



Витрина по месяцам

Цель: построить витрину данных, представляющую собой набор агрегатов по месяцам на основе лога таможни

Описание задания: на текущий момент у вас есть лог импорта/экспорта товаров через таможню, объединенный с текстовыми описаниями категорий.

Вам потребуется развернуть PostgreSQL, создать в нем таблицу (давайте назовем ее `customs_log`) и залить туда данные этого лога (можно полностью, можно частично, для экспериментов).

Основной задачей является создание витрины (таблицы) с агрегированным отчетом по месяцам на основе данных лога. Витрина должна иметь следующие колонки:

`month`: месяц в исходном формате из лога (уууу/мм)

`total_import_value`: общий объем импорта в единицах (см. примечание ниже) за этот месяц

`total_export_value`: общий объем экспорта в единицах (см. примечание ниже) за этот месяц

`trade_balance`: разница между `total_export_value` и `total_import_value`

`top_import_category`: категория, товары из которой чаще всего импортировались в этом месяце

`top_export_category`: категория, товары из которой чаще всего экспортировались в этом месяце

`major_import_partner`: код страны, откуда было больше всего импорта в этом месяце

`major_export_partner`: код страны, куда было больше всего экспорта в этом месяце

`total_transactions`: общее количество операций экспорта и импорта за этот месяц

ПРИМЕЧАНИЕ: для упрощения задачи можно игнорировать поле `measure` и использовать только поле `quantity`, считая все в единицах

ЕЩЕ ПРИМЕЧАНИЕ: Результирующая таблица должна называться `monthly_trade_summary` и создаваться одним SQL-запросом на основе исходной таблицы.

Ответом на задание является ссылка на Git-репозиторий, где лежит SQL-файл `monthly_trade_summary.sql` с кодом всех использованных в задании запросов, т.е.:

- создание таблицы `customs_log`
- (опционально) создание дополнительных индексов на таблице `customs_log`
- заливка всех данных из файла в таблицу `customs_log` (даже если вы использовали только часть данных при подготовке задачи)
- (опционально) создание MATERIALIZED VIEW для агрегации данных для расчета витрины
- создание таблицы `monthly_trade_summary`
- один запрос на заполнение таблицы `monthly_trade_summary` на основе таблицы `customs_log`

Инструменты, которые пригодятся для выполнения: SQL, PostgreSQL



Критерии, по которым будет оцениваться задание: созданные таблицы должны иметь правильные имена, а финальный SQL-запрос должен генерировать корректную витрину по представленным данным. **Если при подготовке витрины вы использовали промежуточные трюки, вроде создания индексов на `customs_log` или промежуточных `materialized view` для финального запроса - SQL-файл должен содержать комментарии около соответствующих запросов, описывающих мотивацию их создания и какой профит в итоге от этого получился.** Грамотное использование этих подходов идет в плюс при оценке.

Ваше решение

Отправить на проверку

Сохранить как черновик

Далее >>

