EST. 1785 UNB
UNIVERSITY OF NEW BRUNSWICK

Faculty of Computer Science

CS1103

Introduction to Databases

Lecture 2

Relational Database Fundamentals

Part I – Table Design

Learning Objectives

By the end of this topic, you will be able to:

PART I:

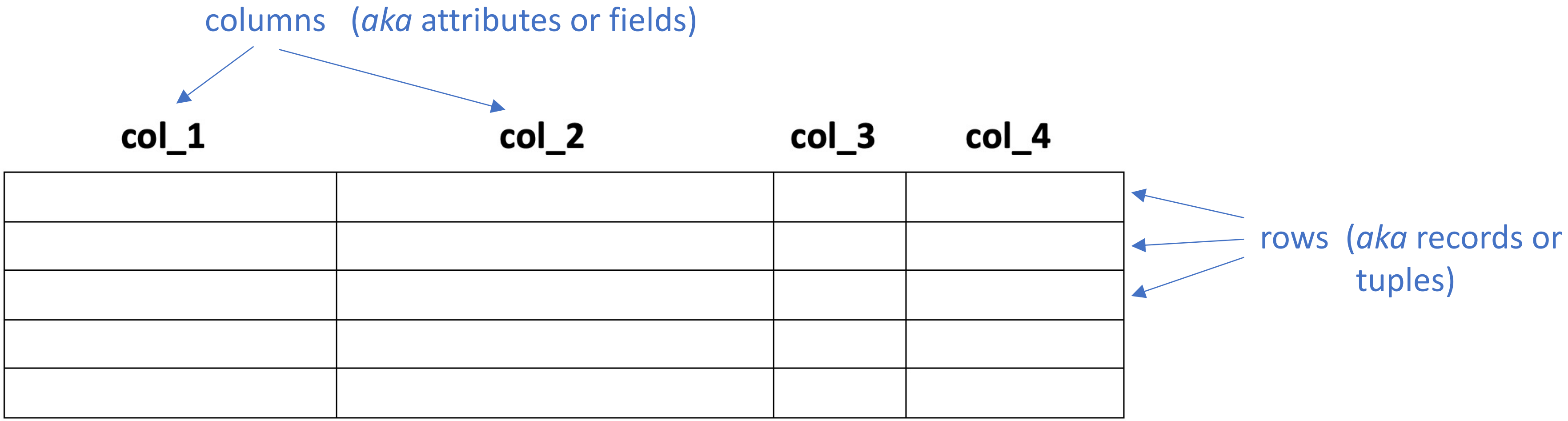
- Describe how a relational database is organized
- Login remotely to MariaDB
- Design a simple database schema

  Schema: Defines the components of a specific database

PART II:

- Use Data Definition Language (DDL) to create a database
- Insert data into database tables
- Display database table contents

Database Tables (Relations)



columns (*aka* attributes or fields)

col_1 col_2 col_3 col_4

rows (*aka* records or tuples)

A table is used to store a LIST of something

- Each row is an item in the list
- Each column is a value that describes an item

A “month” Table

month_name	day_qty	avg_high	avg_low
January	31	-4.0	-13.0
February	28	-2.0	-12.0
…			
August	31	25.0	15.0
September	30	21.0	11.0
…			

This is a list of months

- Each row is one item in the list
- Each cell is a value that describes one month

A Few Rules for Database Tables

No duplicate column names
(within a given table)

No duplicate rows

- Each row can be uniquely identified

month_name	day_qty	avg_high	avg_low
January	31	-4.0	-13.0
February	28	-2.0	-12.0
…			
August	31	25.0	15.0
September	30	21.0	11.0
…			

The order of the rows doesn't matter

Each "cell" stores at most one value

- NULL values are also possible

All values in a given column have the same data type

Primary Key

A **primary key** is one or more columns

Primary key values must uniquely identify every row

In MySQL, every table can have **only one primary key** (even if the PK involves multiple columns)

Superkey: These two columns also uniquely identify rows, but day_qty is unnecessary

month_name	day_qty	avg_high	avg_low
January	31	-4.0	-13.0
February	28	-2.0	-12.0
…			
August	31	25.0	15.0
September	30	21.0	11.0
…			

Candidate key: A superkey with no unnecessary columns

A Second Example: Music Albums

Each album has:

- An album name
- An artist name
- Several tracks … so I can’t list albums like this:

album_name	artist_name	tracks	(etc.)
Greatest Hits	Queen		
Greatest Hits	Boston		
Rumours	Fleetwood Mac		
Second Helping	Lynyrd Skynyrd		

Remember … only one value per cell

Also … is **album_name** a primary key?

Image by Andrew McAllister

Several tracks per album = Separate tables

album

album_id	album_name	artist_name
100	Greatest Hits	Queen
101	Greatest Hits	Boston
102	Rumours	Fleetwood Mac
103	Second Helping	Lynyrd Skynyrd

What are the primary keys?

track

album_id	track_number	song_title	length
100	1	Killer Queen	3:01
100	2	Bohemian Rhapsody	5:55
101	3	More Than a Feeling	null
101	12	Rock and Roll Band	null
102	1	Second Hand News	2:55
102	2	Dreams	4:32
103	1	Sweet Home Alabama	4:43
103	8	Call Me The Breeze	5:09

A Foreign Key

album

album_id	album_name	artist_name
100	Greatest Hits	Queen
101	Greatest Hits	Boston
102	Rumours	Fleetwood Mac
103	Second Helping	Lynyrd Skynyrd

track

album_id	track_number	song_title	length
100	1	Killer Queen	3:01
100	2	Bohemian Rhapsody	5:55
101	3	More Than a Feeling	null
101	12	Rock and Roll Band	null
102	1	Second Hand News	2:55
102	2	Dreams	4:32
103	1	Sweet Home Alabama	4:43
103	8	Call Me The Breeze	5:09

track.album_id is a foreign key that points to **album.album_id**

Referential Integrity: When a value points to a record in another table … that record better be there!

Remote login to mysql

1. Connect to UNB's VPN
2. ssh <username>@<machineid>.cs.unb.ca

   example: ssh andrewm@id414m16.cs.unb.ca

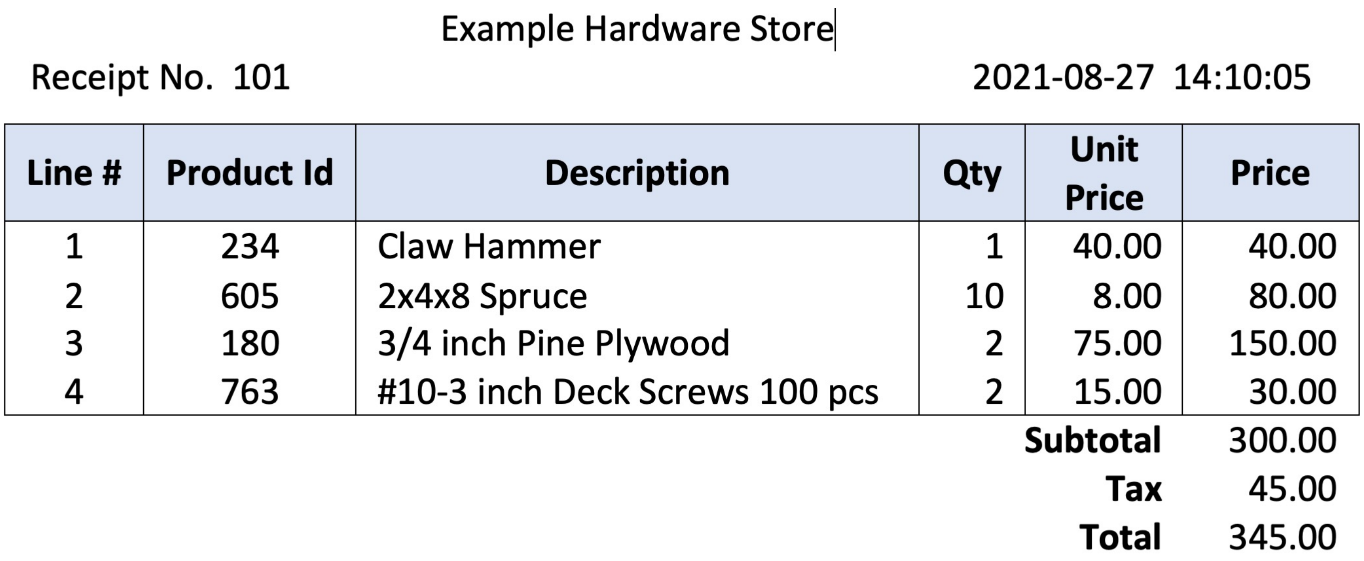
   Instructions for selecting a lab and a machine within the lab will change periodically and are provided separately. See the document entitled "**Lab Access Instructions CS1103.pdf**" provided on D2L
3. Enter your UNB login password
4. mysql -h cs1103.cs.unb.ca -u <username> -p
5. Enter your mysql password
6. use <database>; -- You have your own database for this course

A few commands

- show databases;
  What databases can you access?
- use <database>
  Pick one to work with
- show tables;
  What tables are in the current database?
- show columns from <table>;
  What columns are in a table?
- select * from <table>;
  Show the contents of a table

Do a live demo...

Sales Receipt - Sample #1

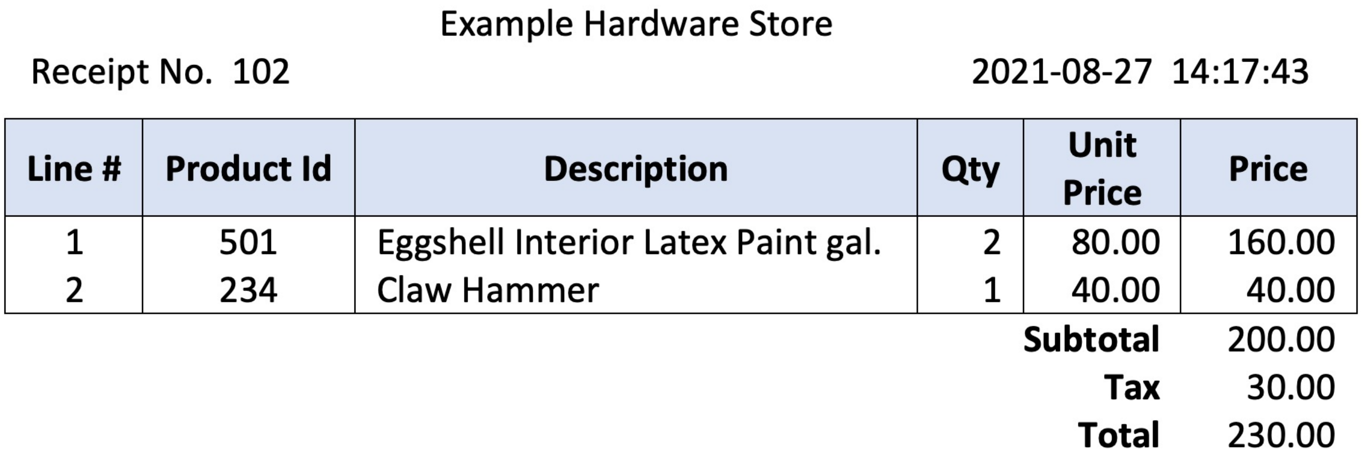


Example Hardware Store

Receipt No. 101 2021-08-27 14:10:05

Line #	Product Id	Description	Qty	Unit Price	Price
1	234	Claw Hammer	1	40.00	40.00
2	605	2x4x8 Spruce	10	8.00	80.00
3	180	3/4 inch Pine Plywood	2	75.00	150.00
4	763	#10-3 inch Deck Screws 100 pcs	2	15.00	30.00
				Subtotal	300.00
				Tax	45.00
				Total	345.00

Sales Receipt - Sample #2

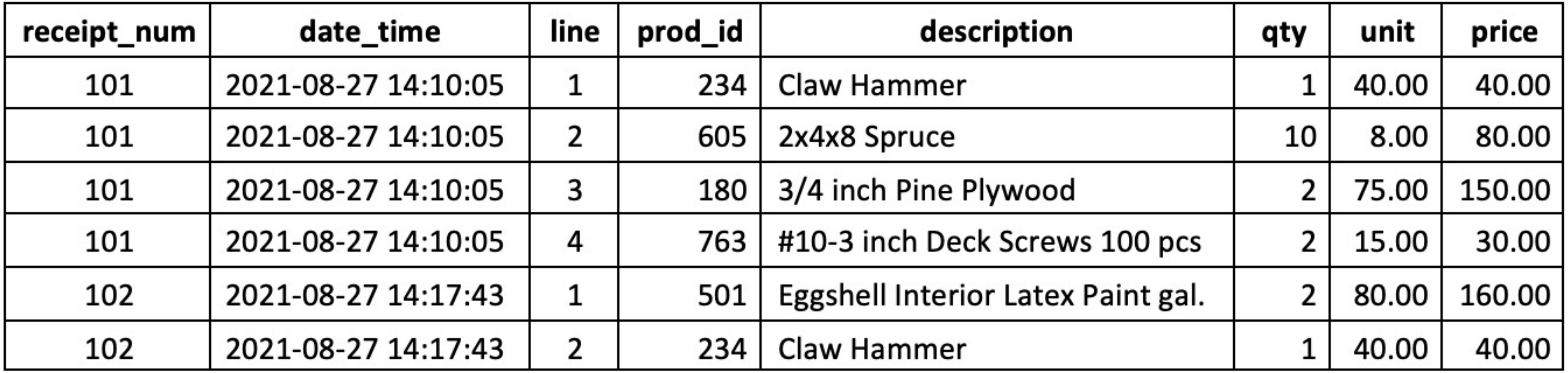


Example Hardware Store

Receipt No. 102 2021-08-27 14:17:43

Line #	Product Id	Description	Qty	Unit Price	Price
1	501	Eggshell Interior Latex Paint gal.	2	80.00	160.00
2	234	Claw Hammer	1	40.00	40.00
				Subtotal	200.00
				Tax	30.00
				Total	230.00

One Possible Database Table



receipt_num	date_time	line	prod_id	description	qty	unit	price
101	2021-08-27 14:10:05	1	234	Claw Hammer	1	40.00	40.00
101	2021-08-27 14:10:05	2	605	2x4x8 Spruce	10	8.00	80.00
101	2021-08-27 14:10:05	3	180	3/4 inch Pine Plywood	2	75.00	150.00
101	2021-08-27 14:10:05	4	763	#10-3 inch Deck Screws 100 pcs	2	15.00	30.00
102	2021-08-27 14:17:43	1	501	Eggshell Interior Latex Paint gal.	2	80.00	160.00
102	2021-08-27 14:17:43	2	234	Claw Hammer	1	40.00	40.00

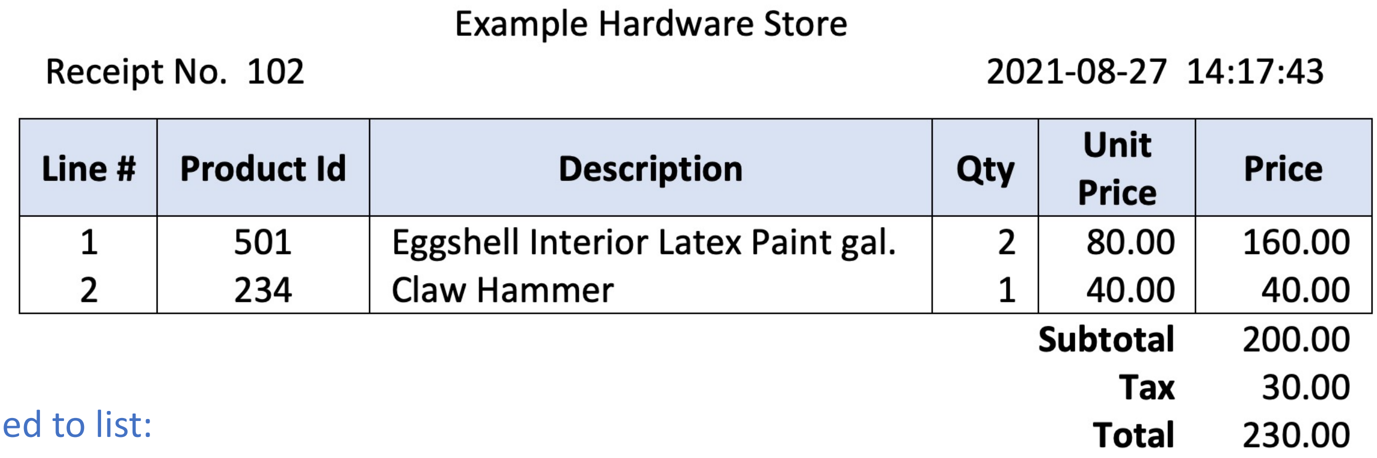
The date/time for receipt 101 is stored multiple times ***(We want to store each fact only once)***

- As are the name & unit price for product 234 - Wasted space
- What if I change unit price in only one spot - Inconsistent data

Where to store data for a product no one has purchased yet?

Do I need the “price” column? - Derived data!

What Lists Do We Need To Keep?



Example Hardware Store

Receipt No. 102 2021-08-27 14:17:43

Line #	Product Id	Description	Qty	Unit Price	Price
1	501	Eggshell Interior Latex Paint gal.	2	80.00	160.00
2	234	Claw Hammer	1	40.00	40.00
				Subtotal	200.00
				Tax	30.00
				Total	230.00

We need to list:

- Products – our inventory: things we have available for sale
- Receipts – each one is a record of a sales transaction
- Line items – we have several of these for each receipt

So ... three tables!

Several Lines per Receipt (and per Product)

The **receipt** table

- One row for each receipt
- Only the “one value per receipt” columns

receipt_num	date_time
101	2021-08-27 14:10:05
102	2021-08-27 14:17:43

The **product** table

prod_id	description	unit
234	Claw Hammer	40.00
605	2x4x8 Spruce	8.00
180	3/4 inch Pine Plywood	75.00
763	#10-3 inch Deck Screws 100 pcs	15.00
501	Eggshell Interior Latex Paint gal.	80.00

The **line_item** table

receipt_num	line	prod_id	qty
101	1	234	1
101	2	605	10
101	3	180	2
101	4	763	2
102	1	501	2
102	2	234	1

price (per line), subtotal, tax, total ... these are all **derived values**. They can be calculated and thus are not stored

Primary Keys

receipt

receipt_num	date_time
101	2021-08-27 14:10:05
102	2021-08-27 14:17:43

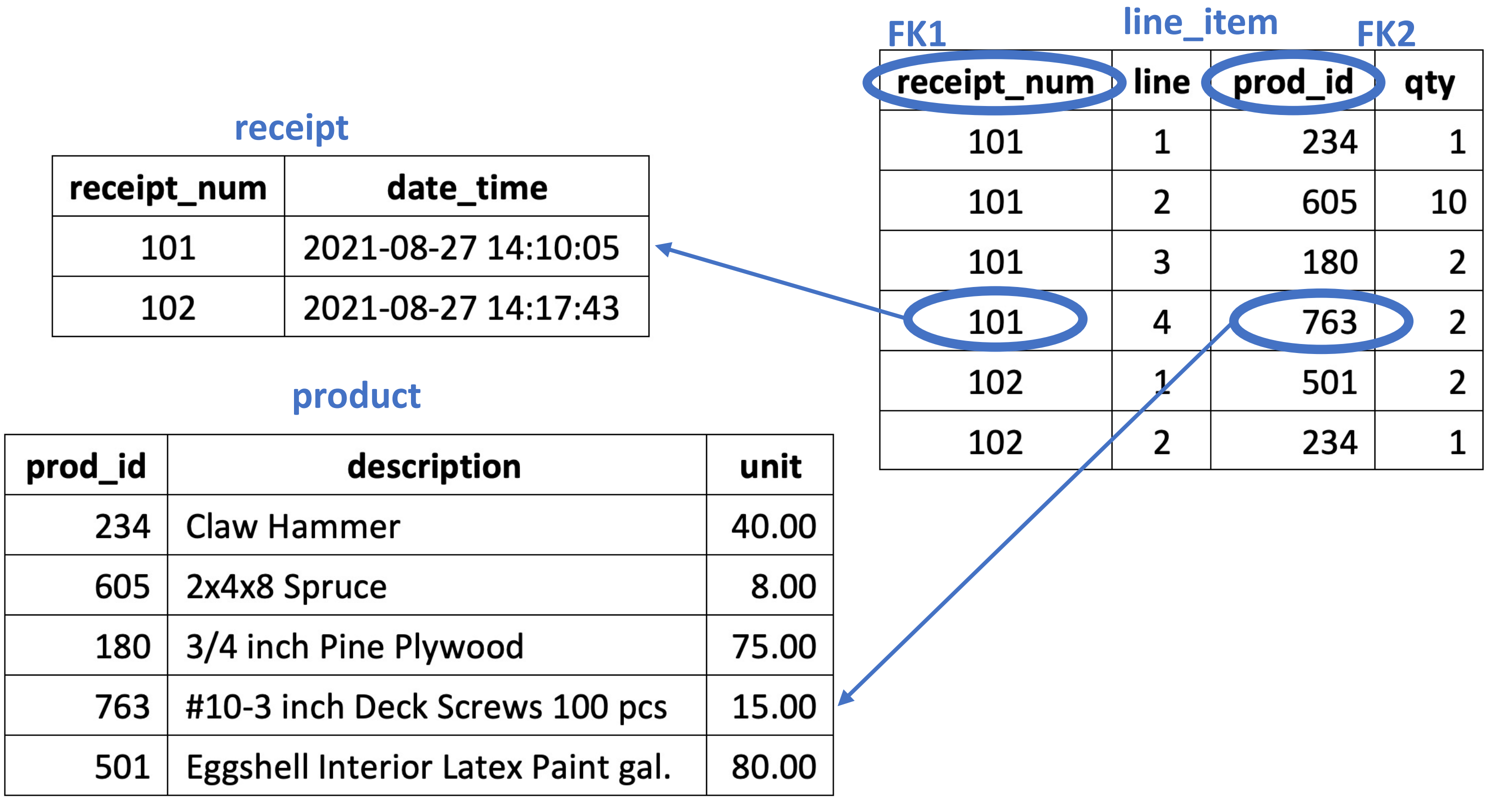
product

prod_id	description	unit
234	Claw Hammer	40.00
605	2x4x8 Spruce	8.00
180	3/4 inch Pine Plywood	75.00
763	#10-3 inch Deck Screws 100 pcs	15.00
501	Eggshell Interior Latex Paint gal.	80.00

line_item

receipt_num	line	prod_id	qty
101	1	234	1
101	2	605	10
101	3	180	2
101	4	763	2
102	1	501	2
102	2	234	1

Foreign Keys



line_item

receipt_num	line	prod_id	qty
101	1	234	1
101	2	605	10
101	3	180	2
101	4	763	2
102	1	501	2
102	2	234	1

receipt

receipt_num	date_time
101	2021-08-27 14:10:05
102	2021-08-27 14:17:43

product

prod_id	description	unit
234	Claw Hammer	40.00
605	2x4x8 Spruce	8.00
180	3/4 inch Pine Plywood	75.00
763	#10-3 inch Deck Screws 100 pcs	15.00
501	Eggshell Interior Latex Paint gal.	80.00

Coming up in Part II…

Let's create our first database!

The End

☺