



PA220: Database systems for data analytics

Home Assignment 2

Vlastislav Dohnal

Prepare Analytical Queries

The queries will be executed on the whole data provided in the first home assignment.

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 in network traffic.
2. List connection counts in individual operator networks per day of the week and month.
 - i.e., see the trends of connections per day in week for each month separately
 - Print also the country name of the operator (extend the dimension), sort descendingly by connection counts.
 - <https://is.muni.cz/auth/el/fi/podzim2022/PA220/um/data/mcc-mnc-table.csv>
3. Report on oldest and newest app version for all device models
 - e.g., for “HUAWEI CUN-L21” and not a particular IMSI/car key.
4. Report on the reliability of device models – number of app restarts without device restart (aka app crashes).
 - List top-10 for the combination of app version and device model
5. Identify “connectivity” problems
 - An app instance connecting too often (i.e., next connection in much shorter time than 5 mins), and
 - also, the same for app version and device model.
 - List raw results (particular app instance (sim_imsi)), as well as aggregates by device model (e.g., “HUAWEI CUN-L21”) 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, including execution time,
 - the output after the execution (if too many rows, list the top 20 only),
 - your comment if required.
- Grading
 - 2 pts per query
 - total 10 pts