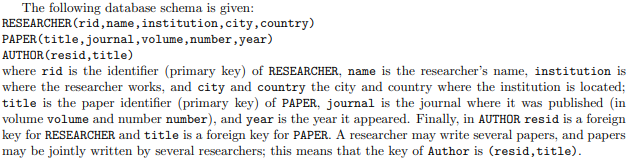
Homework 2

Jacob Taylor Cassady

CECS 535: Introduction to Databases

February 19, 2020



# List the names of authors of any paper published in journal “Databases” in 2019 in volume 12.

**SELECT** R.name

**FROM** RESEARCHER **AS** R, AUTHOR **AS** A, PAPER **AS** P

**WHERE** R.rid = A.resid **and** P.title = A.title **and** P.journal = “Databases” **and** P.year = 2019 **and** P.volume=12;

# List the names of authors who have published a paper in both “Nature” and in “Science” (journals).

**SELECT** name

**FROM** (**SELECT** name, rid

**FROM** RESEARCHER **AS** R, AUTHOR **AS** A, PAPER **AS** P

**WHERE** R.rid = A.resid **and** P.title = A.title **and** R.journal = “Nature”)

**INTERSECT**

(**SELECT** name, rid

**FROM** RESEARCHER **AS** R, AUTHOR **AS** A, PAPER **AS** P

**WHERE** R.rid = A.resid **and** P.title = A.title **and** R.journal = “Science”);

# List the names of authors who have published two or more papers in “Nature” (journal).

**SELECT** name

**FROM** RESEARCHER **AS** R, AUTHOR **AS** A

**WHERE** R.rid = A.resid **and** R.rid **IN (SELECT** A.resid

**FROM** PAPER **AS** P, AUTHOR **AS** A

**WHERE** P.title = A.title

**GROUP BY** A.resid

**HAVING** sum(CASE journal=”Nature” 1 else 0) >= 2);

# List the names of authors who have published a paper in “Nature” but never in “Science” (journals).

**SELECT** name

**FROM** **(SELECT** name, rid

**FROM** RESEARCHER **AS** R, AUTHOR **AS** A, PAPER **AS** P

**WHERE** R.rid = A.resid **and** P.title = A.title **and** R.journal = “Nature”)

**EXCEPT**

**(SELECT** name, rid

**FROM** RESEARCHER **AS** R, AUTHOR **AS** A, PAPER **AS** P

**WHERE** R.rid = A.resid **and** P.title = A.title **and** R.journal = “Science”);

# List the names of authors who have published a paper in “Nature” (journal) alone (i.e. without co-authors).

**SELECT** name

**FROM** (**SELECT** name, resid

**FROM** RESEARCHER **AS** R, AUTHOR **AS** A

**WHERE** R.rid = A.resid **and** A.title **IN** (**SELECT** title

**FROM** (**SELECT** title, **COUNT**(resid) **AS** aCount

**FROM** AUTHOR

**GROUP BY** title)

**WHERE** aCount = 1;

# List the title of papers where all authors are from the same institution. Note: this includes sole authors too.

**SELECT** P.title

**FROM** RESEARCHER **AS** R, AUTHOR **AS** A, PAPER **AS** P

**WHERE** R.rid = A.resid **and** P.title = A.title **and** A.institution = (**SELECT**

**FROM** RESEARCHER **AS** R1,

AUTHOR **AS** A1, PAPER **AS** P1

**WHERE** P.title = P1.title **and**

R1.rid = A1.resid);

# Find out the average number of authors per paper.

**SELECT** A.title, **AVG**(resid)

**FROM** AUTHOR **AS** A, PAPER **AS** P

**WHERE** A.title = P.title

**GROUP BY** A.title

# Find out the number of papers per country (count each paper once, regardless of the number of authors).

**SELECT** R.country, **COUNT**(**DISTINCT** P.title)

**FROM** AUTHOR **AS** A **and** PAPER **AS** P **and** RESEARCHER **AS** R

**WHERE** A.title = P.title **and** R.rid = P.resid

**GROUP BY** R.country;

# List the number of papers per author (identified by author id) but count each paper in proportion to the number of authors. That is, if a paper has n authors, split the credit equally among them (1/n units of credit). Example: if researcher Ullman has 3 papers, one alone, one with somebody else, and one with two other co-authors, he gets 1, .5 and .3 units of credit respectively, for a total of 1.8 papers.

**CREATE VIEW** TitleResearcherCount(title, researchercount) **AS**

**SELECT** A.title, **COUNT**(A.resid)

**FROM** AUTHOR **AS** A

**GROUP BY** A.title;

**CREATE VIEW** ResearcherTitleScore(rid, title, authorscore) **AS**

**SELECT** R.rid, A.title, 1/**COUNT**(A.\*)

**FROM** RESEARCHER **AS** R, AUTHOR **AS** A, TitleResearcherCount **AS** TRC

**WHERE** R.rid = A.resid **and** TRC.title = A.title

**GROUP BY** R.rid;

**SELECT** R.rid, **SUM**(RTS.authorscore)

**FROM** RESEARCHER **AS** R, ResearcherTitleScore **AS** RTS

**WHERE** R.rid = RTS.rid

**GROUP BY** R.rid;

# List the journals that published more papers in 2019 than in 2018.

**SELECT** P.journal, P.year

**FROM** PAPER **AS** P

**WHERE** P.year = 2019

**GROUP BY** P.journal, P.year

**HAVING** **COUNT**(**DISTINCT** P.title) > (**SELECT** 2018\_COUNT

**FROM** (**SELECT** **COUNT**(**DISTINCT** P1.title) **AS** 2018\_COUNT,

P1.journal, P1.year

**FROM** PAPER **AS** P1

**WHERE** P1.journal = P.journal **and** P1.year = 2018));