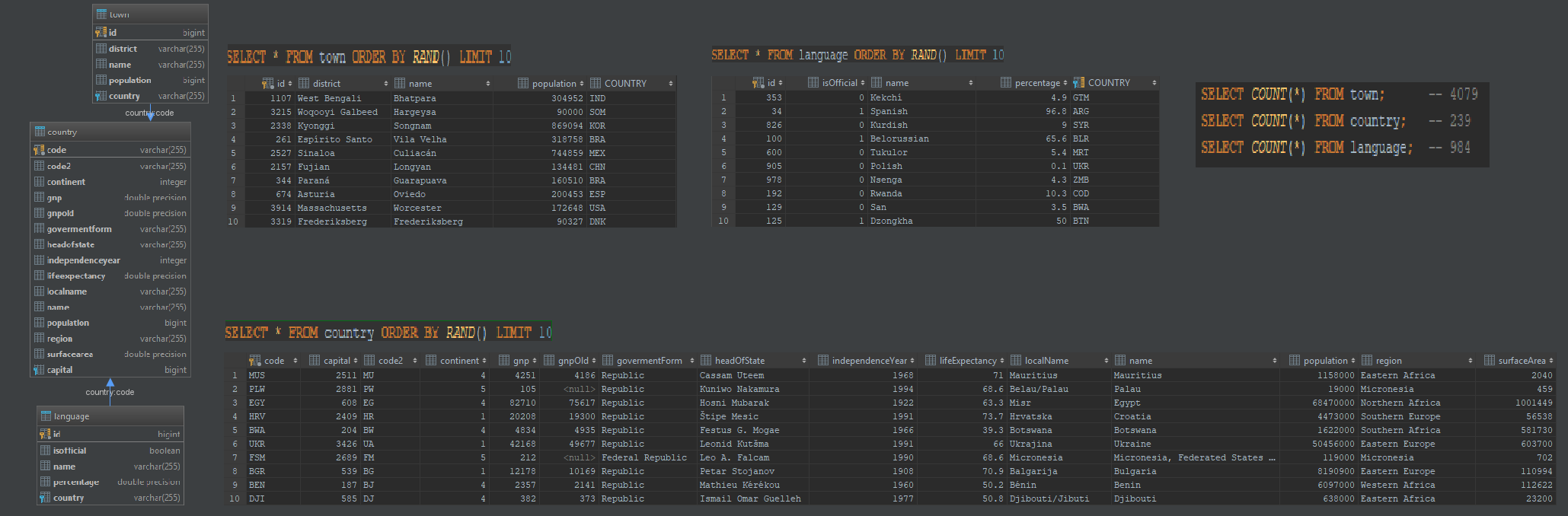
**Test report**

1. **Database schema:**
2. **Queries**

* **Most popular languages by region**

SELECT language.name AS language\_name, MAX(language\_aggregate.region) AS region, MAX(language\_aggregate.language\_population) AS language\_population FROM (

SELECT language\_sum.region, MAX(language\_sum.language\_population) AS language\_population FROM (

SELECT lang.name, cntr.region, sum((lang.percentage / 100) \* cntr.population) AS language\_population

FROM language lang

JOIN country cntr

ON lang.country = cntr.code

GROUP BY lang.name, cntr.region) language\_sum

GROUP BY language\_sum.region) language\_aggregate

JOIN country

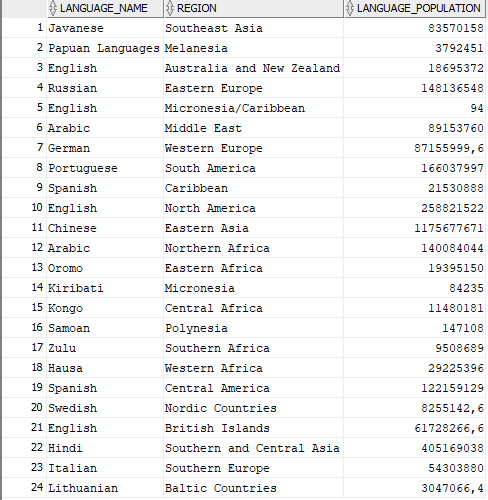
ON country.region = language\_aggregate.region

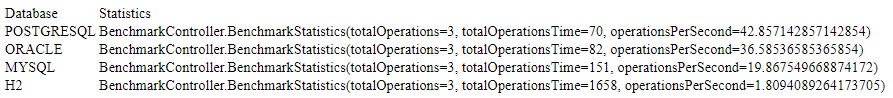
JOIN language

ON language.country = country.code

GROUP BY language.name, country.region

HAVING SUM((language.percentage / 100) \* country.population) = MAX(language\_aggregate.language\_population);





* **Cities from 10 biggest countries**

SELECT town.\*

FROM country

JOIN town

ON town.country = country.code

WHERE (

SELECT COUNT(\*)

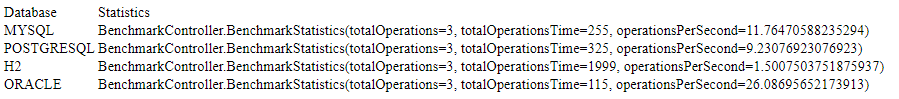
FROM country cntr

WHERE cntr.surfacearea >= country.surfacearea

) <= 10

ORDER BY town.population DESC;





* **Top 3 cities from every region based on population**

SELECT town.\*, country.region

FROM town

JOIN country

ON town.country = country.code

WHERE (

SELECT COUNT(\*)

FROM town twn

JOIN country cntr

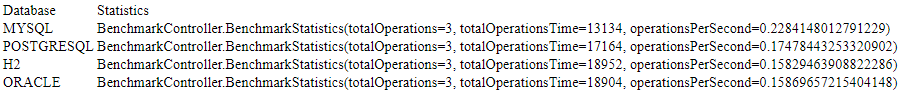
ON twn.country = cntr.code

WHERE twn.population >= town.population AND country.region = cntr.region

) <= 3

ORDER BY country.region;





1. **Conclusions**

Three of the testes databases can be used in production environment, Oracle, MySQL, PostgreSQL. H2 is here only for contrast. As you can see H2 have worst execution times for every query, in first case H2 was about 20 times worst than other databases.

Let’s then focus on these three databases mentioned before, in each case different database won, but in general, Oracle had the most stable results, it of course depends on many factors, for example all of these engines works simultaneously on my local machine, which of course is not the best testing environment.

Another case are indexes, in each queries there was some subqueries, this means at least square complexity, indexes on COUNTRY.POPULATION, COUNTRY.SURFACEAREA and TOWN.POPULATION would for sure speed up a bit all the queries, there can be more indexes of course, but I think, these on mentioned numeric values, are most valuable, as I often searched using these values. Additionally, it is difficult to find two same values for mentioned columns, and that’s good from index point of view.