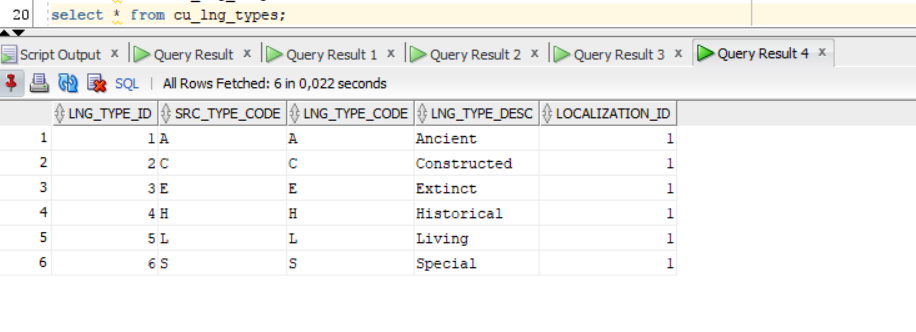
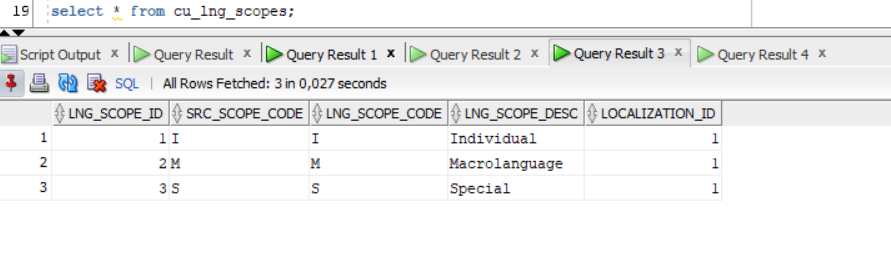
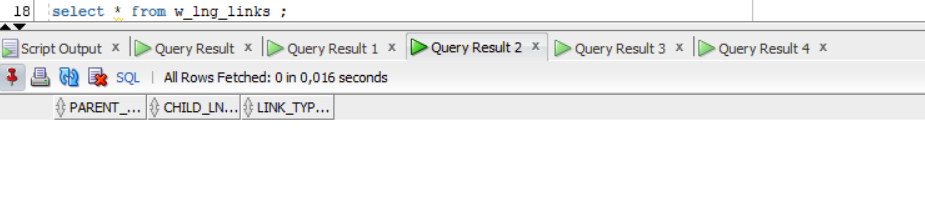
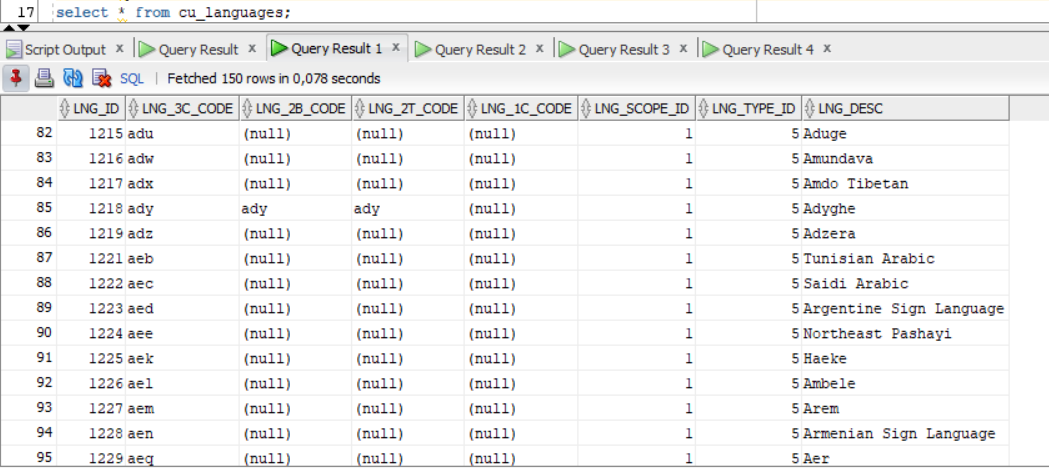
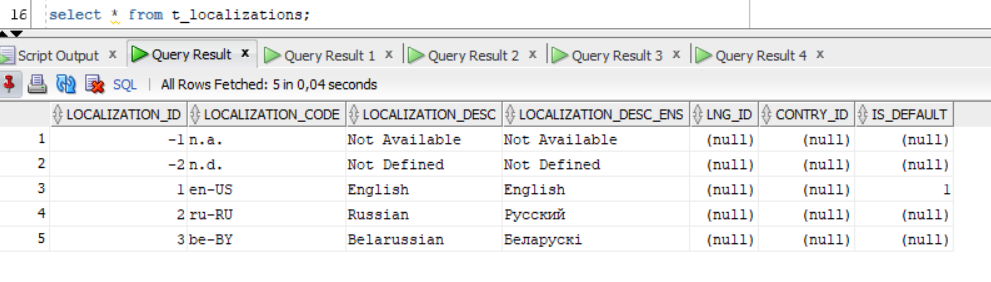
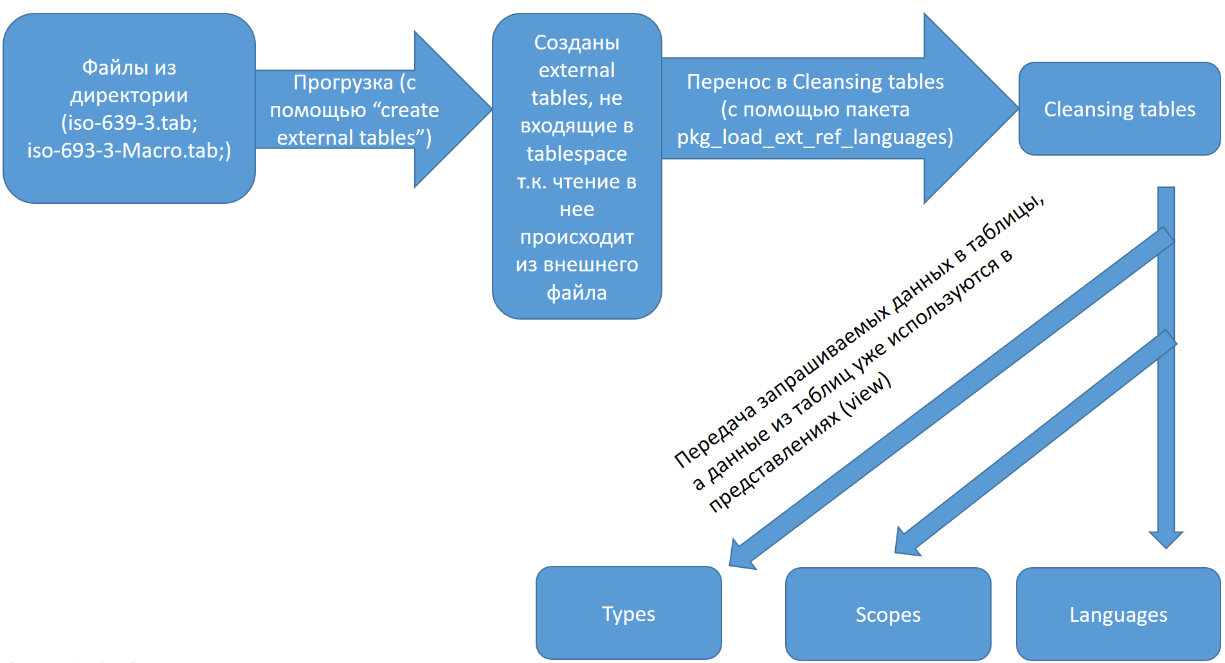
**Лабораторная работа №6**

**Невейков Андрей, 2022**

/\*Task\_1\*/

****

DataFlow:

****

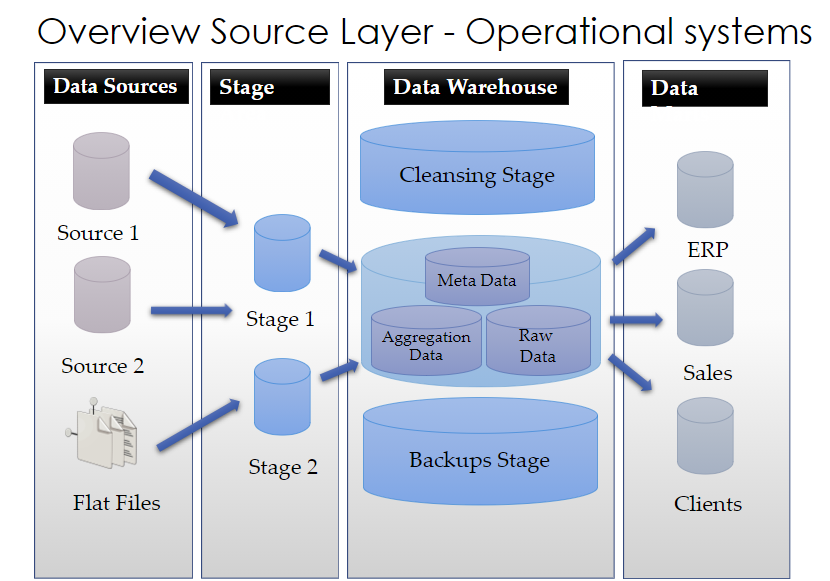
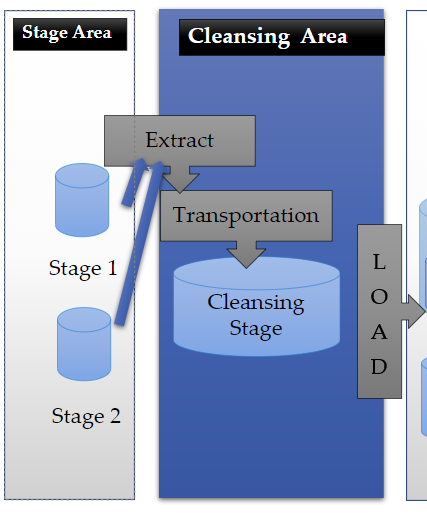
Фактически происходит то же поэтапное разделение, что и на слайде

Презентации №8 (скрин ниже): у нас есть Data Sources (файлы iso-….tab), в которых хранится информация. Далее Stage Area, когда мы создали external

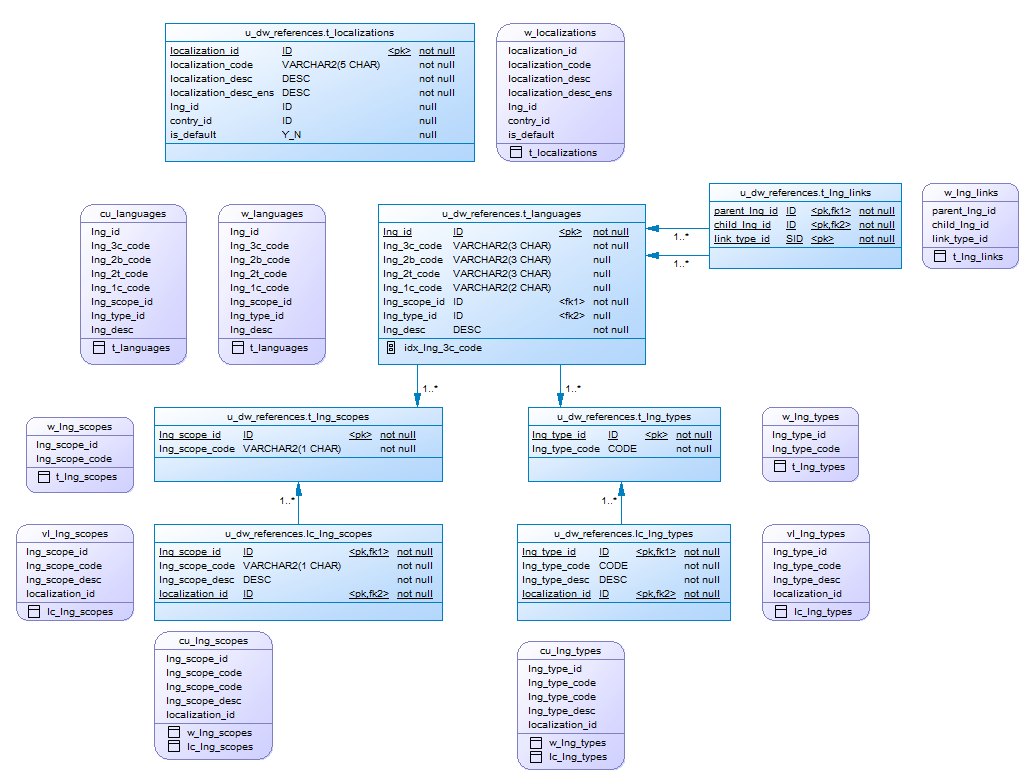
Tables, далее Cleansing Area: с помощью процедур извлекаем данные из

external tables, транспортируем на cleansing stage и распределяем по

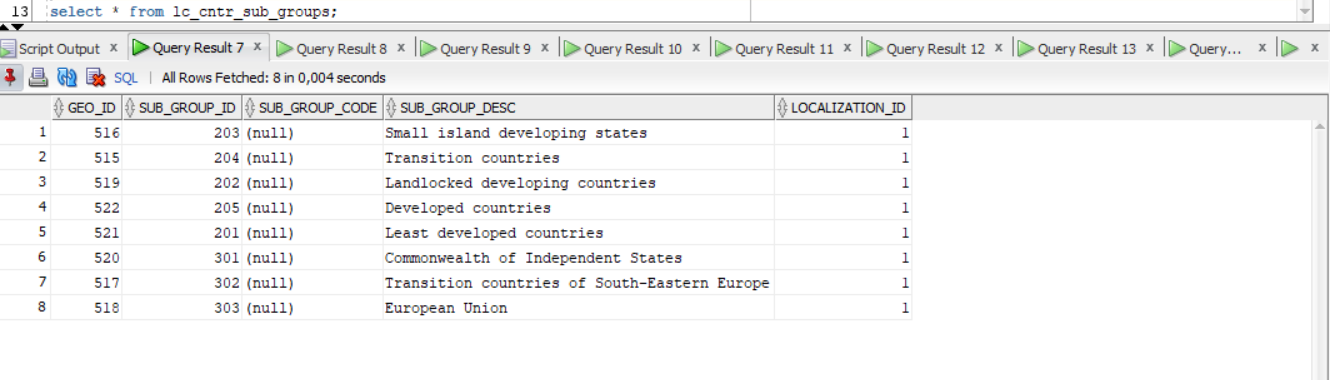
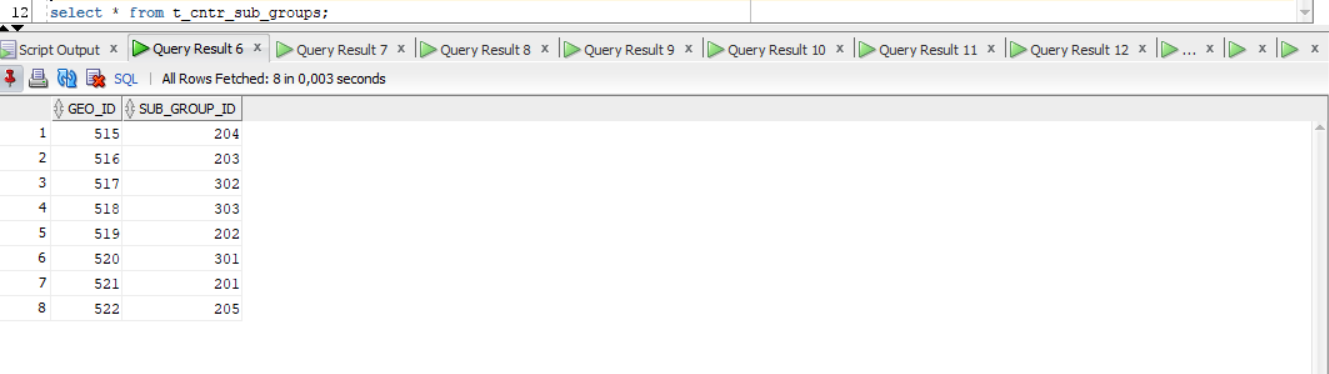
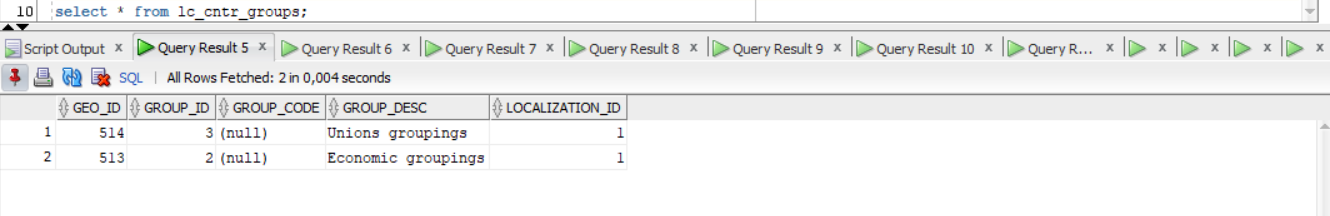
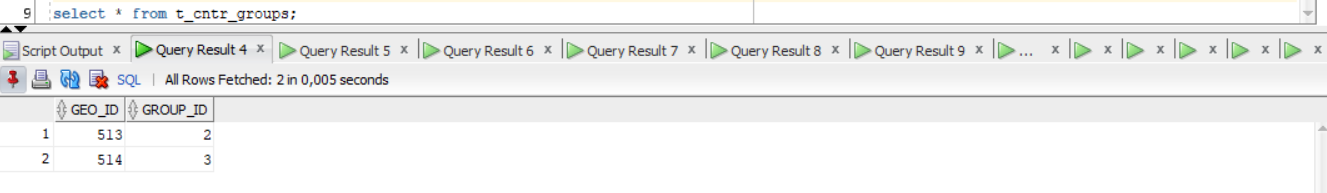
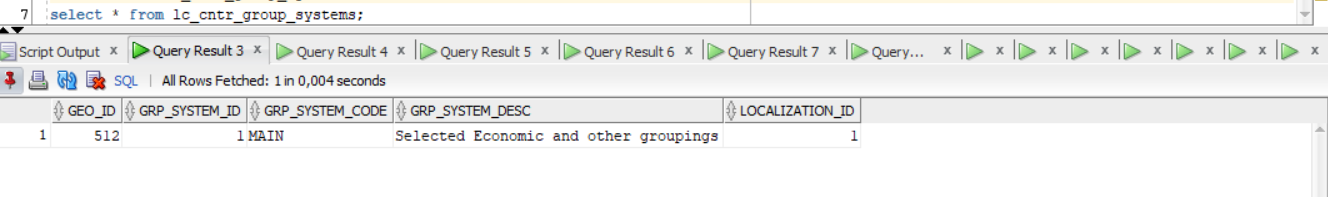
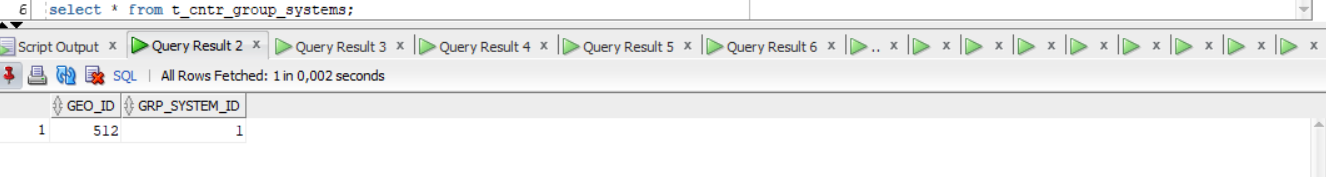
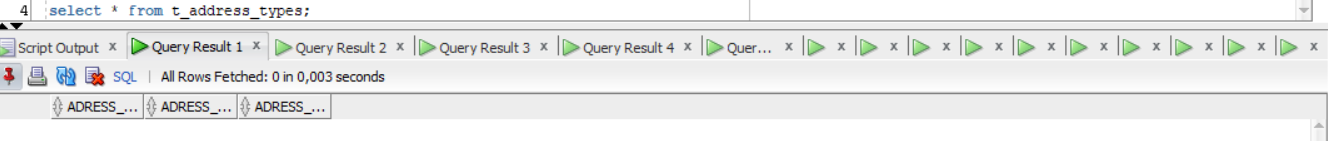
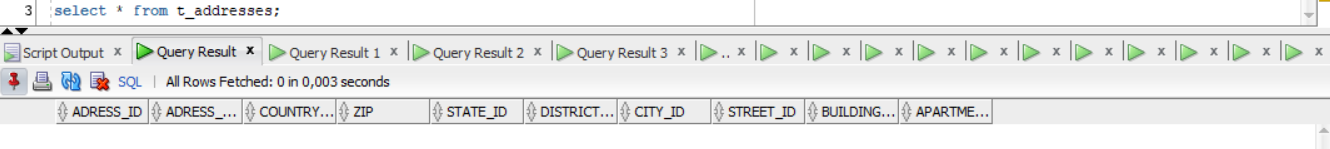
таблицам.

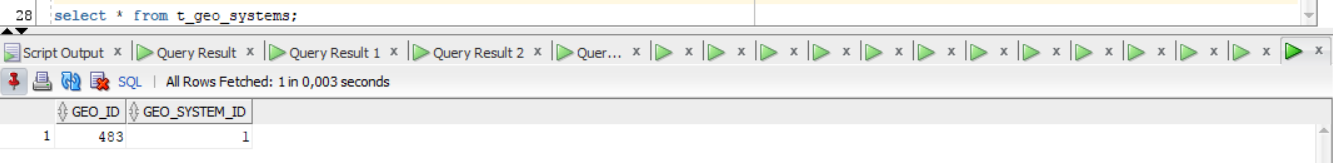
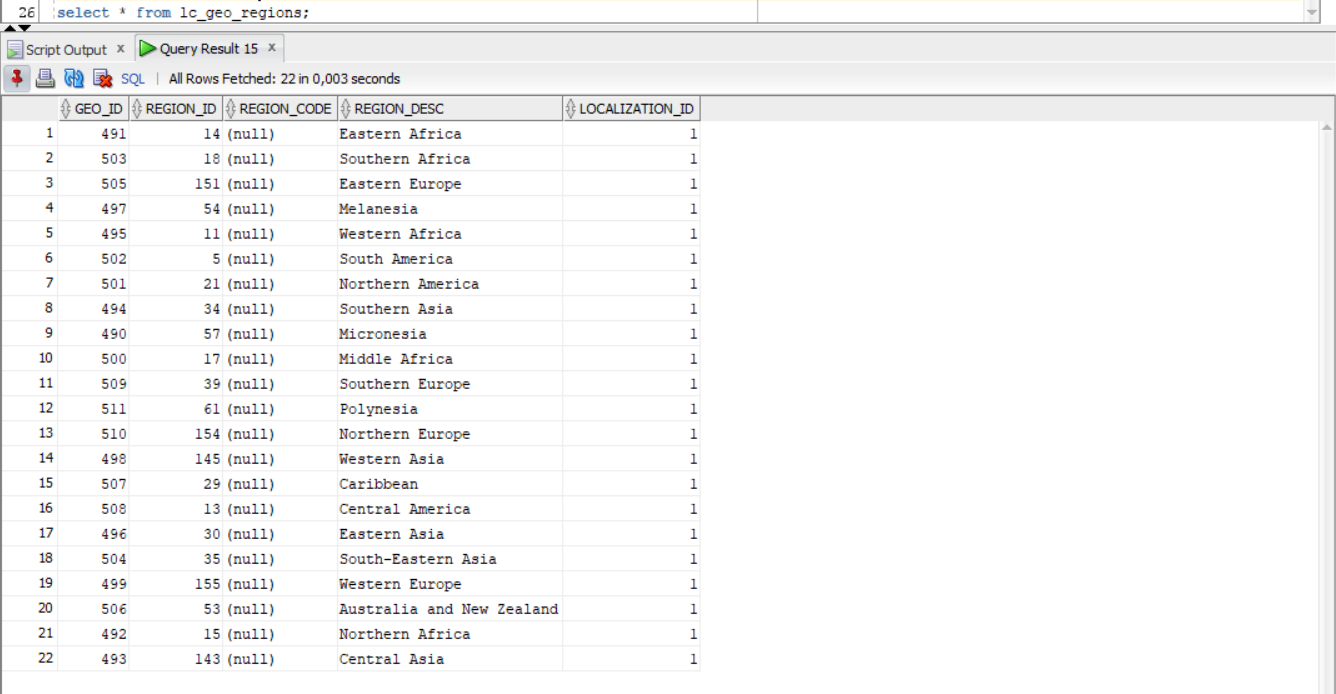
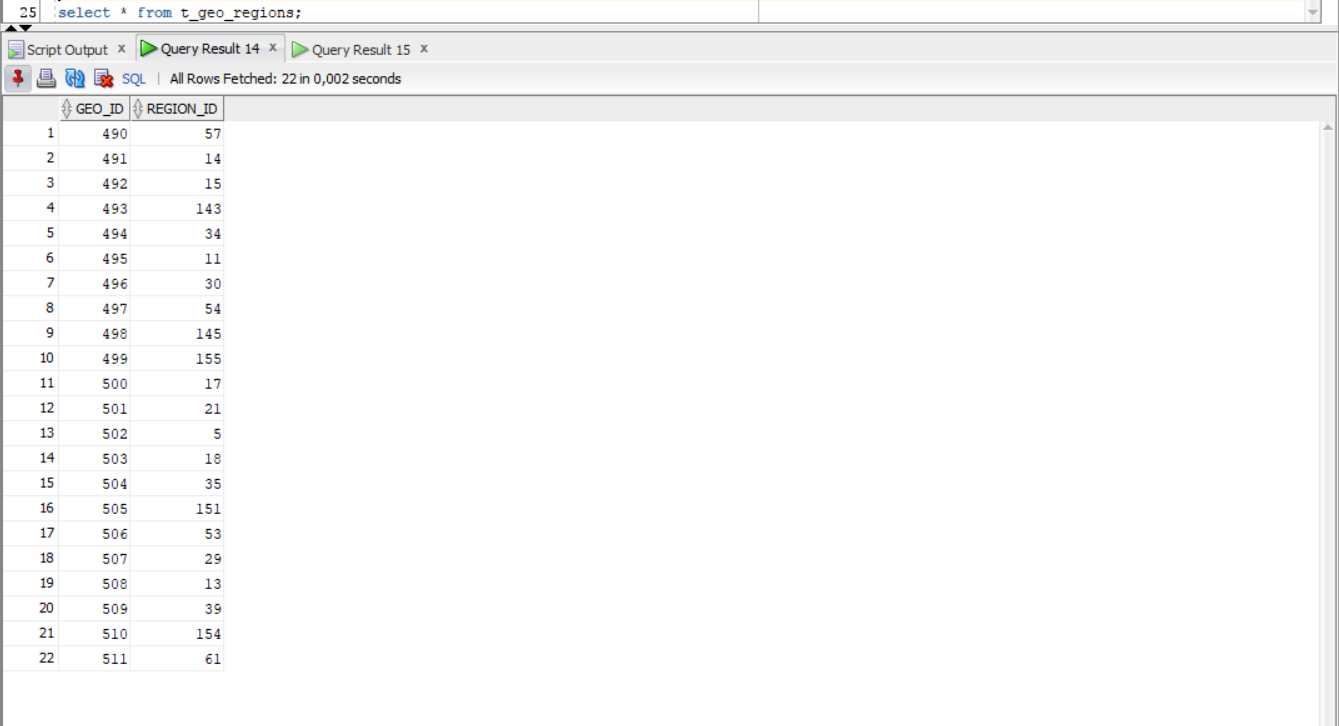
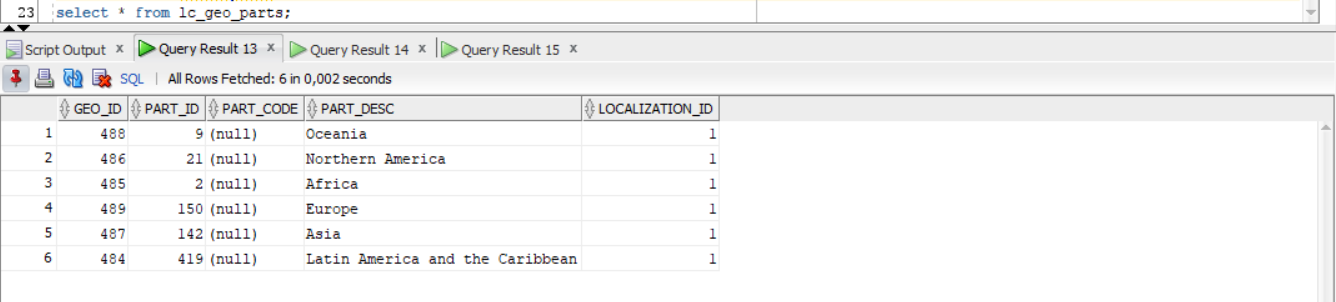
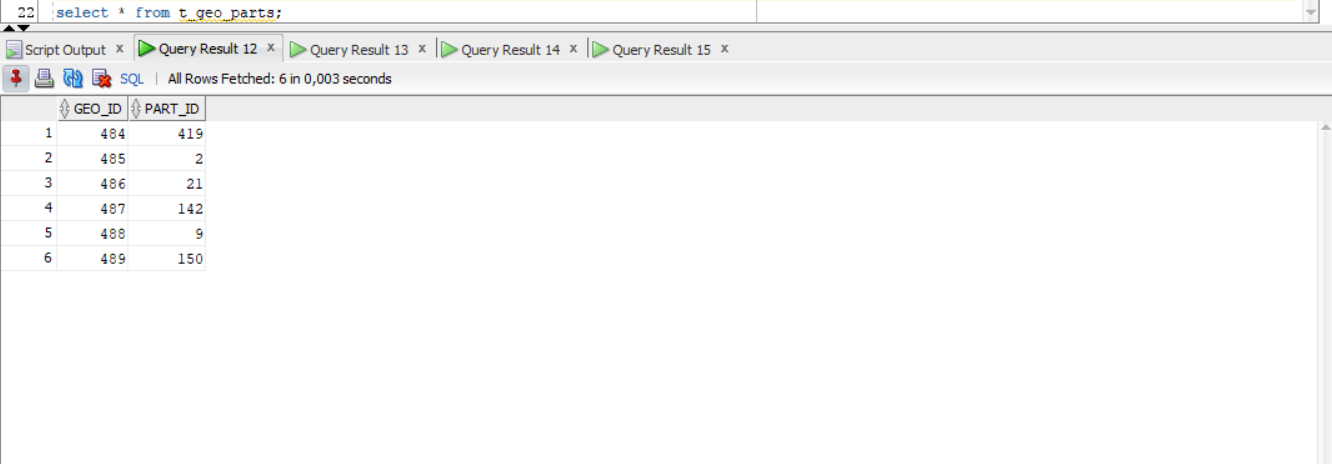
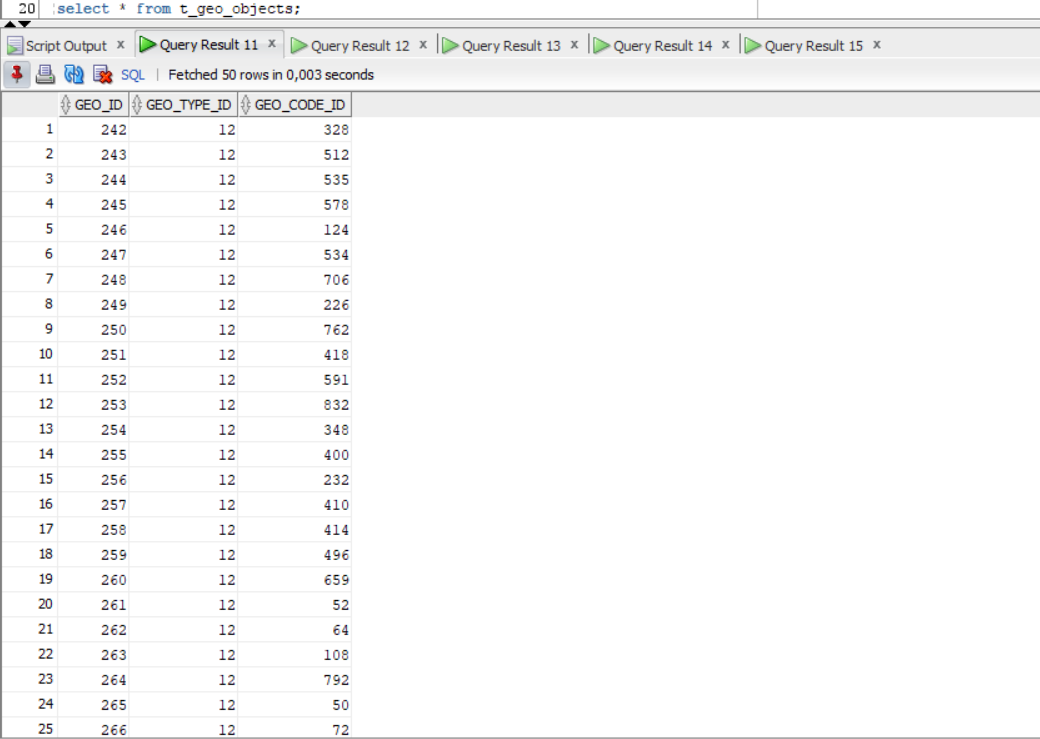
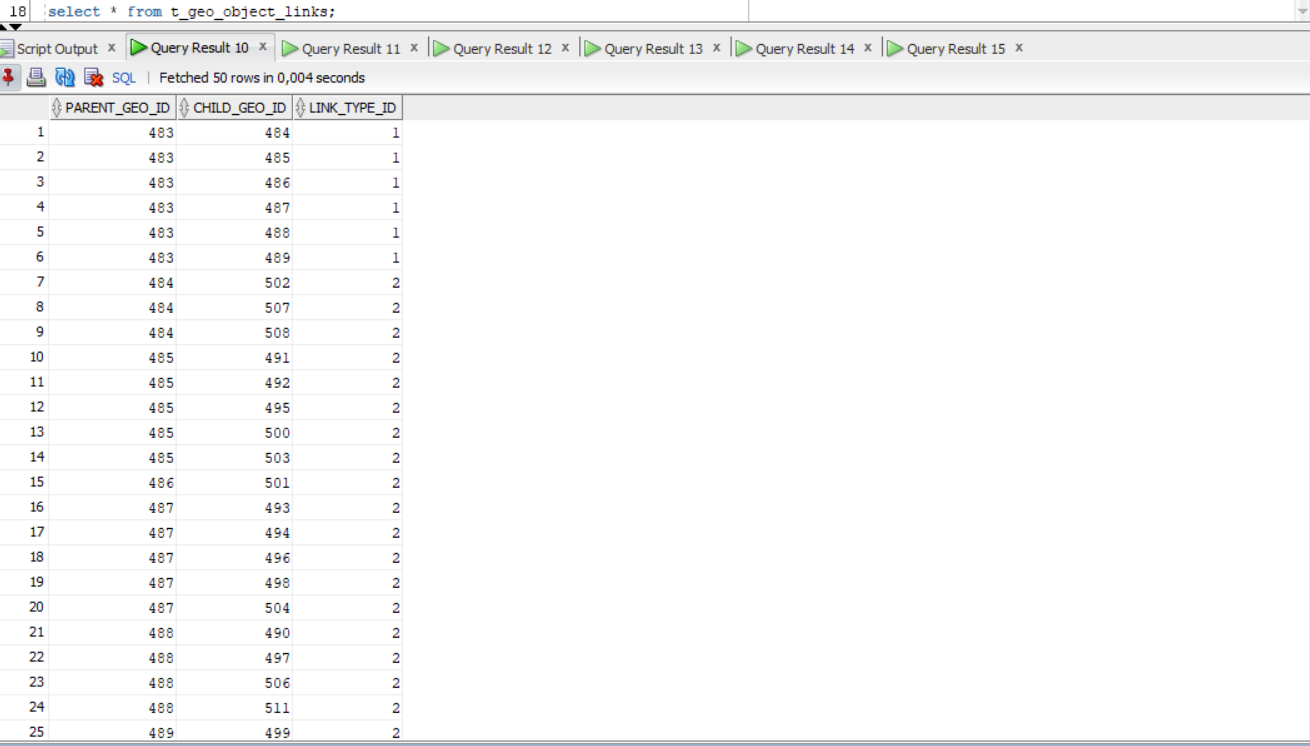
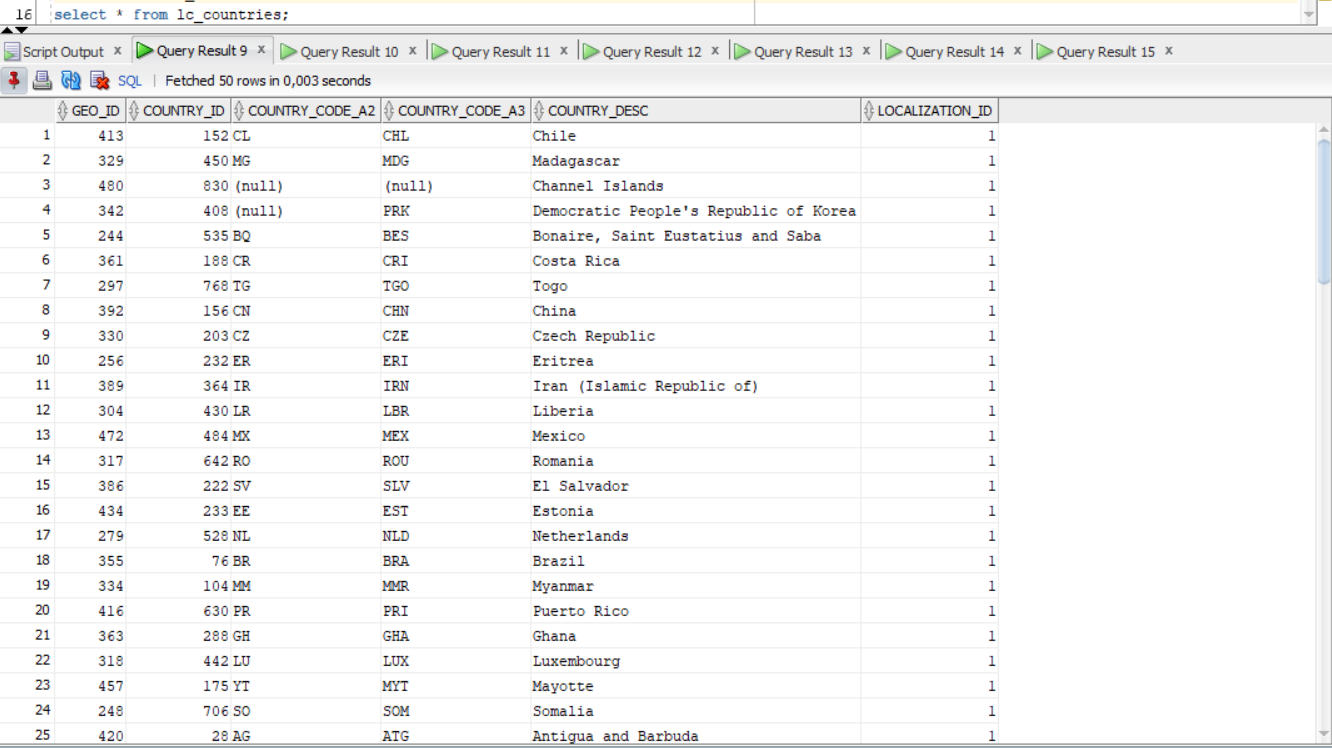
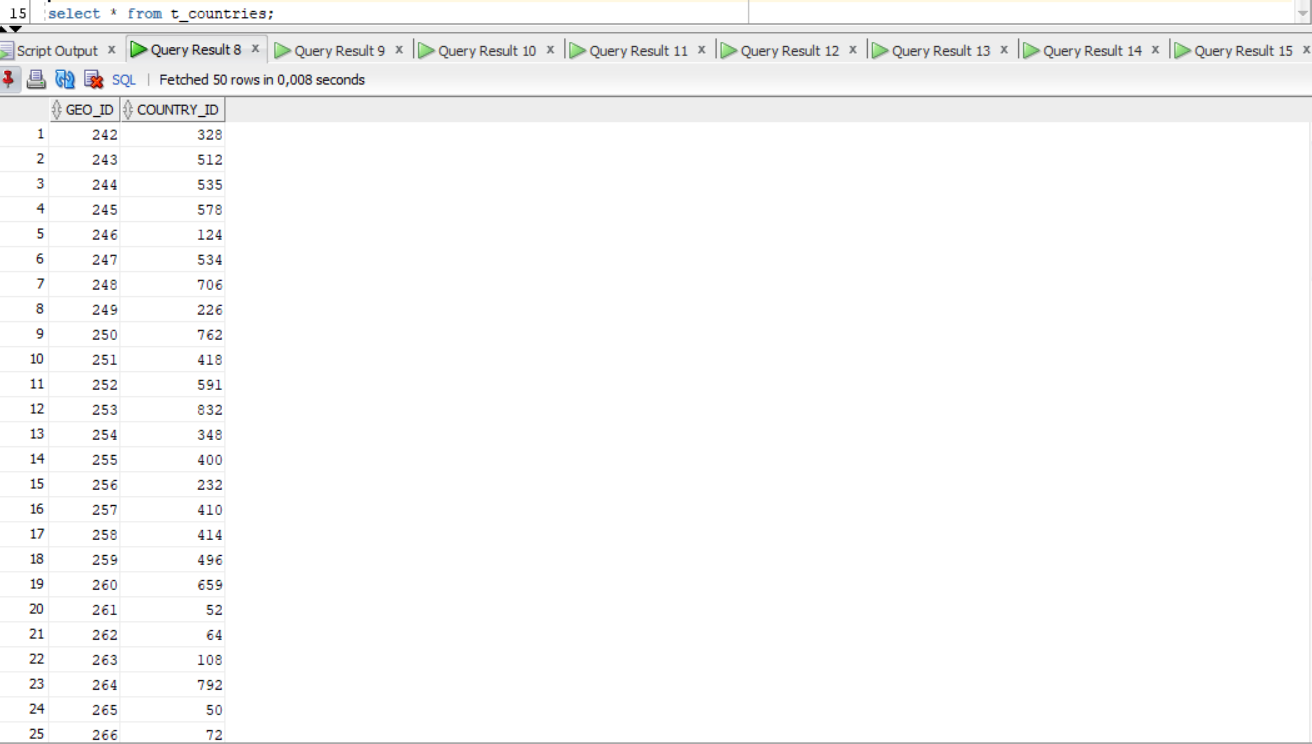


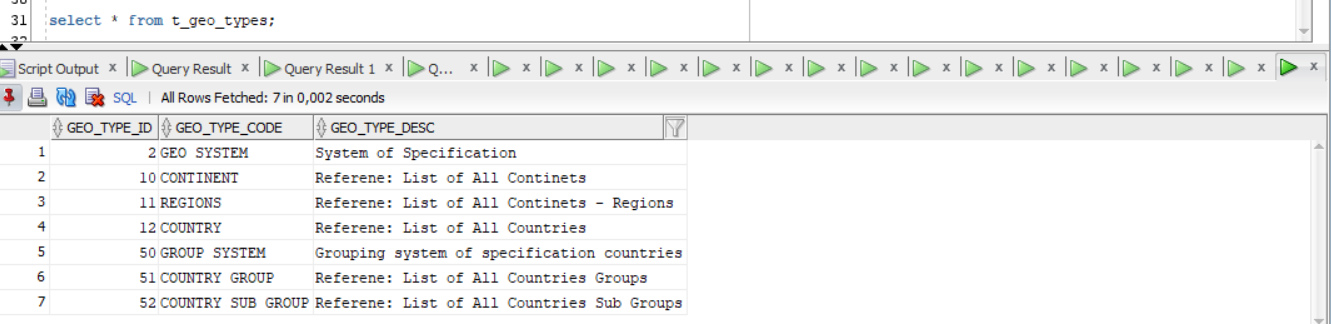
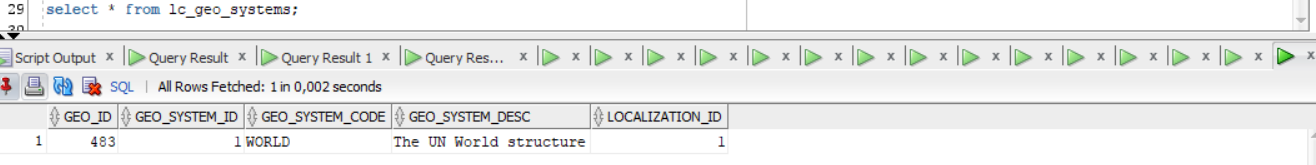
Физическая модель базы task\_1:



/\*Task\_2\*/

****

****

****

DataFlow:



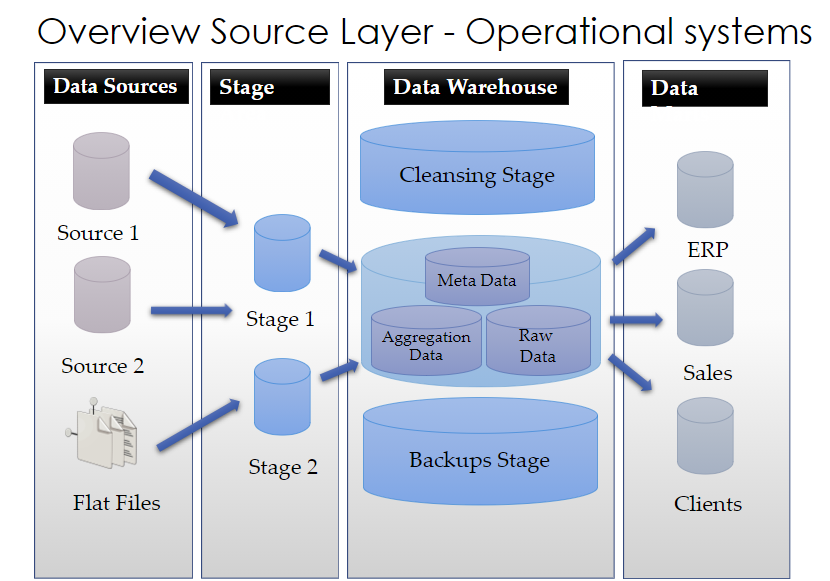
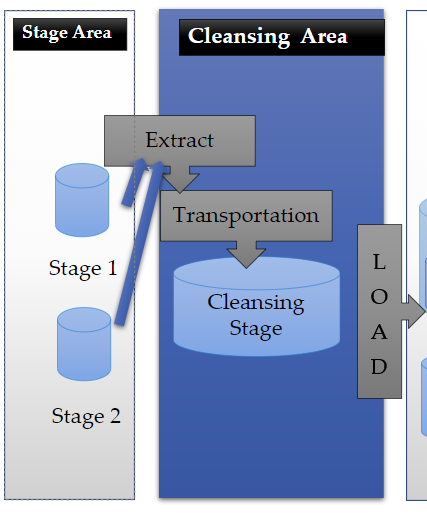
Фактически происходит то же поэтапное разделение, что и на слайде

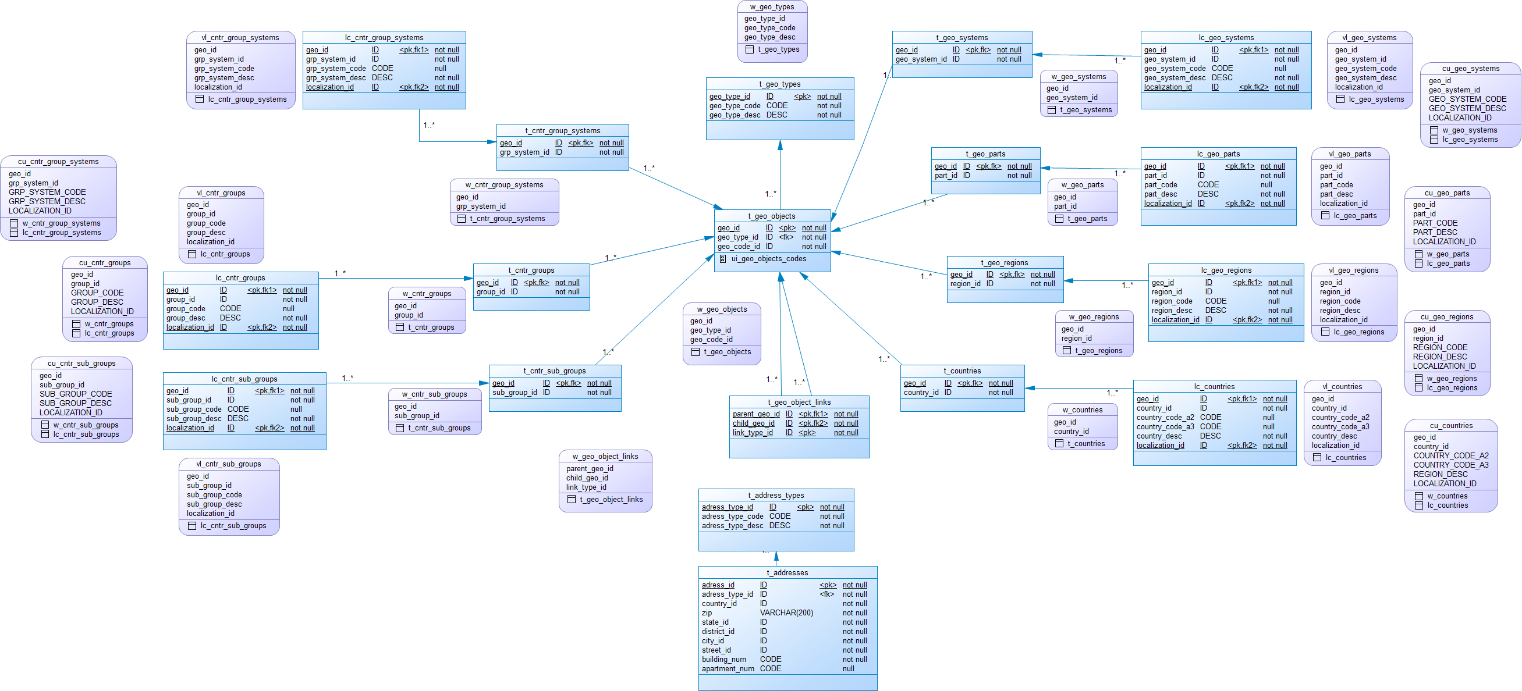
Презентации №8 (скрин ниже): у нас есть Data Sources (файлы iso-….tab), в которых хранится информация. Далее Stage Area, когда мы создали external

Tables, далее Cleansing Area: с помощью процедур извлекаем данные из

external tables, транспортируем на cleansing stage и распределяем по

таблицам.



Физическая модель базы task\_2:

**SUMMARY “TASK\_1-2”**

1. **Я научился** разворачивать базу по физической схеме, понимать последовательность выполнения скриптов по коду (находить ссылки на другие таблицы), создавать для пользователей область видимости, использовать external таблицы для взаимодействия с файлами с данными, создавать таблицы от

имени нужных пользователей, получать скрипт из SAP Power Designer схемы, тестить код и дебажить скрипты по коду ошибок(я и раньше это делал,

но тут приходилось прям часто, так что навык развивался). Отточил навык объяснять «как это работает», «почему это не работает», «как сделать

чтобы работало», искать решение в коллективе, понятно и логически верно доносить мысли.

1. **Я узнал**, что делать, если клиенту надо будет быстро с нуля

переразвернуть базу, как собирать декомпозированное хранилище данных, как построить хранилище данных согласно best practice Data Warehouse

Architecture, как правильно вытягивать информацию из файлов с данными в базу, полезные SQL селекты для понимания, что база строится правильно.(Select \* from dba\_tablespaces; Select \* from all\_tables where owner = ‘some\_owner’;

1. **Я понял**, принципы Data Warehouse Architecture, как поределяется очередность развертывания базы данных, что надо записывать шаги т.к. с первого раза может не получиться и надо будет создавать заново, процесс как данные переходят из файлов с данными в таблицы бд.

**FINAL SUMMARY**

Данная лабораторная работа …

1. … показывает как развернуть бд по схеме.
2. … показывает как логично распределить область видимости между пользователями
3. … учит писать правильные имена скриптам (очень удобно, когда во всех ветках схемы порядок запуска “t\_...”, “w\_...”, “lc\_...”, “vl\_...”, “cu\_...”. Следовательно надо и самому так проектировать с заботой о пользователях)
4. … учит искать баги по тексту ошибок (вдумчиво читая, а не просто гугля код ошибки (просто гуглить не работало)).
5. … учит получать скрипты из схемы SAP Power Designer.

/\*Task\_3\*/

Overview

Business Background

Бизнес по импорту иностранной продукции под заказ. В странах Европейского союза, при вывозе товаров за рубеж, при повторном посещении страны можно вернуть НДС в виде налогового вычета и цены на многие типы товаров ниже. 50 кг продукции можно вывезти без пошлины.

Benefit

Клиент получает товар по более низкой цене. А вычет НДС – доход компании с товара.

Requirements

Business Requirements

Нужна система для хранения данных о заказах, компаниях, которые предоставляют транспорт или могут осуществить доставку, курсах валют (т.к. заказывают в BYN, а покупка происходит в EUR, PLN или другой валюте), могут хранить данные о клиентах как о физических, так и юридических лицах, хранить данные о расположении товара, для расчета стоимости с доставкой и сроках поставок.

Technical Requirements

Обрабатывать данные о примерно 1000 заказов в неделю с возможностью увеличения вблизи праздников и мировых распродаж (черных пятниц). Работать с округлением валют (валюта покупается на бирже, где точность выше на несколько порядков, чем на банковском счету). Должны осуществляться логические проверки на сроки поставок (невозможность заказать на вчера или на 3022 год), возможность вносить заметки по поводу заказов (бесконтактная доставка, нерабочий домофон, просьба предупредить за 5 мин до приезда и т.д.), иметь максимально быстрый доступ к полям, необходимым для покупки за границей (сохранить анонимность клиентов, люди, непосредственно покупающие товар, должны знать сколько штук купить, к какому числу дедлайн поставки и сколько клиент уже ждет). Организовать иерархию доступа для безопасности целостности данных. Интегрировать базу с сайтом компании. На основе способа оплаты (одним платежом/ в кредит) рассчитывать цену (одним платежом дешевле).

Solution Sketch

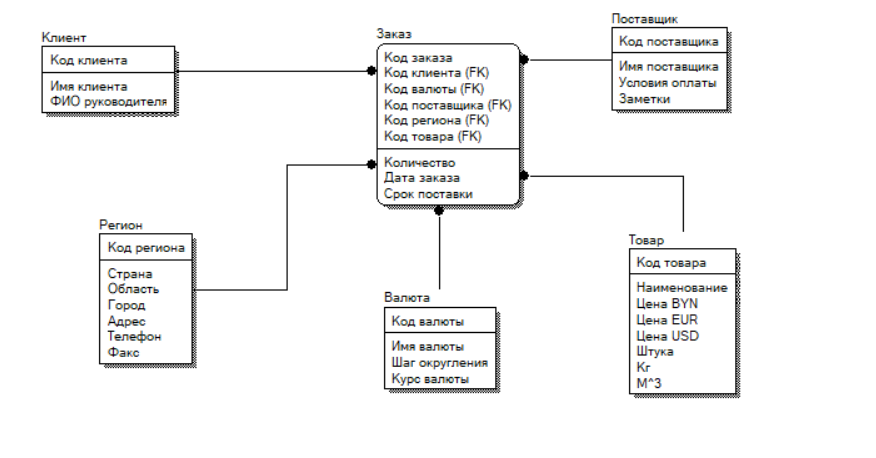
Source Tables structure

Данные в основную базу будут попадать через csv-документ, получающий данные с формы из сайта компании, имеет стандартный разделитель “;”. Пользователь заполняет свои данные (Имя (или название компании + ФИО руководителя), Страна, Область, Город, Адрес, телефон, факс при наличии), нужный товар, кол-во товара, дата заказа, ожидаемая дата поставки, способ оплаты, опционально примечания к заказу.

Summarize Data Plan

Проанализировать данные, для товаров ввести единицы измерения, рассчитать стоимость после валютной переоценки, отправить исполнителям необходимую информацию для покупки .

/\*Task\_4\*/



**Схема «звезды»** (звёздная схема) — специальная организация реляционных таблиц, удобная для хранения многомерных показателей. Лежит в основе реляционного OLAP.

Модель данных состоит из двух типов таблиц: одной таблицы фактов — центр «звезды» — и нескольких таблиц измерений по числу измерений в модели данных — лучи «звезды».

**Таблица фактов** обычно содержит одну или несколько колонок типа DECIMAL, дающих числовую характеристику какому-то аспекту предметной области (например, объём продаж для торговой компании или сумма платежей для банка), и несколько целочисленных колонок-ключей для доступа к таблицам измерений. В таблицы фактов данные должны оперативно записываться в случае изменений.

**Таблицы измерений** расшифровывают ключи, на которые ссылается таблица фактов; например, таблица «products» измерения «товары» базы данных торговой компании может содержать сведения о названии товара, его производителе, типе товара. За счёт использования специальной структуры таблицы измерений реализуется иерархия измерений, в том числе ветвящаяся.

Обычно данные в таблицах-измерениях денормализованы: ценой несколько неэффективного использования дискового пространства удается уменьшить число участвующих в операции соединения таблиц, что обычно приводит к сильному уменьшению времени выполнения запроса.

SQL-запрос к схеме «звезда» обычно содержит в себе:

* одно или несколько соединений таблицы фактов с таблицами измерений;
* несколько фильтров (SQL-оператор WHERE), применяемых к таблице фактов или таблицам измерений;
* группировку и агрегирование по требуемым элементам иерархии измерений (*dimension elements*).

/\*Task\_5\*/

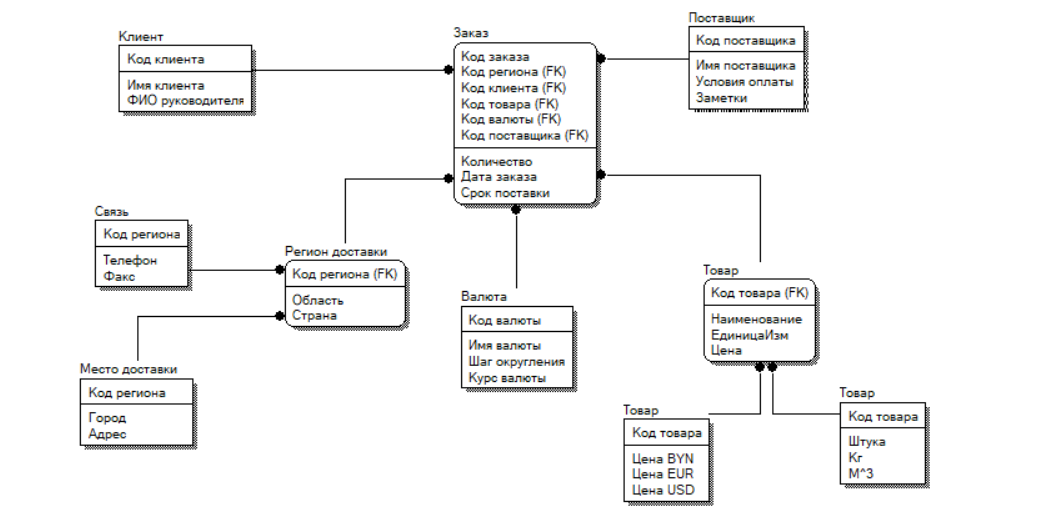
****

Схема «Снежинка» — вид логической структуры реляционного хранилища данных, которая использует два типа таблиц: единственную таблицу фактов и несколько таблиц измерений. Обычно на схеме таблицу фактов изображают в центре, а таблицы измерений – в виде расходящихся лучей.

Основной особенностью схемы «Снежинка» является возможность хранения иерархических измерений, кода одно или несколько измерений связаны с иерархически подчинёнными измерениями. Пример иерархии измерений: «Группа товаров» и «Товар»: каждому значению измерения «Группа товаров» будет соответствовать отдельная таблица товаров данной группы.

Таким образом, в схеме «Снежинка» таблицы измерений могут быть соединены с таблицами измерений других иерархических уровней непосредственно, минуя таблицу фактов.

Если поддержка иерархии измерений не требуется, то лучше использовать схему «Звезда», в которой все измерения связаны с таблицей фактов и не связаны друг с другом непосредственно.

**Недостатками** схемы «Снежинка» по сравнению со схемой «Звезда» являются:

* более трудная для понимания и реализации структура данных;
* более сложная процедура добавления значений в измерения.

К **преимуществам** схемы «Снежинка» можно отнести:

* лучшее соответствие представлению данных в многомерной модели;
* намного более низкая вероятность появления ошибок несоответствия данных;
* компактность представления иерархических данных, поскольку все значения измерений представляются только один раз.