

Дипломный проект на тему: «ETL процесс для Anti-fraud системы в банковской сфере»

Слушатель:

Осипов Антон Владимирович

Актуальность темы и ее проблематика

1. Сегодня данные – основа успеха всех преуспевающих компаний
2. Данных много и их необходимо собирать, обрабатывать
3. Данные - это основа моделей, которые приносят пользу бизнесу
4. Одно из применений ETL – поиск мошеннических операций
5. Для всего этого необходимы специалисты в сфере Data Engineer

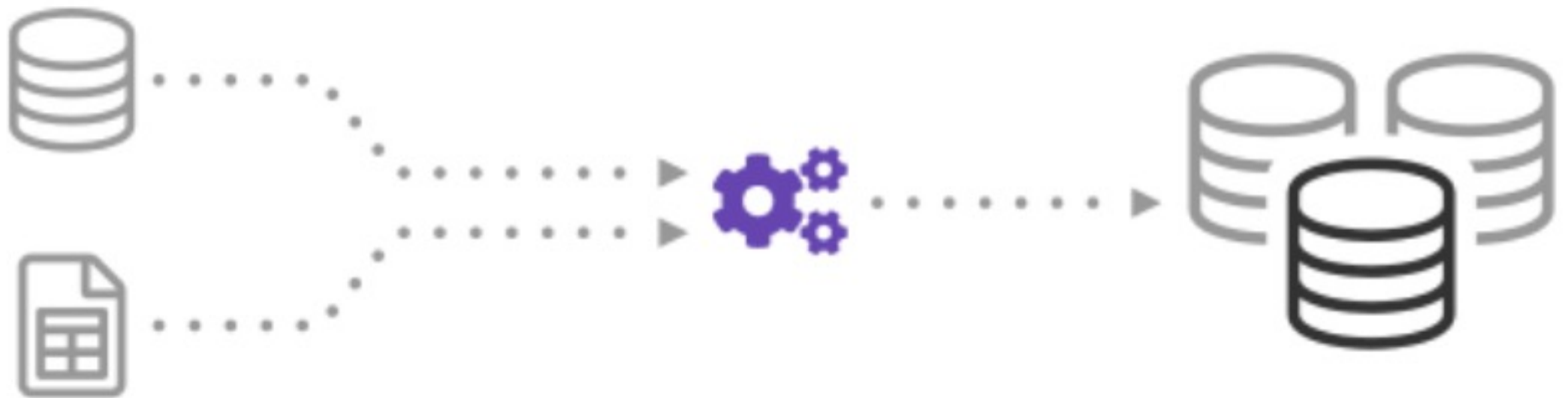
Что такое ETL

(Извлечение Преобразование загрузка)?

Extract

Transform

Load



Слушатель:
Осипов Антон Владимирович

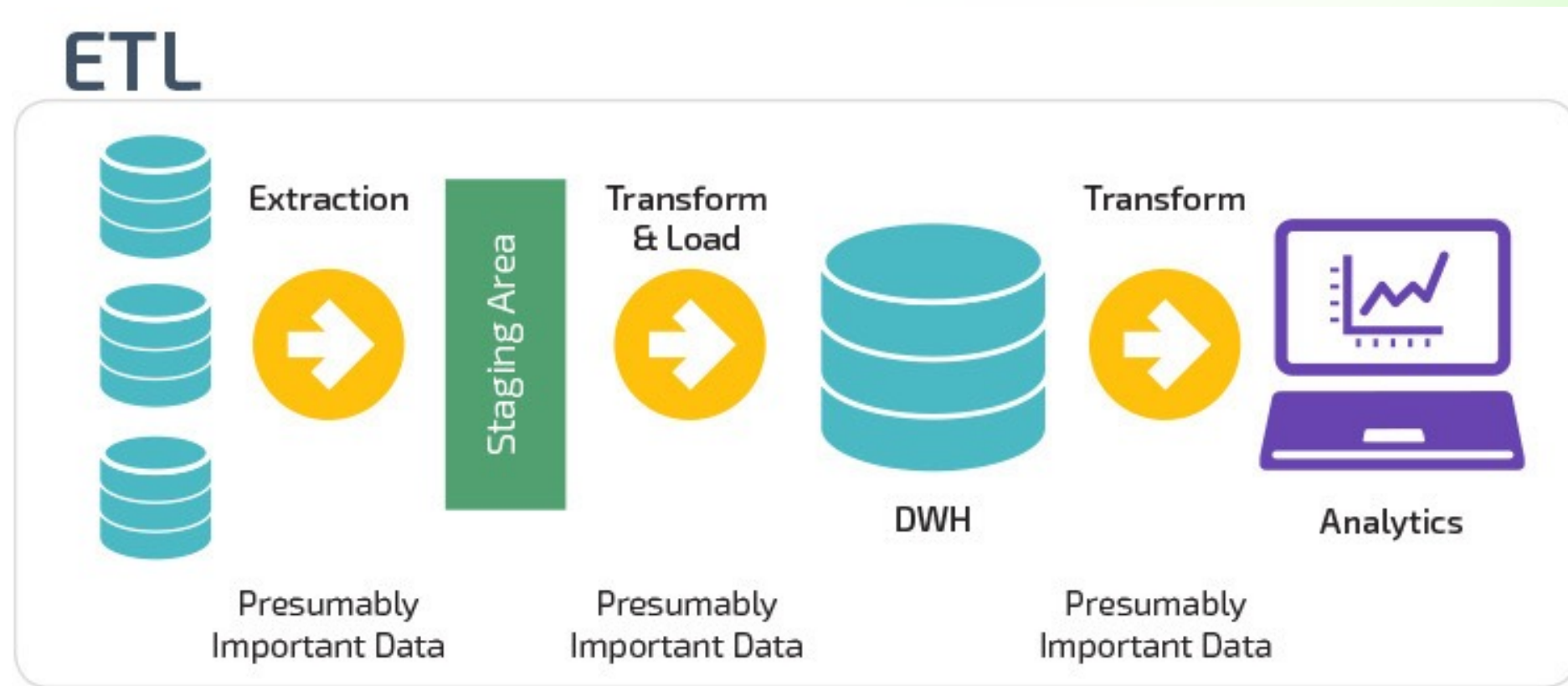
Инфраструктура хранилища данных Data Warehouse

Инкрементальная загрузка (захват изменённых данных)

Трансформация и промежуточная подготовка данных из источника

Слушатель:
Осипов Антон Владимирович

ETL (Извлечение Преобразование загрузка)



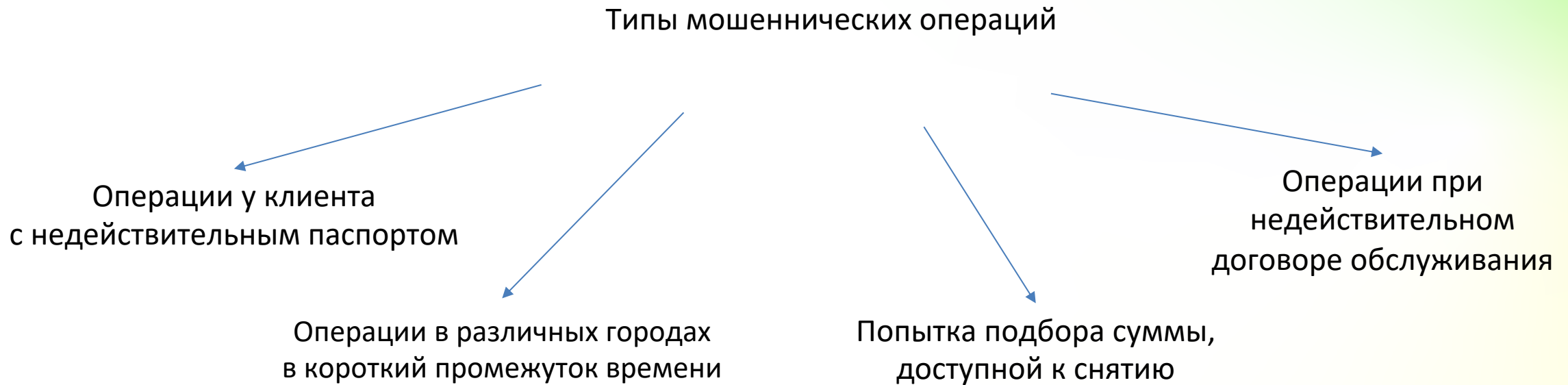
Слушатель:
Осипов Антон Владимирович

Схема взаимодействия языка Python с обработкой данных в базах SQL ORACLE



- обработка данных из источников скриптом на языке Python
- автоматическое выполнение по расписанию
- сбор данных из файлов разных форматов
- загрузка в целевые таблицы

Поиск мошеннических операций в условном банке



Выводы

1. Анализ собранных данных приносит пользу обществу
2. Раскрыт потенциал по обеспечению безопасности банковских карт
3. Противостояние мошенническим операциям
4. Спокойствие для клиентов банка

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. The Data Warehouse Toolkit. Ralph Kimball, Mary Ross
2. Pro Oracle SQL Development. Jon Heller
3. Beginning Oracle SQL. Lex de Haan, Tim Gorman
4. T-SQL Fundamentals, Third Edition. Itzik Ben-Gan