tersebut didapat hasil sebagai berikut:



	buyer_id	dpt_promotion_id	consecutive_count	transaction_duration_days	misuse_reason
1	3077819ec94241590c88a38ed75fa3ef	promotion-219036467	5	23	Penggunaan promosi berulang melebihi ambang batas
2	c799c01d50a17e52ffa074b3a6bad0f5	promotion-188676794	5	101	Penggunaan promosi berulang melebihi ambang batas
3	0fe36baa803c6718b95994af1e4a31cf	promotion-219036467	4	85	Penggunaan promosi berulang melebihi ambang batas
4	860f276a6f6d37a9e28fe96b5712e4d2	promotion-219036467	3	24	Penggunaan promosi berulang melebihi ambang batas
5	3367265f12b5b265841a9105f93b7de	promotion-219036467	3	25	Penggunaan promosi berulang melebihi ambang batas

Buyer_id pada nomor satu menggunakan promo yang sama berturut-turut selama lima kali hanya dalam 23 hari. Dapat dikatakan bahwa bahwa buyer_id ini telah melakukan penyalahgunaan promosi.

BAB VI

[Python] Advanced Fraud Analysis and Network Insights

6.1 Social Network Analysis

Menggunakan query SQL untuk membuat jaringan sosial yang menggambarkan hubungan antara pembeli dan penjual. Data transaksi inidigunakan untuk mengidentifikasi pola interaksi antara keduanya. Analisis ini membantu dalam memahami dinamika hubungan dan potensi risiko atau peluang bisnis, Adapaun analisis nya sebagai berikut:

6.1.1 Analisis Hubungan Pembeli-Penjual untuk Mengidentifikasi Pola Interaksi dan Terlibat Fraud (using SQL)

Query SQL telah dirancang untuk menghasilkan data pasangan pembeli-penjual yang terlibat dalam transaksi fraud. Suatu transaksi diklasifikasikan sebagai fraud apabila terdapat pembeli atau penjual yang ditandai dengan nilai 1 pada kolom fraud_flag atau blacklist_flag. Analisis ini bertujuan untuk memahami pola interaksi dalam jaringan sosial antara pembeli dan penjual, sekaligus mendeteksi risiko terkait aktivitas penipuan

Kode Program 3.1. *Exploratory Data Analysis* pada Python.

```
use pbl_paper_id;


-- Membuat edge_list untuk social network analisis user yang terlibat transaksi fraud
WITH transaction_with_buyer_flags AS (
    SELECT
        t.*,
        u.user_fraud_flag AS user_fraud_flag_buyer,
        u.blacklist_account_flag AS blacklist_account_flag_buyer
    FROM transaction t
    LEFT JOIN user u
        ON t.buyer_id = u.company_id
),
transaction_with_seller_flags AS (
    SELECT
        tb.*,
        COALESCE(us.user_fraud_flag, 1) AS user_fraud_flag_seller,
        COALESCE(us.blacklist_account_flag, 1) AS blacklist_account_flag_seller
    FROM transaction_with_buyer_flags tb
    LEFT JOIN user us
        ON tb.seller_id = us.company_id
),
filtered_transactions AS (
    SELECT *
    FROM transaction_with_seller_flags
    WHERE
        user_fraud_flag_buyer = 1 OR
        blacklist_account_flag_buyer = 1 OR
        user_fraud_flag_seller = 1 OR
        blacklist_account_flag_seller = 1
),
final_filtered_transactions AS (
    SELECT *
    FROM filtered_transactions
    WHERE buyer_id != seller_id
)
SELECT buyer_id, seller_id
FROM final_filtered_transactions;
```

Hasil dari query ini menunjukkan terdapat 5.107 baris data yang mencakup pasangan pembeli-penjual beserta transaksi yang berpotensi fraud.

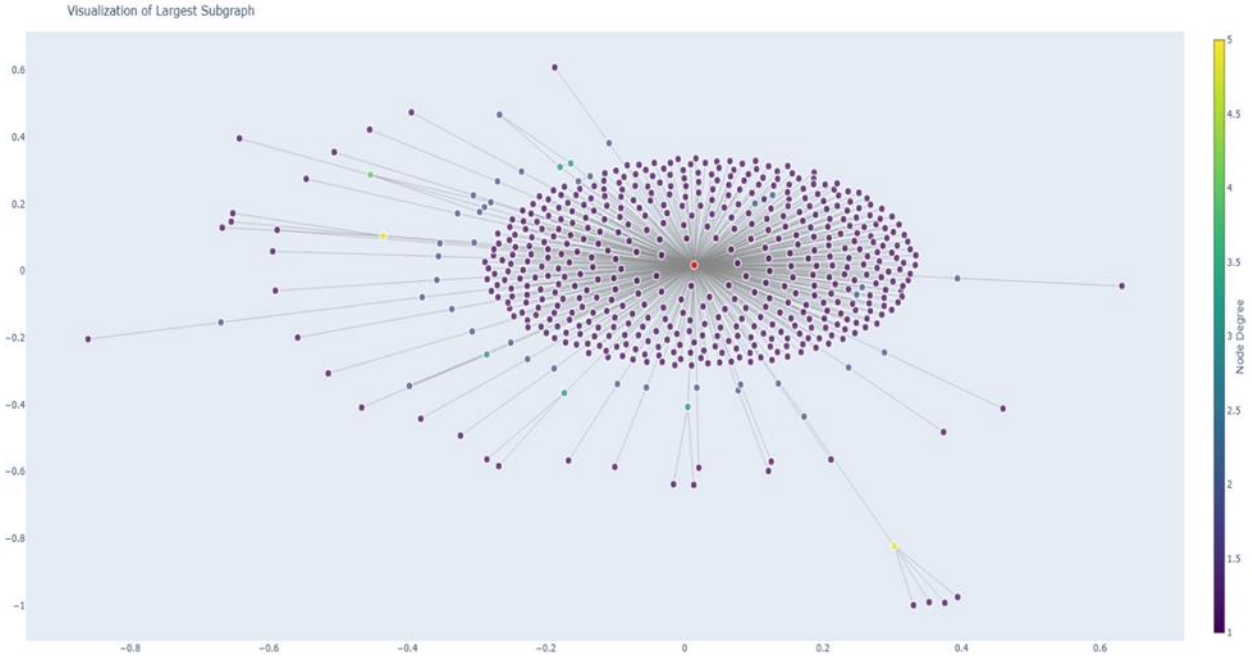
CALL PromoMisuseDetection() | Enter a SQL expression to filter results (use Ctrl+Space)

	buyer_id	dpt_promotion_id	consecutive_count	transaction_duration_days	misuse_reason
1	3077819ec94241590c88a38ed75fa3ef	promotion-219036467	5	23	Penggunaan promosi berulang melebihi ambang batas
2	c799c01d50a17e52ffa074b3a6bad0f5	promotion-188676794	5	101	Penggunaan promosi berulang melebihi ambang batas
3	0fe36baa803c6718b95994af1e4a31cf	promotion-219036467	4	85	Penggunaan promosi berulang melebihi ambang batas
4	860f276a6f6d37a9e28fe96b5712e4d2	promotion-219036467	3	24	Penggunaan promosi berulang melebihi ambang batas
5	3367265f12b5b265841a9105f93b7de	promotion-219036467	3	25	Penggunaan promosi berulang melebihi ambang batas

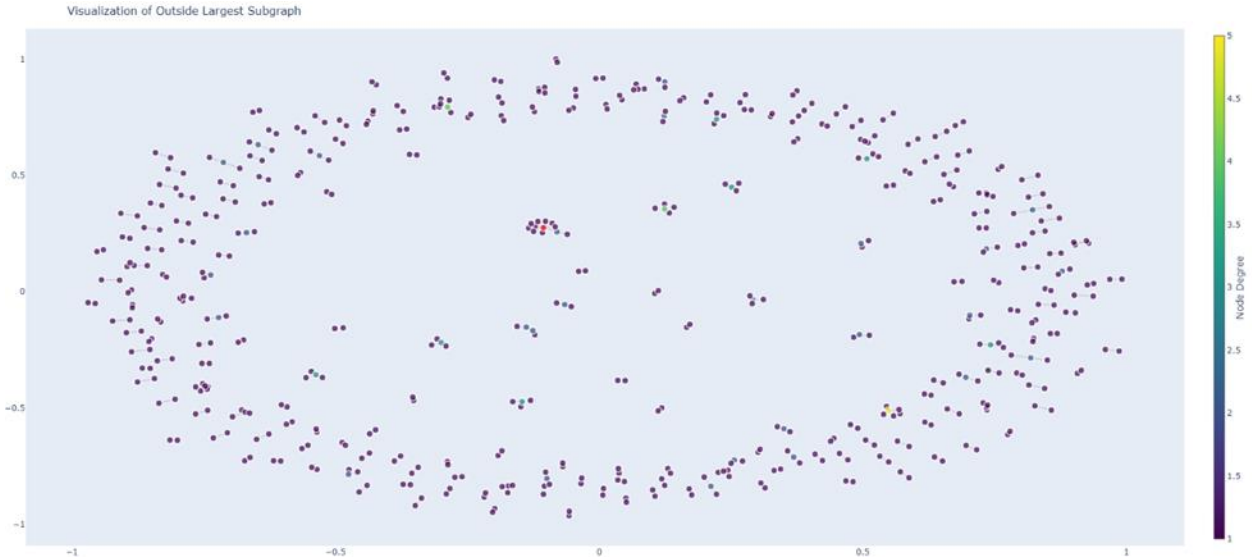
Data ini mencerminkan frekuensi dan pola hubungan antar entitas, termasuk pengelompokan entitas yang sering berinteraksi. Pengelompokan ini dapat membantu mengidentifikasi kluster pembeli dan penjual yang memiliki kemungkinan lebih tinggi untuk terlibat dalam aktivitas mencurigakan.Selanjutnya, data hasil query ini dieksport dalam format CSV untuk analisis lebih lanjut menggunakan Python.

6.1.2 Analisis Hubungan Pembeli-Penjual untuk Mengidentifikasi Pola Interaksi dan Terlibat Fraud (Visualization using Phyton)

Pada tahap ini digunakan beberapa library sebagian diantaranya yaitu networkx untuk membuat graf dan plotly untuk menampilkan graf yang responsif.
Data hasil queri sebelumnya diubah menjadi graf G. Graf ini mempunyai 859 edges dan 1089 nodes. Dari 1089 nodes yang terlibat dalam transaksi fraud ternyata 318 nodes tidak terdaftar di data user. Data transaksi fraud membentuk satu graf besar dengan 552 nodes dan banyak graf-graf kecil yang kebanyakan hanya terdiri dari 2 hingga 3 nodes. Oleh karena itu, visualisasinya dibagi menjadi dua, largest subgraph dan outside largest subgraph seperti di bawah ini:



Node berwarna merah merupakan central node yang terkoneksi ke 511 nodes lainnya pada graf utama. Selain central node kebanyakan node hanya terkoneksi ke satu atau dua node lainnya, tetapi ada juga beberapa node yang terkoneksi ke 3 hingga 5 nodes lain. Jumlah nodes tidak terdaftar pada graf utama G_largest hanya 25 nodes atau sekitar 8% dari total nodes tidak terdaftar. Setelah dilakukan pengecekan pada data user ternyata central node tidak ditandai fraud atau pun tidak diblacklist.



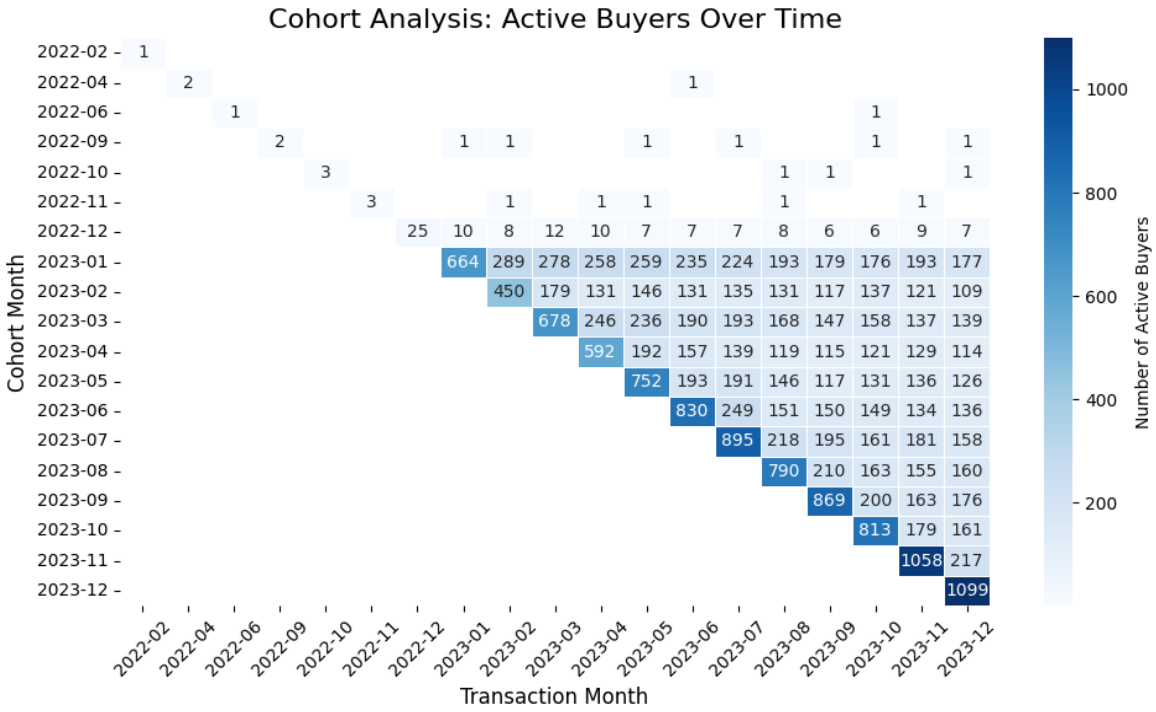
Total nodes pada G_ outside_largest adalah 537 nodes dan 293 diantaranya adalah akun tidak terdaftar (92% dari total akun tidak terdaftar). Pada kumpulan graf-graf kecil ini terdapat satu graph yang lebih banyak nodes-nya dibanding yang lain, totalnya ada 12 nodes. Central node-nya

terkoneksi ke 10 nodes yang lainnya. Setelah di lakukan pengecekan ternyata central node-nya adalah akun testing. Selain itu ada satu lagi akun testing pada graf ini, sementara yang 10 lainnya tidak terdaftar di data user.

6.2 Cohort Analysis

Melakukan analisis kohort untuk melacak **transaksi berulang** antara pembeli dan penjual:

6.2.1 Kelompokkan pembeli berdasarkan tanggal transaksi pertama mereka dan mengukur aktivitas berkelanjutan mereka dari waktu ke waktu.



Analisis kohort menunjukkan bahwa jumlah pembeli aktif menurun seiring waktu dalam setiap kohort, mengindikasikan adanya *churn* atau penurunan retensi pelanggan. Kohort awal, seperti Februari 2022, memiliki retensi yang lebih rendah dibandingkan dengan kohort yang lebih baru. Namun, terlihat tren positif pada pertumbuhan jumlah pembeli baru di bulan-bulan akhir tahun 2023, seperti Desember 2023, yang mencatat 1.099 pembeli aktif pada bulan pertama. Hal ini menunjukkan keberhasilan dalam menarik pelanggan baru, meskipun keberlanjutan aktivitas mereka perlu dipantau lebih lanjut.

Selain itu, beberapa kohort di pertengahan 2023 menunjukkan retensi yang lebih baik, seperti kohort April dan Mei 2023, yang masih memiliki banyak pembeli aktif di bulan-bulan berikutnya. Ini bisa jadi hasil dari strategi pemasaran atau promosi yang efektif pada periode tersebut. Namun, penurunan aktivitas di bulan-bulan selanjutnya tetap menjadi perhatian, sehingga penting untuk meningkatkan retensi melalui pendekatan seperti program loyalitas, personalisasi penawaran, atau peningkatan pengalaman pelanggan.

6.2.2 Identifikasi apakah pembeli tertentu terlibat dalam perilaku penipuan setelah periode tidak aktif atau berulang kali berinteraksi dengan penjual yang sama.

```
=== Suspicious Buyers Analysis ===
                                                buyer_id  inactive_period  unique_seller_count
0      bbce610a3267808752a7ec263a7ecfbe76a4987d529bcb...              0.0                    1
1      09eb3b80abae1238ef39d50b66215e02e1ac9891ad6e8f...              0.0                    1
2      25d0774533d69564d0deca724a55a76c693ed5f7ffa12a...              0.0                    1
3      5b846313375cb4f4d065e50a05833dc3ac20ba3f532bbe...              0.0                    1
4      5c19a13a9b229340b584f621b648f4dec7491e12368392...              0.0                    1
...                                                  ...              ...                  ...
49985  81046b351fb34308b3d2352537f2b7bfb512ec1ca217e8...              6.0                    1
49988  e21565709ff2c4e04a04021bdcc1e857790e65f7d7e7fc...             52.0                    4
49990  4755af5b9840de287023a2e3e7fc73248a3d1a6aed4516...            248.0                    1
49993  4ef8dc3bcaccab5ab5e5c6ffe9ba7ab53917e6e10ffce9...              1.0                    1
49994  f4a949d70d0cc714dca624a15d9f26947c423e4c473d23...              0.0                    1
```

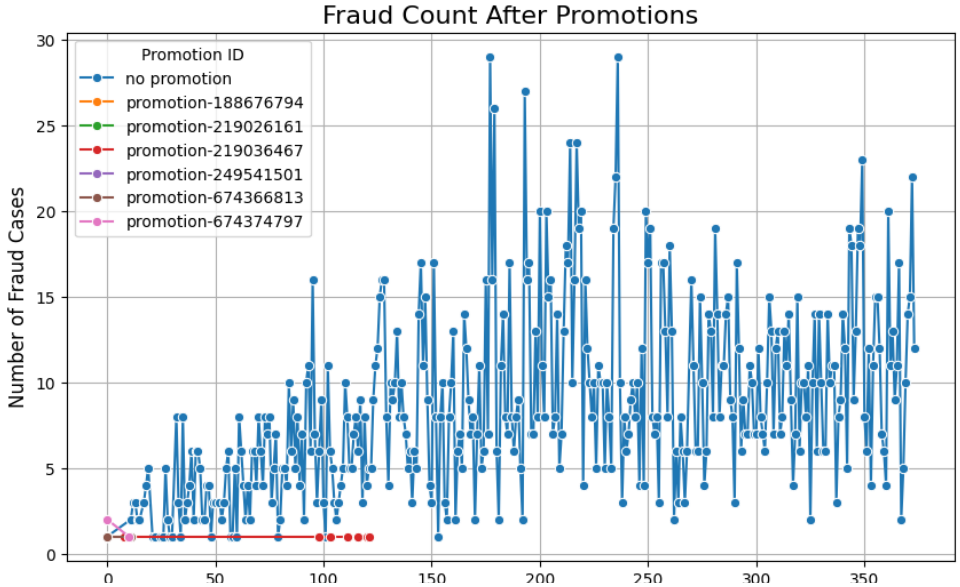
```
unique_seller_count  seller_interaction_count  burst_flag  suspicious_seller_flag
                  1                        18       False                    True
                  1                        13       False                    True
                  1                         7       False                    True
                  1                        11       False                    True
                  1                         7       False                    True
                ...                       ...         ...                     ...
                  1                         5       False                   False
                  4                        54        True                    True
                  1                        32        True                    True
                  1                        11       False                    True
                  1                         1       False                   False
```

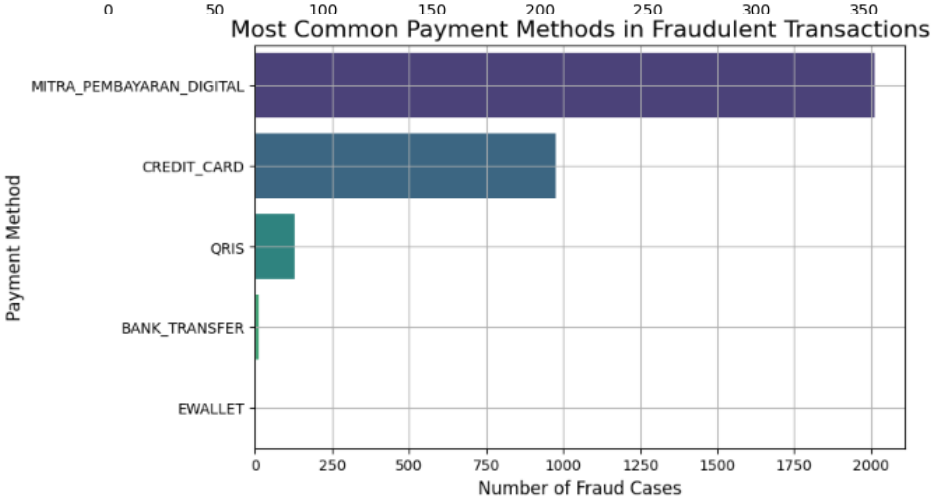
Analisis menunjukkan bahwa pembeli dengan periode tidak aktif yang panjang atau yang sering berinteraksi dengan penjual yang sama cenderung memiliki perilaku mencurigakan. Misalnya, pembeli dengan **inactive_period** tinggi yang tiba-tiba kembali dengan jumlah interaksi yang signifikan atau memiliki **unique_seller_count** rendah dapat mengindikasikan potensi kolusi atau penipuan. Kehadiran **burst_flag**, yang menandakan lonjakan aktivitas mendadak, juga dapat menjadi indikator perilaku tidak wajar, terutama jika diikuti oleh transaksi dengan penjual yang ditandai sebagai mencurigakan (**suspicious_seller_flag**).

Langkah penting berikutnya adalah mengevaluasi apakah pola ini meningkat setelah kampanye promosi tertentu, seperti diskon besar atau event spesial. Hal ini karena pelaku penipuan sering memanfaatkan momen dengan aktivitas tinggi untuk menyembunyikan aktivitas mereka. Dengan memahami korelasi ini, dapat diterapkan langkah pencegahan, seperti pemantauan lebih ketat selama periode promosi atau penerapan sistem peringatan untuk aktivitas yang tidak biasa.

6.3 Insight Generation

Menganalisis perilaku kelompok untuk mendeteksi pola penipuan yang berulang. Misalnya, apakah ada lonjakan penipuan setelah kampanye promosi?






```
Top Promotions with Most Fraud Cases:
      dpt_promotion_id  fraud_count
0         no promotion         3109
3  promotion-219036467            8
5  promotion-674374797            3
5  promotion-674366813            2
1  promotion-188676794            1
```

Dari grafik terlihat bahwa promosi tertentu cenderung meningkatkan jumlah kasus penipuan, meskipun kasus terbanyak justru terjadi tanpa promosi. Lonjakan penipuan signifikan terlihat pada promosi seperti *promotion-219036467* dan *promotion-674347497*, mengindikasikan adanya pola penyalahgunaan promosi oleh pelaku. Di sisi lain, metode pembayaran seperti *MITRA_PEMBAYARAN_DIGITAL* dan *CREDIT_CARD* lebih sering digunakan dalam transaksi penipuan dibandingkan metode lain seperti QRIS atau bank transfer.

Insight ini menunjukkan bahwa promosi dapat menjadi celah bagi pelaku penipuan, sementara tingginya kasus tanpa promosi mengindikasikan perlunya peningkatan keamanan secara sistemik. Rekomendasi utamanya adalah memperkuat pengamanan pada promosi yang rentan dan metode pembayaran populer, sekaligus menerapkan langkah pencegahan lebih ketat untuk transaksi reguler.

BAB VII

[Visualization] Tableau for Fraud Monitoring and Dashboard Creation

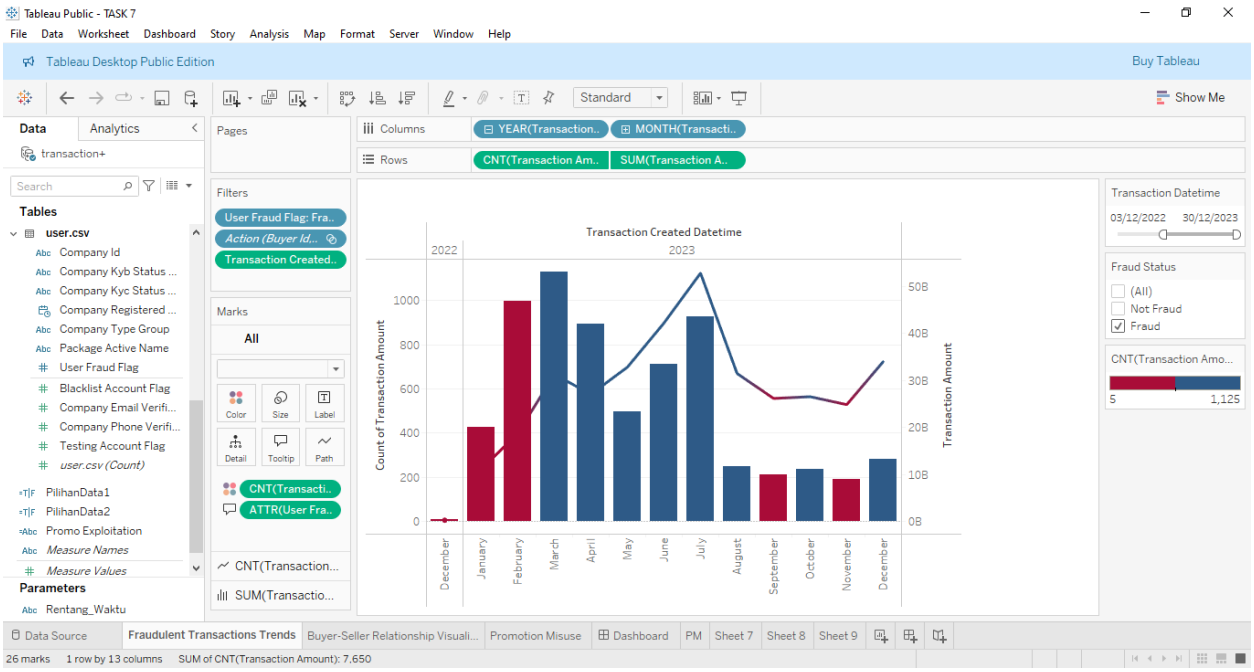
7.1 Interactive Fraud Detection Dashboards

Dashboard Deteksi Fraud Interaktif memungkinkan tim pemantauan untuk memvisualisasikan dan menganalisis transaksi mencurigakan secara real-time. Dengan tampilan

yang interaktif, tim dapat menyesuaikan filter dan parameter untuk menggali data lebih dalam, seperti periode waktu, jenis transaksi, atau status fraud. Dashboard ini memberikan kemudahan dalam memantau pola transaksi dan mengidentifikasi potensi fraud secara cepat. Berikut dashboard yang dapat ditampilkan berdasarkan hasil analisis:

7.1.1 Tableau Dashboards

7.1.1.1Tren Transaksi Penipuan

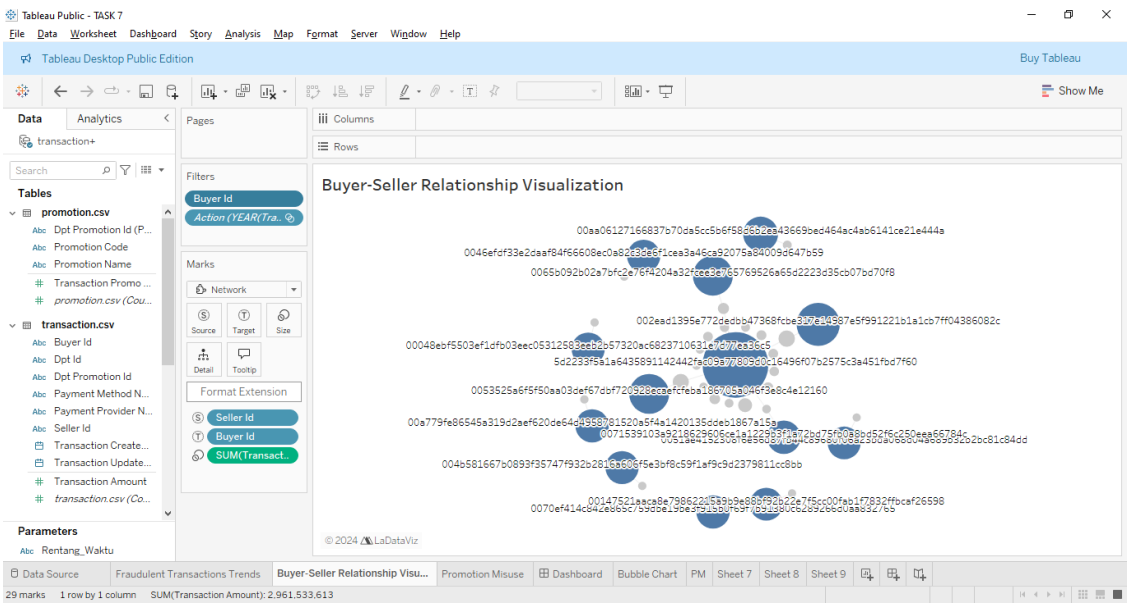


Dari visualisasi tren transaksi penipuan, dapat terlihat bahwa kejadian penipuan cenderung meningkat pada periode tertentu, khususnya pada bulan Juli 2023 yang menunjukkan jumlah transaksi penipuan tertinggi. Selain itu, fenomena penipuan juga lebih sering terjadi pada waktu-waktu tertentu, seperti pada saat akhir bulan. Grafik yang disajikan menunjukkan adanya lonjakan jumlah transaksi yang mencurigakan, terutama pada waktu-waktu tersebut.

Pada visualisasi ini, warna biru pada grafik menunjukkan jumlah transaksi yang tinggi, sementara warna merah menunjukkan nilai transaksi yang lebih rendah. Perbedaan warna ini memudahkan identifikasi waktu dan jumlah transaksi yang mencurigakan, sehingga tim pemantau penipuan dapat dengan cepat mengetahui tren penipuan yang terjadi. Tanda-tanda transaksi

mencurigakan terdeteksi di transaksi dengan status penipuan, memberikan indikasi bahwa ada pola yang perlu diwaspadai.

7.1.1. 2 Visualisasi Hubungan Pembeli-Penjual

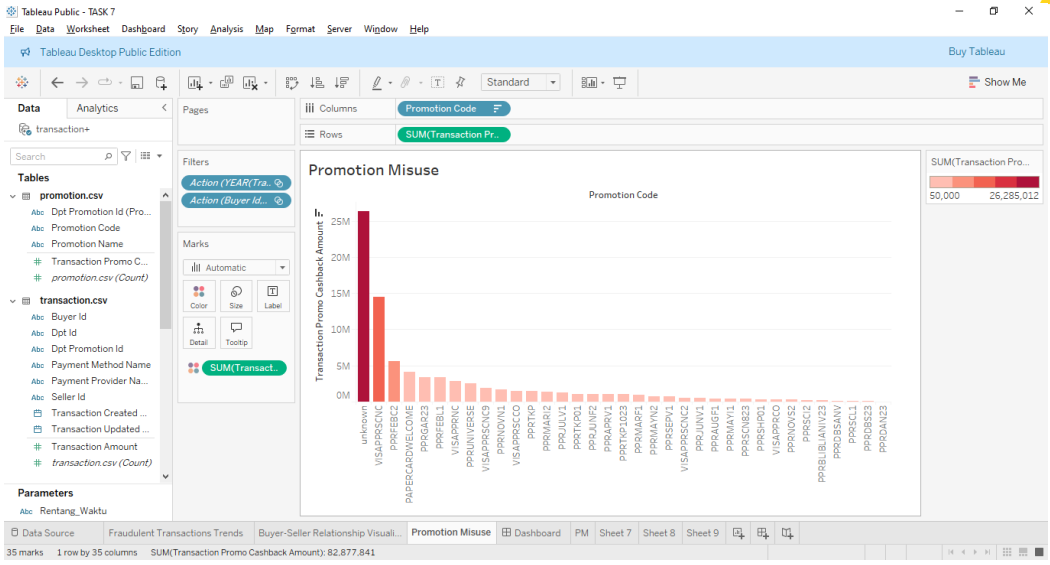


Dalam visualisasi Hubungan Pembeli-Penjual, terdapat dua pola yang perlu dicermati. Pertama, terlihat adanya hubungan yang sangat intens antara satu penjual dan satu pembeli, yang bisa menjadi indikasi adanya potensi penipuan. Kedua, ditemukan pula pola hubungan antara satu penjual dan banyak pembeli, yang menunjukkan adanya kemungkinan transaksi yang lebih luas, namun tetap patut dicurigai.

Dengan demikian, transaksi yang terjadi antara pembeli dan penjual dengan pola hubungan yang sangat sering, terutama jika dihubungkan dengan status penipuan yang terdeteksi, menunjukkan adanya potensi fraud yang perlu ditindaklanjuti.

7.1.1.3 Penyalahgunaan Promosi

Visualisasikan bagaimana promosi dieksploitasi oleh pengguna yang melakukan penipuan.

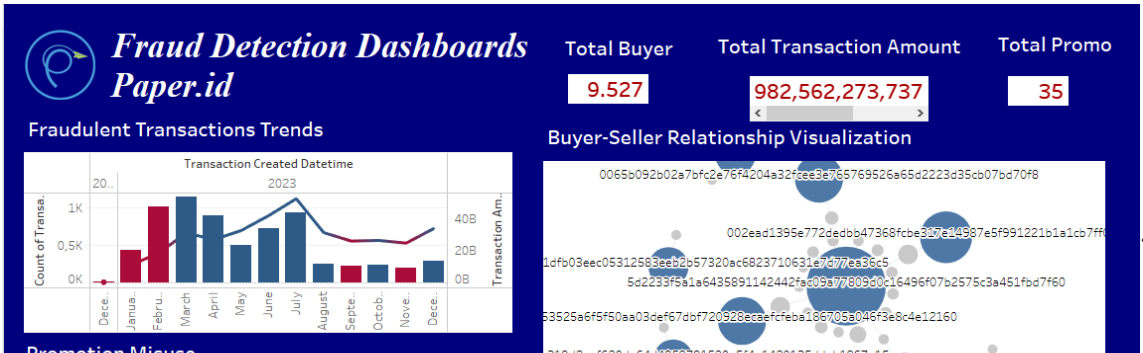


Visualisasi yang telah dibuat menunjukkan hubungan antara kode promosi dan jumlah transaksi serta total cashback yang diterima. Dalam konteks eksploitasi promosi, terlihat bahwa kode promosi tertentu memiliki transaksi dengan cashback yang sangat besar, yang dapat menjadi indikasi adanya penyalahgunaan. Kode promosi dengan jumlah cashback tinggi, seperti yang tercatat pada kategori tertentu, mungkin menunjukkan bahwa promosi tersebut dieksploitasi secara berlebihan oleh pengguna untuk memaksimalkan keuntungan mereka.

Selain itu, status "Unknown" pada beberapa transaksi juga dapat menjadi sinyal adanya eksploitasi, karena transaksi dengan status tersebut mungkin tidak tercatat atau teridentifikasi dengan jelas. Hal ini dapat menunjukkan bahwa promosi digunakan secara tidak transparan atau bahkan oleh akun yang mencurigakan. Dengan demikian, visualisasi ini memberikan gambaran awal tentang potensi penyalahgunaan promosi yang perlu ditindaklanjuti lebih lanjut.

7.2 Dynamic Filtering and Drill-Downs

Dengan mengaktifkan filter dinamis yang memungkinkan tim untuk menyaring data berdasarkan berbagai parameter, seperti periode waktu, hubungan pengguna, dan status flag fraud. Adapun hasilnya sebagai berikut:



Dasbor Deteksi Penipuan yang dibuat di Tableau memberikan gambaran mendalam mengenai tren transaksi penipuan, hubungan antara pembeli dan penjual yang mencurigakan, serta penyalahgunaan promosi dalam transaksi digital. Visualisasi pertama menampilkan tren transaksi penipuan berdasarkan waktu, dengan analisis terhadap jumlah transaksi dan total nilai transaksi yang menunjukkan periode dengan volume penipuan yang lebih tinggi. Ini memberikan wawasan tentang titik rawan penipuan yang terjadi pada bulan-bulan tertentu, yang dapat digunakan untuk memperketat pengawasan pada periode tersebut.

Selanjutnya, visualisasi hubungan pembeli-penjual menunjukkan jaringan transaksi yang mencurigakan melalui analisis hubungan antar pengguna. Pembeli dan penjual yang terlibat dalam transaksi penipuan dapat dikenali dengan lebih mudah, memberikan tim pemantauan penipuan alat yang efektif untuk mengidentifikasi pola dan anomali. Terakhir, visualisasi penyalahgunaan promosi menunjukkan bagaimana kode promosi digunakan berlebihan oleh pengguna yang berpotensi melakukan penipuan, dengan total cashback yang tidak wajar. Dasbor ini juga dilengkapi dengan pemfilteran dinamis, memungkinkan tim untuk menelusuri dan melakukan drill-down pada data sesuai dengan periode waktu, status penipuan, dan hubungan pengguna, memberikan analisis yang lebih tajam dan responsif terhadap ancaman penipuan yang muncul.

BAB VIII

Insights and Recommendations

8.1 Key Fraud Insights

Gunakan dasbor untuk mengidentifikasi:

- Pendorong penipuan teratas dan pengguna penipuan yang paling terhubung.

  Dasbor ini memberikan wawasan penting mengenai pendorong penipuan, termasuk kode promosi yang dieksploitasi secara berulang dan transaksi yang dilakukan selama periode tertentu dengan angka penipuan yang tinggi. Kode promosi yang teridentifikasi sebagai sering digunakan oleh pengguna yang melakukan penipuan menyoroti area yang perlu diperhatikan oleh tim pemantauan. Selain itu, dengan

menggunakan visualisasi hubungan pembeli-penjual, kita dapat mengidentifikasi pengguna yang memiliki keterhubungan tinggi dengan banyak pihak, baik sebagai pembeli atau penjual. Mereka ini lebih cenderung terlibat dalam jaringan penipuan yang lebih luas. Pengguna-pengguna ini, dengan berbagai keterhubungan, menjadi titik fokus utama dalam upaya pencegahan lebih lanjut.

- Hubungan pembeli-penjual utama yang terlibat dalam penipuan.

  Hubungan pembeli-penjual yang terlibat dalam penipuan dapat dengan jelas diidentifikasi melalui visualisasi jaringan yang ada. Dalam visualisasi ini, pembeli dan penjual yang terhubung satu sama lain dengan banyak transaksi mencurigakan menonjol sebagai titik-titik utama dalam jaringan penipuan. Pembeli yang terhubung dengan banyak penjual atau sebaliknya, menunjukkan adanya pola transaksi yang mencurigakan yang melibatkan lebih dari satu pihak. Analisis ini memberikan gambaran yang lebih jelas tentang siapa yang berperan aktif dalam transaksi yang tidak sah. Mengidentifikasi hubungan ini memungkinkan untuk melakukan penyaringan yang lebih mendalam terhadap pembeli dan penjual yang terlibat, serta memfokuskan upaya pencegahan penipuan terhadap transaksi yang dilakukan oleh aktor-aktor ini. Seiring berjalannya waktu, memonitor hubungan ini dapat membantu mengurangi dampak penipuan yang lebih luas dan mencegah pelaku penipuan yang mungkin berulang.

8.2 Action Plans:

Memberikan rekomendasi untuk pencegahan penipuan berdasarkan temuan:

- Terapkan proses verifikasi pengguna yang lebih ketat.

  Sebagai langkah pencegahan penipuan yang lebih efektif, sangat penting untuk memperketat proses verifikasi pengguna, terutama pada tahap pendaftaran dan transaksi pertama. Implementasi prosedur verifikasi identitas melalui multi-faktor otentikasi atau verifikasi dokumen dapat membantu memastikan bahwa pengguna yang terlibat dalam transaksi benar-benar sah. Proses ini bisa melibatkan verifikasi melalui email, SMS, atau aplikasi otentikasi untuk meminimalkan risiko pendaftaran akun palsu yang dapat digunakan untuk penipuan. Dengan melakukan

verifikasi yang lebih ketat, platform dapat meminimalisir pelaku penipuan yang mencoba masuk ke dalam sistem dengan identitas palsu atau menggunakan informasi yang tidak valid.

Selain itu, proses verifikasi yang lebih ketat dapat diimplementasikan pada pengguna yang terlibat dalam transaksi dengan volume tinggi atau transaksi yang mencurigakan. Pengguna yang melakukan transaksi besar atau yang melibatkan kode promosi yang sering dieksploitasi dapat diminta untuk memberikan informasi tambahan sebagai langkah pencegahan. Ini akan membantu mengurangi potensi kerugian yang timbul dari pengguna yang berniat buruk. Dengan pengawasan yang lebih ketat pada transaksi yang mencurigakan dan pengguna baru, platform dapat lebih proaktif dalam mendeteksi dan mencegah penipuan sebelum terjadi, meningkatkan keamanan dan kepercayaan pengguna secara keseluruhan.

- Pantau dan tandai aktivitas pembeli-penjual yang mencurigakan secara real-time.

  Untuk menangani penipuan dengan lebih cepat dan efektif, sistem pemantauan real-time harus diterapkan untuk memonitor aktivitas pembeli dan penjual. Dengan menggunakan teknologi analisis data yang canggih, seperti machine learning dan algoritma prediktif, platform dapat mendeteksi perilaku mencurigakan yang mungkin terlewatkan oleh pengawasan manual. Aktivitas mencurigakan, seperti pembelian berulang dengan kode promo yang sama atau transaksi yang tidak wajar dalam periode tertentu, dapat segera ditandai dan diteruskan untuk pemeriksaan lebih lanjut. Hal ini tidak hanya membantu dalam mendeteksi penipuan lebih awal tetapi juga dapat mencegah kerugian finansial yang lebih besar dengan menghentikan transaksi yang mencurigakan sebelum diproses lebih lanjut.

  Pemantauan secara real-time juga memungkinkan untuk mendeteksi jaringan hubungan pembeli dan penjual yang terlibat dalam penipuan. Pembeli atau penjual yang memiliki banyak transaksi dengan pengguna lain yang tercatat sebagai penipu dapat segera diidentifikasi dan diberi peringatan atau bahkan diblokir. Dengan pendekatan ini, sistem akan lebih siap dalam menangani aktivitas mencurigakan, memungkinkan tim pemantauan untuk bertindak dengan lebih cepat dan meminimalkan dampak penipuan. Pemantauan real-time yang tepat dapat

mengurangi keterlambatan dalam deteksi dan mempercepat langkah-langkah pencegahan yang diperlukan.

- Tetapkan kebijakan promosi yang lebih ketat untuk mencegah penyalahgunaan.

  Kebijakan promosi yang lebih ketat sangat penting untuk mencegah penyalahgunaan promosi yang sering kali menjadi pendorong utama penipuan. Salah satu langkah yang dapat diambil adalah dengan membatasi penggunaan kode promosi oleh pengguna yang terdeteksi sering menggunakannya dalam jumlah yang tidak wajar. Misalnya, mengatur pembatasan seperti jumlah maksimum penggunaan kode per pengguna atau periode waktu tertentu dapat mengurangi potensi penyalahgunaan. Selain itu, platform dapat mengenakan kebijakan untuk mengkonfirmasi apakah transaksi dengan kode promosi melibatkan transaksi yang sah atau hanya digunakan untuk memanfaatkan sistem secara tidak adil. Kebijakan promosi yang lebih selektif dan terkontrol dapat membantu meminimalkan celah yang dapat dimanfaatkan oleh penipu.

  Selain membatasi jumlah penggunaan, kebijakan promosi yang lebih ketat juga dapat melibatkan analisis perilaku pengguna yang menerima promosi. Promosi yang diberikan kepada pengguna dengan riwayat transaksi mencurigakan atau yang terhubung dengan pembeli atau penjual yang terlibat dalam penipuan dapat dibatasi atau dibatalkan untuk menghindari potensi kerugian lebih lanjut. Langkah-langkah preventif ini akan lebih memperkecil ruang bagi pengguna yang berniat menyalahgunakan promosi untuk melakukan penipuan, sekaligus menjaga kredibilitas dan integritas platform. Dengan kebijakan yang lebih ketat, perusahaan dapat menciptakan lingkungan transaksi yang lebih adil bagi pengguna yang sah.

BAB IX

Kesimpulan dan Saran

A. Kesimpulan

Perusahaan fintech yang memproses jutaan transaksi digital menghadapi tantangan besar akibat meningkatnya kasus penipuan yang merugikan baik secara finansial maupun kepercayaan pelanggan. Berdasarkan analisis data, pola penipuan menunjukkan bahwa pelaku sering memanfaatkan kode promosi secara berlebihan dan memiliki jaringan hubungan antara pembeli dan penjual yang saling terkait. Penggunaan SQL tingkat lanjut memungkinkan penggalian data mendalam untuk mengidentifikasi perilaku yang tidak normal, seperti lonjakan transaksi dalam waktu singkat dan pola penggunaan promosi yang tidak wajar. Analisis mendalam ini menegaskan bahwa kurangnya pengawasan real-time dan kebijakan promosi yang lemah berkontribusi pada peningkatan kasus penipuan. Temuan ini memberikan gambaran penting bagi perusahaan untuk mengidentifikasi area kelemahan dalam sistem dan meningkatkan mekanisme keamanan mereka.

B. Saran

Untuk mengurangi risiko penipuan, perusahaan perlu menerapkan langkah-langkah strategis berbasis data. Pertama, adopsi algoritma machine learning untuk mendeteksi anomali

transaksi secara real-time dapat membantu dalam mengidentifikasi pola mencurigakan sejak dini. Kedua, memperketat proses verifikasi pengguna dengan menggunakan autentikasi multi-faktor akan mengurangi peluang pendaftaran akun palsu. Selanjutnya, kebijakan promosi perlu diperbaiki dengan membatasi penggunaan kode promo pada perilaku tertentu, misalnya dengan menerapkan deteksi pengguna yang sering menggandakan akun untuk memanfaatkan promosi. Dengan implementasi langkah-langkah ini, perusahaan dapat mengurangi kerugian akibat penipuan dan memulihkan kepercayaan pelanggan terhadap platform.

LAMPIRAN

A. Online Diagram BPMN:

Business Process BPMN :
https://drive.google.com/file/d/1ystxfq1UZwjYS9Dha4MCD6jh-zHReCMZ/view

B. Python Code

2.1 Data Cleaning :
https://colab.research.google.com/drive/15d9THxmybdyF7BAVqykP9Uf-1J92l71c?usp=drive_link
2.2 Featuring Engineering :
https://colab.research.google.com/drive/1DILx9sPzDnBP_t3ojtTNEioDySU1izP5?usp=drive_link
2.3 Scaling and Normalization :
https://colab.research.google.com/drive/1MYCN3P0NdeOlhVvNYWbTaywE7wsjja-k?usp=drive_link
3.1 Exploratory Data Analysis :
https://colab.research.google.com/drive/1tsI1EAvnbMZs3dmw7wuptRoQAXFfmf17?usp=drive_link
3.2 Visualization : https://colab.research.google.com/drive/1Tc77-V_2DSuYzSJ8PKAcZcFmH2t0FUp7?usp=drive_link

6.1.2 Analisis Hubungan Pembeli-Penjual untuk Mengidentifikasi Pola Interaksi dan Terlibat Fraud (Visualization using Phyton) : https://colab.research.google.com/drive/1Uzx7Z1xqsoTPQT7VycDPe9Fa6ygH8C-0?usp=drive_link

6.2 Cohort Analysis : https://colab.research.google.com/drive/10HCHdiao3NzFGjORCnhfgIND3h3Y0RyX#scrollTo=f6beTAKDRlyF

C. Notulen Canva (PPT) dan Tableau

Tableau: https://public.tableau.com/app/profile/syahirotul.maulidiyah/viz/FraudDetectionDashboardsPaperID_17336703277340/Dashboard?publish=yes

Notulen Canva PPT : https://www.canva.com/design/DAGYt5mxznk/b-MFeZreZmkBF3XzOPcI_g/edit?utm_content=DAGYt5mxznk&utm_campaign=designshare&utm_medium=link2&utm_source=sharebutton

D. Recording

Link Youtube Recording : https://youtu.be/ggbildpcUFc?si=50XWObYJFlkdS6oH