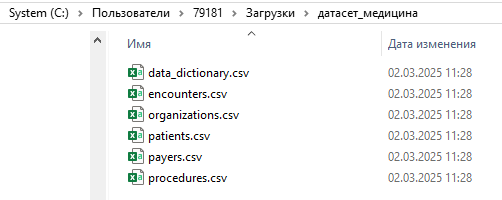
### **Описание**

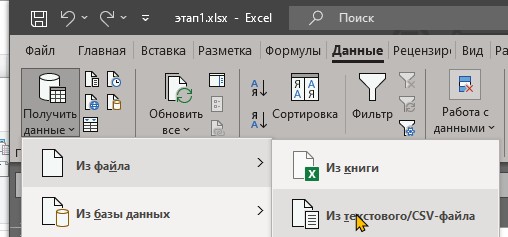
* Вам предстоит выступить в роли аналитика данных интернет-магазина.
* Вам предоставлены сырые данные о продажах, товарах, клиентах и маркетинговых кампаниях.
* Ваша задача - провести комплексный анализ данных, выявить ключевые показатели эффективности (KPI) и визуализировать полученные результаты для руководства компании.

### **Данные**

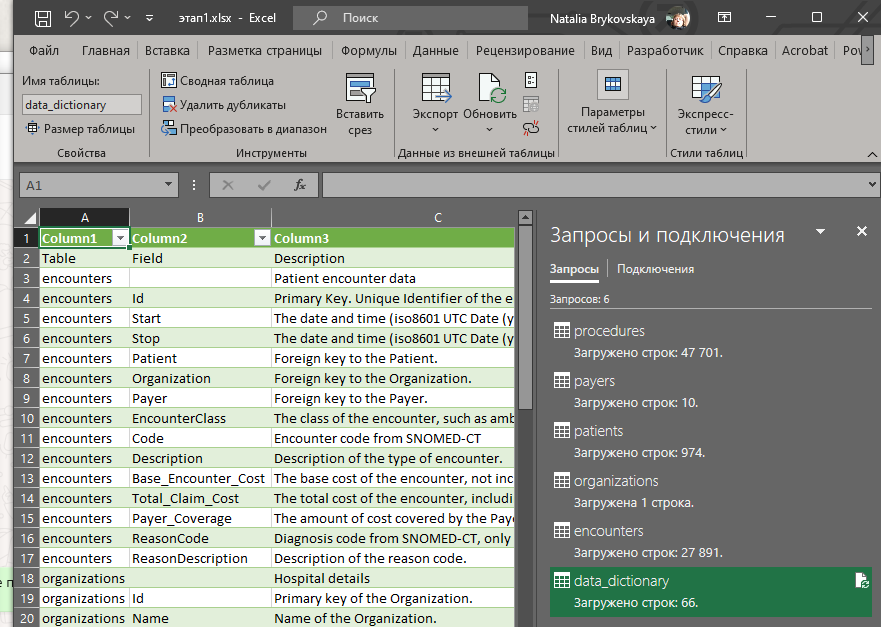


### **Этап 1: Подготовка данных (Excel + Power Query)**

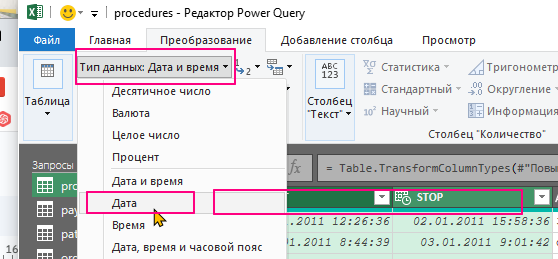
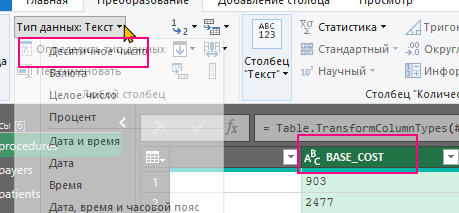
* Импорт данных: Импортируйте предоставленные CSV файлы в Excel.



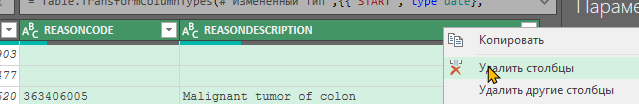
пока только загружаем все таблицы



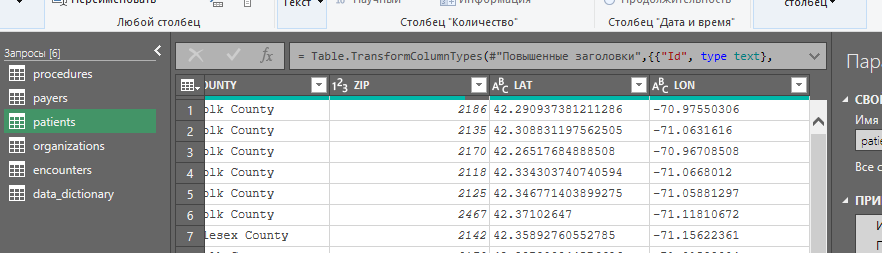
теперь проведем очистку данных в Power Query 

* Очистка данных: Используя Power Query, проведите очистку данных:
  + Удалите дубликаты.
  + Обработайте пропущенные значения.
  + Приведите данные к нужным типам (дата, число, текст).
  + 
  + 

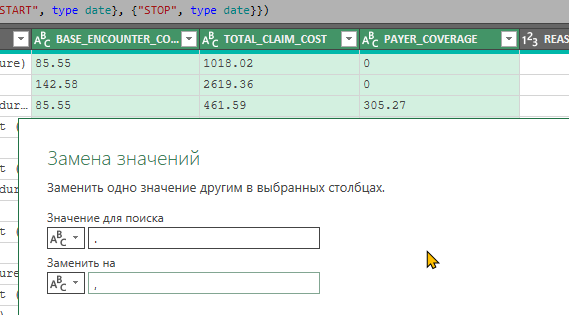
в рамках исследования не все столбцы нам нужны - удалим лишнюю информацию

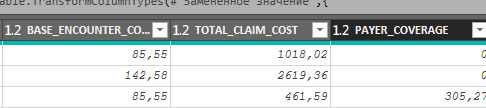


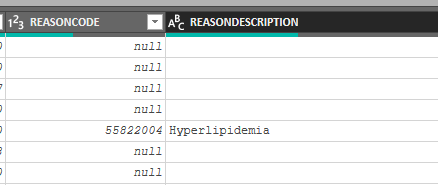
также удалим ненужные столбцы из таблицы с пациентами



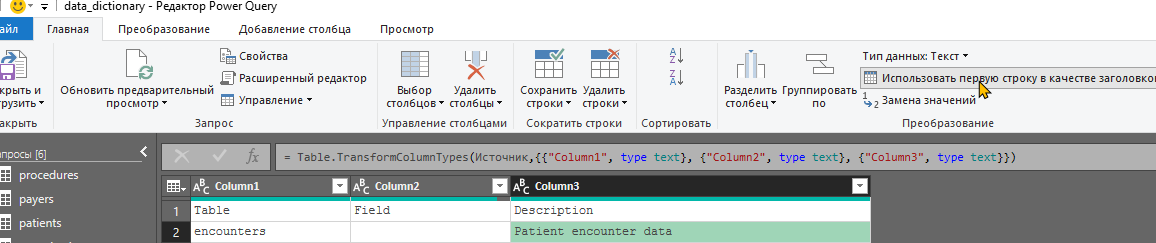
меняем тип данных предварительно нужно заменить точку на запятую



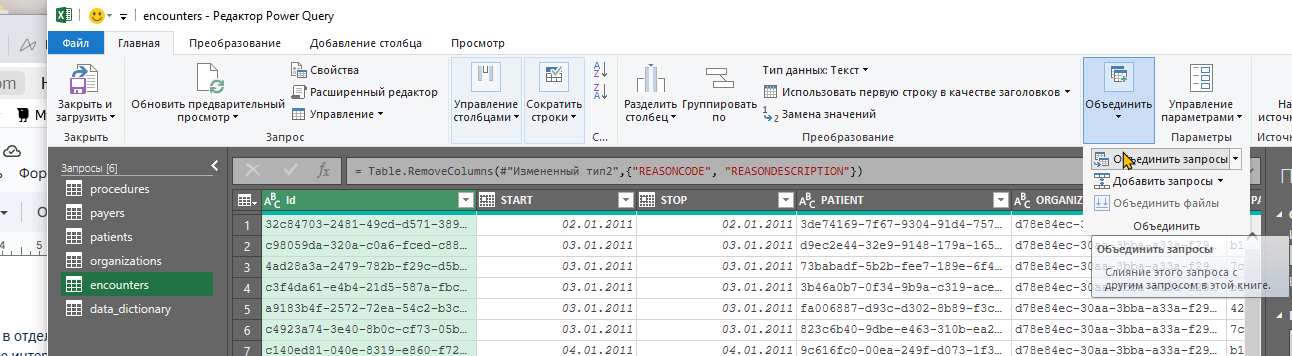


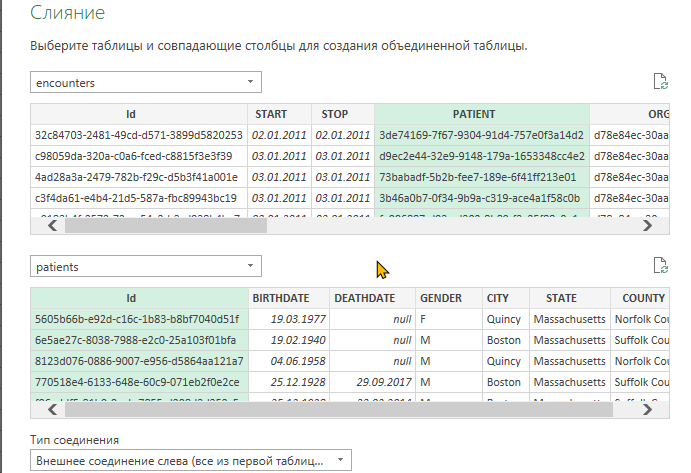
удаляем избыточную для анализа информацию\

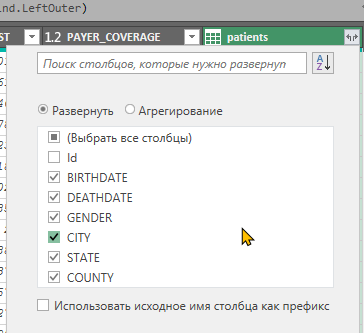
также оставим описание данных как отдельную таблицу. тут надо поправить названия столбцов.

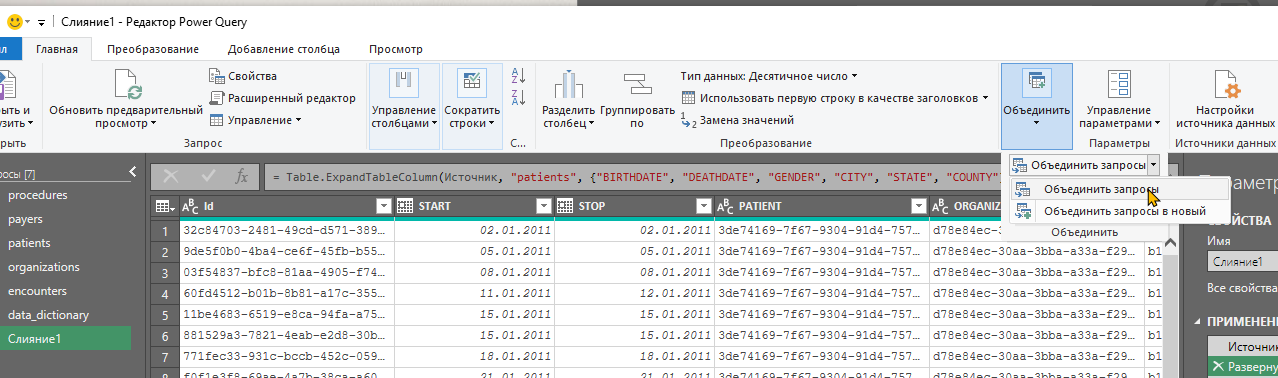


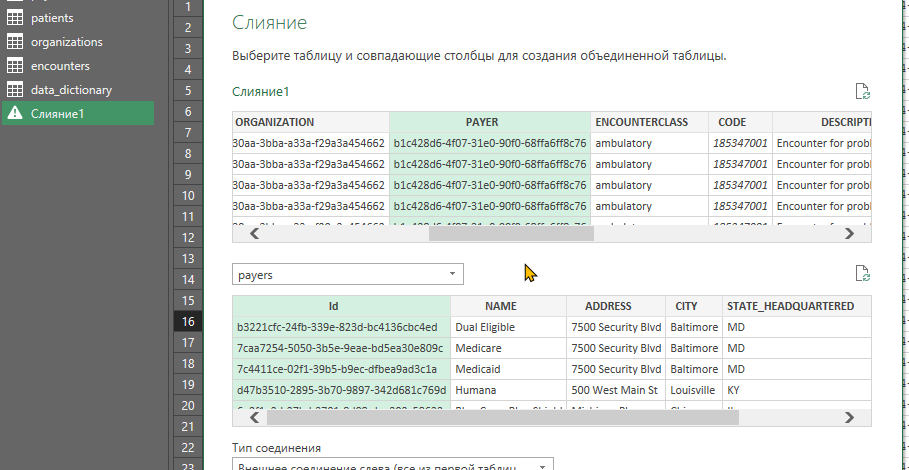
ОБЪЕДИНИМ данные в отдельную таблицу. сейчас данные нормализованы и содержатся по разным таблицам. добавим все интересующие данные в одну таблицу -ее потом и будем анализировать.

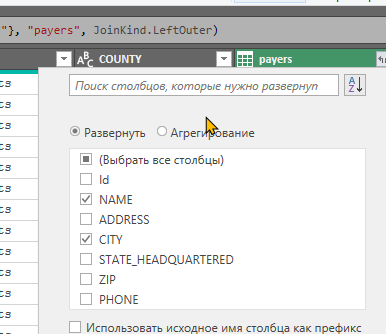






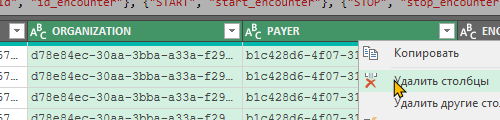




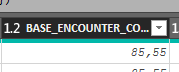


поменяем названия столбцов на понятные у таблицы слияние1



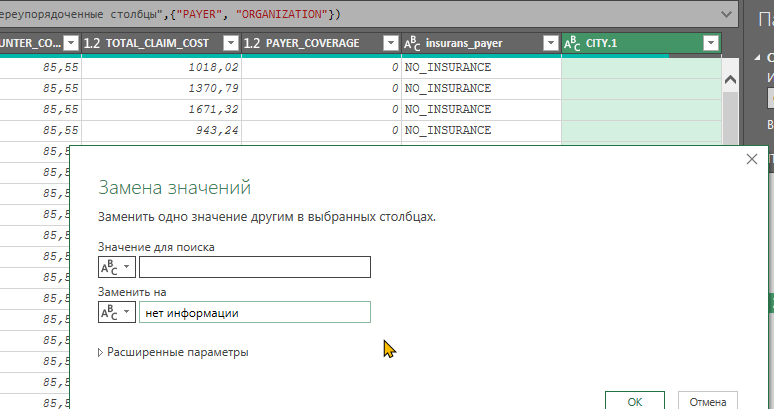
попутно удалим столбцы  организация у нас одна - для анализа сравнений не получится

а плательщика-страховщика будем выводить по названию

удалим  и это 

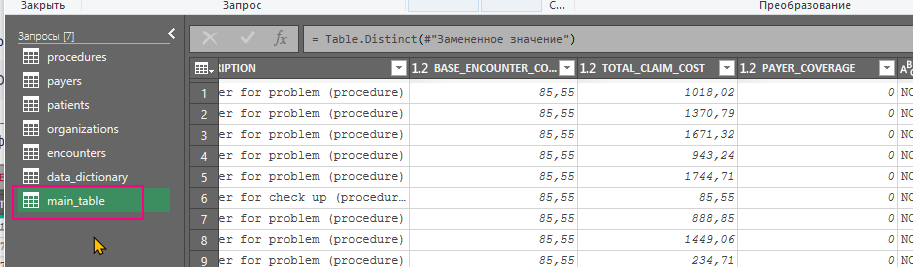
также переместим некоторые столбцы/

заменим отсутствующую информацию о городе соответствующей отметкой

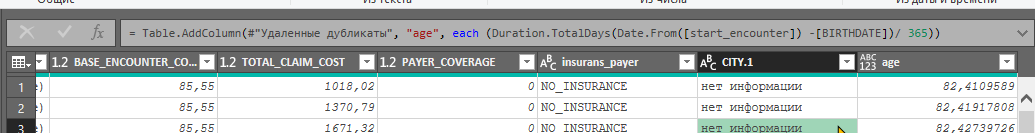


* Трансформация данных:Создайте вычисляемые столбцы

переименуем таблицу

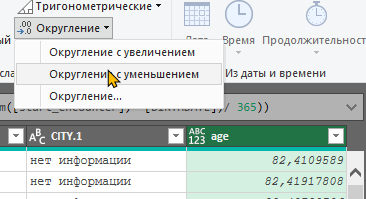


вычислим возраст пациента на дату посещения



такой грубый подсчет для наших целей оправдан.

округлим возраст в сторону уменьшения (нужны полные года пациента)



переместим возраст к дню рождения

пометка:

PAYER\_COVERAGE - сумма, уплаченная плательщиком (страховой компанией) - страховое покрытие

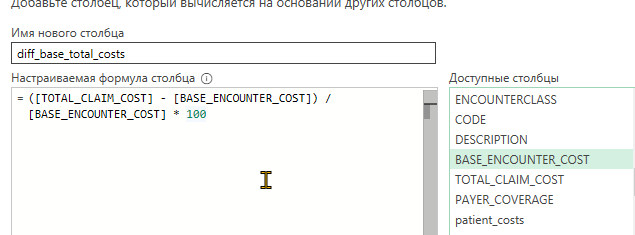
Base\_Encounter\_Cost - "Базовая стоимость посещения, не включающая расходы по отдельным статьям, связанные с лекарствами, прививками, процедурами или другими услугами".

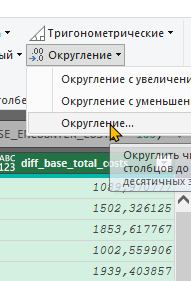
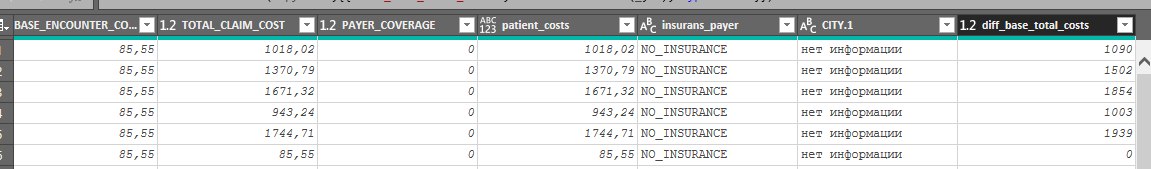
Total\_Claim\_Cost - "Общая стоимость посещения , включая все расходники и позиции".

вычислим сумму, которую не покрыли страховые компании - значит ее покрыть должен пациент.

добавим столбец patient\_costs

а также столбец разницу между базовой стоимостью и стоимостью итоговой в процентах diff\_base\_total\_costs



* + Сгруппируйте данные по различным параметрам (например, продажи по месяцам, продажи по категориям товаров).
* Связывание данных: Создайте связи между таблицами, используя соответствующие ID (например, ID товара в таблице "Продажи" и ID товара в таблице "Товары").

сохраним данные и приступим ко 2 этапу

