CHANDIGARH UNIVERSITY

Discover. Learn. Empower.

A Project Report on

"SQL for Data Science"

Submitted in the partial fulfilment of the requirements for the award of

Summer Training

Submitted By:

Mohit Kumar Gupta

Group – 2

Section – C

Under the support and guidance of

Sadie St. Lawrence

UIC

Department of University Institute of Computing

Batch: 2019-2021

DECLARATION

I hereby declare that I have completed my six weeks summer training at **Coursera** (one of the world's leading online certification training providers) from **21st April, 2020** to **26th May, 2020** under the guidance of **Sadie St. Lawrence.**

I declared that I have worked with full dedication during these six weeks of training and my learning outcome fulfil the requirements of training for the award of degree of **Masters of Computer Applications (MCA), Chandigarh University, Gharuan.**

Date: _____

_____

(Signature of Student)

Name of Student: - Mohit Kumar Gupta

UID: - 19MCA8265

ACKNOWLEDGEMENT

The success and final outcome of learning SQL for Data Science for everybody required a lot of guidance and assistance from many people and I am extremely privileged to have got this all along the completion of my course and few of the projects. All that I have done is only due to such supervision and assistance and I would not forget to thank them.

I respect and thank **Coursera**, for providing me an opportunity to do the course and project work and giving me all support and guidance, which made me complete the course duly. I am extremely thankful to the course advisor Charles Severance.

I am thankful to and fortunate enough to get constant encouragement, support and guidance from all Teaching staffs of Coursera who helped me in successfully.

We pay our respects and love to our parents and all other family members and friends for their love and encouragement throughout our career. Last but not the least we express our thanks to our friends for their cooperation and support.

INDEX

INTRODUCTI
ON

About this Specialization

This Specialization covers how to write syntactically correct SQL and SQL SERVER, and how to create and store data in SQL that can manage with different servers. Mastering this range of technologies allowed me to develop high quality Software's that, work seamlessly on mobile, tablet, and large screen browsers accessible with enhancing the accuracy of the fetching the data from SQL server.

As data collection has increased exponentially, so has the need for people skilled at using and interacting with data; to be able to think critically, and provide insights to make better decisions and optimize their businesses. This is a data scientist, "part mathematician, part computer scientist, and part trend spotter" (SAS Institute, Inc.). According to Glassdoor, being a data scientist is the best job in America; with a median base salary of $110,000 and thousands of job openings at a time. The skills necessary to be a good data scientist include being able to retrieve and work with data, and to do that you need to be well versed in SQL, the standard language for communicating with database systems.

This course is designed to give you a primer in the fundamentals of SQL and working with data so that you can begin analyzing it for data science purposes. You will begin to ask the right questions and come up with good answers to deliver valuable insights for your organization. This course starts with the basics and assumes you do not have any knowledge or skills in SQL. It will build on that foundation and gradually have you written both simple and complex queries to help you select data from tables. You'll start to work with different types of data like strings and numbers and discuss methods to filter and pare down your results.

What is SQL?

This course is designed to give you a primer in the fundamentals of **SQL** and working with data so that you can begin analyzing it for data science purposes. You will begin to ask the right questions and come up with good answers to deliver valuable insights for your organization.

This course starts with the basics and assumes you do not have any knowledge or skills in SQL. It will build on that foundation and gradually have you written both

simple and complex queries to help you select data from tables. You'll start to work with different types of data like strings and numbers and discuss methods to filter and pare down your results. You will create new tables and be able to move data into them. You will learn common operators and how to combine the data. You will use case statements and concepts like data governance and profiling. You will discuss topics on data, and practice using real-world programming assignments. You will interpret the structure, meaning, and relationships in source data and use SQL as a professional to shape your data for targeted analysis purposes. Although we do not have any specific prerequisites or software requirements to take this course, a simple text editor is recommended for the final project. So, what are you waiting for? This is your first step in landing a job in the best occupation in the US and soon the world!

SQL Overview

SQL is Structured Query Language, which is a computer language for storing, manipulating and retrieving data stored in a relational database.

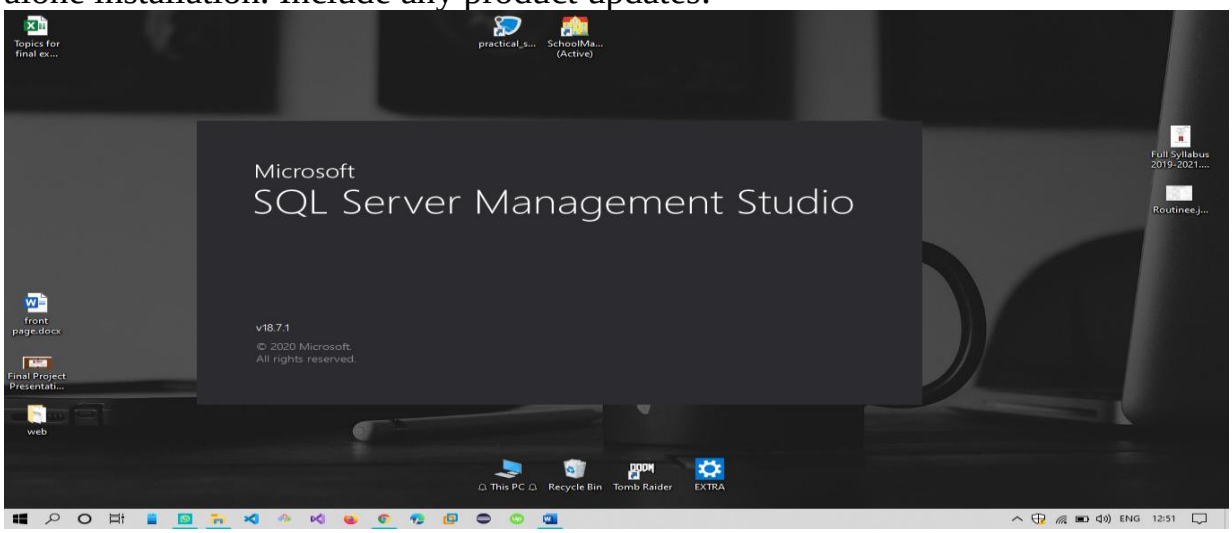
SQL is the standard language for Relational Database System. All the Relational Database Management Systems (RDMS) like MySQL, MS Access, Oracle, Sybase, Informix, Postgres and SQL Server use SQL as their standard database language.

Also, they are using different dialects, such as −

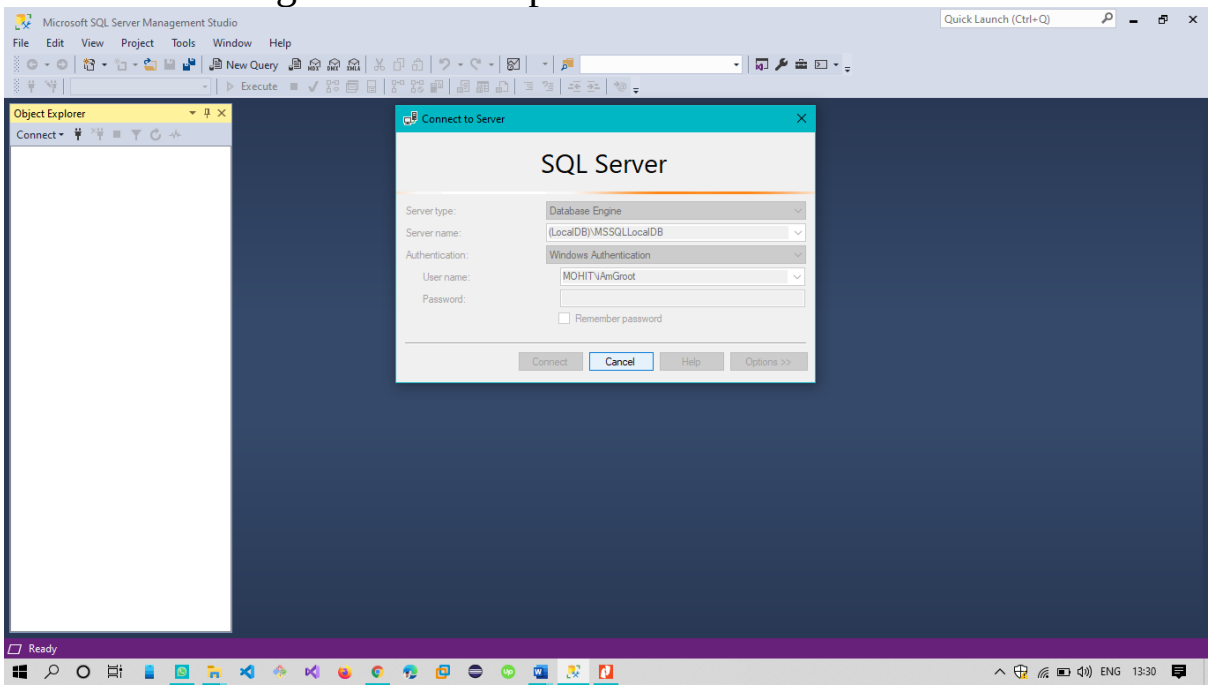
- MS SQL Server using T-SQL,
- Oracle using PL/SQL,
- MS Access version of SQL is called JET SQL (native format) etc.

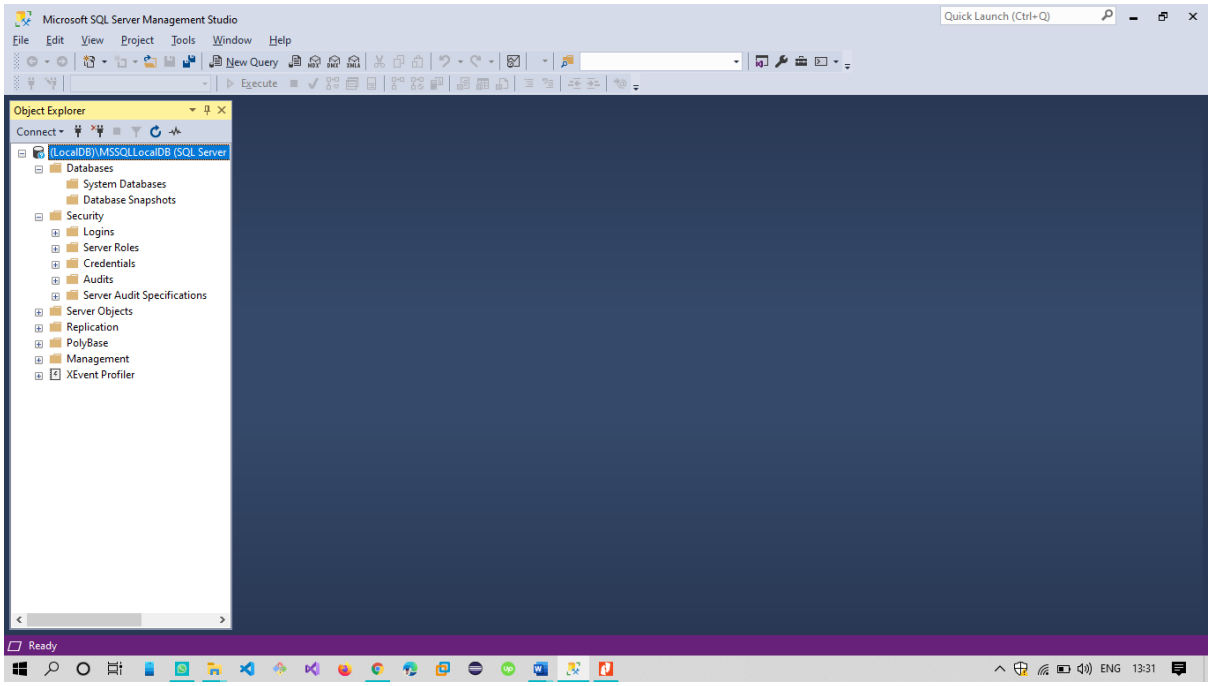
SQL Installation

For that we need SQL server management tool for managing the servers and data tables. Install SQL. Check compatible versions. Choose New SQL Server stand-alone installation. Include any product updates.



Connecting to the server point





Several Contents for the course are:

Getting Started and Selecting & Retrieving Data with SQL

In this module, you will be able to define SQL and discuss how SQL differs from other computer languages. You will be able to compare and contrast the roles of a database administrator and a data scientist, and explain the differences between one-to-one, one-to-many, and many-to-many relationships with databases. You will be able to use the SELECT statement and talk about some basic syntax rules. You will be able to add comments in your code and synthesize its importance.

Quiz will be taken at the end of the topic.

- 9 videos
- 1 reading
- 3 quizzes

Filtering, Sorting, and Calculating Data with SQL

In this module, you will be able to use several newer clauses and operators including WHERE, BETWEEN, IN, OR, NOT, LIKE, ORDER BY, and GROUP BY. You will be able to use the wildcard function to search for more specific or parts of records, including their advantages and disadvantages, and how best to use them. You will be able to discuss how to use basic math operators, as well as aggregate functions like AVERAGE, COUNT, MAX, MIN, and others to begin analyzing our data.

Quiz will be taken at the end of the topic.

- 1 reading
- 3 practice papers include.
- 9 videos

Subqueries and Joins in SQL

In this module, you will be able to discuss subqueries, including their advantages and disadvantages, and when to use them. You will be able to recall the concept of a key field and discuss how these help us link data together with JOINs. You will be able to identify and define several types of JOINs, including the Cartesian join, an inner join, left and right joins, full outer joins, and a self-join. You will be able to use aliases and pre-qualifiers to make your SQL code cleaner and efficient.

Quiz will be taken at the end of the topic.

- 10 videos
- 2 reading
- 3 quizzes

Modifying and Analyzing Data with SQL

In this module, you will be able to discuss how to modify strings by concatenating, trimming, changing the case, and using the substring function. You will be able to discuss the date and time strings specifically. You will be able to use case statements and finish this module by discussing data governance and profiling. You will also be able to apply fundamental principles when using SQL for data science. You'll be able to use tips and tricks to apply SQL in a data science context.

Quiz will be taken at the end of the topic.

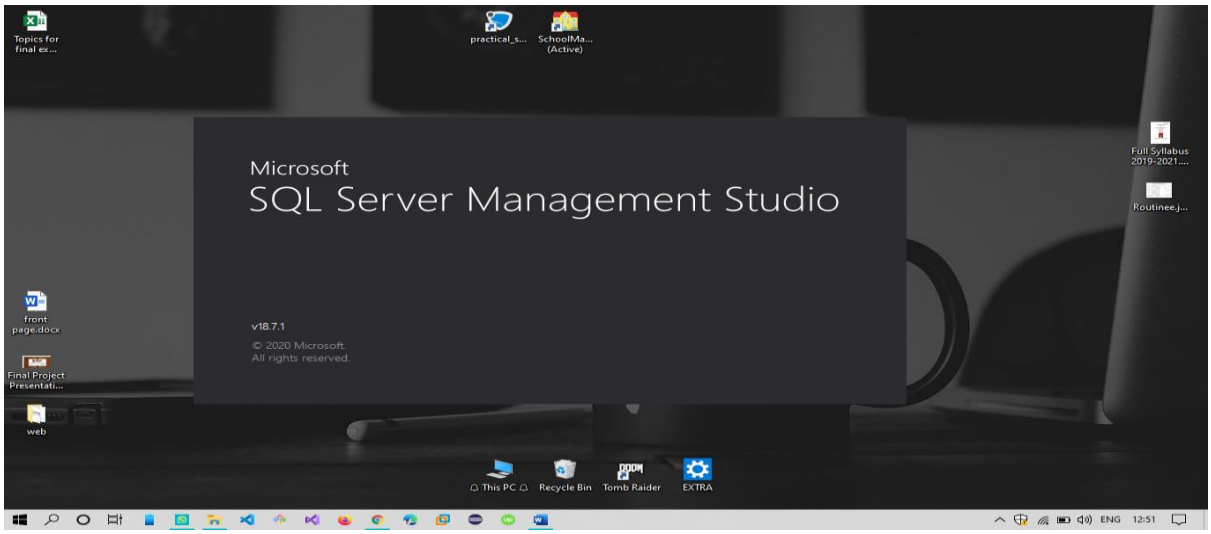
- 10 videos
- 3 readings
- 3 quizzes

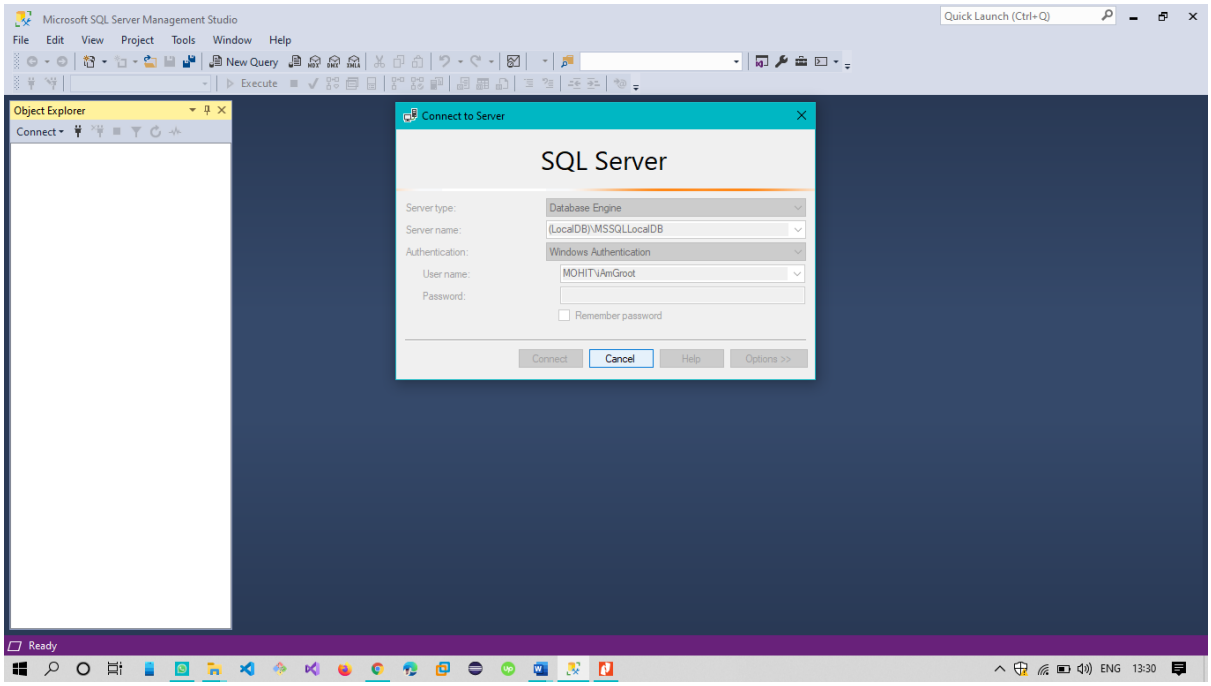
Project made during the training

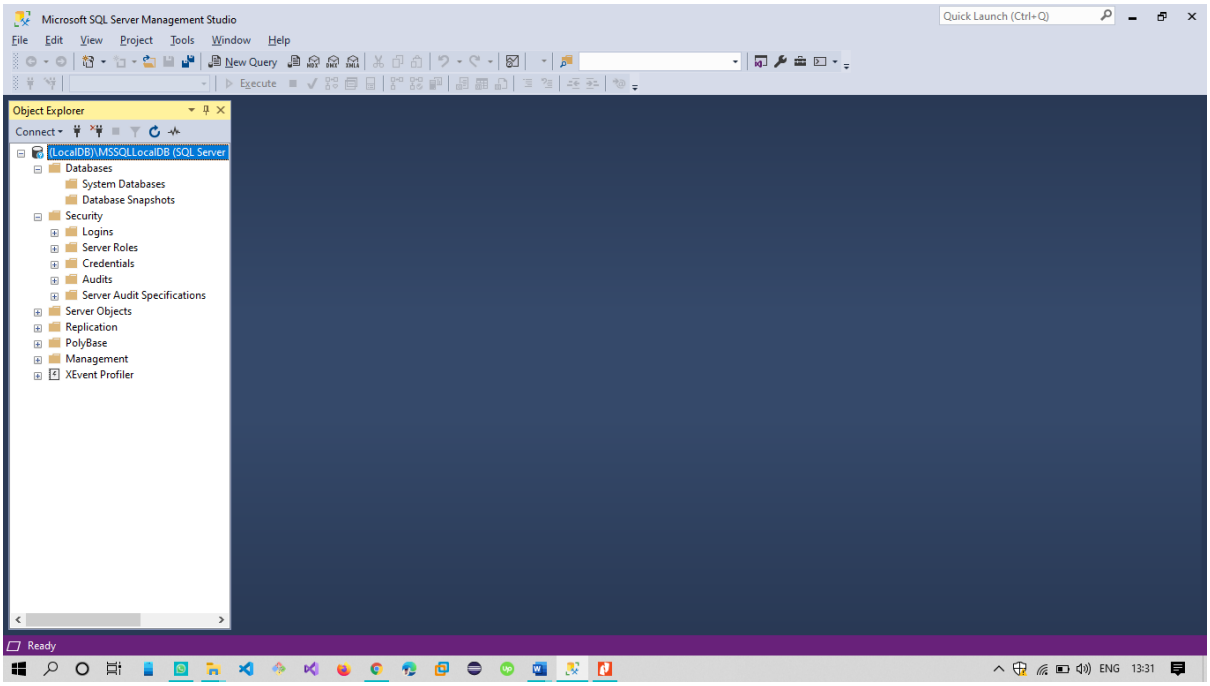
School Management system

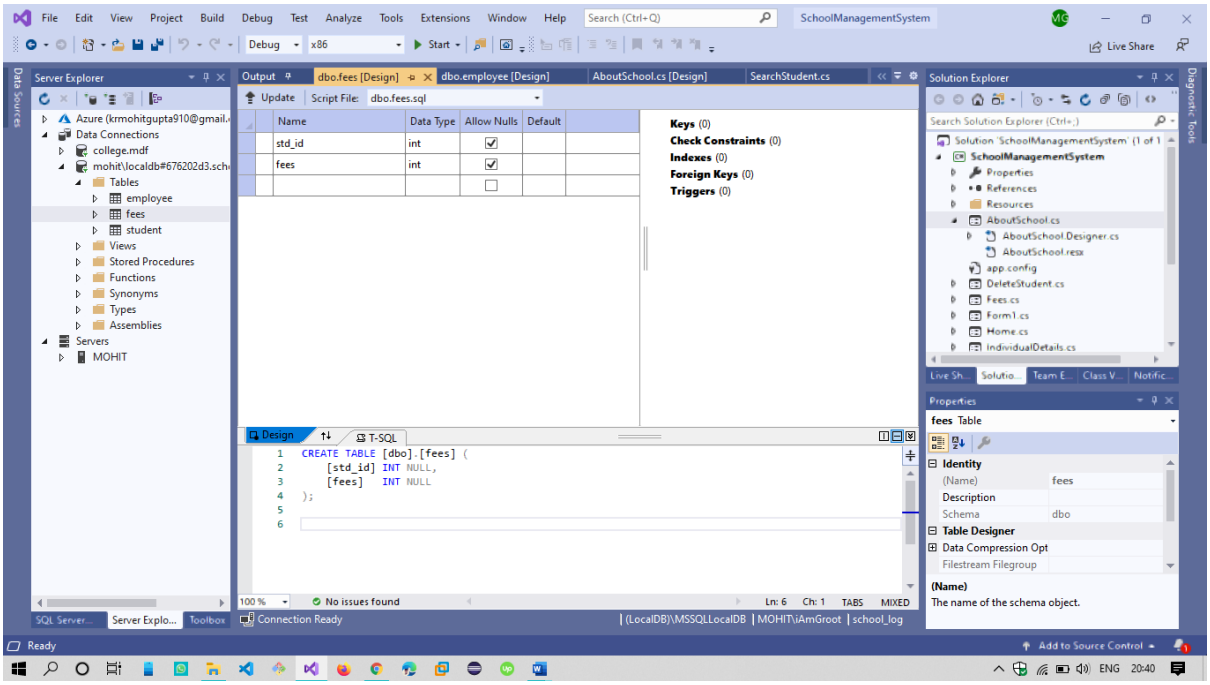
- **Visual studio 2019**
- **SQL Server 2019**

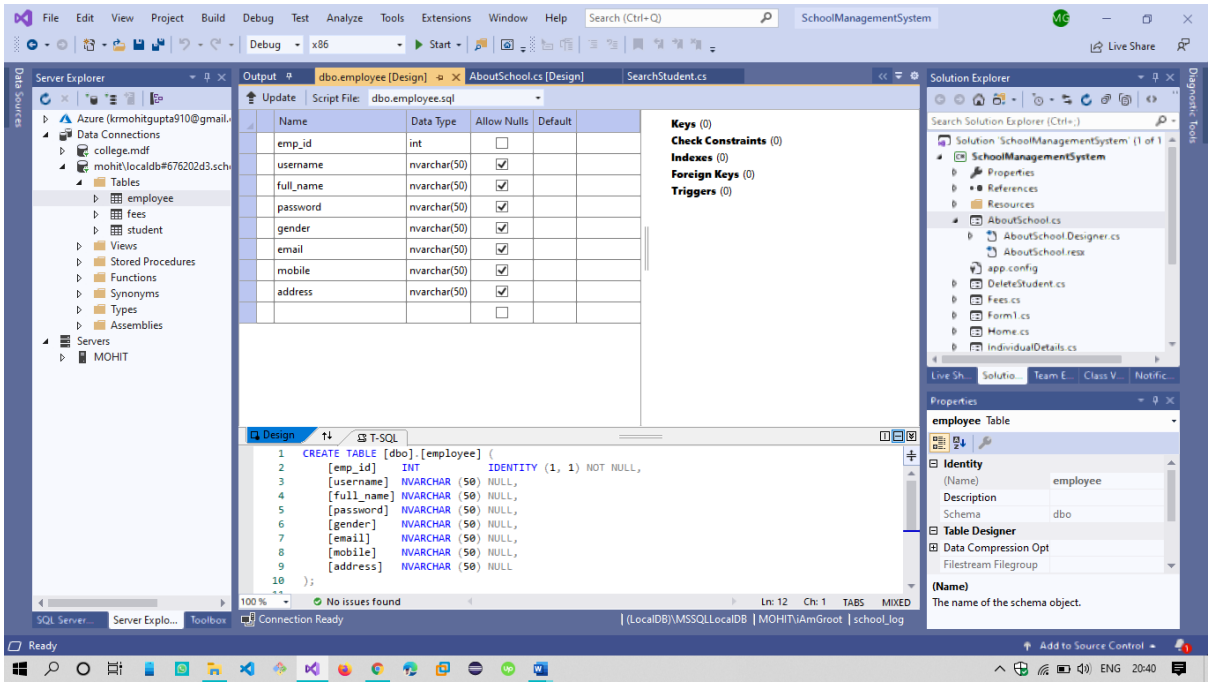
Output for project after execution:

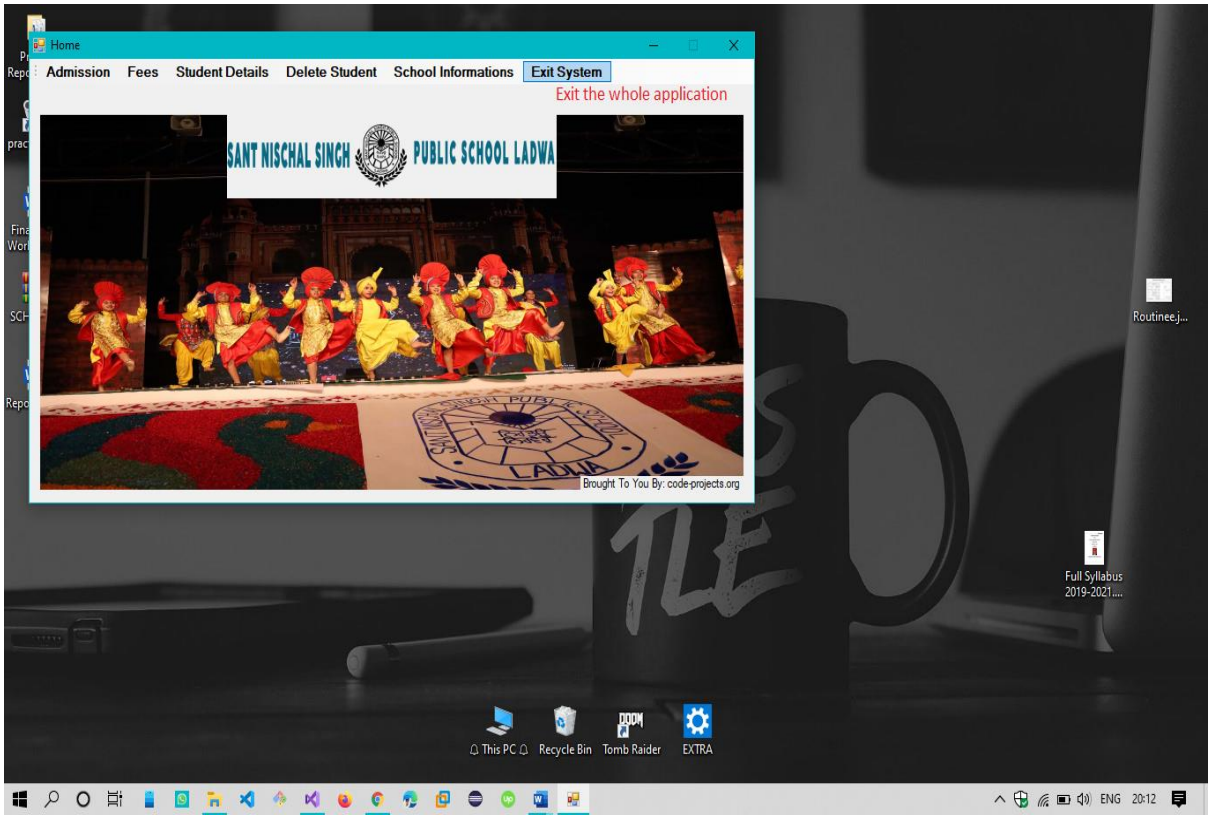


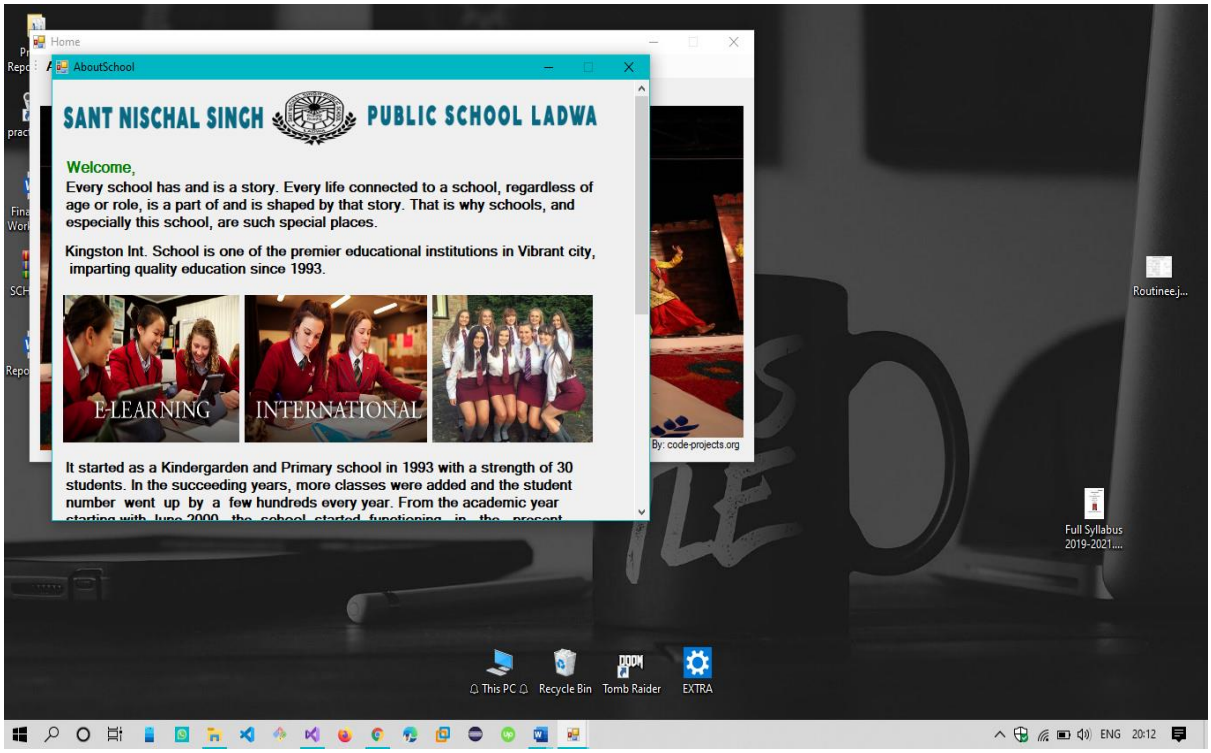


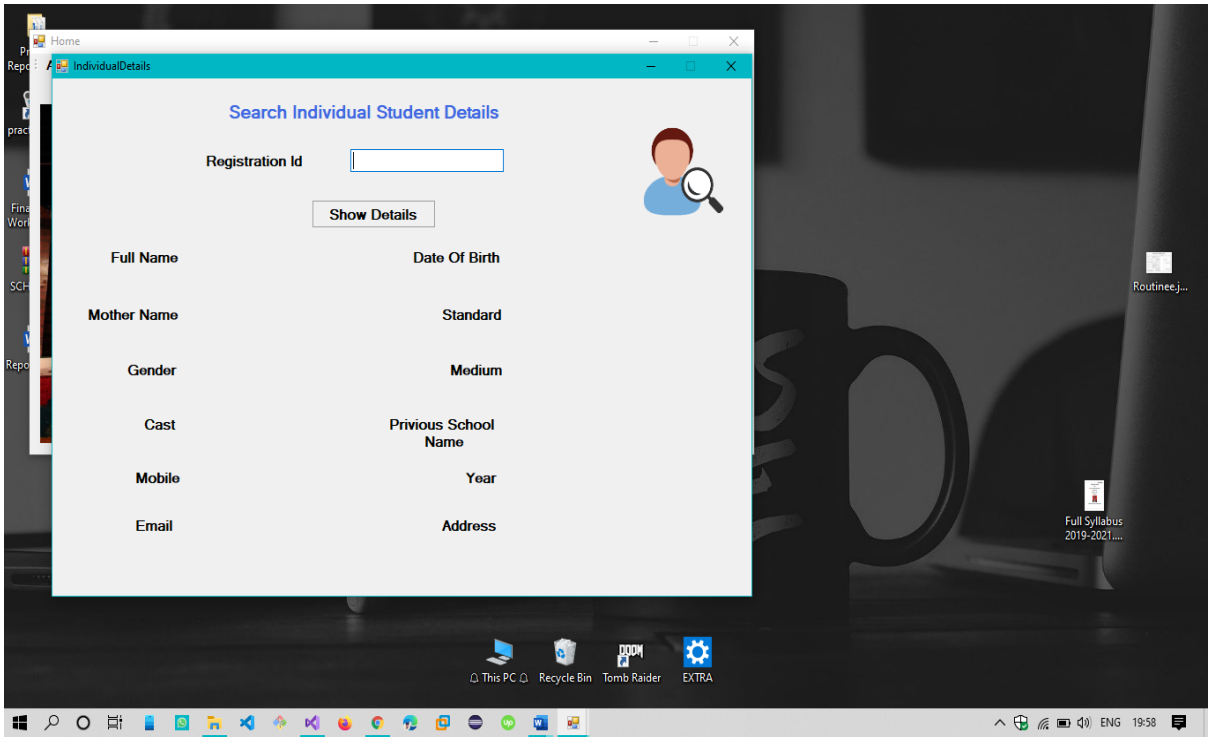

Microsoft SQL Server Management Studio
File Edit View Project Tools Window Help
New Query
Execute
Quick Launch (Ctrl+Q)
Object Explorer
Connect
(LocalDB)\MSSQLLocalDB (SQL Server
Databases
System Databases
Database Snapshots
Security
Logins
Server Roles
Credentials
Audits
Server Audit Specifications
Server Objects
Replication
PolyBase
Management
XEvent Profiler
Ready
ENG 13:31

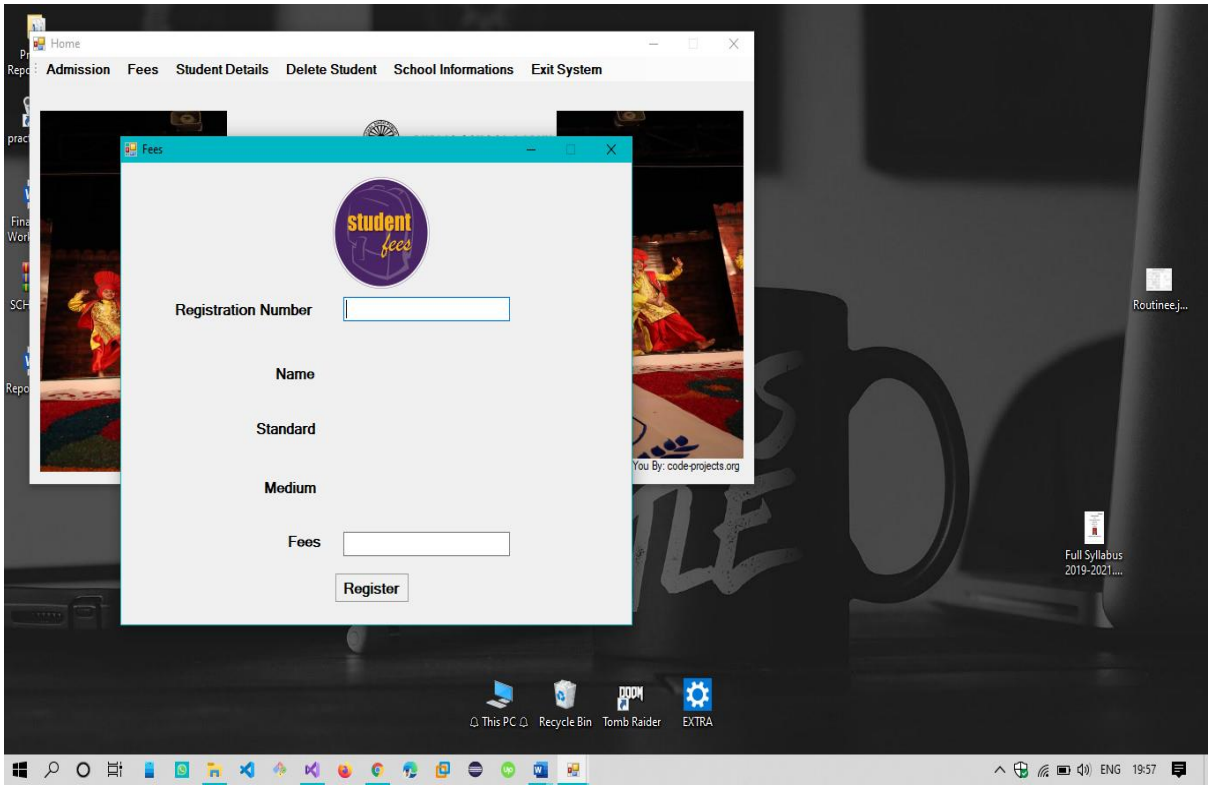

File Edit View Project Build Debug Test Analyze Tools Extensions Window Help
Search (Ctrl+Q)
SchoolManagementSystem
Debug
x86
Start
Live Share
Server Explorer
Azure (krmohitgupta910@gmail.
Data Connections
college.mdf
Tables
employee
fees
student
Views
Stored Procedures
Functions
Synonyms
Types
Assemblies
Servers
MOHIT
Output
dbo.fees [Design]
dbo.employee [Design]
AboutSchool.cs [Design]
SearchStudent.cs
Update
Script File: dbo.fees.sql
Name
Data Type
Allow Nulls
Default
std_id
int
fees
int
Keys (0)
Check Constraints (0)
Indexes (0)
Foreign Keys (0)
Triggers (0)
Design
T-SQL
1 CREATE TABLE [dbo].[fees] (
2 [std_id] INT NULL,
3 [fees] INT NULL
4);
5
6
100 %
No issues found
Ln: 6 Ch: 1 TABS MIXED
Connection Ready
(LocalDB)\MSSQLLocalDB | MOHIT\iAmGroot | school_log
Solution Explorer
Solution 'SchoolManagementSystem' (1 of 1
SchoolManagementSystem
Properties
References
Resources
AboutSchool.cs
AboutSchool.Designer.cs
AboutSchool.resx
app.config
DeleteStudent.cs
Fees.cs
Form1.cs
Home.cs
IndividualDetails.cs
Properties
fees Table
Identity
(Name) fees
Description
Schema dbo
Table Designer
Data Compression Opt
Filestream Filegroup
(Name)
The name of the schema object.
Ready
Add to Source Control
ENG 20:40

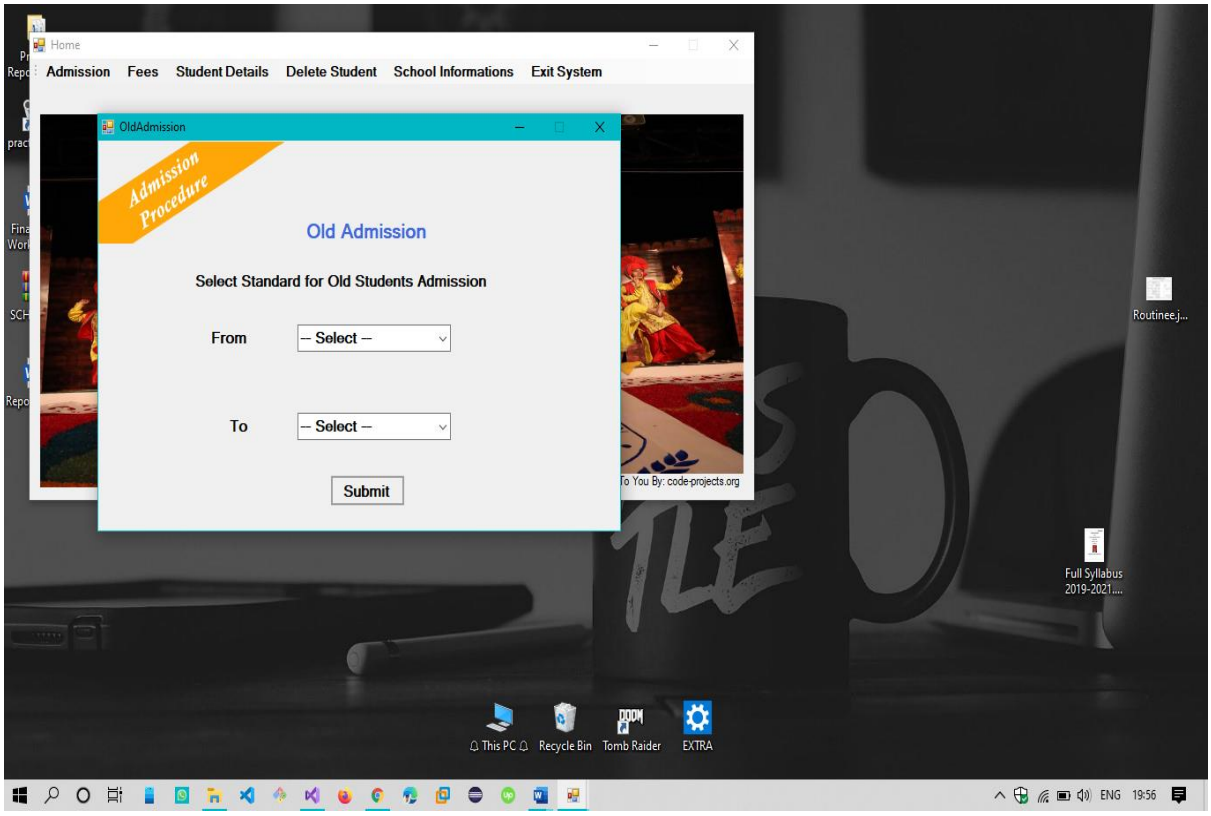

File Edit View Project Build Debug Test Analyze Tools Extensions Window Help
SchoolManagementSystem
Server Explorer
Output
dbo.employee [Design]
AboutSchool.cs [Design]
SearchStudent.cs
Solution Explorer
Name Data Type Allow Nulls Default
emp_id int
username nvarchar(50)
full_name nvarchar(50)
password nvarchar(50)
gender nvarchar(50)
email nvarchar(50)
mobile nvarchar(50)
address nvarchar(50)
Keys (0)
Check Constraints (0)
Indexes (0)
Foreign Keys (0)
Triggers (0)
CREATE TABLE [dbo].[employee] (
[emp_id] INT IDENTITY (1, 1) NOT NULL,
[username] NVARCHAR (50) NULL,
[full_name] NVARCHAR (50) NULL,
[password] NVARCHAR (50) NULL,
[gender] NVARCHAR (50) NULL,
[email] NVARCHAR (50) NULL,
[mobile] NVARCHAR (50) NULL,
[address] NVARCHAR (50) NULL
);
No issues found
Connection Ready
Properties
employee Table
Identity
Table Designer
(Name)
The name of the schema object.
Ready
Add to Source Control

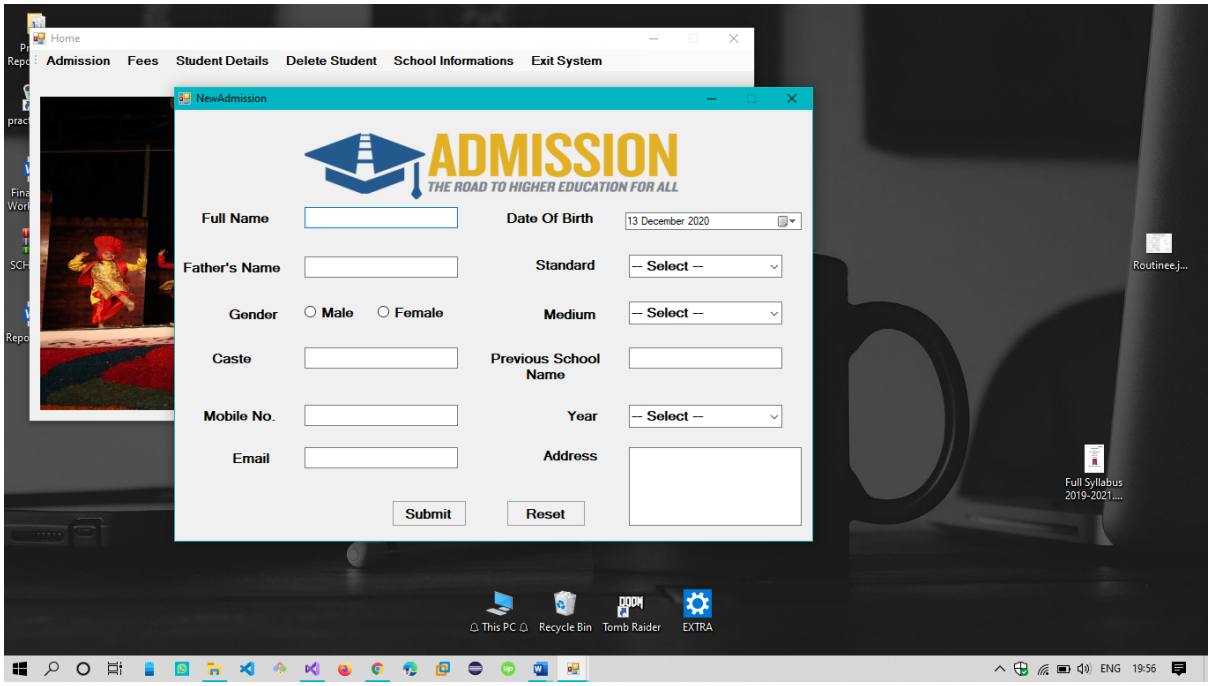

Home
Admission Fees Student Details Delete Student School Informations Exit System
Exit the whole application
SANT NISCHAL SINGH PUBLIC SCHOOL LADWA
Brought To You By: code-projects.org
Routinee.j...
Full Syllabus 2019-2021....
This PC Recycle Bin Tomb Raider EXTRA

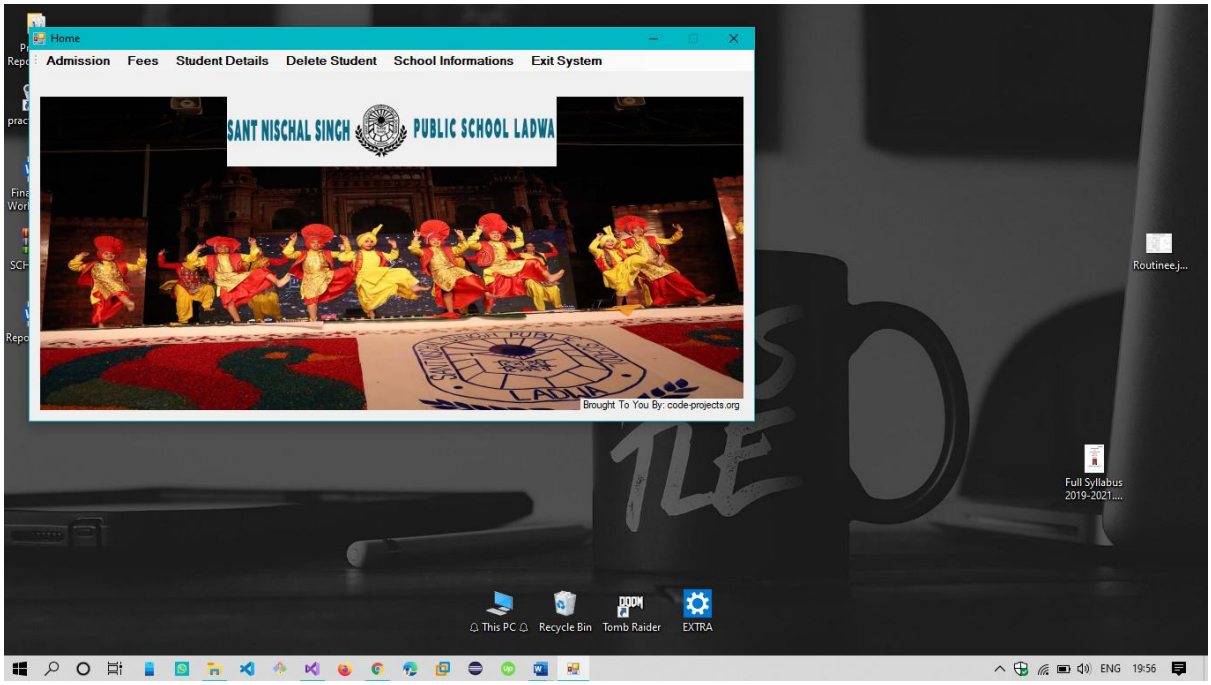

Home
AboutSchool
SANT NISCHAL SINGH PUBLIC SCHOOL LADWA
Welcome,
Every school has and is a story. Every life connected to a school, regardless of age or role, is a part of and is shaped by that story. That is why schools, and especially this school, are such special places.
Kingston Int. School is one of the premier educational institutions in Vibrant city, imparting quality education since 1993.
E-LEARNING
INTERNATIONAL
By: code-projects.org
It started as a Kindergarden and Primary school in 1993 with a strength of 30 students. In the succeeding years, more classes were added and the student number went up by a few hundreds every year. From the academic year
Routinee.j...
Full Syllabus 2019-2021....
This PC
Recycle Bin
Tomb Raider
EXTRA
ENG
20:12

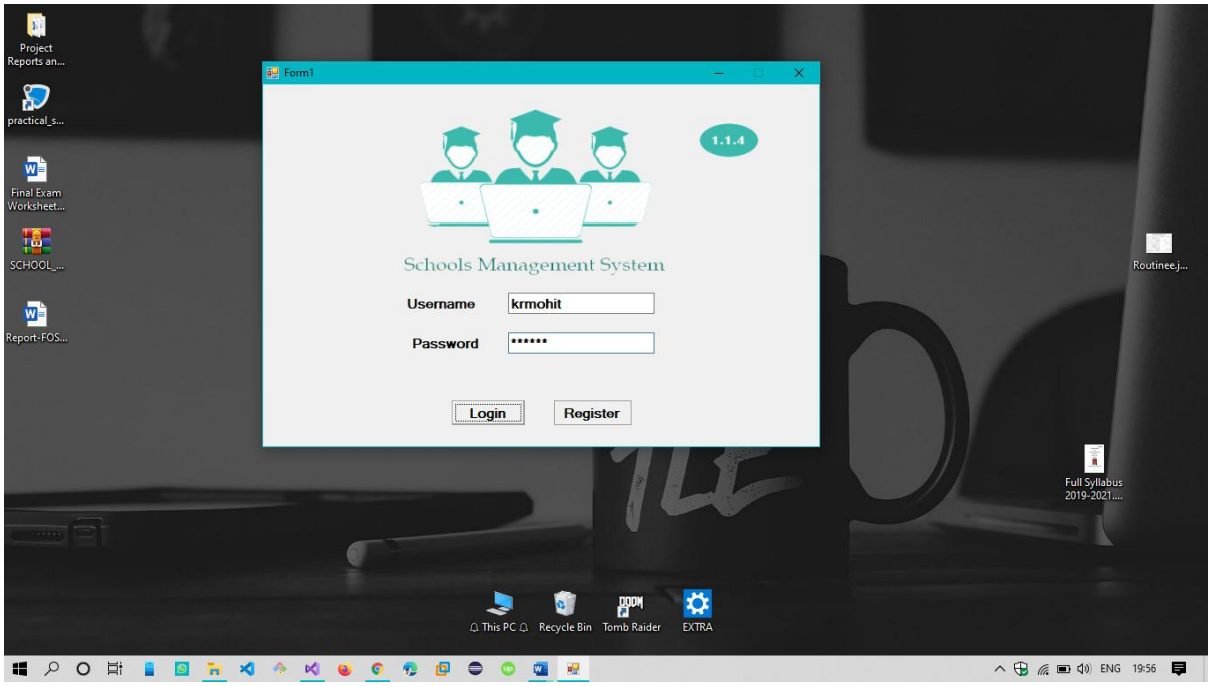

Home
IndividualDetails
Search Individual Student Details
Registration Id
Show Details
Full Name
Date Of Birth
Mother Name
Standard
Gender
Medium
Cast
Privious School Name
Mobile
Year
Email
Address
Routinee.j...
Full Syllabus 2019-2021....
This PC
Recycle Bin
Tomb Raider
EXTRA
ENG
19:58

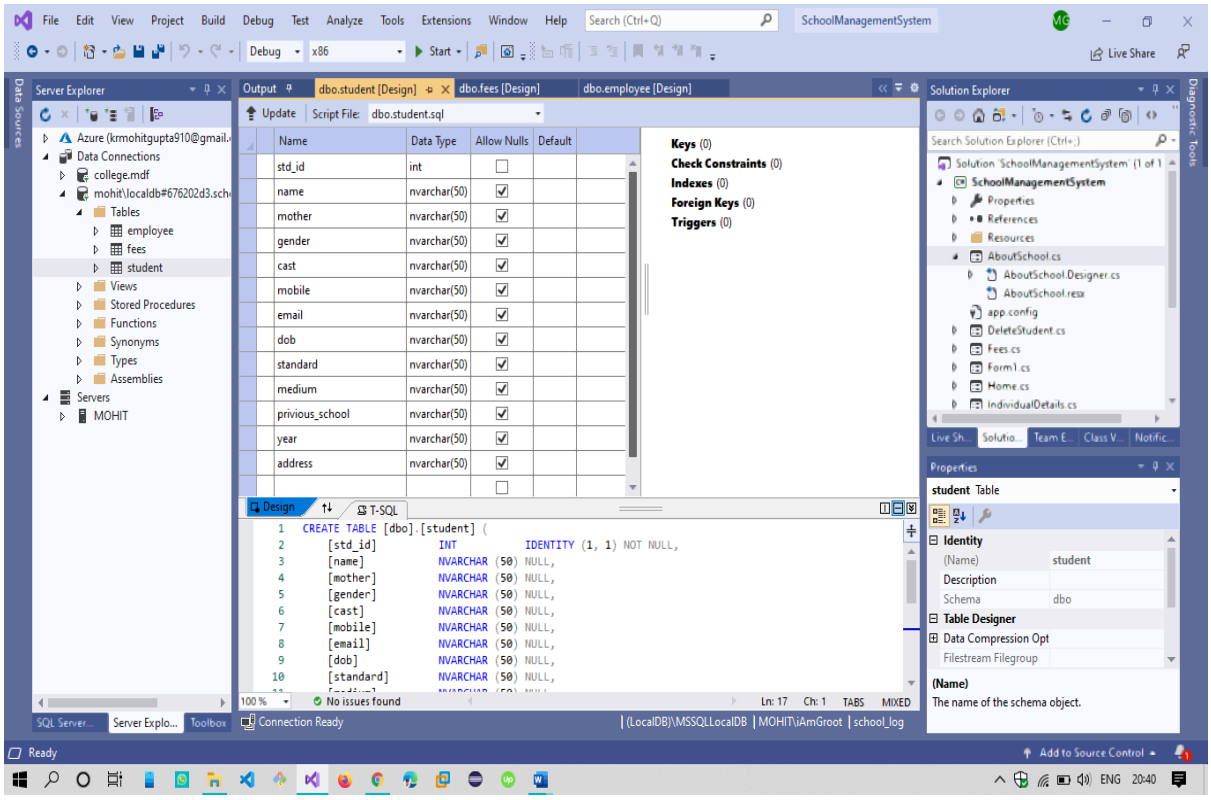

Home
Admission
Fees
Student Details
Delete Student
School Informations
Exit System
Fees
student fees
Registration Number
Name
Standard
Medium
Fees
Register
You By: code-projects.org
Routinee.j...
Full Syllabus 2019-2021....
This PC
Recycle Bin
Tomb Raider
EXTRA
ENG
19:57


Home
Admission
Fees
Student Details
Delete Student
School Informations
Exit System
OldAdmission
Admission Procedure
Old Admission
Select Standard for Old Students Admission
From
-- Select --
To
-- Select --
Submit
To You By: code-projects.org
Routinee.j...
Full Syllabus 2019-2021....
This PC
Recycle Bin
Tomb Raider
EXTRA
ENG
19:56


Home
Admission
Fees
Student Details
Delete Student
School Informations
Exit System
NewAdmission
ADMISSION
THE ROAD TO HIGHER EDUCATION FOR ALL
Full Name
Date Of Birth
13 December 2020
Father's Name
Standard
-- Select --
Gender
Male
Female
Medium
-- Select --
Caste
Previous School Name
Mobile No.
Year
-- Select --
Email
Address
Submit
Reset
This PC
Recycle Bin
Tomb Raider
EXTRA
Routinee.j...
Full Syllabus 2019-2021....
ENG
19:56


Home
Admission
Fees
Student Details
Delete Student
School Informations
Exit System
SANT NISCHAL SINGH
PUBLIC SCHOOL LADWA
Brought To You By: code-projects.org
This PC
Recycle Bin
Tomb Raider
EXTRA
Routinee.j...
Full Syllabus 2019-2021....
ENG
19:56


Form1
1.1.4
Schools Management System
Username
krmohit
Password
Login
Register
Project Reports an...
practical_s...
Final Exam Worksheet...
SCHOOL_...
Report-FOS...
Routinee.j...
Full Syllabus 2019-2021....
This PC
Recycle Bin
Tomb Raider
EXTRA
ENG 19:56


File Edit View Project Build Debug Test Analyze Tools Extensions Window Help
SchoolManagementSystem
Server Explorer
Azure (krmohitgupta910@gmail...
Data Connections
college.mdf
Tables
employee
fees
student
Views
Stored Procedures
Functions
Synonyms
Types
Assemblies
Servers
MOHIT
dbo.student [Design]
dbo.fees [Design]
dbo.employee [Design]
Update
Script File: dbo.student.sql
Name
Data Type
Allow Nulls
Default
std_id int
name nvarchar(50)
mother nvarchar(50)
gender nvarchar(50)
cast nvarchar(50)
mobile nvarchar(50)
email nvarchar(50)
dob nvarchar(50)
standard nvarchar(50)
medium nvarchar(50)
privious_school nvarchar(50)
year nvarchar(50)
address nvarchar(50)
Keys (0)
Check Constraints (0)
Indexes (0)
Foreign Keys (0)
Triggers (0)
Design
T-SQL
CREATE TABLE [dbo].[student] (
[std_id] INT IDENTITY (1, 1) NOT NULL,
[name] NVARCHAR (50) NULL,
[mother] NVARCHAR (50) NULL,
[gender] NVARCHAR (50) NULL,
[cast] NVARCHAR (50) NULL,
[mobile] NVARCHAR (50) NULL,
[email] NVARCHAR (50) NULL,
[dob] NVARCHAR (50) NULL,
[standard] NVARCHAR (50) NULL,
No issues found
Ln: 17 Ch: 1 TABS MIXED
Connection Ready
(LocalDB)\MSSQLLocalDB | MOHIT\iAmGroot | school_log
Solution Explorer
Solution 'SchoolManagementSystem' (1 of 1
SchoolManagementSystem
Properties
References
Resources
AboutSchool.cs
AboutSchool.Designer.cs
AboutSchool.resx
app.config
DeleteStudent.cs
Fees.cs
Form1.cs
Home.cs
IndividualDetails.cs
Properties
student Table
Identity
(Name) student
Description
Schema dbo
Table Designer
Data Compression Opt
Filestream Filegroup
(Name)
The name of the schema object.
Ready
Add to Source Control
ENG 20:40

Bibliography

- Coursera
  - https://www.coursera.org/learn/sql-for-data-science/home/welcome
- SQL server tutorial
  - http://www.sqlservertutorial.net
- Microsoft SQL server
  - https://www.hub.docker.com/
- Cloud SQL server
  - https://www.cloud.google.com/