Advanced Topics in Databases

1st Deliverable

Demertzoglou Efstratios | TH20580

Tatsios Ilias | TH20566

Table of Contents

The csv file I will be working with

The Universal table is a group of attributes from the *races.csv*

UNIVERSAL(raceid,year,driverid,points,circuitid,alt)

Functional dependencies:

raceid --> year;

raceid, points --> driverid;

raceid --> circuitid;

circuited --> alt;

Initial Relational Schema

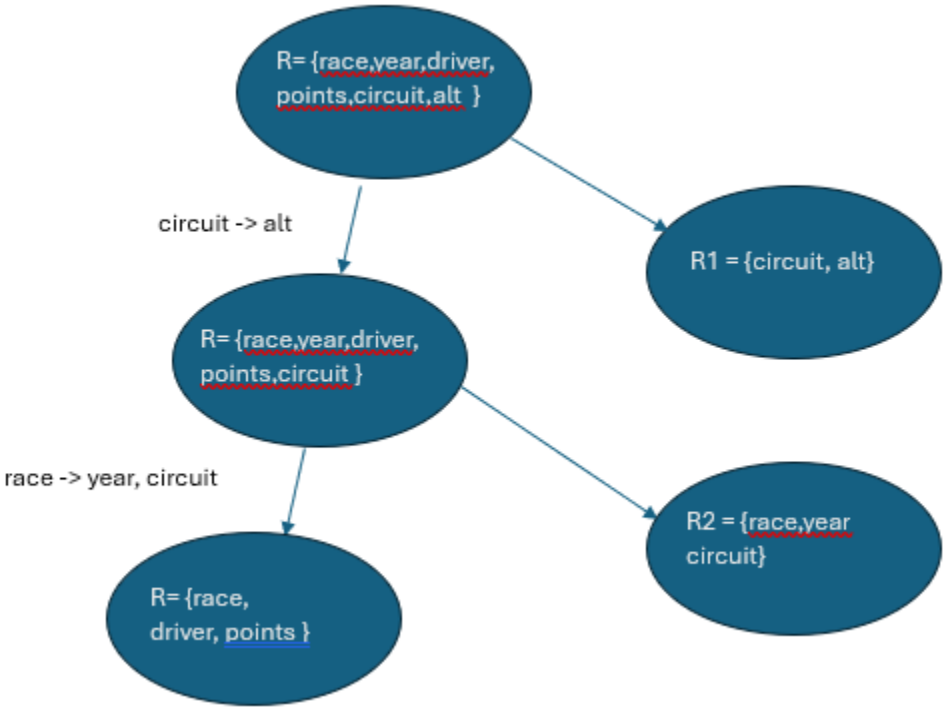
```
DROP TABLE IF EXISTS DRIVER_STANDINGS CASCADE;

DROP TABLE IF EXISTS RACES CASCADE;

DROP TABLE IF EXISTS CIRCUITS CASCADE;


CREATE TABLE CIRCUITS(

circuitId INT NOT NULL,

alt INT NOT NULL,

PRIMARY KEY(circuitId));


COPY CIRCUITS FROM 'C:\uni\8x\ATD\archive\circuits_modified.csv' DELIMITER ',' CSV HEADER;


CREATE TABLE RACES (

raceId INT NOT NULL,

circuitId INT NOT NULL,

FOREIGN KEY (circuitId) REFERENCES CIRCUITS (circuitId),

PRIMARY KEY (raceId));


COPY RACES FROM 'C:\uni\8x\ATD\archive\races_modified.csv' DELIMITER ',' CSV HEADER;


CREATE TABLE DRIVER_STANDINGS (

raceId INT NOT NULL,

driverId INT NOT NULL,

points INT NOT NULL,

FOREIGN KEY (raceId) REFERENCES RACES (raceId),

PRIMARY KEY(driverId,raceId));


COPY RESULTS FROM 'C:\uni\8x\ATD\archive\driver_standings_modified.csv' DELIMITER ',' CSV HEADER;


COPY (

  SELECT * FROM CIRCUITS NATURAL JOIN RACES NATURAL JOIN RESULTS

) TO 'C:\uni\8x\ATD\archive\NATURALJOINTABLE.csv' DELIMITER ',' CSV HEADER;
```

By creating a natural join between these 3 tables as:

"*SELECT * FROM CIRCUITS NATURAL JOIN RACES NATURAL JOIN DRIVER_STANDINGS*"

The output projected should be the same to "*SELECT * FROM UNIVERSAL*"

Propose a Joinless Decomposition



Candidate Key

UNIVERSAL(raceid,year,driverid,points,circuitid,alt)

raceid --> year;

raceid, points --> driverid;

raceid --> circuitid;

circuited --> alt;

We have to prove that {raceid, points} is a candidate key

X: desired candidate key = { raceid, points }

1st step is to set X+ = { raceid, points }

1*st Iteration*

F = { raceid → year }
$Old\_X+ = \{raceid, points\}$
raceid → year $\in F$
if raceId $\subseteq \{raceid, points\}$

$Then\ X+ = \{raceid, points\} \cup (year) = \{raceid, points, year\}$

$\{raceid, points\} \neq \{raceid, points, year\}$

2*nd Iteration*

F = { raceid → year , raceid →circuit}

$Old\_X+ = \{raceid, points, year\}$

raceid –> circuitid $\in F$
if year $\subseteq \{raceid, points, year\}$

$Then\ X+ = (raceid, points, year) \cup (circuitid) = (raceid, points, year, circuitid)$

$(raceid, points, year) \neq (raceid, points, year, circuitid)$

$3rd\ Iteration$

F = { raceid → year , raceid →circuitid, circuitid →alt}

$Old\_X+ =$ {raceid, points, year, $circuit$id}

circuitid –> alt ∈ F

$if\ circuit$id ⊆ {raceid, points, year, $circuit$id}

$Then\ X+ =$ (raceid, points, year, $circuit$id) ∪ $(alt) =$ (raceid, points, year, $circuit$id, alt)

(raceid, points, year, $circuit$id) ≠ (raceid, points, year, circuitid, alt)

$4th\ Iteration$

F = { raceid → year , raceid →circuitid, circuitid →alt, raceid -> alt}

$Old\_X+ =$ {raceid, points, year, $circuit$id, alt}

raceid –> alt ∈ F

$if\ circuit$ ⊆ {raceid, points, year, $circuit$id, alt}

$Then\ X+ =$ (raceid, points, year, $circuit$id, alt) ∪ $(alt) =$ (raceid, points, year, $circuit$id, alt)

(raceid, points, year, $circuit$id, alt) = (raceid, points, year, circuitid, alt)

Therefore {raceid, points} is candidate key

Link to all necessary CSVs and sql scripts:

https://drive.google.com/drive/folders/1zmYyTNyI_6Miez2muLYqYLkN5V4PmFNU?usp=sharing

R1 = {circuit, alt}