Normalization

- Why?
- 1. Remove data redundancy
- 2. Data Anomaly
- 3. Data integrity
- 4. Query easiness
- 5. Performance (Update/Delete)
- 6. Maintenance
- 7. Scalability
- 8. Storage

Needs

- Anomaly
  - Insert Anomaly
  - Update anomaly
  - Delete anomaly

Students

Roll	Class	Section	Subj	Teacher
121	10	C	**Math**	**Mr Abul**
545	10	C	**Math**	**Mr Abul**
646	8	B	Bangla	Ms Nargis
545	10	A	Eng	Mr XYZ
646	8	B	Bangla	Ms Nargis

Anomalies (Insert)

Students

Roll	Class	Section	Subj	Teacher
121	10	C	**Math**	**Mr Abul**
545	10	C	**Math**	**Mr Abul**
646	8	B	Bangla	Ms Nargis
545	10	A	Eng	Mr XYZ
646	8	B	Bangla	Ms Nargis
987	10	A	**Math**	**Mr Abul**

Students

Anomalies (Updation Anomaly)

Roll	Class	Section	Subj	Teacher
121	10	C	**Math**	**Mr Abul**
545	10	C	**Math**	**Mr Abul**
646	8	B	Bangla	Ms Nargis
545	10	A	Eng	Mr XYZ
646	8	B	Bangla	Ms Nargis
987	10	A	**Math**	**Mr Abul**

Roll	Class	Section	Subj	Teacher
121	10	C	**Math**	**Shahadat**
545	10	C	**Math**	**Shahadat**
646	8	B	Bangla	Ms Nargis
545	10	A	Eng	Mr XYZ
646	8	B	Bangla	Ms Nargis
987	10	A	**Math**	**Shahadat**

Students

Anomalies (DeletionAnomaly)

Roll	Class	Section	Subj	Teacher
121	10	C	**Math**	**Mr Abul**
545	10	C	**Math**	**Mr Abul**
646	8	B	Bangla	Ms Nargis
545	10	A	Eng	Mr XYZ
646	8	B	Bangla	Ms Nargis
987	10	A	**Math**	**Mr Abul**

Roll	Class	Section	Subj	Teacher

Normalization

- 1 NF
  - Has a PK
  - Each column should have unq values
  - Duplicate rows not allowed
- 2NF
  - Must be in 1nF
  - No partial dependency/ No Non prime attrbute
- 3NF
  - Must be in 2NF
  - No transtitive dependency
- …
- …
- …
- …

Students

1NF

Roll	Class	Section	Subj	Teacher
121	10	C	**Math, Eng, Bang**	**Mr Abul, Nargis, Mr X**
545	8	A	**Eng, Bang**	**Mr Abul, Nargis**

Roll (PK)	Class	Section	Subj (PK)	Teacher
545	8	A	**Eng**	**Mr Abul**
545	8	A	**Bang**	**Nargis**

2NF

- Must be in 1NF
- All non-key attribute should be fully dependent on PK / No Partial dependency

Roll (PK)	Class	Section	Subj (PK)	Teacher
545	8	A	**Eng**	**Mr Abul**
545	8	A	**Bang**	**Nargis**

Roll (PK)	Class	Section
545	8	A
545	8	A

Roll (PK)	Subj (PK)	Teacher
545	**Eng**	**Mr Abul**
545	**Bang**	**Nargis**

3NF

- Must be in 2NF
- No Transitive dependency

Roll (PK)	Subj	Teacher
545	Eng	Mr Abul
646	Bang	Nargis

StudentSubject

Roll (PK)	Subj
545	Eng
646	Bang

TeacherSubject

Subj (PK)	Teacher
Eng	Mr Abul
Bang	Nargis

3.5 NF/ BCNF

- Must be in 3NF
- X→ Y (X must be Super key)

Emp_ID (PK)	BOSS (PK)	Email
X	A	empx@aiquest
A	B	empa@aiquest
B	C	ampb@aiquest

Emp_ID	Email
X	empx@aiquest
A	empa@aiquest
B	ampb@aiquest

Emp_ID	BOSS (PK)
X	A
A	B
B	C

4NF

- Already in BCNF
- Multi values dependency

ID	Roll	Subj	Teacher	Email	Father name
1	545	Eng	Mr Abul	jhfhg@jhfhg	Mr GG
2	545	Bang	Nargis	jhfhg@jhfhg	Mr GG
3	545	Eng	Mr babu	jhfhg@jhfhg	Mr GG
4	545	Bang	Ms Z	jhfhg@jhfhg	Mr GG

Roll (PK)	Id	Subj (PK)
545	1	Eng
545	2	Bang

Subj (PK)	Teacher
Eng	Mr Abul
Bang	Nargis
Eng	Mr babu
Bang	Ms Z

STS	Teacher
1	Mr Abul
1	Babu
2	Nargis
2	Ms Z

KEYs

- Candidate key: Unqiuely any record can be identified
- **Primary Key**: A specific candidate key,which is selected to identify a row
- Foreign key: A pk of a table refers as fk in another table to identify a record
- Superkey: minimal keys , that can identify uniquely any record
- Alternative key: Candidatekey except PK
- Composite key: two or more columns that can unq identify rows
- **Unique key**: Uniquely can identify a row

Q&A