

## Where are we?

- Completed
  - Quizzes: 10%
  - Midterm: 30%
- Partial
  - Assignment: 27% completed
  - Group Presentation: 10% (three groups left)
  - Class participation (Labs etc): 5%
- Left: SQL
  - Two lectures: lab and assignment 8 (8%; bonus questions)
  - Final Take home exam on SQL: 10%
  - SQL is a more online self-study + ask questions as office hours topic

2

2

## Midterm (30%)

- Mean: 87%
- Median: 89.6
- Std: 8.1
- Max/Min: 97/66.8 = 29.1/20 out of 30
- As long as everyone completes learning activities as done so far A or B
  - Sql 18% = 8% (assignment 8)+ 10% (final take home)
  - + 10% (group presentation) + 5% (class participation)

3

3



Anything ?

4

## Relational Databases: SQL 1

Hye-Chung Kum (kum@tamu.edu)

Associate Professor

Population Informatics Lab (<https://pinformatics.org/>)

Course URL: <http://pinformatics.org/phpm631>

Health Information Technology by Hye-Chung Kum is licensed under a  
Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License

License:



5

5

## Outline

- What is a database?
- What is a database management system?
- An Introduction to SQL
  - How to retrieve data from a database
  - How to create a database (optional)

6

## Why are we learning this?

- Databases are perhaps the single most important class of corporate applications
- Databases are surprisingly powerful data modeling and analysis tools in situations where spreadsheets fall short
  - Students who plan to work in management consulting will soon find this out
- MS Access is a great example of how easy it is to build powerful applications without the need of a background in technology
  - But we are going to learn SQL directly

7

## What is a database ?

- Boring answer
  - A structured collection of data
  - Example: A telephone directory
- Insightful answer
  - Digital representation of the real world
  - A data-centered mirror of an organization's business processes
  - Structure of data reflects organizational processes
  - Content of data reflects organization's history

8

## Representing the Real World as Data: What Data is needed?

- Entity
  - a person, place, thing, or event on which we maintain information
  - Examples: Employees, Patients, Products, Services, Warehouses
- Attribute
  - characteristic or quality of particular entity
  - Examples: Employee's SSN, Patient's Address, Product's Unit Price
- Relationships Among Entities
  - Examples:
    - Patients -Visits -Services(s)
    - Patients -Seen by -Doctors

9

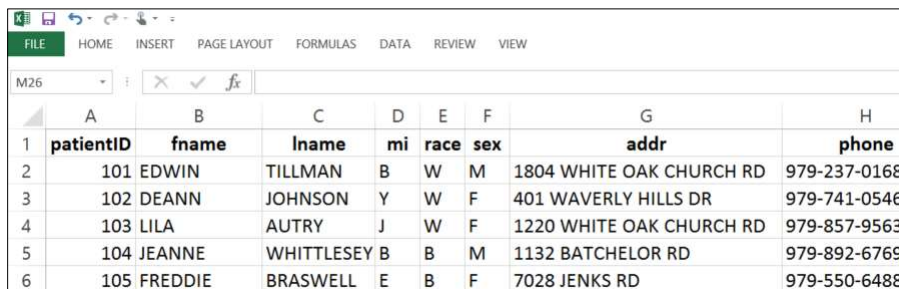
## From Spreadsheets to Databases

- Spreadsheets are great for keeping track of data for one type of entities
  - Participants of a conference
  - Students of a class
  - Customers of a company
  - ....

10

## What is the basic spreadsheet “data model”?

- Each row stores data about one entity
- Each column stores data about an attribute
- Each cell stores data about an attribute of an entity



	A	B	C	D	E	F	G	H
1	<b>patientID</b>	<b>fname</b>	<b>lname</b>	<b>mi</b>	<b>race</b>	<b>sex</b>	<b>addr</b>	<b>phone</b>
2	101	EDWIN	TILLMAN	B	W	M	1804 WHITE OAK CHURCH RD	979-237-0168
3	102	DEANN	JOHNSON	Y	W	F	401 WAVERLY HILLS DR	979-741-0546
4	103	LILA	AUTRY	J	W	F	1220 WHITE OAK CHURCH RD	979-857-9563
5	104	JEANNE	WHITTLESEY	B	B	M	1132 BATCHELOR RD	979-892-6769
6	105	FREDDIE	BRASWELL	E	B	F	7028 JENKS RD	979-550-6488

11

# Spreadsheets



- PRO: Human friendly = flexible
  - Data for people (notes)
- CON: not computer friendly
  - Difficult to keep the data strictly following rules (protocols/standards)
  - If rules are not strictly followed, computers can not do much with the data
  - Keeping multiple sheets consistent (relationships) is difficult
  - Concurrent access by many users is difficult
  - Tracking change is difficult
    - Trackback, if something is not right, is difficult

12

The image displays four overlapping spreadsheet tables, each with a callout box identifying a unique ID:

- Patient Unique ID:** Points to the 'patientID' column in the top table.
- Visit Unique ID:** Points to the 'visitID' column in the second table.
- Medication Unique ID:** Points to the 'medID' column in the third table.
- Provider Unique ID:** Points to the 'providerID' column in the bottom table.

**Table 1: Patient Information**

patientID	fname	lname	mi	race	sex	addr	phone
101	EDWIN	TILLMAN	B	W	M	1804 WHITE OAK CHURCH RD	979-237-0168
		JOHNSON					979-741-0546
		UTRY	J				979-857-9563

**Table 2: Visits**

visitID	patientID	providerID	dtl_dos	dtl_qtr	diag1	diag2	diag3	proc
50	86	101	11/29/2012	2012QTR4	4019	2720		G0154
51	87		12/7/2012	2012QTR4	4019	2720		G0154
52	88		12/14/2012	2012QTR4	4019	2720		G0163
53	140		1/1/2012	2012QTR1	78830			73130

**Table 3: Medications**

medID	patientID	ahfs_cd	ahfs_desc	rx_qua
483612	102	282408	BENZODIAZEPINES (ANXIOLYTIC,SEDATIV/HYP)	
483613	102	200404	IRON PREPARATIONS	
483614		282408	BENZODIAZEPINES (ANXIOLYTIC,SEDATIV/HYP)	

**Table 4: Providers**

providerID	fname	lname	mi	type
1	LAURIE	FRY	L	Doctor
2	WILLIAM	GAINES	W	Doctor
3	DANA	MAGOON	L	Doctor
4		BONELLO	L	Nurse
5		CHANG	Y	Nurse
6		DUBAY	R	Nurse
7	ROBERT	MACDDE	M	Nurse

13

## Relational Databases

- A relational DB supports storage of data as a set of inter-related tables
  - Each table stores data about a set of Entities
  - Each table row is a record about one such Entity
  - Each record column is a field specifying an attribute of this Entity
  - Each record has a field that acts as a unique identifier of an entity (usually an ID number)
  - Relationships among entities are specified by referring to this unique identifier from other tables
  - DB Schema: blueprint of the DB
    - Table (entity, attributes) & relationships

14

## Primary Key & Foreign key

- Combinations of foreign and primary keys are used to implement & enforce relationships
- Primary Key
  - Every table must have a field whose value uniquely identifies each record in the table (e.g. the SSN for people, the ISBN for books, etc.)
  - That field is called the primary key and is marked in the fields list
  - This “represents” the entity of the table
- Foreign Key
  - Used to identify records (entities) in other table according to relation
  - An attribute in a table that uses primary key in another table
- Referential Integrity
  - Based on types of relations keep DB consistent
  - Example: Cascading deletes, not allow inserts

15

patientID	fname	lname	mi	race	sex	addr	phone
101	EDWIN	TILLMAN	B	W	M	1804 WHITE OAK CHURCH RD	979-237-0168
		HNSON					979-741-0546
		UTRY	J			CH RD	979-857-9563

Patient Unique ID

Reference to a patient

visitID	patientID	providerID	dtl_dos	dtl_qtr	diag1	diag2	diag3	proc
50	86	101	11/29/2012	2012QTR4	4019	2720		G0154
51			12/12/2012					G0154
52			12/14/2012					G0154
53	140	102	1/1/2013	2013QTR1	78830			73130

Primary Key: Visit Unique ID

Foreign Keys: patientID & providerID

medID	patientID	ahfs_cd	ahfs_desc	rx_qua
483612	102	282408	BENZODIAZEPINES (ANXIOLYTIC,SEDATIV/HYP)	
483613	102	200404	IRON PREPARATIONS	
483614	102	282408	BENZODIAZEPINES (ANXIOLYTIC,SEDATIV/HYP)	

Medication Unique ID

providerID	fname	lname	mi	type
1	LAURIE	FRY	L	Doctor
2	WILLIAM	GAINES	W	Doctor
3	DANA	MAGOON	L	Doctor
4	ALAN	BONELLO	L	Nurse
5	CHANG	Y		Nurse
6	DUBAY	R		Nurse
7	ROBERT	MACDDE	M	Nurse

Provider Unique ID

16

## Schema

- What tables?
- What variables in a table?
- How are they related? (keys)

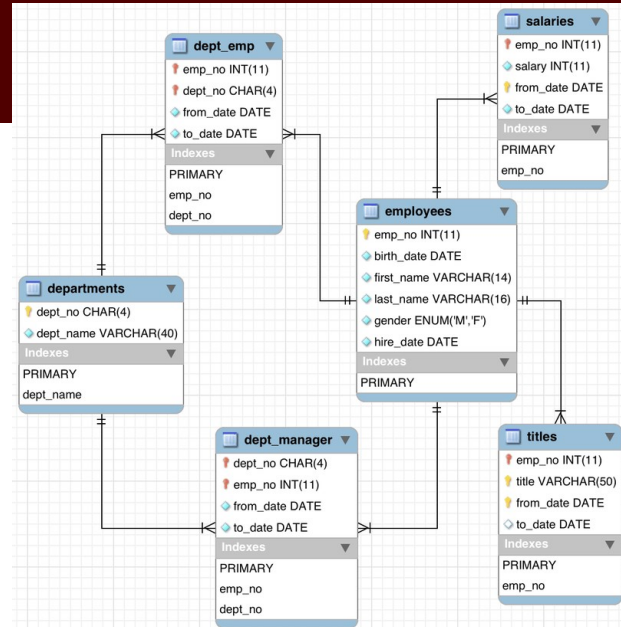
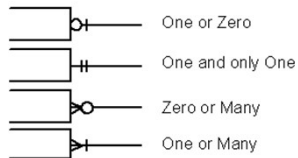
17

17



## Entity Relationship Diagram Crow's Foot Notation

### Summary of Crow's Foot Notation



18

## Outline



- What is a database?
- What is a database management system?
- An Introduction to SQL
  - How to retrieve data from a database
  - How to create a database (optional)

19

## Relational Database Management Systems (DBMS)

- A software that allows the creation of relational databases
- Supports specialized languages (e.g., SQL) for easy retrieval of data from a set of inter-related tables
- Supports easy construction of a Graphical User Interface (GUI) on top of the database
- Allows very large table sizes
- Provides security, fault tolerance, multi-user support, etc.

20

## DBMS

- Different DBMS
  - SQLite
  - MySQL (backend) + php (frontend: web GUI)
  - MS SQL Server, MS Access
  - Oracle
  - DB2 (IBM)
- ODBC: Open DB Connectivity
  - Middleware API to talk to DBs
- Standard Language: SQL
  - Each software has variations to the standard
  - Import/export via standard sql
- Most universal file format
  - Comma separate values: csv
  - Some can handle first line being header

21

## Outline



- What is a database?
- What is a database management system?
- An Introduction to SQL
  - Conditional logic
  - variables
  - How to retrieve data from a database
  - How to create a database (optional)

22

- What Tables?

1	A	B	C	D	E	F	G	H
1	patientID	fname	lname	mi	race	sex	addr	phone
2	101	EDWIN	TILLMAN	B	W	M	1804 WHITE OAK CHURCH RD	979-237-0168
3			JOHNSON					979-741-0546
4			UTRY	J			CH RD	979-857-9563

Patient Unique ID

Reference to a patient

1	2	3	4	5	6	7	8	9	10
1	visitID	patientID	providerID	dtl_dos	dtl_qtr	diag1	diag2	diag3	proc
5	50	86	101	1	11/29/2012	2012QTR4	4019	2720	G0154
7	51	87	101	1	12/7/2012	2012QTR4	4019	2720	G0154
8	52	88	101	1	12/14/2012	2012QTR4	4019	2720	G0163
9	53	140	102	3	1/1/2012	2012QTR1	78830		73130

Visit Unique ID

1	A	B	C	D	E
1	medID	patientID	ahfs_cd	ahfs_desc	rx_qua
14	483612	102	282408	BENZODIAZEPINES (ANXIOLYTIC,SEDATIV/HYP)	
15	483613	102	200404	IRON PREPARATIONS	
16	483614	102	282408	BENZODIAZEPINES (ANXIOLYTIC,SEDATIV/HYP)	

Medication Unique ID

1	A	B	C	D	E
1	providerID	fname	lname	mi	type
2	1	LAURIE	FRY	L	Doctor
3	2	WILLIAM	GAINES	W	Doctor
4	3	DANA	MAGOON	L	Doctor
5	4	ANDREA	BONELLO	L	Nurse
6	5	CHANG	Y	Nurse	
7	6	DUBAY	R	Nurse	
8	7	ROBERT	MACDDE	M	Nurse

Provider Unique ID

23

## Extracting Information from DB

- Download sql\_handout.xlsx (save as)
  - [https://pinformatics.org/info/phpm631/sql\\_handout.xlsx](https://pinformatics.org/info/phpm631/sql_handout.xlsx)
- Look at the excel and answer questions in the exercise
  - You must be logged in at TAMU email ID
  - You have to not be logged in as any other account
  - <https://forms.gle/gJ787gR5JYcXwnx28>

24

24

## SQL - Structured Query Language (The next two weeks)

- Create tables
- Enter values
- Query: ask questions
  - Every statement yields a table of values as output
    - Sometimes there's only one row in the table!
    - semicolon: Don't Forget.
      - ✓ Tells the computer I am done writing my statement

Keyword	parameters
<b>select</b>	lname, fname
<b>from</b>	patients
<b>where</b>	gender='F'
<b>group by</b>	group rows together
<b>order by</b>	lname, fname
<b>;</b>	

25

## SQL: Final Take home Given Tables (hand out)



- Given the real world questions
  - Write the SQL queries
  - Write the expected outcome from the SQL query
  - AND answer the real world question
- Read SQL queries
  - Write the expected outcome from the SQL query
  - Write the real world question the query is answering (in plain English)
  - AND the answer to the question
- Write your own real world question, and the three items above
- Boolean Logic:  $x \ \& \ y$ ,  $x \ | \ y$ ,  $\sim x$
- Extra Credit [+10] Constructive, specific, and concrete course feedback
  - Reasonable effort will get +5 with reasons for more allowed for up to +10
- Questions?

Use following handout slides...

How many pediatric patients  
(age<18)?

Select fname, age  
from patients  
where age<18;

**Brittany Greene, 17**

**Michael Blane, 13**

2

26

26

## Handout Patients Table



Patients

	PatientName	PatientID	DOB	Age	Sex	Height	Weight	Phone	Doctor	Allergy
1	Jane Doe	16577661	3/5/1972	46	F	5'4"	130	512-630-9999	Dr. Alexandria Knight	None
2	John Black	16577680	7/7/1993	25	M	5'9"	180	512-098-7654	Dr.Colby Jay	None
3	Allison Bellame	16577623	11/28/1994	24	F	5'4"	135	512-989-7685	Dr.Mathew Britt	None
4	Chase Hugh	16577650	5/9/1996	22	M	5'8"	182	512-765-4568	Dr. Mathew Britt	Latex
5	Karsen Dixon	16577660	1/28/1959	60	M	5'10"	190	512-453-1324	Dr. Alexandria Knight	None
6	Courtney Jones	16577635	12/15/1987	31	F	5'6"	152	512-398-0137	Dr.Colby Jay	Amoxicillin
7	Ashley Martin	16577645	5/13/1995	23	F	5'5"	136	512-047-8283	Dr.Angela Wright	None
8	Audrey Kramer	16577678	6/24/1990	28	F	5'3"	132	512-367-9020	Dr.Angela Wright	None
9	Brittany Greene	16577690	10/18/2001	17	F	5'7"	143	512-746-5687	Dr.Matew Britt	None
10	Jessica Lange	16577685	9/30/1999	19	F	5'8"	155	512-218-9589	Dr.Alexandria Knight	None
11	Blake Noah	16577601	8/21/2000	18	M	5'11"	186	512-216-4637	Dr.Colby Jay	None
12	Christopher Bell	16577679	10/23/1988	30	M	5'9"	196	512-039-8050	Dr.Colby Jay	Asprin
13	Michael Blane	16577615	2/6/2005	13	M	6'0"	194	512-462-9758	Dr.Angela Wright	Atracurium
14	Scott Peters	16577620	3/15/1965	53	M	5'10"	181	512-369-8564	Dr.Alexandria Knight	None
15	Eric Sanders	16577655	4/30/1975	43	M	5'8"	197	512-384-5038	Dr.Mathew Britt	None

27

27

## Practice Questions



- How many patients does Dr. Colby Jay have?
  - SQL query
  - Query result
  - Answer to question
  
  - Select doctor From patients
  - Where doctor="Dr. Colby Jay"

28

28

## Practice Questions Answers More than one correct answer



- How many patients does Dr. Colby Jay have?
  - SQL query
    - select patientname from patients
    - where doctor="Dr.Colby Jay"
  - Query result
  - Answer to question: 4
- Another correct answer
  - select \* from patients
  - where doctor="Dr.Colby Jay"
  - Query results ?
  - Answer to question?

### PatientName

John Black

Courtney Jones

Blake Noah

Christopher Bell

29

29

## More complicated queries

- May need to (learn next week)
  - Sort
  - Aggregate
  - Calculate new variables using variables given
  - Combine tables
  - Require multiple queries (e.g. views)

30

30

## Take Away I SQL - Structured Query Language

- Every statement yields a table of values as output
  - Sometimes there's only one row in the table!
  - Semi colon is very important to signal to the computer you are done !

Keyword	parameters
<b>select</b>	columns and/or expressions
<b>from</b>	Tables
<b>where</b>	conditions on the rows
<b>group by</b>	group rows together
<b>order by</b>	order the rows
<b>;</b>	

31

## Take Away II

### Boolean Logic: Truth Tables (1=T; 0=F)



- WHERE gender='Male' AND age>18;
- WHERE race='Hispanic';
- WHERE race=Hispanic;
- WHERE race=5;

y	NOT
	~ y
F	T
T	F

x	y	AND	OR
		x & y	x   y
F	F	F	F
F	T	F	T
T	F	F	T
T	T	T	T

32

## Take Away III

### Variable Types



Type	Stored value	Interpreted value	Label Interpreted Value
int	1000001 (65)	65	65 or older
Char/string (ASCII)	1000001 (65)	A	Asian
date	1000001 (65)	1960/3/6 (SAS)	

- 1 0 0 0 0 0 1 =64+1=65
- 64 32 16 8 4 2 1
- Understand variables types? What is the difference below?
  - SELECT \* FROM patients where race=H;
  - SELECT \* FROM patients where race='H';
  - SELECT \* FROM patients where race='h';

33



## Choose Columns

- Choosing a subset of columns is sometimes called a "project" operation
- Display first and last name of all patients
  - SELECT fname, lname FROM patients;
- TRY: Display diagnosis and visitDate for all patients

```
SELECT Column1, Column2  
FROM Table;
```

34

## Display an Entire Table

- Wildcard
- \* : means all columns
- Regular Expression

```
SELECT *  
FROM patients;
```

35

## Choose Rows

- Find Hispanic patients
  - SELECT fname, lname, race
  - FROM patients
  - WHERE race = "H";
- TRY: Find visits before Feb 1, 2012
  - YYYY-MM-DD: standard SQL (e.g. '2012-02-01')

```
SELECT Column1, Column2
FROM Table
WHERE Condition;
```

36

## Choose Rows: conditional statements

- String: WHERE race = "Hispanic";
- String: WHERE fname LIKE "s%";
  - starting with letter s
- Number: WHERE age > 18;
  - Date: Where svc\_dt > '1993-01-01';
  - YYYY-MM-DD: standard SQL
  - Find patients born after Jan. 1, 1993

```
SELECT Column1, Column2
FROM Table
WHERE Condition;
-- variable operator
value;
```

37

## Outline



- What is a database?
- What is a database management system?
- An Introduction to SQL
  - Conditional logic
  - Variables
  - How to retrieve data from a database
  - How to create a database (optional)

38

## Conditional Operators



Operator	Description
=	Equal
<>	Not equal. Note: In some versions of SQL this operator may be written as !=
>	Greater than
<	Less than
>=	Greater than or equal
<=	Less than or equal
BETWEEN	Between an inclusive range
LIKE	Search for a pattern (approximately same)
IN	To specify multiple possible values for a column

39

## Boolean Logic: Truth Tables (1=T; 0=F)

- WHERE gender='Male' AND age>18;
- WHERE race='Hispanic';
- WHERE race=Hispanic;
- WHERE race=5;

y	NOT
	$\sim y$
F	T
T	F

x	y	AND	OR
		$x \& y$	$x   y$
F	F	F	F
F	T	F	T
T	F	F	T
T	T	T	T

40

## Practice

- $\sim(x \& y) \text{ or } (x \text{ or } y)$

x	y	$x \& y$	$\sim(x \& y)$	$x \text{ or } y$	$\sim(x \& y) \text{ or } (x \text{ or } y)$
F	F	F	T	F	T
F	T	F	T	T	T
T	F	F	T	T	T
T	T	T	F	T	T

41

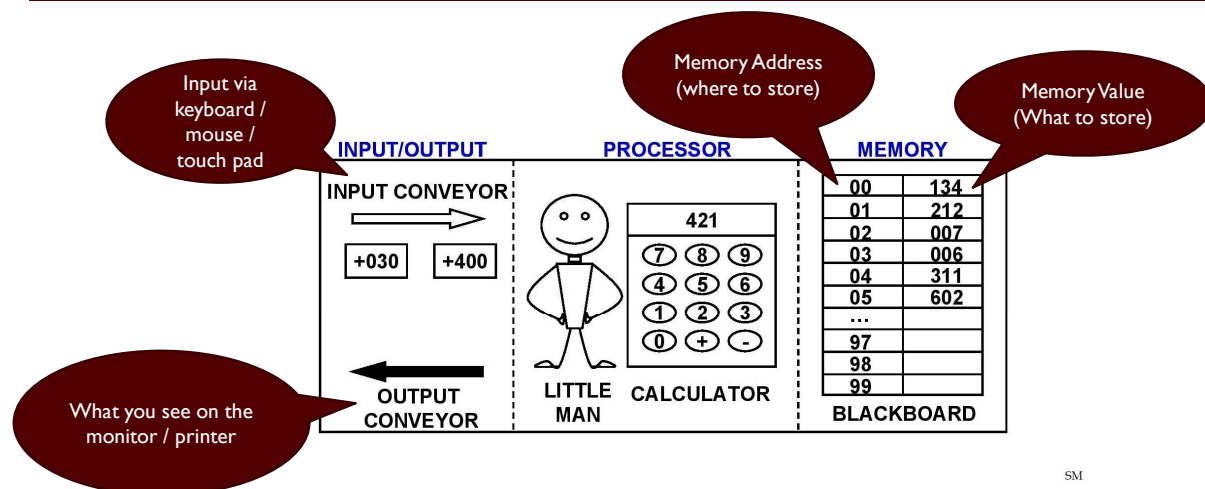
41

## Outline

- What is a database?
- What is a database management system?
- An Introduction to SQL
  - Conditional logic
  - Variables
  - How to retrieve data from a database
  - How to create a database (optional)

42

## A Simplified Computer



43

## What is a Variable?


- A user defined name to represent a piece of memory for storing evaluated value(s). A variable consists of 5 items
  - Name:
    - meaningful human readable name
    - How the user refers to variable
  - Data Type:
    - How to interpret variable for data representation
  - Size:
    - How much storage memory is needed to store data value
    - Can be inferred from data type
  - Value:
    - Actual value associated with variable
    - stored in memory
  - Storage location:
    - Usually hidden from user by the interpreter or compiler
    - How the computer refers to a variable
  - For Our Purposes: Columns
    - Many variables. A columns of variables

44

## Variable

Name	Data Type	Size	Memory Location (hidden from user)	Value
Radius	float32	4 bytes	0x1800F040	3.23
currKey	char	1 byte	0x1800F049	'k'
firstName	string	6 bytes	0x1800B0E0	"morgan"
width	int32	4 bytes	0x1800CCE8	800
type	int8	1 byte	0x1800CCE7	27

45



## ASCII: character encoding

0	<NUL>	32	<SPC>	64	@	96	`	128	À	160	†	192	¿	224	#
1	<SOH>	33	!	65	A	97	a	129	Á	161	°	193	¡	225	·
2	<STX>	34	"	66	B	98	b	130	Â	162	±	194	ª	226	,
3	<ETX>	35	#	67	C	99	c	131	Ã	163	ë	195	»	227	~
4	<EOT>	36	\$	68	D	100	d	132	Ä	164	§	196	¼	228	%
5	<ENQ>	37	%	69	E	101	e	133	Å	165	•	197	½	229	^
6	<SO>	38	&	70	F	102	f	134	Ä	166	¶	198	¾	230	_
7	<BEL>	39	'	71	G	103	g	135	Ö	167	ß	199	¸	231	À
8	<BS>	40	(	72	H	104	h	136	Ø	168	©	200	¹	232	Á
9	<TAB>	41	)	73	I	105	i	137	Ù	169	®	201	º	233	Â
10	<LF>	42	*	74	J	106	j	138	Ú	170	™	202	»	234	Ã
11	<VT>	43	+	75	K	107	k	139	Û	171	·	203	¼	235	Ä
12	<FF>	44	,	76	L	108	l	140	Ü	172	ˆ	204	½	236	Å
13	<CR>	45	-	77	M	109	m	141	Ý	173	¸	205	¾	237	Ä
14	<SO>	46	.	78	N	110	n	142	ÿ	174	Æ	206	¸	238	Ö
15	<SI>	47	/	79	O	111	o	143	è	175	Ø	207	¹	239	Ù
16	<DLE>	48	0	80	P	112	p	144	é	176	Ù	208	º	240	Ú
17	<DC1>	49	1	81	Q	113	q	145	ê	177	Ú	209	»	241	Û
18	<DC2>	50	2	82	R	114	r	146	ë	178	Û	210	¼	242	Ü
19	<DC3>	51	3	83	S	115	s	147	ì	179	≥	211	½	243	Ý
20	<DC4>	52	4	84	T	116	t	148	í	180	¥	212	¾	244	ÿ
21	<NAK>	53	5	85	U	117	u	149	î	181	µ	213	¸	245	À
22	<SYN>	54	6	86	V	118	v	150	ï	182	¶	214	¹	246	Á
23	<ETB>	55	7	87	W	119	w	151	ò	183	¸	215	º	247	Â
24	<CAN>	56	8	88	X	120	x	152	ó	184	ˆ	216	»	248	Ã
25	<EN>	57	9	89	Y	121	y	153	ô	185	π	217	¼	249	Ä
26	<SUB>	58	:	90	Z	122	z	154	õ	186	∏	218	½	250	Å
27	<ESC>	59	;	91	[	123	{	155	ö	187	ª	219	¾	251	Ä
28	<FS>	60	<	92	\	124		156	ó	188	°	220	¸	252	Ö
29	<GS>	61	=	93	]	125	}	157	ü	189	Ω	221	¹	253	Ù
30	<RS>	62	>	94	^	126	~	158	ú	190	æ	222	º	254	Ú
31	<US>	63	?	95	_	127	<DEL>	159	û	191	ø	223	»	255	Û

65 A

66 B

67 C

68 D

48 0


49 1

50 2

51 3

52 4

46



## Variable Types

Type	Stored value	Interpreted value	Label Interpreted Value
int	1000001 (65)	65	65 or older
Char/string (ASCII)	1000001 (65)	A	Asian
date	1000001 (65)	1960/3/6 (SAS)	

- 1 0 0 0 0 0 1 =64+1=65
- 64 32 16 8 4 2 1

47

## Variable Type

- Number
  - Int (long), real (double, float), date time
- String/Character
  - Length matters
- Missing
  - . ‘.’
  - “
  - SAS: .<0

48

## Try on w3schools

- Which is the correct SQL statement below?
- What is the problem with the incorrect SQL statements ?
  - SELECT \* FROM patients where race=H;
  - SELECT \* FROM patients where race='H';
  - SELECT \* FROM patients where race='h';
- You may also try the free trials at
  - <https://academy.vertabelo.com>

49



## Take Away I

### SQL - Structured Query Language



- Every statement yields a table of values as output
  - Sometimes there's only one row in the table!
  - Semi colon is very important to signal to the computer you are done !

Keyword	parameters
<b>select</b>	columns and/or expressions
<b>from</b>	Tables
<b>where</b>	conditions on the rows
<b>group by</b>	group rows together
<b>order by</b>	order the rows
<b>;</b>	

50

## Take Away II

### Boolean Logic: Truth Tables (1=T; 0=F)



- WHERE gender='Male' AND age>18;
- WHERE race='Hispanic';
- WHERE race=Hispanic;
- WHERE race=5;

y	NOT
	$\sim y$
F	T
T	F

x	y	AND	OR
		$x \& y$	$x   y$
F	F	F	F
F	T	F	T
T	F	F	T
T	T	T	T

51

## Take Away III Variable Types



Type	Stored value	Interpreted value	Label Interpreted Value
int	I00000I (65)	65	65 or older
Char/string (ASCII)	I00000I (65)	A	Asian
date	I00000I (65)	I960/3/6 (SAS)	

- 1 0 0 0 0 0 0 1 =64+1=65
- 64 32 16 8 4 2 1

52


## Resources



- <http://www.w3schools.com/sql/default.asp>
- <https://academy.vertabelo.com>
- <https://www.sqlite.org/lang.html>
- SQLite DB Browser: portableapps
  - [http://portableapps.com/apps/development/sqlite\\_database\\_browser\\_portable](http://portableapps.com/apps/development/sqlite_database_browser_portable)

53

---




Break  
Read Assignment 8

54

---

# GROUP PRESENTATION



55

55

## Next week

- Complete Lab 8
- Do assignment 8
  - At least to midpoint
- No additional quiz or readings
  - Two full weeks to focus on learning basics of SQL
- Office Hours
  - Michelle - Thur 3:30-5:30 pm on zoom
  - Dr. Kum - Wed 3-5pm on zoom

56

56

## Assignment 8

- Get your feet wet: Learn SQL
  - W3C online tutorial (individual activity)
  - Do quiz and submit results
- Assignment 8: work on problems
  - Group activity: work together online
  - BUT both must type and submit (no copy/paste)
  - Remember you will have to do this on the final
  - Download excel & kum.db
  - DUE:
    - Mid point EMAIL in one week
    - Sunday before Last day of class
  - No Late Assignment

57