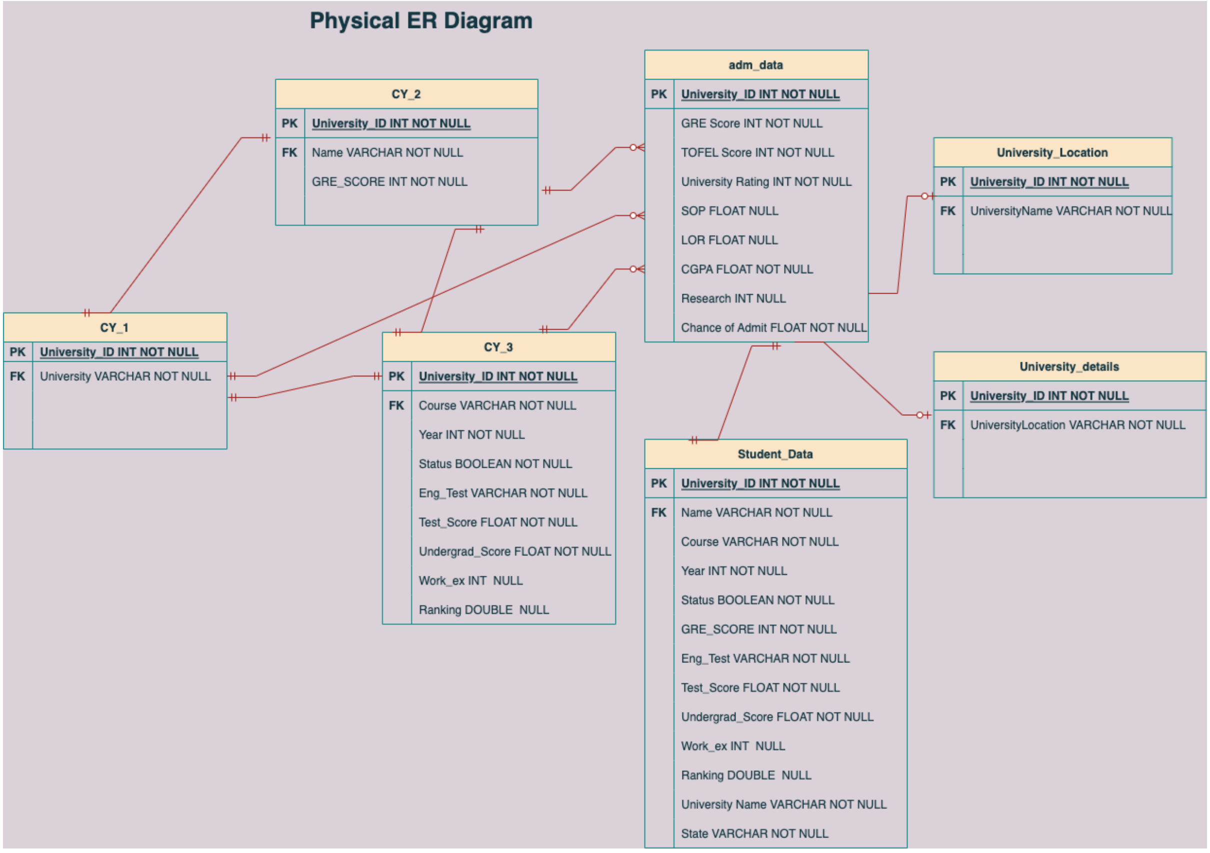
Team Members:
DARSHIT SHAH - 002762097
VIDISH KALE - 002194756
DIVYA MYNENI - 002746471

DATABASE NORMALIZATION

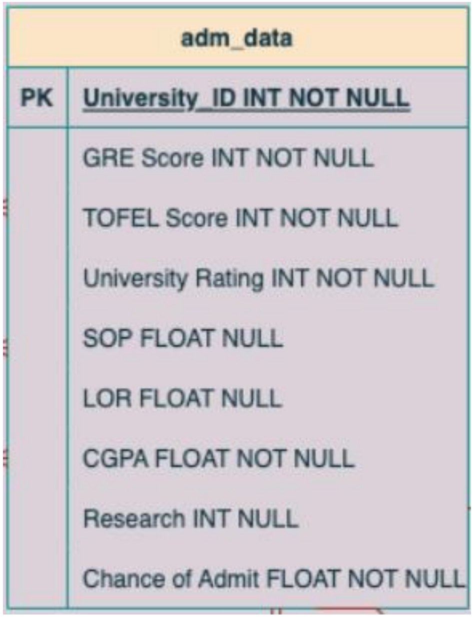
Github Link: Darshit98/DMDD_Assignment_4 (github.com)

ER- Diagram


Physical ER Diagram
CY_2
PK University_ID INT NOT NULL
FK Name VARCHAR NOT NULL
GRE_SCORE INT NOT NULL
adm_data
PK University_ID INT NOT NULL
GRE Score INT NOT NULL
TOFEL Score INT NOT NULL
University Rating INT NOT NULL
SOP FLOAT NULL
LOR FLOAT NULL
CGPA FLOAT NOT NULL
Research INT NULL
Chance of Admit FLOAT NOT NULL
University_Location
PK University_ID INT NOT NULL
FK UniversityName VARCHAR NOT NULL
CY_1
PK University_ID INT NOT NULL
FK University VARCHAR NOT NULL
CY_3
PK University_ID INT NOT NULL
FK Course VARCHAR NOT NULL
Year INT NOT NULL
Status BOOLEAN NOT NULL
Eng_Test VARCHAR NOT NULL
Test_Score FLOAT NOT NULL
Undergrad_Score FLOAT NOT NULL
Work_ex INT NULL
Ranking DOUBLE NULL
University_details
PK University_ID INT NOT NULL
FK UniversityLocation VARCHAR NOT NULL
Student_Data
PK University_ID INT NOT NULL
FK Name VARCHAR NOT NULL
Course VARCHAR NOT NULL
Year INT NOT NULL
Status BOOLEAN NOT NULL
GRE_SCORE INT NOT NULL
Eng_Test VARCHAR NOT NULL
Test_Score FLOAT NOT NULL
Undergrad_Score FLOAT NOT NULL
Work_ex INT NULL
Ranking DOUBLE NULL
University Name VARCHAR NOT NULL
State VARCHAR NOT NULL

Team Members:
DARSHIT SHAH - 002762097
VIDISH KALE - 002194756
DIVYA MYNENI - 002746471

Table_1: adm_data (Admission Data)



adm_data	
PK	**University_ID INT NOT NULL**
	GRE Score INT NOT NULL
	TOFEL Score INT NOT NULL
	University Rating INT NOT NULL
	SOP FLOAT NULL
	LOR FLOAT NULL
	CGPA FLOAT NOT NULL
	Research INT NULL
	Chance of Admit FLOAT NOT NULL

1st Normal Form:

- It has primary key with minimal attributes (PK = University_id).
- A relation is said to be in 1NF if it contains atomic value. This Normal Form addresses the atomicity issue. Atomicity here denotes that the values in the table shouldn't be subdivided further.
- One cell cannot contain more than one value. A table can violate the First Normal Form if it has a composite or multivalued attribute.
- If it has primary key with minimal attributes.
- The values in each column of a table are atomic (No multi-valued attributes are present).
- There are no repeating groups.

As the table, adm_data satisfies all of the above criteria. It is said to be in 1st Normal form.

2nd Normal Form:

- For the table to be in 2nd Normal Form, it must satisfy the conditions of 1st Normal Form.
- Partial dependencies should not be present in the table either.
- In this case, partial dependency means that a non-prime attribute is determined by the right subset of the candidate key.
- This table does not contain partial dependency.

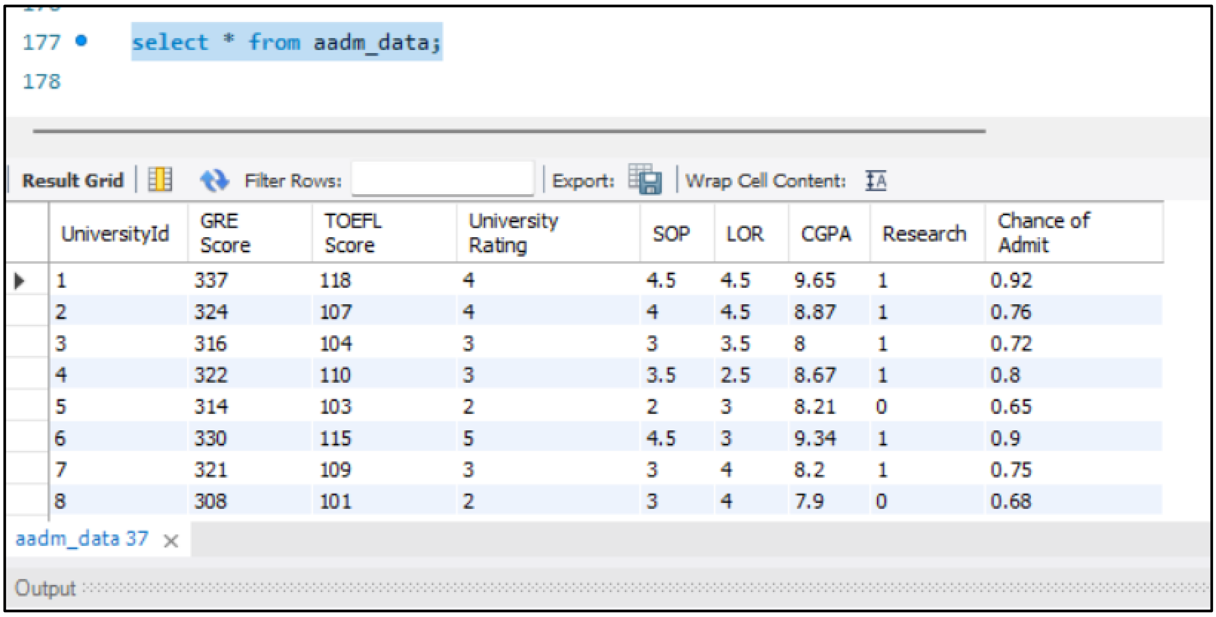
As this table, fulfills all these above criteria, it is said to be in 2nd Normal Form.

Team Members:
DARSHIT SHAH - 002762097
VIDISH KALE - 002194756
DIVYA MYNENI - 002746471

3rd Normal Form:

- For the table to be considered in 3rd Normal Form, it has to be in 1st NF and 2nd NF.
- It should have no partial or transitive dependencies.

As the adm_table is already in 1NF and 2NF, and is free from partial and transitive dependencies, it is already said to be in 3rd Normal Form.



```
select * from aadm_data;
```

UniversityId	GRE Score	TOEFL Score	University Rating	SOP	LOR	CGPA	Research	Chance of Admit
1	337	118	4	4.5	4.5	9.65	1	0.92
2	324	107	4	4	4.5	8.87	1	0.76
3	316	104	3	3	3.5	8	1	0.72
4	322	110	3	3.5	2.5	8.67	1	0.8
5	314	103	2	2	3	8.21	0	0.65
6	330	115	5	4.5	3	9.34	1	0.9
7	321	109	3	3	4	8.2	1	0.75
8	308	101	2	3	4	7.9	0	0.68

Team Members:
DARSHIT SHAH - 002762097
VIDISH KALE - 002194756
DIVYA MYNENI - 002746471

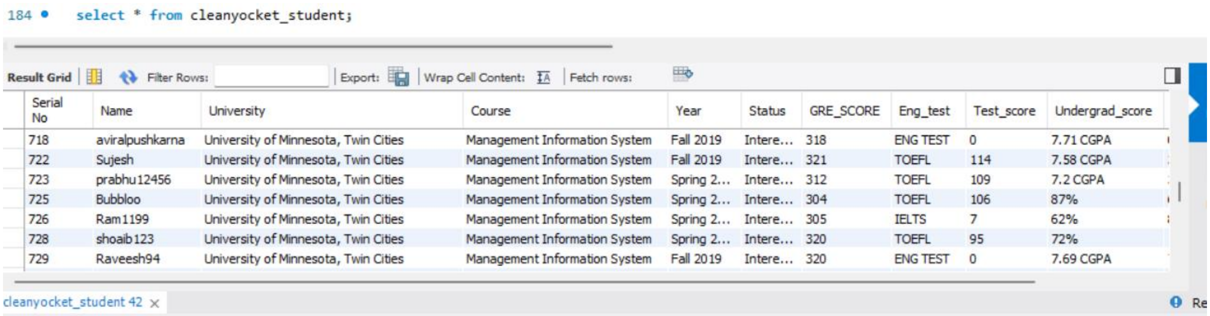
Table_2: Student_Data

1st Normal Form:

- It has primary key with minimal attributes (PK = University_id).
- A relation is said to be in 1NF if it contains atomic value. Atomicity here denotes that the values in the table shouldn't be subdivided further.
- One cell cannot contain more than one value. It is violating the First Normal Form if it has a composite or multivalued attribute.
- If it has primary key with minimal attributes.
- The values in each column of a table are atomic (No multi-valued attributes are present).
- There are no repeating groups.

The student_data table is not in 1st Normal Form. So it has been converted to 1st Normal Form. The University table has been spitted into University_Name and state.

Table before Normalization:

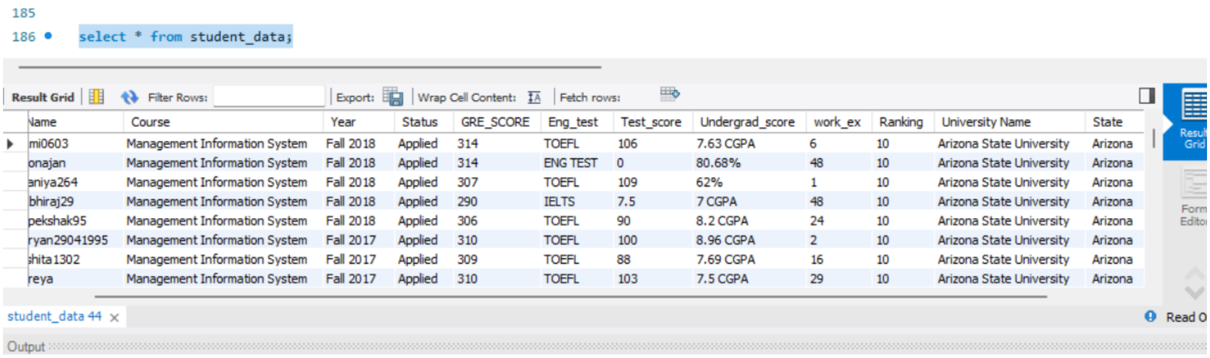


```
184 ●   select * from cleanyocket_student;
```

Serial No	Name	University	Course	Year	Status	GRE_SCORE	Eng_test	Test_score	Undergrad_score
718	aviralpushkarna	University of Minnesota, Twin Cities	Management Information System	Fall 2019	Intere...	318	ENG TEST	0	7.71 CGPA
722	Sujesh	University of Minnesota, Twin Cities	Management Information System	Fall 2019	Intere...	321	TOEFL	114	7.58 CGPA
723	prabhu12456	University of Minnesota, Twin Cities	Management Information System	Spring 2...	Intere...	312	TOEFL	109	7.2 CGPA
725	Bubbloo	University of Minnesota, Twin Cities	Management Information System	Spring 2...	Intere...	304	TOEFL	106	87%
726	Ram1199	University of Minnesota, Twin Cities	Management Information System	Spring 2...	Intere...	305	IELTS	7	62%
728	shoaib123	University of Minnesota, Twin Cities	Management Information System	Spring 2...	Intere...	320	TOEFL	95	72%
729	Raveesh94	University of Minnesota, Twin Cities	Management Information System	Fall 2019	Intere...	320	ENG TEST	0	7.69 CGPA

Team Members:
DARSHIT SHAH - 002762097
VIDISH KALE - 002194756
DIVYA MYNENI - 002746471

Table after Normalization:



```
select * from student_data;
```

Name	Course	Year	Status	GRE_SCORE	Eng_test	Test_score	Undergrad_score	work_ex	Ranking	University Name	State
mi0603	Management Information System	Fall 2018	Applied	314	TOEFL	106	7.63 CGPA	6	10	Arizona State University	Arizona
onajan	Management Information System	Fall 2018	Applied	314	ENG TEST	0	80.68%	48	10	Arizona State University	Arizona
aniya264	Management Information System	Fall 2018	Applied	307	TOEFL	109	62%	1	10	Arizona State University	Arizona
bhiraj29	Management Information System	Fall 2018	Applied	290	IELTS	7.5	7 CGPA	48	10	Arizona State University	Arizona
pekshak95	Management Information System	Fall 2018	Applied	306	TOEFL	90	8.2 CGPA	24	10	Arizona State University	Arizona
ryan29041995	Management Information System	Fall 2017	Applied	310	TOEFL	100	8.96 CGPA	2	10	Arizona State University	Arizona
shita1302	Management Information System	Fall 2017	Applied	309	TOEFL	88	7.69 CGPA	16	10	Arizona State University	Arizona
reya	Management Information System	Fall 2017	Applied	310	TOEFL	103	7.5 CGPA	29	10	Arizona State University	Arizona

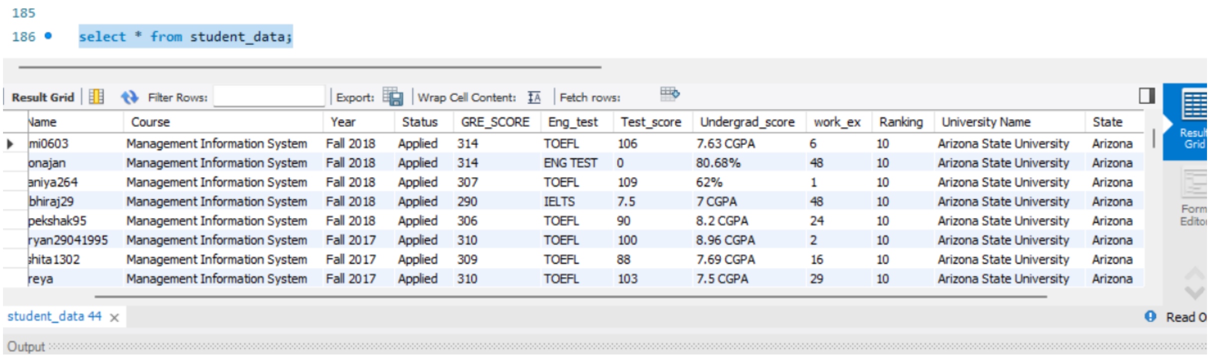
2nd Normal Form:

- For the table to be in 2nd Normal Form, it must satisfy the conditions of 1st Normal Form.
- The table has now been converted into 2nd Normal Form
- This table does not contain partial dependency.

As this table, fulfils all the above criteria, it is said to be in 2nd Normal Form.

3rd Normal Form:

- For the table to be in 3rd Normal Form, it must satisfy the conditions of 1st Normal Form and 2nd Normal Form.
- The table is now satisfying the 2nd Normal Form.
- There are no partial and transitive dependencies in the table. So the table is now in 3rd Normal Form.



```
select * from student_data;
```

Name	Course	Year	Status	GRE_SCORE	Eng_test	Test_score	Undergrad_score	work_ex	Ranking	University Name	State
mi0603	Management Information System	Fall 2018	Applied	314	TOEFL	106	7.63 CGPA	6	10	Arizona State University	Arizona
onajan	Management Information System	Fall 2018	Applied	314	ENG TEST	0	80.68%	48	10	Arizona State University	Arizona
aniya264	Management Information System	Fall 2018	Applied	307	TOEFL	109	62%	1	10	Arizona State University	Arizona
bhiraj29	Management Information System	Fall 2018	Applied	290	IELTS	7.5	7 CGPA	48	10	Arizona State University	Arizona
pekshak95	Management Information System	Fall 2018	Applied	306	TOEFL	90	8.2 CGPA	24	10	Arizona State University	Arizona
ryan29041995	Management Information System	Fall 2017	Applied	310	TOEFL	100	8.96 CGPA	2	10	Arizona State University	Arizona
shita1302	Management Information System	Fall 2017	Applied	309	TOEFL	88	7.69 CGPA	16	10	Arizona State University	Arizona
reya	Management Information System	Fall 2017	Applied	310	TOEFL	103	7.5 CGPA	29	10	Arizona State University	Arizona

Team Members:
DARSHIT SHAH - 002762097
VIDISH KALE - 002194756
DIVYA MYNENI - 002746471

```sql
use assignment4_schema;

select * from student_data;

ALTER TABLE student_data
RENAME COLUMN MyUnknownColumn TO Student_id;

ALTER TABLE student_data
DROP COLUMN Serial_No;

CREATE table university_details
(
universityId int NOT NULL AUTO_INCREMENT,
universityName varchar(500) DEFAULT NULL,
Primary key (universityId)
);

insert into university_details (universityName)
select distinct `University Name` from student_data;

select distinct `University Name` from student_data;

select * from university_details;

ALTER TABLE student_data
ADD universityId int;
```

Team Members:
DARSHIT SHAH - 002762097
VIDISH KALE - 002194756
DIVYA MYNENI - 002746471

```
-- to update UniversityId in student table
update student_data , university_details u
set universityId = (
select u.universityId from university_details u
join student_data s on u.universityName=s.`University Name`
where u.universityName=s.`University Name`
)
where u.universityName=student_data.`University Name`;


-- to update UniversityId in student table
update student_data , university_details u
set universityId = (
select u.universityId from university_details u
join student_data s on u.universityName=s.`University Name`
where u.universityName=s.`University Name`
)
where u.universityName=student_data.`University Name`;


CREATE table university_location
(
universityId int NOT NULL AUTO_INCREMENT,
universityLocation varchar(500) DEFAULT NULL,
Primary key (universityId)
);



insert into university_location (universityLocation)
select distinct `State` from student_datanormalized;
```

Team Members:
DARSHIT SHAH - 002762097
VIDISH KALE - 002194756
DIVYA MYNENI - 002746471

```
select * from university_location;
```

```
ALTER TABLE student_datanormalized

DROP COLUMN State;
```

Team Members:
DARSHIT SHAH - 002762097
VIDISH KALE - 002194756
DIVYA MYNENI - 002746471

```
-- Creating views

    1. CREATE VIEW `admit300plus` AS(
SELECT c2.GRESCORE,c2.Name,c3.Status FROM cy3 c3
JOIN cy2 c2 on c3.SerialNo = c2.SerialNo
WHERE c3.Status = 'Admit' AND c2.GRESCORE > 300);


select * from admit300plus;


    2. CREATE VIEW `admit100Toefl` AS(
SELECT c3.Status, c2.Name FROM cy3 c3
JOIN cy2 c2 on c3.SerialNo = c2.SerialNo
WHERE c3.Status = 'Admit' AND c3.TestScore like '%100%');


select * from admit100Toefl;


    3. CREATE VIEW `admitInformationSystems` AS(
SELECT c3.Status, c2.Name FROM cy3 c3
join cy2 c2 on c3.SerialNo = c2.SerialNo
join cy3 on c2.SerialNo = c3.SerialNo
where c3.Status = 'Admit' and c3.Course = 'Information Systems');


select * from admitInformationSystems;


    4. CREATE VIEW `noWorkExp` AS(
SELECT Distinct c3.Course, c1.University from cy3 c3
join cy1 c1 on c3.SerialNo = c1.SerialNo
where c3.WorkExp = 0);
```

Team Members:
DARSHIT SHAH - 002762097
VIDISH KALE - 002194756
DIVYA MYNENI - 002746471

```
select * from noWorkExp;
```

```
    5. CREATE VIEW `ManagementInfoCourse` AS(
SELECT University from cy1 c1
join cy3 c3 on c1.SerialNo = c3.SerialNo
where c3.Course like '%Management Information System%');

select * from ManagementInfoCourse;
```

```
    6. CREATE VIEW `cgpaAndGre` AS(
Select `Chance of Admit` from adm_data
where CGPA >= 7.5 and `GRE Score` < 310);

select `Chance of Admit`*100 from cgpaAndGre;
```

```
    7. CREATE VIEW `cgpaAndToefl` AS(
Select `Chance of Admit`*100 from adm_data
where CGPA >= 7 and `TOEFL Score` >= 100);

select * from cgpaAndToefl;
```

```
    8. CREATE VIEW `cgpaAndUniRating` AS(
Select `Chance of Admit`*100 from adm_data
where CGPA >= 6 and `University Rating` >= 3);

select * from cgpaAndUniRating;
```

Team Members:
DARSHIT SHAH - 002762097
VIDISH KALE - 002194756
DIVYA MYNENI - 002746471

```
9.  CREATE VIEW `cgpaToeflUniRating` AS(
Select `Chance of Admit`*100 from adm_data
where CGPA >= 8 and `TOEFL Score` > 80 and `University Rating` >= 4);

select * from cgpaToeflUniRating;
```

```
10. CREATE VIEW `cgpaAndResearch` AS(
Select `Chance of Admit` from adm_data
where CGPA = 8 and `Research` = 1);

select * from cgpaAndResearch;
```

```
11. CREATE VIEW `Toefl100Plus` AS(
Select `Status` from cy3
where `TestScore` >= 100);

select * from Toefl100Plus;
```

```
12. CREATE VIEW `admitAppTest` AS(
SELECT c3.Status, c2.Name FROM cy3 c3
join cy2 c2 on c3.SerialNo = c2.SerialNo
where c3.Status = 'Admit' or 'Applied' and c3.TestScore >= '90' or '6.5');

select * from admitAppTest;
```

Team Members:
DARSHIT SHAH - 002762097
VIDISH KALE - 002194756
DIVYA MYNENI - 002746471

13.
```
CREATE VIEW `admitUndergrad` AS(
SELECT c3.Status, c2.Name FROM cy3 c3
join cy2 c2 on c3.SerialNo = c2.SerialNo
where c3.Status = 'Admit' or 'Applied' and UndergradScore = 8);

select * from admitUndergrad;
```

14.
```
CREATE VIEW `admitUndergradTestWorkExp` AS(
SELECT c3.Status, c2.Name FROM cy3 c3
join cy2 c2 on c3.SerialNo = c2.SerialNo
where c3.Status = 'Admit' or 'Applied' and UndergradScore = 8.96 and TestScore = 100 and WorkExp = 2);

select * from admitUndergradTestWorkExp;
```

15.
```
CREATE VIEW `admitRanking` AS(
SELECT c3.Status, c2.Name FROM cy3 c3
join cy2 c2 on c3.SerialNo = c2.SerialNo
where c3.Status = 'Admit' or 'Applied' and Ranking <= 20);

select * from admitRanking;
```

16.
```
CREATE VIEW `admitMgmtInfoSys` AS(
SELECT c3.Status, c2.Name FROM cy3 c3
join cy2 c2 on c3.SerialNo = c2.SerialNo
where c3.Status = 'Admit' or 'Applied' and Course = 'Management Information Systems');

select * from admitMgmtInfoSys;
```