Session 5 - September 30 - Query performance

In this session

- how PostgresSQL handles the execution of queries
- analyze a query performance with EXPLAIN

EXPLAIN and query performance

Last time

We saw

- normalizing the WorldHits database
- window functions
- simple CTEs

Today

- A lab on windows functions and CTEs (loosely graded)
  https://forms.gle/Y2i6ShyzQkVcpZzr6
- A look at how PostgresSQL optimizes queries
  - query optimizer,
  - cost functions,
  - algorithms
- `EXPLAIN` and `EXPLAIN ANALYZE` to analyze the performance of queries

Forgot

We can use `row_number()` to delete dups see https://blog.devgenius.io/10-postgresql-techniques-i-find-the-most-useful-1f98b5c7c38e

EOD

- you can write CTEs and use window functions in your queries
- you understand what makes a query efficient for the postgreSQL engine
- you can interpret EXPLAIN query plans

Before we start

- if you have already cloned the repo, cd into the project directory and refresh it with `git pull origin master`
- if you haven't cloned the repo yet: `git clone git@github.com:SkatAI/epitadb.git` and `cd epitadb`
- you can also download the repo as a zip file

Resources

- a good post on PostgresSQL functions https://blog.devgenius.io/10-postgresql-techniques-i-find-the-most-useful-1f98b5c7c38e. A pdf version is on the github