

**Дипломный проект на тему:
«Разработка ETL-процесса: построение
витрины отчетности по
мошенническим операциям»**

Слушатели:

**Рохина Марина Евгеньевна
Калинин Сергей Владимирович
Новиков Денис Александрович**

Актуальность темы и ее проблематика

Данная тема является актуальной, т.к. согласно статистике:

1. В 2021 году Россияне потеряли 3,15 млрд рублей из-за мошенничества с фейковыми платёжными системами (tadviser.ru)
2. Пандемия спровоцировала увеличение активности мошенников. Рост объема и количества несанкционированных операций связан с активным переходом граждан на дистанционный формат потребления продуктов и услуг, в том числе финансовых (rbc.ru)
3. Объем рынка продаж краденых данных банковских карт приблизился к \$2 млрд

Ежедневное отслеживание мошеннических операций поможет в кратчайшие сроки выявить обман и отменить выполненные транзакции

Разработка ETL-процесса: построение витрины отчетности по мошенническим операциям

Задача : необходимо разработать ETL процесс, получающий выгрузку данных, загружающий ее в хранилище данных и строящий отчет. Данные должны быть загружены в хранилище BANK (рис. 1). По результатам загрузки необходимо построить витрину отчетности по мошенническим операциям.

Рохина Марина Евгеньевна
Калинин Сергей Владимирович
Новиков Денис Александрович

Разработка ETL-процесса: построение витрины отчетности по мошенническим операциям

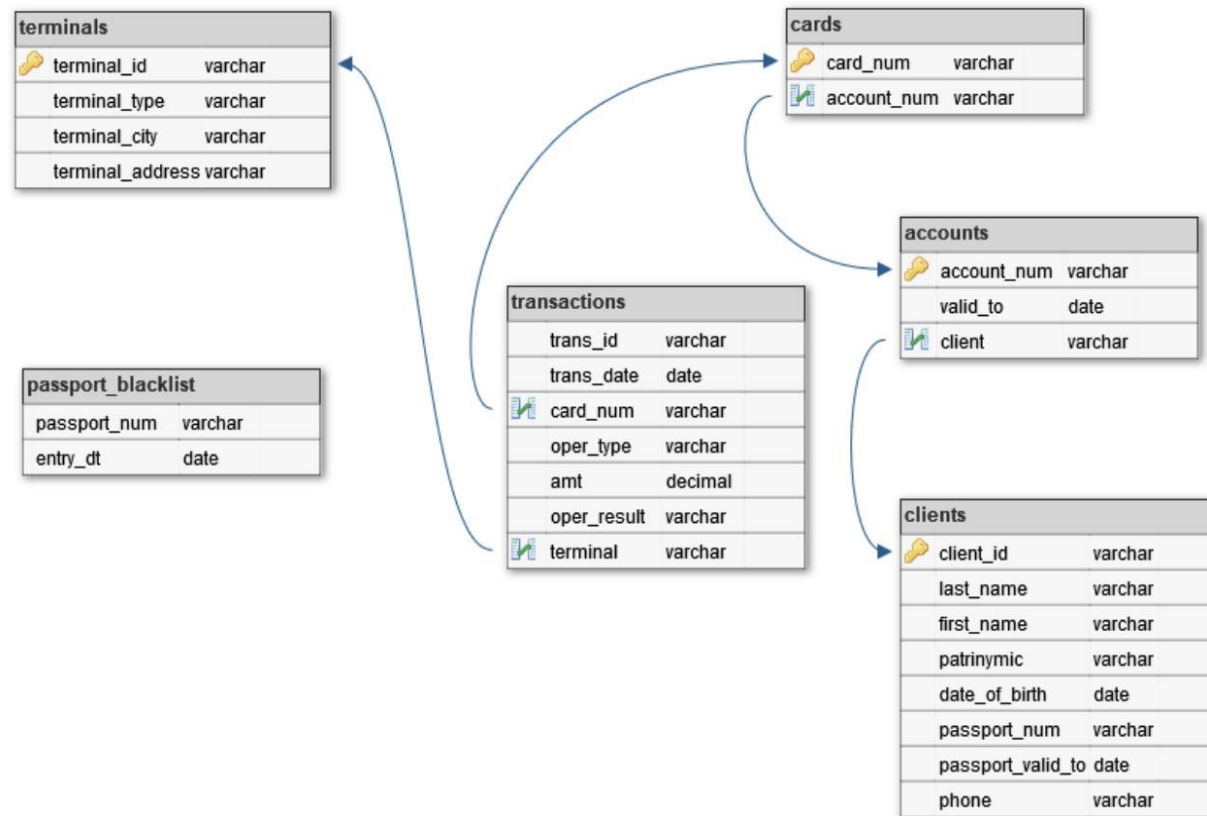


рис.1 ER-диаграмма схемы BANK

Рохина Марина Евгеньевна
Калинин Сергей Владимирович
Новиков Денис Александрович

Разработка ETL-процесса: построение витрины отчетности по мошенническим операциям

Признаки мошеннических операций:

- Совершение операции при просроченном или заблокированном паспорте
- Совершение операции при недействующем договоре
- Совершение операций в разных городах в течение одного часа
- Попытка подбора суммы. В течение 20 минут проходит более 3х операций со следующим шаблоном – каждая последующая меньше предыдущей, при этом отклонены все кроме последней. Последняя операция (успешная) в такой цепочке считается мошеннической.

Рохина Марина Евгеньевна
Калинин Сергей Владимирович
Новиков Денис Александрович

Разработка ETL-процесса: построение витрины отчетности по мошенническим операциям

Обработка и загрузка данных в БД :

Начальным этапом ETL процесса является процедура извлечения записей из поступающих файлов и последующей их обработки с передачей в таблицы БД ORACLE. Передача файлов осуществляется с помощью библиотеки «jaudebeari» . В процессе обработки файлов производится очистка и инициализация данных с помощью библиотеки «Pandas».

Рохина Марина Евгеньевна
Калинин Сергей Владимирович
Новиков Денис Александрович

Разработка ETL-процесса: построение витрины отчетности по мошенническим операциям

Формирование исторических таблиц:

В ходе формирования исторических таблиц созданы функции для обновления данных в исторической таблице, удаления всех таблиц (кроме исторической) перед началом работы, вывода таблиц в терминал при необходимости, обработки запроса. Так же инициализированы таблицы для новых и удаленных записей. В таблице «Terminals» добавлены технические поля `effective_from`, `effective_to`, `deleted_flg`.

Рохина Марина Евгеньевна
Калинин Сергей Владимирович
Новиков Денис Александрович

Разработка ETL-процесса: построение витрины отчетности по мошенническим операциям

Формирование отчета определения мошеннических действий:

На основании признаков мошеннических операций определены следующие правила их идентификации:

- При просроченном или заблокированном паспорте, происходит соединение через Inner Join таблиц Transaction и Clients с последующем соединением по полю passport_num из таблицы Passport_blacklist
- При недействующем договоре, происходит проверка вхождений операций из таблицы Transaction в пределы значения поля valid_to из таблицы Accounts

Рохина Марина Евгеньевна
Калинин Сергей Владимирович
Новиков Денис Александрович

Выводы

В ходе работы над дипломным проектом разработан ETL-процесс построения витрины отчетности по мошенническим операциям:

1. Совершение операции при просроченном или заблокированном паспорте;
2. Совершение операции при недействующем договоре;

Рохина Марина Евгеньевна
Калинин Сергей Владимирович
Новиков Денис Александрович

Список использованных источников

1. Грофф, Джеймс . SQL : Энциклопедия : пер. с англ. / Д. Р. Грофф, П. Н. Вайнберг. — 3-е изд. — СПб. : Питер.
2. Уэс Маккинли Python и анализ данных [Электронный ресурс]/ Уэс Маккинли— Электрон. текстовые данные.— Саратов: Профобразование, 2017.— 482 с.
3. Кузнецов, Сергей Дмитриевич. Основы баз данных : курс лекций : учебное пособие / С. Д. Кузнецов.
4. <https://pythontutor.ru/>
5. www.sql-ex.ru

Рохина Марина Евгеньевна
Калинин Сергей Владимирович
Новиков Денис Александрович