PRINCIPLES OF GOOD DATABASE DESIGN

PRINCIPLES OF GOOD DATABASE DESIGN

THERE IS SO MUCH THAT CAN BE SAID ABOUT THIS - ENTIRE TEXTBOOKS CAN BE FILLED ONLY ON THIS TOPIC!

WE'LL LOOK AT A FEW OF THE MOST IMPORTANT PRINCIPLES OF GOOD DATABASE DESIGN

THIS IS NOT A COMPREHENSIVE LIST OF WHAT GOES INTO DESIGNING A DATABASE WHICH STANDS THE TEST OF TIME

BUT IT'S A GOOD STARTING POINT FOR HOW TO START THINKING ABOUT TABLES AND THEIR RELATIONSHIPS WITH EACH OTHER

TABLES

IDENTIFY THE PROBLEM THAT YOU'RE TRYING TO MODEL OR CAPTURE

HAVE A GOOD IDEA OF THE TYPES OF QUERIES THAT YOU WANT TO RUN ON THE DATABASE

BREAK UP THE INFORMATION INTO LOGICAL UNITS WHERE EACH LOGICAL UNIT IS ONE TABLE IN YOUR DATABASE

EACH LOGICAL COMPONENT WILL PROBABLY CONNECT WITH OTHER LOGICAL COMPONENTS

CAPTURE THE RELATIONSHIPS EITHER VIA CONSTRAINTS OR OTHER TABLES DEVOTED TO HOLDING RELATIONSHIP SPECIFIC DATA

MORE ON THIS LATER...

WITHIN A TABLE...

TABLES TYPICALLY HOLD INFORMATION
ABOUT AN ENTITY OR A RELATIONSHIP

THINK OF WHETHER YOU NEED A WAY TO
UNIQUELY IDENTIFY EACH ROW IN A TABLE

YES?

YOU NEED A PRIMARY KEY!

WITHIN A TABLE...

YOU NEED A PRIMARY KEY!

A PRIMARY KEY CAN NEVER HAVE A NULL VALUE

A PRIMARY KEY WILL BE USED TO REPRESENT THAT ROW - OTHER TABLES CAN REFERENCE THAT ROW BY SPECIFYING THE PRIMARY KEY VALUE

IF THERE IS NO SUCH VALUE IN THE FIELDS YOU WANT TO STORE THEN GENERATE A VALUE WHICH WILL BE USED AS A UNIQUE IDENTIFIER

E.G. GENERATE A CUSTOMER_ID TO REPRESENT A CUSTOMER, DO NOT RELY ON EMAIL, PHONE NUMBER ETC WHICH CAN CHANGE

WITHIN A TABLE...

YOU NEED A PRIMARY KEY!

NUMERIC VALUES TEND TO BE MORE EFFICIENT IN REPRESENTATION AND LOOKUP - PREFER THOSE FOR PRIMARY KEYS

USE AUTO_INCREMENT TO GET THE DBMS TO GENERATE UNIQUE IDENTIFIERS FOR YOU

HAVING AN IDENTIFIER WHOSE ONLY JOB IS TO BE THE KEY IS STANDARD PRACTICE

THEN THERE IS NO REASON FOR THE KEY TO CHANGE FOR A SPECIFIC ROW

WITHIN A TABLE...

YOU NEED A PRIMARY KEY!

REMEMBER MOST DATABASE SYTEMS WILL AUTOMATICALLY CREATE AN INDEX ON THE PRIMARY KEY FOR FASTER LOOKUP AND RETRIEVAL

WITHIN A TABLE...

DO COLUMNS HAVE OTHER CONSTRAINTS WHICH THEY NEED TO SATISFY?

ARE NULL VALUES ALLOWED FOR ANY COLUMNS?

NO?

USE "NOT NULL"

SHOULD THEY HAVE HAVE DEFAULT VALUES?

YES?

USE "DEFAULT"

WITHIN A TABLE...

DO COLUMNS HAVE OTHER CONSTRAINTS WHICH THEY NEED TO SATISFY?

ARE THE DATA TYPES FOR EACH COLUMN CORRECT?

DON'T USE STRINGS FOR STORING DATE INFORMATION

IF THE DATA IS SINGLE CHARACTER DON'T USE A 30 CHAR STRING TO STORE IT