



PA220: Database systems for data analytics

Home Assignment 2

Vlastislav Dohnal

Prepare Analytical Queries

The queries will be run for at least for the period Sept 1 – Oct 31, if not stated otherwise. I do prefer execution on the whole data.

1. List connection counts per app version and day during Oct 2020.
 - Print also totals per day, per app and grand total.
 - Sort the result by day and app both descending.
 - Add a commentary about possible anomalies.
2. List the connections counts in individual operator networks per weekday and month.
 - Print also the country name of the operator (extend the dimension)
3. Report on oldest and newest app version for all devices, e.g., for “HUAWEI CUN-L21” and not a particular IMSI/IMEI/car_key.
4. Report on the reliability of devices – number of app restarts without device restart (aka app crashes).
 - List top-10 for the combination of app version and device (if any)
5. Identify “connectivity” problems – an app instance connecting too often (i.e., next connection in much shorter time than 5 mins).
 - List raw results (particular devices (IMEIs)), aggregate by device name (Huawei ...) and by app version.

Submission of Assignment 2

- Hand in to the IS vault:
 - ZIP archive with txt files per query,
 - the txt file named `queryN.txt` will contain
 - SELECT command,
 - result of EXPLAIN ANALYZE (plan plus execution time),
 - the first 20 rows (if too many) of output after the execution,
 - your comment if required.
- Grading
 - 2 pts per query
 - total 10 pts