DATA607_Project 3 - Data Science Skills

Team 3: David Simbandumwe, Thomas Buonora, Charles Ugiagbe, Jaya Veluri

2021-10-17

```
usr <- keyring::key_list("DATA607")[1,2]
pwd <-  keyring::key_get("DATA607", usr)
con = dbConnect(MySQL(), user=usr, password=pwd, dbname='DATA607', host='localhost')
```

```
# Job Opening
query <- "DROP TABLE IF EXISTS JobOpenings CASCADE;"
results <- dbSendQuery(con, query)
query<-"CREATE TABLE JobOpenings (
    job_id              INT,
    min_experience      INT,
    max_experience      INT,
    skill               TEXT,
    location            TEXT,
    min_salary          INT,
    max_salary          INT,
    company_id          INT,
    PRIMARY KEY (job_id)
);"
results <- dbSendQuery(con, query)
dbClearResult(results)
```

```
## [1] TRUE
```

```
query <- "DROP TABLE IF EXISTS JobSeekers CASCADE;"
results <- dbSendQuery(con, query)
query <- "CREATE TABLE JobSeekers (
    resp_id             INT,
    gender              TEXT,
    age                 INT,
    location            TEXT,
    education           TEXT,
    major               TEXT,
    title               TEXT,
    industry            TEXT,
    experience          TEXT,
    dataScientist       TEXT,
    primarySkill        TEXT,
```

```
    skill               TEXT,
    PRIMARY KEY (resp_id)
);"
results <- dbSendQuery(con, query)
dbClearResult(results)
```

```
## [1] TRUE
```

```
query <- "DROP TABLE IF EXISTS JobSalary CASCADE;"
results <- dbSendQuery(con, query)
query <- "CREATE TABLE JobSalary     (
    id                  INT,
    min                 INT,
    max                 INT,
    PRIMARY KEY (id)
);"
results <- dbSendQuery(con, query)
dbClearResult(results)
```

```
## [1] TRUE
```

```
query <- "DROP TABLE IF EXISTS JobLocation CASCADE;"
results <- dbSendQuery(con, query)
query <- "CREATE TABLE JobLocation   (
    id              INT,
    location_id     INT
);"
results <- dbSendQuery(con, query)
dbClearResult(results)
```

```
## [1] TRUE
```

```
query <- "DROP TABLE IF EXISTS JobRequirements CASCADE;"
results <- dbSendQuery(con, query)
query <- "CREATE TABLE JobRequirements (
    id              INT,
    skill_id        INT
);"
results <- dbSendQuery(con, query)
dbClearResult(results)
```

```
## [1] TRUE
```

```
query <- "DROP TABLE IF EXISTS JobSeekerSkills CASCADE;"
results <- dbSendQuery(con, query)
query <- "CREATE TABLE JobSeekerSkills (
    id              INT,
    skill_id        INT,
    PRIMARY KEY (id)
);"
results <- dbSendQuery(con, query)
dbClearResult(results)
```

```
## [1] TRUE
```

```
query <- "DROP TABLE IF EXISTS JobSeeker CASCADE;"
results <- dbSendQuery(con, query)
query <- "CREATE TABLE JobSeeker (
    id                  INT,
    location            TEXT,
    education_level_id  INT,
    major_id            INT,
    title_id            INT,
    industry_id         INT,
     PRIMARY KEY (id)
);
"
results <- dbSendQuery(con, query)
dbClearResult(results)
```

```
## [1] TRUE
```

```
query <- "DROP TABLE IF EXISTS SkillsMeta CASCADE;"
results <- dbSendQuery(con, query)
query <- "CREATE TABLE SkillsMeta (
    id                  INT,
    key_skills          TEXT,
    PRIMARY KEY (id)
);
"
results <- dbSendQuery(con, query)
dbClearResult(results)
```

```
## [1] TRUE
```

```
dbGetQuery(con, "insert into SkillsMeta
                    (id,key_skills)
                values
                    (1,'analytics'),
                    (2,'big data'),
                    (3,'big data analytics'),
                    (4,'data privacy'),
                    (5,'data science'),
                    (6,'effective communication'),
                    (7,'fraud analytics'),
                    (8,'hadoop'),
                    (9,'machine learning'),
                    (10,'machine learning engineer'),
                    (11,'marketing automation'),
                    (12,'matlab'),
                    (13,'model development'),
                    (14,'natural language processing'),
                    (15,'predictive analytics'),
                    (16,'python'),
                    (17,'r'),
```

```
                    (18,'regression testing'),
                    (19,'sql'),
                    (20,'team leading')
            ;")
```

```
## data frame with 0 columns and 0 rows
```

```
dbDisconnect(con)
```

```
## [1] TRUE
```