EXAMINATION GUIDE: NORMALIZATION PROBLEM AND SOLUTION

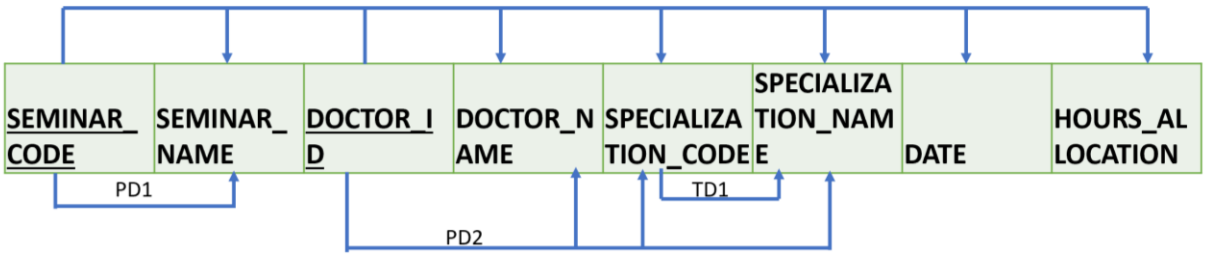
Convert the table below to 2NF and 3NF while showing all the steps of normalization. Resulting tables must contain data at each stage of the normal form.

SEMINAR_CODE	SEMINAR_NAME	DOCTOR_ID	DOCTOR_NAME	SPECIALIZATION_CODE	SPECIALIZATION_NAME	DATE	HOURS_ALLOCATION
s001	Homeopathic Medicine	D_0108	Samuel Jones	CAR01	Cardiology	01/01/2010	8
s001	Homeopathic Medicine	D_0124	Julia Cortez	NEUR1	Neurology	01/01/2010	16
s001	Homeopathic Medicine	D_0140	Mark Harris	ONCO2	Oncology	01/01/2010	8
s002	Cellfood Benefits	D_0110	Anna Harris	STIN01	Ent	01/07/2010	8
s002	Cellfood Benefits	D_0140	Mark Harris	ONCO2	Oncology	01/01/2011	8
s002	Cellfood Benefits	D_0124	Julia Cortez	NEUR1	Neurology	01/07/2010	16

In 1NF, we identify the following.

- Primary key (SEMINAR_CODE + DOCTOR_ID)
- All functional dependencies (partial and transitive)

Using a dependency diagram IN 1NF to identify functional dependencies.



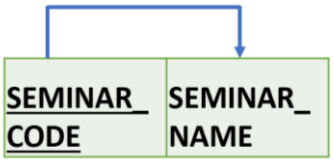
In 2NF, we need to ensure the following.

- The table in 1NF
- Remove all partial dependencies to form new tables/relations.

Partial dependency 1 (PD1)

PD1 : SEMINAR_CODE → SEMINAR_NAME

Dependency diagram for PD1



Relation schema:

SEMINAR (**SEMINAR_CODE**, SEMINAR_NAME)

Table name: SEMINAR, Primary key: SEMINAR_CODE. Foreign key: NONE

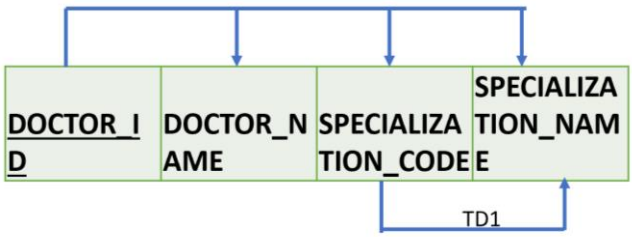
SEMINAR_CODE	SEMINAR_NAME
s001	Homeopathic Medicine
s002	Cellfood Benefits

Partial dependency 2 (PD2)

PD2: DOCTOR_ID → DOCTOR_NAME, SPECIALIZATION_CODE, SPECIALIZATION_NAME

Note (PD2: has a transitive dependency: TD1: SPECIALIZATION_CODE → SPECILIZATION_NAME)

Dependency diagram for PD2



Relation schema:

DOCTOR (**DOCTOR_ID,** DOCTOR_NAME, SPECILIZATION_CODE, SPECIALIZATION_NAME)

Table name: DOCTOR, Primary key: DOCTOR_ID, Foreign key: NONE

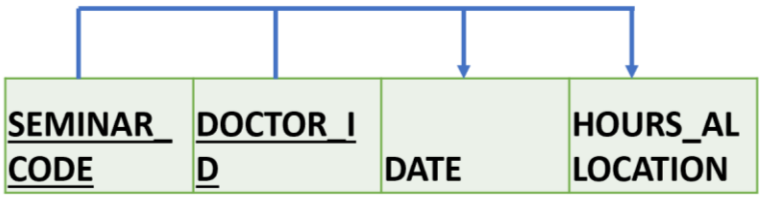
DOCTOR_ID	DOCTOR_NAME	SPECIALIZATION_CODE	SPECIALIZATION_NAME
D_0108	Samuel Jones	CAR01	Cardiology
D_0124	Julia Cortez	NEUR1	Neurology

D_0140	Mark Harris	ONCO2	Oncology
D_0110	Anna Harris	STIN01	Ent

We create a new table/relation to contain DATE, ALLOCATION with DOCTOR_ID and SEMINAR_CODE as primary key.

SEMINAR_CODE, DOCTOR_ID → DATE, ALLOCATION_HOURS

Dependency diagram



Relation schema:

SEMINAR_DOCTOR (**SEMINAR_CODE, DOCTOR_ID,** DATE, HOURS_ALLOCATION)

Table name: SEMINAR_DOCTOR, **Primary key**: SEMINAR_CODE + DOCTOR_ID, **Foreign keys**: SEMINAR_CODE, DOCTOR_ID

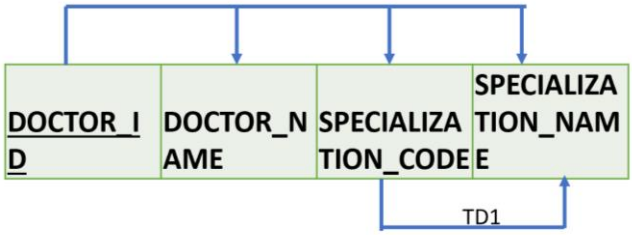
SEMINAR_CODE	DOCTOR_ID	DATE	HOURS_ALLOCATION
s001	D_0108	01/01/2010	8
s001	D_0124	01/01/2010	16
s001	D_0140	01/01/2010	8
s002	D_0110	01/07/2010	8
s002	D_0140	01/01/2011	8
s002	D_0124	01/07/2010	16

In 3NF, we need to ensure the following.

- The tables are in 2NF.
- Remove all transitive dependencies to form new tables/relations.

In this example, we have one transitive dependency (TD1) that exists in PD2.

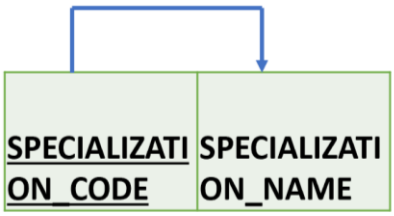
PD2: DOCTOR_ID → DOCTOR_NAME, SPECIALIZATION_CODE, SPECIALIZATION_NAME



TD1: SPECIALIZATION_CODE → SPECILIZATION_NAME

Note: TD1 forms a new table of its own to the SPECIALIZATION_CODE as Primary.

Dependency diagram for TD1



Relation schema for TD1:

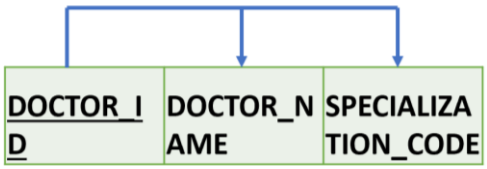
SPECIALIZATION (**SPECIALIZATION_CODE**, SPECILIZATION_NAME)

Table name: SPECIALIZATION, **Primary key**: SPECIALIZATION_CODE, **Foreign keys**: NONE

SPECIALIZATION_CODE	SPECIALIZATION_NAME
CAR01	Cardiology
NEUR1	Neurology
ONCO2	Oncology
STIN01	Ent

Note: PD2 relation schema changes to include as a foreign key, as we remove the transitive dependency, whereas the rest of the tables in 2NF remain unchanged.

Dependency diagram for PD2



Partial dependency for PD2:

PD2: DOCTOR_ID → DOCTOR_NAME, SPECIALIZATION_CODE

Relation schema for PD2

DOCTOR(**DOCTOR_ID**, DOCTOR_NAME, SPECIALIZATION_CODE)

Table name: DOCTOR, **Primary key**: DOCTOR_ID, **Foreign keys**: SPECIALIZATION_CODE

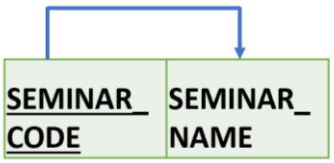
DOCTOR_ID	DOCTOR_NAME	SPECIALIZATION_CODE
D_0108	Samuel Jones	CAR01
D_0124	Julia Cortez	NEUR1
D_0140	Mark Harris	ONCO2
D_0110	Anna Harris	STIN01

The rest of the tables (DOCTOR_SEMINAR and SEMINAR) remain unchanged (same as in 2NF). Therefore, we rewrite them as they appear in 2NF.

Partial dependency 1 (PD1)

PD1 : SEMINAR_CODE → SEMINAR_NAME

Dependency diagram for PD1



Relation schema:

SEMINAR (**SEMINAR_CODE**, SEMINAR_NAME)

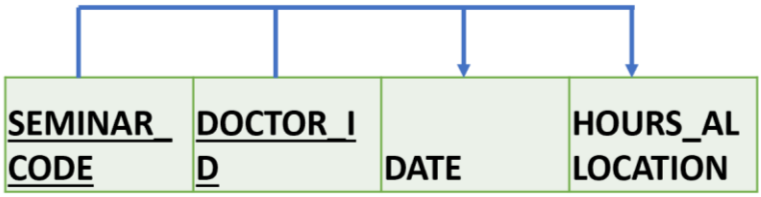
Table name: SEMINAR, Primary key: SEMINAR_CODE. Foreign key: NONE

SEMINAR_CODE	SEMINAR_NAME
s001	Homeopathic Medicine

s002	Cellfood Benefits

SEMINAR_CODE, DOCTOR_ID → DATE, ALLOCATION_HOURS

Dependency diagram showing: SEMINAR_CODE, DOCTOR_ID → DATE, ALLOCATION_HOURS



Relation schema:

SEMINAR_DOCTOR (**SEMINAR_CODE, DOCTOR_ID,** DATE, HOURS_ALLOCATION)

Table name: SEMINAR_DOCTOR, **Primary key**: SEMINAR_CODE + DOCTOR_ID, **Foreign keys**: SEMINAR_CODE, DOCTOR_ID

SEMINAR_CODE	DOCTOR_ID	DATE	HOURS_ALLOCATION
s001	D_0108	01/01/2010	8
s001	D_0124	01/01/2010	16
s001	D_0140	01/01/2010	8
s002	D_0110	01/07/2010	8
s002	D_0140	01/01/2011	8
s002	D_0124	01/07/2010	16

In summary, 3NF consists of 4 tables with following schemas:

- SEMINAR (**SEMINAR_CODE**, SEMINAR_NAME)
- DOCTOR(**DOCTOR_ID**, DOCTOR_NAME, SPECIALIZATION_CODE)
- SPECIALIZATION (**SPECIALIZATION_CODE**, SPECILIZATION_NAME)
- SEMINAR_DOCTOR (**SEMINAR_CODE, DOCTOR_ID,** DATE, HOURS_ALLOCATION)