Group 02-4
CSC 315-02

Stage IV – Elaboration: Database Design

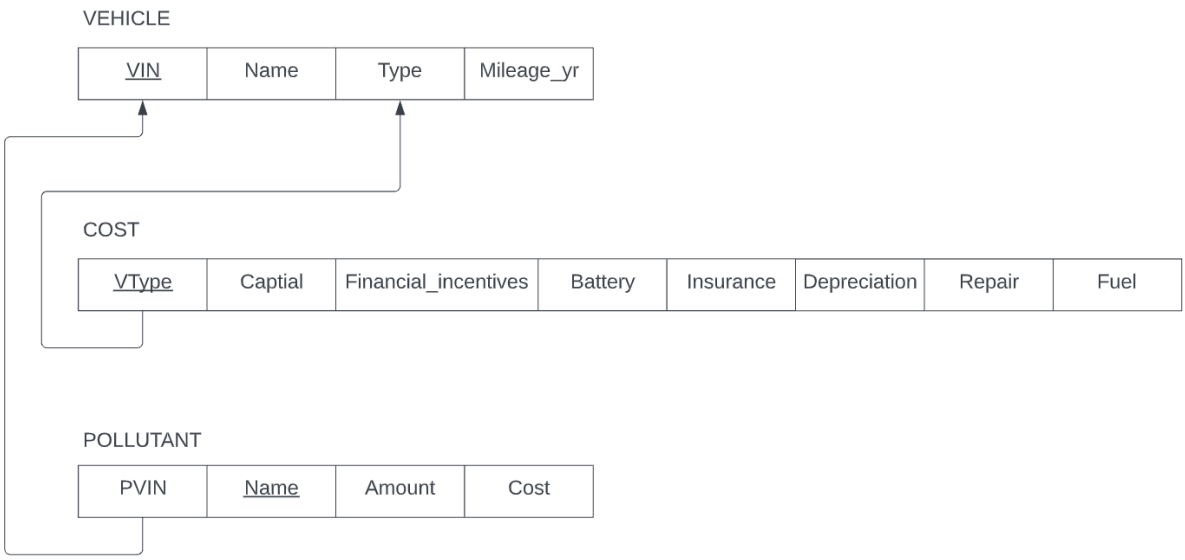
Team 02-04
Jenna Stiesi, Parvathi Krishnan, Joseph Carmichael

Modifications are shown in red.

1. Demonstrate that all the relations in the relational schema are normalized to Boyce–Codd normal form (BCNF).
    a. For each table, specify whether it is in BCNF or not, and explain why.
    b. For each table that is not in BCNF, show the complete process that normalizes it to BCNF.
2. Define the different views (virtual tables) required. For each view list the data and transaction requirements. Give a few examples of queries, in English, to illustrate.
3. Design a complete set of SQL queries to satisfy the transaction requirements identified in the previous stages, using the relational schema and views defined in tasks 2 and 3 above.

Old schema, not in BCNF:

RELATIONAL SCHEMA



1. All the relations are in BCNF besides the Vehicle Relation

    1NF - there are only atomic values and no composite or multivalued attributes so the relation is in 1NF.

    2NF - there are no partial dependencies and every non-prime attribute is fully dependent on the primary key

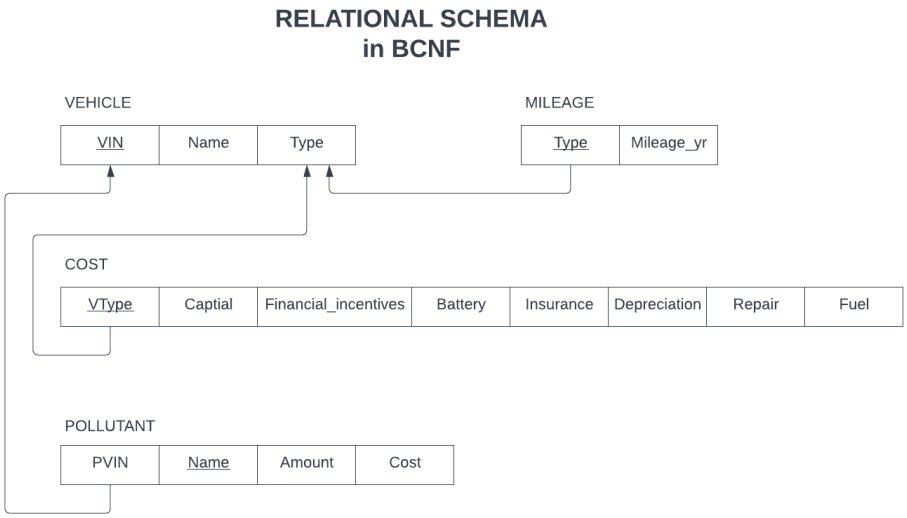
Group 02-4
CSC 315-02

3NF- there are no transitive dependencies and every non-prime attribute is fully dependent on the primary key

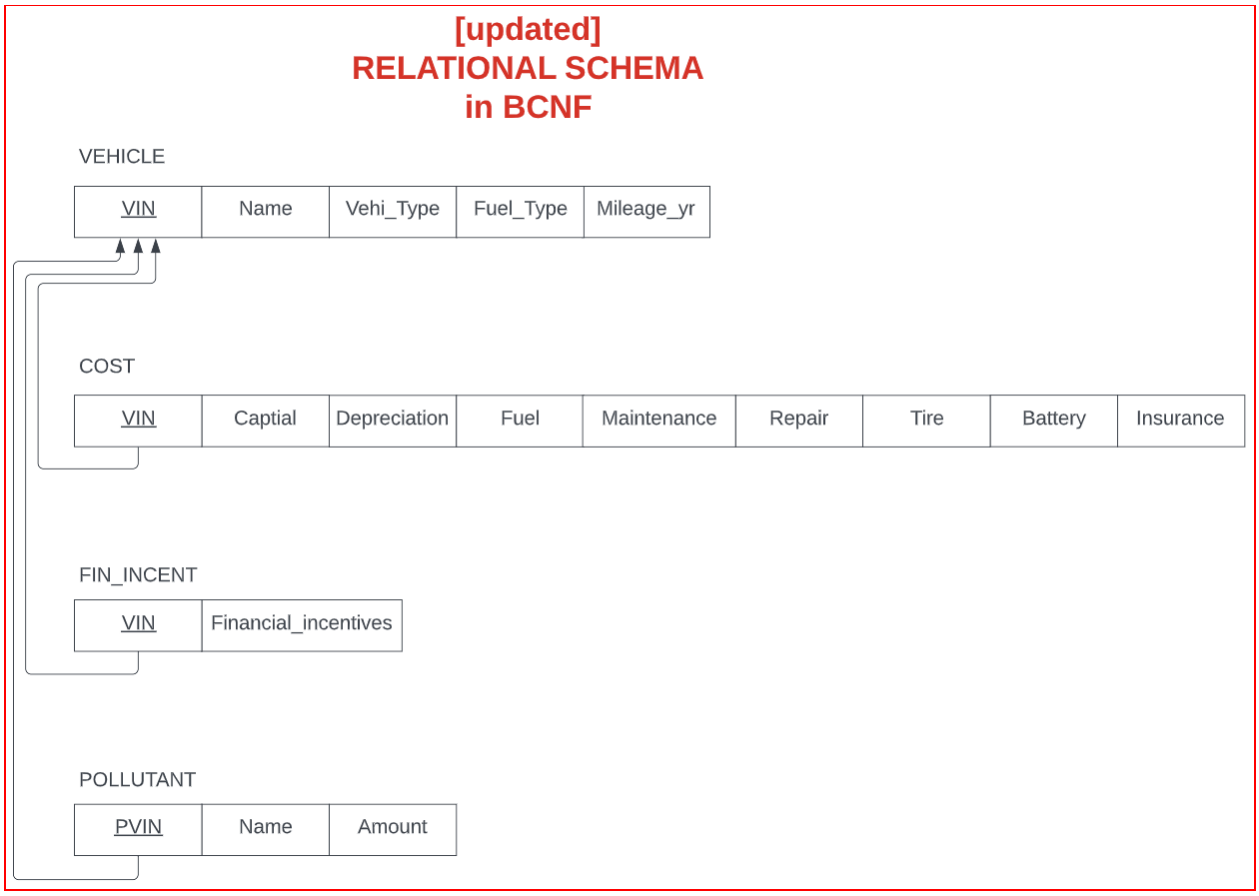
BCNF - every relation has a superkey and so it is in BCNF.

a. The Vehicle relation is in 1NF and 2NF but not in 3NF because the Mileage_Yr depends on Type which depends on VIN
    i. VIN → Type → Mileage_Yr
    ii. To fix this, we decompose the relation into separate relations.
    iii. Now, the table is in 3NF and BCNF
b. The Cost relation
    i. Is in 1NF as there are only atomic values for each attribute.
    ii. It is in 2NF because all the attributes are dependent on the primary key which is VType.
    iii. It is in 3NF because there are no transitive dependencies and all of them depend on the primary key directly.
    iv. In BCNF because the primary key which any FDs depend upon is a superkey..
c. The Pollutant relation is:
    i. In 1NF because there are only atomic values and no composite or multivalued attributes.
    ii. In 2NF because there are no partial dependencies and every non-prime attribute is fully dependent on the primary key.
    iii. In 3NF because there are no transitive dependencies and every non-prime attribute is fully dependent on the primary key.
    iv. In BCNF because the primary key which any FDs depend upon is a superkey.

New schema in BCNF:



Group 02-4
CSC 315-02



2. ***View:***
   ~~a. Join between VEHICLE and MILEAGE tables:~~

   ~~CREATE VIEW VEHI_MIL~~
   ~~AS SELECT VIN, Name, Type, Mileage_yr~~
   ~~FROM VEHICLE, MILEAGE~~
   ~~WHERE VEHICLE.Type = MILEAGE.Type;~~

   ~~The data and transaction requirements require the VEHI_MIL view to be used in an SQL query in our web interface. These are satisfied because we are using the one view we have created in our web interface queries. Example queries using this view can be seen in the answer to question 3.~~

   After modifying our schema, this view is not needed.

3. ***Queries :***
   a. For economic: Output VIN, name, vehi_type, fuel_type, mileage_yr, capital (eg. if "capital" is chosen from dropdown)
   b. For emissions: Output PVIN, Vname, vehi_type, fuel_type, Pname, mileage_yr, amount/cost (if "amount" or “cost” is chosen from dropdown)

Group 02-4
CSC 315-02

Economic queries:

```
SELECT VEHICLE.VIN, Name, Vehi_type, Fuel_type, Mileage_yr, Capital
FROM VEHICLE
INNER JOIN COST ON VEHICLE.VIN=COST.VIN
WHERE Capital < 20000
ORDER BY Capital ASC;
```

Environmental queries:

```
SELECT POLLUTANT.Name, Amount, VIN, VEHICLE.Name, Vehi_type,
Fuel_type, Mileage_yr
FROM VEHICLE
INNER JOIN POLLUTANT ON VEHICLE.VIN=POLLUTANT.PVIN
WHERE Amount < 1500
ORDER BY Amount ASC;
```