

Assignment_4 Normalization

1. NEU_Program Table:

First Normal Form: Satisfies 1NF requirements

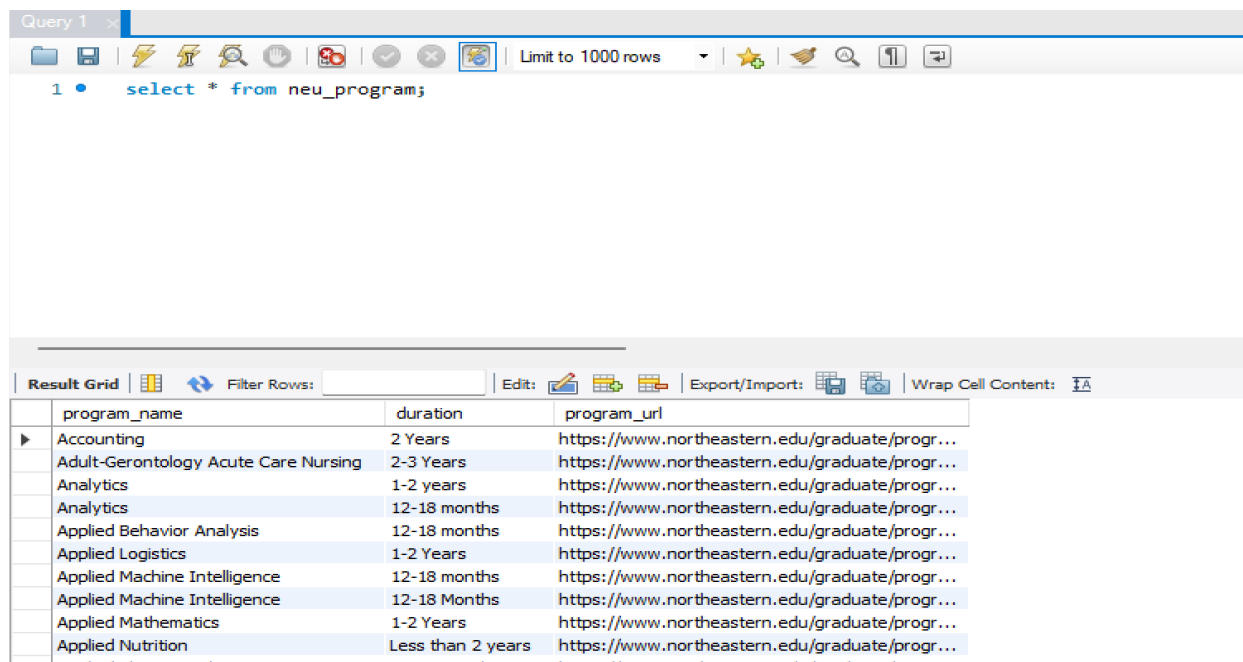
- Table has a primary key attribute identified as program_name
- No multi value attributes present
- No repeating groups

Second Normal Form: Satisfies 2NF requirements

- All requirements for 1NF met
- No partial dependency on any column
- None of the fields have data calculated from other fields

Third Normal Form: Satisfies 3NF requirements

- 2NF requirements satisfied
- No transitive dependency



Query 1

Limit to 1000 rows

1 • `select * from neu_program;`

Result Grid

	program_name	duration	program_url
▶	Accounting	2 Years	https://www.northeastern.edu/graduate/progr...
	Adult-Gerontology Acute Care Nursing	2-3 Years	https://www.northeastern.edu/graduate/progr...
	Analytics	1-2 years	https://www.northeastern.edu/graduate/progr...
	Analytics	12-18 months	https://www.northeastern.edu/graduate/progr...
	Applied Behavior Analysis	12-18 months	https://www.northeastern.edu/graduate/progr...
	Applied Logistics	1-2 Years	https://www.northeastern.edu/graduate/progr...
	Applied Machine Intelligence	12-18 months	https://www.northeastern.edu/graduate/progr...
	Applied Machine Intelligence	12-18 Months	https://www.northeastern.edu/graduate/progr...
	Applied Mathematics	1-2 Years	https://www.northeastern.edu/graduate/progr...
	Applied Nutrition	Less than 2 years	https://www.northeastern.edu/graduate/progr...
	Applied Physics and Engineering	12-18 Months	https://www.northeastern.edu/graduate/progr...

2. NEU_Course_Catalog Table:

First Normal Form: Satisfies 1NF requirements

- Table has a primary key attribute identified as course_id
- No multi value attributes present
- No repeating groups

Second Normal Form: Satisfies 2NF requirements

- All requirements for 1NF met
- No partial dependency on any column
- None of the fields have data calculated from other fields

Third Normal Form: Satisfies 3NF requirements

- 2NF requirements satisfied
- No transitive dependency

Query 1 x

Limit to 1000 rows

```
1 • select * from neu_course_catalog;
```

Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content: |

program_name	course_id	course_name	course_description
Software Engineering Systems	CSYE 6200	Concepts of Object-Oriented Design	Introduces object-oriented design and program...
Software Engineering Systems	CSYE 6202	Concepts of Object-Oriented Design with C#	Introduces object-oriented design and program...
Software Engineering Systems	CSYE 6205	Concepts of Object-Oriented Design with C++	Introduces object-oriented design and program...
Software Engineering Systems	CSYE 6220	Enterprise Software Design	Designed to build on previous experience in con...
Software Engineering Systems	CSYE 6225	Network Structures and Cloud Computing	Offers a practical foundation in cloud computing...
Software Engineering Systems	CSYE 6230	Operating Systems	Covers basic concepts of operating systems an...
Software Engineering Systems	CSYE 6305	Introduction to Quantum Computing with Applic...	Addresses how scientists and engineers can us...
Software Engineering Systems	CSYE 6700	Technical Writing and Professional Development	Emphasizes professional communication skills thr...
Software Engineering Systems	CSYE 6962	Elective	Offers elective credit for courses taken at other...
Software Engineering Systems	CSYE 7105	High-Performance Parallel Machine Learning and...	Explores the parallelization of machine learning ...
Software Engineering Systems	CSYE 7125	Advanced Cloud Computing	Examines a collection of reusable, generic sof...

neu_course_catalog 17 x

Apply

3. NEU_Event Table:

First Normal Form: Satisfies 1NF requirements

- Table has a primary key attribute identified as event_id
- No multi value attributes present
- No repeating groups

Second Normal Form: Satisfies 2NF requirements

- All requirements for 1NF met
- No partial dependency on any column
- None of the fields have data calculated from other fields

Third Normal Form: Satisfies 3NF requirements

- 2NF requirements satisfied
- No transitive dependency

Query 1

```

1  select * from neu_event;
2
3

```

Limit to 1000 rows

Result Grid

	event_id	program_name	event_name
▶	341	Information Systems	Ethics Institute Speaker, Michael Hannon (Frida...
	342	Information Systems	Ethics Institute Speaker, Don Fallis (Friday, De...
	343	Information Systems	Using a Systems-based Culturally Responsive E...
	344	Information Systems	Designing an Engaging Academic Presentation (...)
	345	Information Systems	Listener Research Study (Friday, November 11,...)
	346	Data Analytics Engineering	Problem Play (Friday, November 11, 2022 8pm t...
	347	Data Analytics Engineering	Graduate programs open house
	348	Data Analytics Engineering	Women in STEM: An Evening with Distinguished ...
	349	Data Analytics Engineering	Fall Graduate Programs Open House
	350	Data Analytics Engineering	Northeastern University Mumbai Education Exhi...
	351	Software Engineering	Graduate Studies Information Session

4. Course_Core_Requirement Table:

First Normal Form: Satisfies 1NF requirements

- Table has a primary key attribute identified as core_course_id
- No multi value attributes present
- No repeating groups

Second Normal Form: Satisfies 2NF requirements

- All requirements for 1NF met
- No partial dependency on any column
- None of the fields have data calculated from other fields

Third Normal Form: Satisfies 3NF requirements

- 2NF requirements satisfied
- No transitive dependency

Query 1 x

Limit to 1000 rows

```

1  select * from course_core_requirement;
2
3

```

Result Grid Filter Rows: Edit: Export/Import: Wrap Cell Content:

	core_course_id	program_name	course_id	course_name
▶	1453	Software Engineering Systems	CSYE 6200	Concepts of Object-Oriented Design
	1454	Software Engineering Systems	INFO 6205	Program Structure and Algorithms
	1455	Data Analytics Engineering	IE 5374	Special Topics in Industrial Engineering
	1456	Data Analytics Engineering	IE 6600	Computation and Visualization for Analytics
	1457	Data Analytics Engineering	IE 6700	Data Management for Analytics
	1458	Data Analytics Engineering	or DAMG 6210	Data Management and Database Design
	1459	Data Analytics Engineering	IE 7275	Data Mining in Engineering
	1460	Data Analytics Engineering	IE 7280	Statistical Methods in Engineering
	1461	Data Analytics Engineering	OR 6205	Deterministic Operations Research
	1462	Information Systems	INFO 5100	Application Engineering

5. Resource_Material1 Table:

First Normal Form: Satisfies 1NF requirements

- The table has a primary key attribute identified as core_course_id
- No multi-value attributes present
- No repeating groups

Second Normal Form: Does not satisfy 2NF requirements

- All requirements for 1NF met

- No partial dependency on any column
- None of the fields have data calculated from other fields

Third Normal Form: Does not satisfy the 3rd normal form

- 2NF requirements satisfied.
- There is a transitive dependency on the course_id column in addition to the PK software_name. Different courses may need the same software downloads and the curriculum of different courses may be in alignment for some part leading to similar resource material. To eradicate this we created NEU_Course_Resource table where we have software mapped to specific course IDs.

NEU_Resource_Materials Table:

Query 1

Limit to 1000 rows

```
1 select * from neu_resource_materials;
2
3
```

software_name	software_download_url	professional_certification	online_platform
Apache Hadoop	https://spark.apache.org/downloads.html	Data Analytics certificate	https://powerbi.microsoft.com/en-us/
Azure	https://azure.microsoft.com/en-us/	Microsoft Azure Administrator Certificate	Microsoft Azure
Microsoft Azure	https://azure.microsoft.com/en-us/downloads/	AZ-900 Microsoft Azure Fundamentals	https://www.cloudzero.com/platform
Neural Designer	https://neural-designer.en.softonic.com/	Azure AI Engineer Associate: Certificate from ...	https://playground.tensorflow.org/#activa
Python	https://www.python.org/downloads/	Certified Professional in Python Programming 1	https://www.online-python.com/
Spring MVC	https://www.jetbrains.com/idea/old/promo/?so...	https://www.udemy.com/course/spring-boot-le...	Leetcode
SQL Developer	https://www.oracle.com/database/sqldevelope...	Data Analytics certificate	https://www.jdoodle.com/execute-sql-on
Tableau	https://www.tableau.com/resource/business-in-...	Professional Certificate in Database Manageme...	https://www.programiz.com/java-programi
Ubuntu	https://ubuntu.com/#download	Linux Fundamentals Certification Training	https://www.onworks.net/os-distribution
Visual Studio	https://visualstudio.microsoft.com/	Linux Fundamentals Certification Training	https://www.onworks.net/os-distribution

Neu_Course_Resource Table:

Query 1

```
1 select * from neu_course_resource;
```

Result Grid

course_resource_id	course_id	software_name	course_rating
5233	INFO 6106	Neural Designer	4
5234	DAMG 6105	Python	3
5235	CSYE 6225	Microsoft Azure	5
5236	INFO 5002	Python	3
5237	INFO 6205	Visual studio	4
5238	DAMG 7245	Apache Hadoop	4
5239	CSYE 6230	Ubuntu	2
5240	DAMG 7275	SQL Developer	4
5241	DAMG 7370	Tableau	5
5242	CSYE 6220	Spring MVC	4
5243	CSYE 7125	Apache	2

6. Jobs_Info table:

First Normal Form: Satisfies 1NF requirements

- Table has a primary key attribute identified as job_no
- No multi value attributes present
- No repeating groups

Second Normal Form: Satisfies 2NF requirements

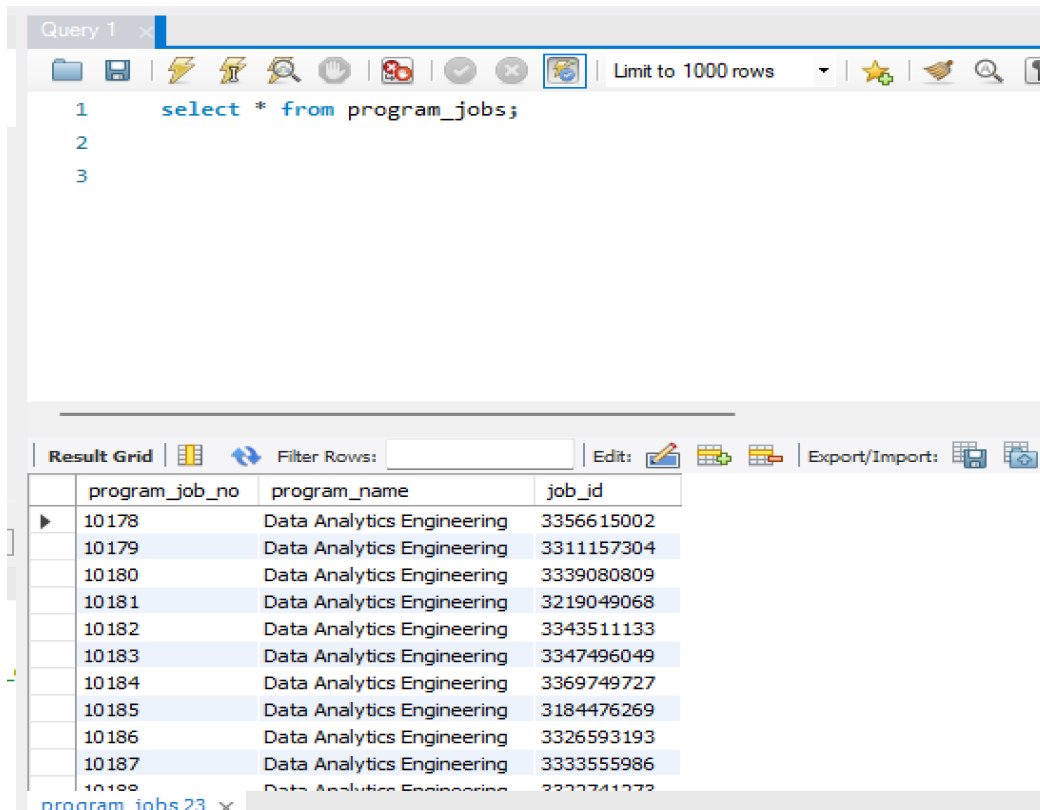
- All requirements for 1NF met
- No partial dependency on any column
- None of the fields have data calculated from other fields

Third Normal Form: Does not satisfy the 3rd normal form

- 2NF requirements satisfied
- There is transitive dependency on job_id. There is a possibility that a program can have same job_id which will generate two different job numbers for same

job_i. Hence, we are creating a bridge table **program_jobs** to remove the transitive dependency.

Program_jobs Table:



The screenshot shows a database query tool interface. At the top, a tab labeled 'Query 1' is active. Below the tab is a toolbar with various icons for file operations, execution, and search. The query editor contains the following SQL statement:

```
1 select * from program_jobs;
2
3
```

Below the query editor is a 'Result Grid' section. It includes a 'Filter Rows' input field, an 'Edit' button, and 'Export/Import' options. The result grid displays the following data:

	program_job_no	program_name	job_id
▶	10178	Data Analytics Engineering	3356615002
	10179	Data Analytics Engineering	3311157304
	10180	Data Analytics Engineering	3339080809
	10181	Data Analytics Engineering	3219049068
	10182	Data Analytics Engineering	3343511133
	10183	Data Analytics Engineering	3347496049
	10184	Data Analytics Engineering	3369749727
	10185	Data Analytics Engineering	3184476269
	10186	Data Analytics Engineering	3326593193
	10187	Data Analytics Engineering	3333555986
	10188	Data Analytics Engineering	3327211272

Job_Info table:

Query 1 x

Limit to 1000 rows

```
1 select * from job_info;
```

Result Grid

	job_no	title	job_id	company_name	location	date_posted	link
▶	601	Associate, Content Analytics & Operations	3356615002	Netflix	Los Angeles, CA	2022-11-16	https
	602	Principal Analyst, Data Science & Analytics	3311157304	Zillow	United States	2022-11-23	https
	603	VP, Data Analytics	3339080809	TechInsights	Denver, CO	2022-11-01	https
	604	Vice President – Data Science & Analytics	3219049068	H&R Block	United States	2022-11-28	https
	605	VP, Data Analytics	3343511133	Propio Language Services	Overland Park, KS	2022-11-08	https
	606	Data Analytics Engineer	3347496049	akdm	New York City Metropolitan Area	2022-11-07	https
	607	Data Analytics Associate	3369749727	First Help Financial	Needham, MA	2022-11-21	https
	608	Software Data Engineer - Analytics Engineering	3184476269	Apple	Seattle, WA	2022-11-13	https
	609	Director of Data & Analytics	3326593193	Lawrence Harvey	United States	2022-11-23	https

7. Neu_Course_Faculty table:

First Normal Form: Satisfies 1NF requirements

- The table has a primary key attribute identified as job_no
- No multi-value attributes present
- No repeating groups

Second Normal Form: Satisfies 2NF requirements

- All requirements for 1NF met
- No partial dependency on any column
- None of the fields have data calculated from other fields

Third Normal Form: Does not satisfy the 3rd normal form

- 2NF requirements satisfied
- There is a transitive dependency on faculty_name. Different courses can have the same faculty. Therefore the faculty_name column is transitively dependent on the course_id column. To resolve this we created two tables NEU_Facultys and NEU_Course_Faculty.

NEU_facultys:

Query 1

```
1 select * from neu_facultys;
2
3
```

Result Grid

	faculty_name	faculty_rating
▶	Krishnamurthy, Srikanth	5
	Daniel Peters	3
	Dehghani Mohammad	2
	Dino Konstantopoulos	4
	Handan Liu	5
	Hillyard, Robin	4
	Jin Xuemin	4
	Khaled Bugrara	3
	Muhammad Noor E Alam	3
	Nasser Fard	3

NEU_Course_faculty Table:

Query 1

```
1 select * from neu_course_faculty;
2
3
```

Result Grid

	fac_course_id	course_id	faculty_name
▶	4280	DAMG 7350	Nick Beauchamp
	4281	INFO 6105	Nik Bear Brown
	4282	CSYE 6305	Prof. Hossein Noorian
	4283	DAMG 7275	Sagar Kamarthi
	4284	DAMG 7290	Muhammad Noor E Alam
	4285	CSYE 6205	Nasser Fard
	4286	CSYE 6700	Jin Xuemin
	4287	DAMG 6210	Radhakrishnan Srinivasan
	4288	DAMG 7978	Dehghani Mohammad
	4289	DAMG 6210	Nik brown
	4290	DAMG 6105	Dino Konstantopoulos

8. NEU_Specialization Table:

First Normal Form: Satisfies 1NF requirements

- The table has a primary key attribute identified as job_no
- No multi-value attributes present
- No repeating groups

Second Normal Form: Satisfies 2NF requirements

- All requirements for 1NF met
- No partial dependency on any column
- None of the fields have data calculated from other fields

Third Normal Form: Does not satisfy the 3rd normal form

- 2NF requirements satisfied
- There is a transitive dependency on course_id column. Multiple specializations can have same courses under them due to which spec_course_id which is the primary key will not be unique for different specializations. To resolve this we created two tables NEU_course_specialization and NEU_Specialization.

NEU_course_Specialization table:

Query 1 x

Limit to 1000 rows

```
1 select * from neu_course_specialization;
2
3
```

Result Grid

	course_spec_id	specialization_id	course_id	credit_hours
▶	777	9100	CSYE 6225	4
	778	9100	DAMG 7245	4
	779	9100	INFO 7250	4
	780	9100	INFO 7390	4
	781	9101	CSYE 7280	4
	782	9101	INFO 7375	4
	783	9101	INFO 7390	4
	784	9101	INFO 7610	4
	785	9102	INFO 7500	4
	786	9102	INFO 7510	4

Neu_Specialization Table:

Query 1 x

Limit to 1000 rows

```
1 select * from neu_specialization;
2
3
```

Result Grid

	specialization_id	specialization_name
▶	9100	Big Data Systems and Analytics Concentration
	9101	Intelligent Systems Concentration
	9102	Smart Contracts Concentration
	9103	User Experience Concentration
*	NULL	NULL

