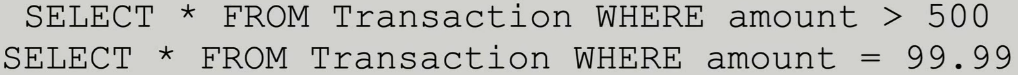
**Индексы в БД**

Предположим, что у нас есть таблица Transaction с большим количеством строк.

  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
Если мы на такой таблицы будем выполнять такого рода запросы:



то на такой таблице с большим количеством строк эти запросы будут очень долго работать.

Чтобы таких медленных запросов не было, чтобы ускорить запросы такого рода мы будем использовать индексы.

**Как у работают эти SQL запросы БЕЗ индекса ?**

У нас есть какие-то значения в колонке **amount** и они расположены хаотично, тоесть они не упорядочены.

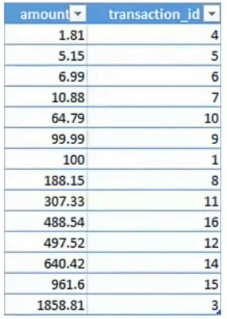
Запросы которые используют эту колонку будут заставлять БД проходиться по всем значениям колонки (full table scan). Тоесть будет перебор всех строк колонки. И так как в таблице огромное количество строк, такой запрос будет выполняться очень долго потому-что базе данных нужно пройтись по всем строкам.

Для ускорения работы на этой колонке нужно создать индекс.

**Что из себя представляет индекс ?**

Так как мы знаем, что будем делать большое количество запросов к таблице где будем фильтровать строки по колонке **amount**, то нам нужно создать индекс на колонке **amount**.

Если мы создадим индекс на колонке **amount**, то будет что то похожее на эту таблицу:



Тоесть **индекс** это еще одна сущность в базе данных. **Индекс** можно представить в качестве отдельной таблицы. В таблице индекса мы видим, что значения **amount** отсоритрованны по возрастанию и каждому значению соответствует какой то **transaction\_id**.

Теперь когда у нас есть таблица где значения **amount** отсортированны, мы можем для запросов к базе данных которые используют колонку amount использовать не обычный линейный поиск, а какие нибудь более умные алгоритмы, например, линейный пиоск.  
  
  
**Сравнение линейного поиска и двоичного:**

**1.** Линейный поиск по 5 млн. строк в зуднем случае займет 5 млн. итераций(в среднем 2.5 млн. итераций)

**2.** Двоичный поиск по 5 млн. строк в зуднем случае займет 23 итерации.

**3.** Упорядочив элементы в колонке amount мы теперь можем гораздо быстрее искать значение в этой колонке.  
  
  
**Недостатки индексов (почему нельзя создавать индекс на каждую колонку):**

**1.** Надо поддерживать индекс в отсортированном виде (при каждой ставке), соответственно будут более медленные вставки в таблицу. Тоесть если используется большое количество вставок в таблицу, но при этом редко происходит поиск по какой то колонке, то на эту колонку не стоит ставить индекс.

**2.** Индексы занимают память на диске.

**Итог:**

Индексы нужно содздавать с умом. Индексы нужно создавать только для тех колонок, по которым чаще всего производиться поиск (с помощью WHERE, например).

**Индексы в postgreSQL:**

Команада для создания индекса:

**Transaction** – имя таблицы  
**amount** – имя колонки  
**???** - тип индекса

Вместо вопросительных знаков указывается тип индекса, индексы могут быть разных типов, например:

• PRIMARY KEY

• B-Tree Index

• Hash Index

• Gin & Gist Index

• BRIN Index