

Ülesanded PostgreSQL ja Power BI kooskasutamiseks

Koostanud: Rene Aleksander Truuts

Sissejuhatus

PostgreSQL-ist

Avatud lähtekoodiga andmebaasi haldussüsteem PostgreSQL on olnud arenduses üle 30 aasta. Selle skaleeritavuse ja SQL-standardite järgimise tõttu on see populaarne valik nii era- kui avaliku sektori ettevõtete seas. PostgreSQL toetab väga paljusid andmeformaate, sealhulgas järjendeid, JSON-i ja geograafilisi andmeid. PostgreSQL on tuntud ka oma kogukonnapõhise arenduse ja regulaarsete värskenduste poolest.

Microsoft Power BI-st

Microsoft Power BI on andmetöötlusprogramm, mis võimaldab kasutajatel visuaalselt esitada ja analüüsida suuri andmemahтусid.

Power BI-ga saab luua dünaamilisi aruandeid, mis aitavad kiirendada otsustusprotsessi. Power BI võimaldab ühendada andmeid erinevatest allikatest, sealhulgas pilvepõhistest teenustest nagu Azure SQL ja Dynamics 365, ning pakub erinevaid funktsioone, nagu masinõpe ja prognoosimine.

Power BI on kasutusel paljudes erinevates tööstusharudes nagu näiteks rahandus, tervishoid, jaemüük ja tootmine, kus suured andmemahud on paratamatud ning dünaamilised aruanded hädavajalikud. Power BI on saadaval nii tasuta kui ka tasulise versioonina, sõltuvalt kasutaja vajadustest.

Enne õppematerjali juurde asumist

Enne õppematerjali juurde asumist tuleb veenduda, et kõik on vastavalt juhendile installeeritud.

Juhendi link:

https://docs.google.com/document/d/1JROVjWE15uKPo_G0Yg2YoQ8Jz7kU9VFI/edit?usp=sharing&rtpof=true&sd=true

Õppematerjali kohta

1) Ülesanded 1-6 koosnevad PostgreSQL skriptide loomisest. Ülesandega kontrollitakse PostgreSQL süntaksi tundmist ja selle kasutusoskust ülesannete lahendamisel;

2) Ülesannetes 7-10 tuleb vastavalt juhisele visualiseerida enda moodustatud skriptide tulemusi.

Õppematerjali läbinu oskab kasutada Power BI graafikute loomise variante ka suurte PostgreSQL andmebaasidega, samuti oskab ta enda tulemusi erinevate graafikutena visualiseerida. Töölehe koostaja arvates oleks parem kasutada eestikeelset varianti, samas peaks hakkama saama ka ingliskeelse variandiga. Internetis on saadaval rohkem abimaterjale inglise keeles. Kui avad programmi, siis logi sisse oma Tartu Ülikooli kasutajaga kujul kasutaja@ut.ee.

ÜLESANDED

Visualiseeringu jaoks soovitab antud juhendi autor tutvuda Richard Nõmme 2020. Aasta bakalaureusetöoga “Ülevaade Microsoft Power BI-st ja õppematerjali koostamine”.

Töö on leitav siin: https://comserv.cs.ut.ee/ati_thesis/datasheet.php?id=69710&year=2020 .

Näidised:

1. Koosta ülevaade mängijate võitudest, kaotustest ja viikidest kõikidel turniiridel kokku. Järjesta mängijad võitude kahanevas järjekorras.
 - CASE ja SUM funktsioonide kasutamine:
 - Kontrollige, kas CASE avaldiste ja SUM funktsiooni kasutamine on korrektne võitude, kaotuste ja viikide arvutamisel. Hindage, kas CASE avaldistes kasutatud tingimused on õigesti määratud vastavalt ülesande nõuetele (1 punkt).
 - GROUP BY ja ORDER BY klauslite kasutamine ning tulemuste esitamine:
 - Veenduge, et päring rühmitab tulemused korrektselt mängija nime järgi, lisades GROUP BY klauslisse kõik SELECT klauslis kasutatud veerud, mis ei ole agregaatfunktsioonide osad. Kontrollige, kas ORDER BY klausel on õigesti kasutatud (võitude arvu järgi kahanevas järjekorras) (1 punkt).
 - Algallikad ja seosed: kontrollige, kas andmed on võetud õigetest tabelitest partiid ja isikud (1 punkt).

```
SELECT
    isikud.eesnimi,
    isikud.perenimi,
    SUM(CASE
        WHEN partiid.valge = isikud.id AND partiid.valge_tulemus = '2' THEN 1
        WHEN partiid.must = isikud.id AND partiid.must_tulemus = '2' THEN 1
        ELSE 0
    END) AS võidud,
```

```

SUM(CASE
    WHEN partiid.valge = isikud.id AND partiid.valge_tulemus = '0' THEN 1
    WHEN partiid.must = isikud.id AND partiid.must_tulemus = '0' THEN 1
    ELSE 0
END) AS kaotused,
SUM(CASE
    WHEN partiid.valge = isikud.id AND partiid.valge_tulemus = '1' THEN 1
    WHEN partiid.must = isikud.id AND partiid.must_tulemus = '1' THEN 1
    ELSE 0
END) AS viigid
FROM
    partiid
JOIN
    isikud ON partiid.valge = isikud.id OR partiid.must = isikud.id
GROUP BY
    isikud.eesnimi,
    isikud.perenimi
order by võidud desc;

```

| | ABC eesnimi | ABC perenimi | 123 võidud | 123 kaotused | 123 viigid |
|----|-------------|--------------|------------|--------------|------------|
| 1 | Maria | Murakas | 10 | 5 | 5 |
| 2 | Malle | Maasikas | 9 | 3 | 2 |
| 3 | Andrei | Sosnov | 9 | 6 | 8 |
| 4 | Maari | Mustikas | 7 | 4 | 5 |
| 5 | Artur | Muld | 7 | 0 | 12 |
| 6 | Maria | Medvedovna | 6 | 2 | 7 |
| 7 | Jelena | Pirn | 6 | 5 | 4 |
| 8 | Toomas | Umnik | 6 | 1 | 10 |
| 9 | Kalle | Kivine | 6 | 7 | 6 |
| 10 | Irys | Sisalik | 5 | 6 | 5 |
| 11 | Mihkel | Maakamar | 5 | 4 | 6 |
| 12 | Arvo | Mets | 5 | 8 | 1 |
| 13 | Pjotr | Pustota | 5 | 1 | 12 |

2. Koosta ülevaade sellest, kuidas jagunesid igal turniiril iga klubi võitude, viikide ja kaotuste arvud. Järjesta tulemus turniiri nime ja võitude järjekorras.
 - Liitumiste ja agregeeritud funktsioonide kasutamine:
 - Kontrollige, kas päring ühendab tabelid õigesti (JOIN klauslid) ja kasutab SUM funktsiooni klubide võitude, viikide ja kaotuste loendamiseks, arvestades mängija kuulumist klubisse (1 punkt).
 - GROUP BY ja ORDER BY klauslite kasutamine:
 - Veenduge, et päring kasutab GROUP BY klauslit turniiri ja klubi nime järgi rühmitamiseks ning ORDER BY klauslit tulemuste esitamiseks turniiri nime alusel. See aitab kaasa selgele ja loogilisele tulemuste esitlusele (1 punkt).
 - Päringu loetavus ja veergude nimede selgus:

- Hinnake päringu loetavust, sealhulgas veerunimede asjakohasust ja selgust. Tuletatud veerud peaksid omama nimetusi, mis peegeldavad nende sisu, näiteks võidud ja kaotused (1 punkt).

```
SELECT turniirid.nimi AS turniir,
       klubid.nimi AS klubi_nimi,
       SUM(CASE WHEN partiid.valge_tulemus = '2' AND valge.klubi = klubid.id
THEN 1 ELSE 0 END) +
       SUM(CASE WHEN partiid.must_tulemus = '2' AND must.klubi = klubid.id THEN
1 ELSE 0 END) AS võidud,
       SUM(CASE WHEN partiid.valge_tulemus = '1' AND valge.klubi = klubid.id
THEN 1 ELSE 0 END) +
       SUM(CASE WHEN partiid.must_tulemus = '1' AND must.klubi = klubid.id THEN
1 ELSE 0 END) AS viigid,
       SUM(CASE WHEN partiid.valge_tulemus = '0' AND valge.klubi = klubid.id
THEN 1 ELSE 0 END) +
       SUM(CASE WHEN partiid.must_tulemus = '0' AND must.klubi = klubid.id THEN
1 ELSE 0 END) AS kaotused
FROM partiid
JOIN isikud AS valge ON partiid.valge = valge.id
JOIN isikud AS must ON partiid.must = must.id
JOIN klubid ON valge.klubi = klubid.id OR must.klubi = klubid.id
JOIN turniirid ON partiid.turniir = turniirid.id
GROUP BY turniirid.nimi, klubi_nimi
order by turniirid.nimi, võidud desc;
```

| ABC turniir | ABC klubi_nimi | 123 võidud | 123 viigid | 123 kaotused |
|------------------------------|------------------------|------------|------------|--------------|
| Eesti meistrivõistlused 2007 | Maletäht | 10 | 10 | 2 |
| Eesti meistrivõistlused 2007 | Ajurebend | 7 | 4 | 1 |
| Eesti meistrivõistlused 2007 | Raudne Ratsu | 6 | 3 | 2 |
| Eesti meistrivõistlused 2007 | Areng | 5 | 7 | 8 |
| Eesti meistrivõistlused 2007 | Võitmatu Valge | 4 | 2 | 5 |
| Eesti meistrivõistlused 2007 | Vabaettur | 4 | 3 | 8 |
| Eesti meistrivõistlused 2007 | Valge Mask | 3 | 2 | 2 |
| Eesti meistrivõistlused 2007 | Pärnu Parimad | 3 | 8 | 5 |
| Eesti meistrivõistlused 2007 | Ruudu Liine | 2 | 7 | 5 |
| Eesti meistrivõistlused 2007 | Chess | 2 | 5 | 6 |
| Eesti meistrivõistlused 2007 | Laudnikud | 1 | 2 | 3 |
| Eesti meistrivõistlused 2007 | Musta kivi kummardajad | 1 | 5 | 1 |
| Kolme klubi kohtumine | Võitmatu Valge | 8 | 8 | 4 |

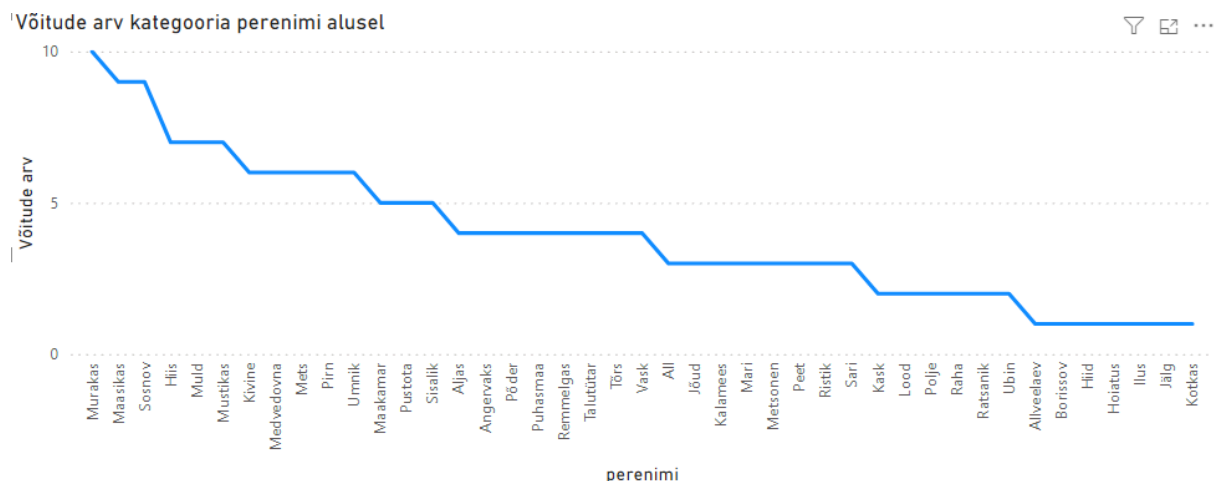
3. Leia kõige edukamad mängijad võiduprotsendi alusel, samuti järjestage need võiduprotsendi järjekorras.. Ühenda ees- ja perenimi ühiseks veeruks mängija_nimi.
NB! Ära unusta arvutusse ka viike arvestada! Järjestage tulemus võiduprotsendi järgi.
4. Leia kõikide mängijate võiduprotsent **mustal pool** ja järjestage nad selle alusel. Ühenda mängijate ees- ja perenimi veeruks mängija_nimi. Kasuta tulemuse välja

võiduprotsent arvutamisel funktsioone ROUND() ja CAST AS (numeric). Kontrolli, et kuvate ainult mängijad, kellel on rohkem kui null mängu. Järjesta tulemus võiduprotsendi järgi.

5. Koosta päring, mis arvutab kõigi mängijate keskmise mänguaaja (algusest lõpuni) minutites igal turniiril eraldi (välja nimeks näiteks keskmine_mänguaeg_minutites). Järjesta mängijad keskmise mänguaaja järgi. Kasuta funktsioone EXTRACT HOUR FROM ja MINUTE FROM. Ümarda keskmine mänguaeg 2 komakohani.
6. Koosta päring, mis loob igale klubile edetabeli, kus on mängijate nimed ja nende võitude arv klubis, järjestades iga klubi kõik mängijad võitude järjekorras, kusjuures klubid on tähestikulises järjekorras.

NB! SQL-e Power BI-sse lisades tuleks seda teha ühe kaupa, samuti tuleks päringu lõpust eemaldada semikoolon, kui olete selle sinna lisanud.

7. Visualiseeri ülesande 1 skript nii, et horisontaalteljel on mängija perenimi ja joondiagrammina on ka iga mängija võitude arv.



8. Visualiseeri ülesande 2 skript vabalt nii, et tegemist ei oleks tulpdiaagrammiga.
9. Kujuta ülesande 3 skript vabalt valitud visualiseeringuga, mille olete leidnud Microsoft Power BI-st.
10. Kujuta punktis 4 leitud võidud joondiagrammina, kus mängija_nimi peab asetsema x-teljel ning võiduprotsent y-teljel.

Vastuste fail:

<https://docs.google.com/document/d/1jmvj0rvufKPDt39oHVpJkPBPtcR5rneL/edit?usp=sharing&oid=100360935309244295279&rtpof=true&sd=true>