



## PSafe Total Daily Features Dimensioned

Calculates daily used features of PSafe Total Android product. All user data is grouped by the following dimensions:

- os,
- version,
- model,
- manufacture,
- sys\_version,
- mcc,
- channel,
- language,
- country,
- date (*rounded to day, e.g. 2014-12-31 00:00:00*)

**Note.** The numbers may differ from unique report due to multiple dimensions. E.g. the day of application update will cause increasing numbers due to version change, so that each user has 2 different versions during the day.

```
set hive.exec.dynamic.partition.mode=nonstrict;
```

```
add jar hdfs://PSBIHNN001/user/bi.user/scripts/csv-serde.jar;
```

```
set CURRENT_TIMEZONE='GMT-02:00';
```

```
-- set LAST_REPORT_DATE_PARAM='2014-12-15';
```

```
set FIRST_REPORT_DATE=date_sub(to_date('${LAST_REPORT_DATE_PARAM}'), 1);
```

```
CREATE TABLE IF NOT EXISTS exp_db.exp_FactAnFeatures
```

```
(
  feture STRING, channel STRING, language STRING, version STRING,
  manufacture STRING, model STRING, sys_version STRING, mcc STRING, country STRING, full_count BIGINT, count BIGINT
)
```

```
PARTITIONED BY (year INT, month INT, day INT)
```

```
ROW FORMAT SERDE 'com.bizo.hive.serde.csv.CSVSerde'
```

```
WITH serdeproperties (
```

```
  "separatorChar" = ",",
```

```
  "quoteChar" = "'",
```

```
  "escapeChar" = "\\")
)
```

```
STORED AS TEXTFILE;
```

```
INSERT OVERWRITE TABLE exp_db.exp_FactAnFeatures PARTITION (year, month, day)
```

```
SELECT feature_id, channel, language, version,
  manufacture, model, sys_version, mcc, country,
  sum(full_count) as full_count, sum(count) as count,
  year(time) year,
  month(time) month,
  day(time) day
```

```
FROM
```

```
(
  SELECT
    to_date(from_utc_timestamp(time, ${hiveconf:CURRENT_TIMEZONE})) time, imei, feature_id, channel, language, version,
    machine.manufacture as manufacture, machine.model as model, machine.sys_version as sys_version, machine.mcc as mcc,
    country, sum(cnt) as full_count, sign(sum(cnt)) as count
```

```
FROM
```

```
  staging_db.mobile_operational_data LATERAL VIEW explode(features) ft AS feature_id, cnt
```

```
WHERE
```

```
  to_date(from_utc_timestamp(time, ${hiveconf:CURRENT_TIMEZONE})) = to_date(${hiveconf:FIRST_REPORT_DATE})
```

```
  AND concat(year, '-', month, '-', day) >= to_date(${hiveconf:FIRST_REPORT_DATE})
```

```
  AND concat(year, '-', month, '-', day) <= to_date('${LAST_REPORT_DATE_PARAM}')
```

```
GROUP BY
```

```
  to_date(from_utc_timestamp(time, ${hiveconf:CURRENT_TIMEZONE})), imei, feature_id, channel, language, version,
  machine.manufacture, machine.model, machine.sys_version, machine.mcc, country
```

```
) s1
```

```
GROUP BY
```

```
feature_id, channel, language, version,  
manufacture, model, sys_version, mcc, country,  
year(time),  
month(time),  
day(time);
```

---

/var/www/html/dokuwiki/data/pages/documentacao/bi/hive\_reports/psafe\_total\_daily\_features.txt · Última modificação: 13/01/2015 10:16 por Mikhail Bragin