Lecture 4: More Microsoft Access® and Relational Databases

All screen shots from MS Access® 23 January 2002

Goal:

Introduction to the relational database model and intermediate queries (aggregation/grouping functions and multi-table queries).

Outline:

Discussion of databases, introduction to the relational model, and examples of the kinds of queries you might be asked to do in lab exercises.

Main Points:

Discuss data format and data structure Basic Characteristic of relational database model Intermediate queries

- aggregation function
- expression builder
- update database

- join multiple tables.

Databases Review:

Data format, data standards:

- Plain text: ASCII
- Formatted text: MSWord, WordPerfect
- Spreadsheet: *.xls, *wk3
- Database: dBase *.dbf

Difference between choice of data model (category of data class) and possible ways of standardizing it.

One way to see what a database is to see it in its ascii plain text format, common means of exchanging data among software programs. This format is also important because lots of information on the web is available in this most basic standard.

Data Structures and data models

- Logic structure to present real world situation
- How things of interest should be represented
- How things can be relatd to each other

Simplest form of database is flat file (for relatively simple and small amounts of data)

Queries with flat file: Department of Transportation employee survey database.

Importing a type *.dbf file. To use a *.dbf file in MS Access, go to **File** then **New Database**... and with **Blank Database** highlighted, click **OK**. At the *File New Database* window, type in a File name such as **my_dot.mdb** and make sure that the **Save in** location is your crlspace private folder so that the window looks something like this:

File New Data	ibase	? ×
Save in:	I my_crispace 💽 🗈 🖄 📰 📰 🎵	
Mmy_dot.md	b	Create
		Cancel
		Exclusive
File <u>n</u> ame:	my_dot.mdb	
Save as type:	Microsoft Access Databases (*.mdb)	

Creating my_dot.mdb in *File New Database* Window

Then click on the **Create** button. You should now see a my_dot: Database window:

my_dot :	Database				
III Tables	De Queries	E Forms	🖪 Reports	🛛 Macros	🐗 Modules
					<u>Open</u>
					Design
a har					New
					Real Providence
Harrana ana ana ana ana ana ana ana ana a		ininininininininininininininininininin	аналарианыныныныны		

new database window for my_dot

Now go to **File**... **Get External Data...** and then **Import**. In the Import window, you must specify where to look for the *.dbf file you want to import. In this case, go to K:\11.208\data\ and then be sure to specify Files of Type as dBase 4 (*.dbf). Now select **employee.dbf** and then click on the **Import** button. Access now displays a window saying "Successfully imported 'employee.' Close the Import window. In the my_dot: Database window, you will see employee listed under tables.

Review of database documentation: data definitions and survey form. Review basic query design

Use of AND versus OR:

 When you type expressions in more than one Criteria cell, Microsoft Access combines them using either the And or the Or operator. If the expressions are in different cells in the same row, Microsoft Access uses the And operator, which means only the records that meet the criteria in all the cells will be returned. If the expressions are in different rows of the design grid, Microsoft Access uses the Or operator, which means records that meet criteria in any of the cells will be returned.

Aggregate functions (count, distinct); group by clauses

Expressions and calculated expressions (create new field first!)

• First a simple query: how many DOT employees drove alone as primary mode of transportation. First let's look at the data dictionary and survey.

Field: MODE1_TO Table: employee Total: Group By Sort: Count Show: Image: Count Criteria: "1"		Parcel and a second		n al se restante. Na se
Total: Group By Count Sort: Image: Count Image: Count Show: Image: Count Image: Count			MODE1_TO	
Total: Group By Count Sort: Image: Count Image: Count Show: Image: Count Image: Count	Table:	employee	employee	
Show:	Total:	Group By	Count	
and the second	Sort:			
Criteria: "1"	Show:			
	Criteria:	"1"		
			1	

Query of employees who drove alone as primary mode

🖃 Query1 : Select Query 📃 🗖 🗙					
MODE1_TO	CountOfMODE				
1	761				
Record: II 4	1 🕨 🕨 💌 of 1				

Result of above query

 \circ $\;$ Notice that we are looking at the record. So counting the unique survey identifier yields the same result.

🖃 Query1 :	Select Qu	ery		
SUF DOT COL	ployee RVEYNO F_BRANCH JNTY FICE			•
Field:	MODE1_TO)	SURVEYNO	
Table:	employee		employee	
Total:	Group By		Count	
Sort:				
Show:	¥	2		
Criteria:	"1"			
or:				
🗐 Query1	: Select Qu	ery		
	E1_TO	And the second se		
1	AND A CONTRACT OF A		761	
	and the second secon		701	
Record: II		1 🕞 🖬	▶ * of 1	

 Notice, however, that the Total functions can provide calculations on the selected field. For example, we can calculate the average commute time for those who drive alone.

	🗐 Query1	Select Query	
		ployee	
	Table:	MODE1_TO employee Group By I I I I I I I I I I I I I	T_COMMUT
	Query1 : Select		
		AvgOfT_COMMUT 41.1051248357424	
Red	cord: 🚺 🕢	<mark>) <u>→ → →</u> of 1</mark>	

Ē

 $\circ~$ Group By functions also allow increasingly finer levels of selection. We can examine average commute times by DOT Branch...

em	ployee			
*				
SUR	RVEYNO -			
DOT	T_BRANCH			
COL	JNTY			
OFF	-ICE			
ZIP	-ICE			
14 (10 A)	-ICE			
8400 A	-ICE			
ZIP Field:	MODE1_TO	T_COMMUT	DOT_BRANCH	
ZIP Field:	MODE1_TO employee	T_COMMUT employee	DOT_BRANCH employee	
Field: Table: Total:	MODE1_TO employee			
Field: Table: Total: Sort:	MODE1_TO employee Group By	employee Avg	employee Group By	
Field: Table: Total:	MODE1_TO employee	employee	employee	

🗐 Query1 : Select	t Query	
MODE1_TO	AvgOfT_COMMUT	DOT_BRANCH
1	44.1167883211679	FAA
1	36.1777777777778	FHWA
1	41.7857142857143	FRA
1	41.6060606060606	MARAD
1	41.8181818181818	NHTSA
1	33.3636363636364	OIG
1	39.2702702702703	OST
1	40	RSPA
1	25	SLSDC
1	38.75	UMTA
1	41.4224137931034	USCG
Record: 14 4	11 🕨 🕨 💌 of 11	

 \circ $\;$ And we can look further into the average commute times by county residence of particular DOT Branch employees...

.	Query1 :	Select	Query						
-	* SUR DOT COL OFF ZIP	NOVEE RVEYNO I_BRANG JNTY ICE	Э						*
	Field:	MODE	1 TO	т_сом	MUT	DO	T_BRANCH	COUNTY	-
	Table:	employ		employe			ployee	employee	
	Total:	Group	Ву	Avg		Gro	oup By	Group By	
	Sort: Show:				✓		✓		
	Criteria:	"1"							
	or:	and and							
e (Query1 :	the second s	and the second se						
	MODE	1_TO	AvgOfT_CO	MMUT	DOT	BRANCH	COUN	ITY ZIP	
	1			55	FAA			02186	
	1			125	FAA			17325	
Sec.	1			120	FAA			19933	
	1				FAA		Alex	22301	
	1				FAA		Alex	22302	
	1				FAA		Alex	22304	
	1				FAA		Alex	22305	
	1				FAA		Alex	22312	
	1				FAA		Alex	22314	
	1				FAA		Anne A.	20751	
	1				FAA		Anne A.	21061	
	1				FAA		Anne A.	21108	
	1				FAA		Anne A.	21401	
	1			15	FAA		Arling	22202	
550	1			15	FAA		Arling	22204	
	1				FAA		Arling	22206	-
Rei	cord: 🔢	•	2 🕨 🕨	* of 45	5				

 \circ $\;$ And, of course, of Zip codes within counties.... $\;$

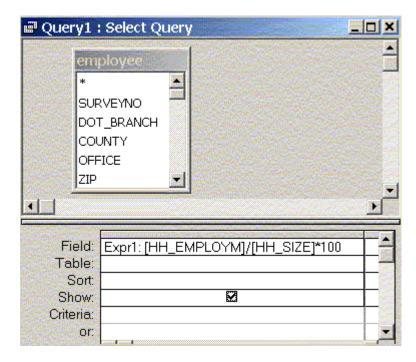
🗐 Query1	: Select	Query					
SUF DO'	ployee RVEYNO T_BRANC JNTY FICE	сн					
Field:	MODE	1 TO	T_COM	MUT	DOT_BRANCH	COUNTY	ZIP
Table:	employ	 /ee	employe		employee	employee	emplo
Total:	Group B	By	Avg		Group By	Group By	Group
Sort: Show:							
Criteria:	"1"						
or:	Terrare Deserved						
I Query1	1 Select	Otterv					
statute of the second se		AvgOfT_CO	MMUT	and the second			
1		44.11678832		and the second se			
1		36.17777777					
1		41.78571428					
1		41.60606060					
1		41.8181818					
1		33.36363636					
1		39.27027027	702703	OST			
1			40	RSPA			
1			25	SLSDC			
1			38.75	UMTA			
1		41.42241379	931034	USCG			
Record: II		11 💽 🖭 🤊	* of 11				

 It may also be useful to calculate expressions using two or more fields. We can number of persons in householod (HH_SIZE) and number of employed persons in household (HH_EMPLOYM). But what is percentage of persons employed?

	THE REPORT OF			
	Select Query			(<u> </u>
Field: Table:	HH_SIZE employee	HH_EMPLOYM employee	HH_EMPLOYM/HH_SIZE*100	
Sort: Show: Criteria: or:				
■ Query1 :	Select Query			- - >
	ployee			- - -
Table:	HH_SIZE employee	HH_EMPLOYM employee	Expr1: [HH_EMPLOYM]/[HH_SIZE]*100	
Sort: Show: Criteria: or:				
		Sector Contraction		

🖃 Query1 : Sel	ect Qu	ery	
HH_SIZ	E I	HH_EMPLOYM	Expr1
	3	2	66.666666667
	5	4	80
	2	2	100
	3	2	66.666666667
	2	2	100
	2	1	50
	2	1	50
	1	1	100
	1	1	100
	5	2	40
	1	1	100
	2	2	100
	2	1	50
	6	1	16.666666667
	4	1	25
Record: H		1 ▶ ▶ ▶ ▶ ★ of 4	/35

• Of course, you only need the expression:



 \circ $\;$ However this does not allow us to use Group By functions. For that, we add Sums:

IRVEYNO OT_BRANCH OUNTY FFICE IP			
	ATTICATION OF A DESCRIPTION OF A DESCRIP	CONTRACTOR AND INCOME AND INCOME AND INCOME AND	CONTRACTOR AND

Service and the service of the servi			
Total:	Group By	Group By	Expression
Sort:			
Show:		\checkmark	
Criteria:			
or			
Sector Contractor	1987 S 1997		

🖃 Query1 : Select Qu	Jery		×
DOT_BRANCH	COUNTY	Expr1	•
FAA		70	
FAA	Alex	79.775280899	
FAA	Anne A.	63.829787234	
FAA	Arling	73.988439306	
FAA	Balt Co.	100	
FAA	Baltimore	64.285714286	
FAA	Berkeley W	64.285714286	
FAA	Calvert	67.647058824	
FAA	Caroline	33.333333333	
FAA	Charles	67.647058824	
FAA	Fairfax	64.748201439	
FAA	Fauquier	70	
FAA	Fredrick	21.428571429	
FAA	Harford	100	
FAA	Howard	64.406779661	
FAA	Jefferson	50	
FAA	Loudon	67.272727273	-
Record: 14	1 • • • • • • of 2	201	1

ADDING A FIELD AND CALCULATED EXPRESSION AND SAVING TO NEW TABLE:

• It also may be helpful to create a new field and then add data to it based on a calculated expression. At the main database window, go to **Tables** tab and then

click on **Design.** Scroll down the end of the employee table and click in the last empty row. Enter a field name (no spaces or odd characters). Then click inside Data Type cell, and a pull down menu appears. Select "Number." Also, in the gray area below under the "General" tab, select "Double." You window should like something like this:

	Fable		
Field	d Name	Data Type)escription 🔺
ANT_MODE1		Text	
NOON_TRIPS		Number	
NOON_PURPS	;	Text	
NOON_MODE	NOON_MODE		
HH_INCOME		Text	
CIV_MIL		Text	
TENURE		Text	
HH_VEHICLS		Number	
HH_SIZE		Number	
HH_EMPLOYM		Number	
CHILD_TR		Text	
D_TRANSIT		Text	815
BIKE_JOG		Text	
PUB_DISC		Text	
PUB_FLEX		Text	
PUB_DAYCAR		Text	
PUB_EMGNCY		Text	
PUB_PKRATE		Text	
PUB_INFO		Text	
CP_ASSIST		Text	
CP_FLEX		Text	
CP_DAYCAR		Text	
CP_EMGNCY		Text	
CP_PKRATE		Text	
CP_PARTTIM		Text	
EXP_FACTOR		Number	
My_New_Field	11	Number	
		Text	-
	Field P	ro Memo	
		Number	
General Look	up	Date/Time	
Field Size	Double	Currency	and a second
Format		AutoNumber	
Decimal Places	Auto	- Yes/No	
	Auw	OLE Object	
Input Mask		Hyperlink	
Caption	_	Lookup Wizard	
Default Value	O		
Validation Rule			
Validation Text	Contractor of the second se		
Required	No		
Indexed			

- \circ $\;$ Then close the table and save it. At this point, your table should now have a new column with no data in it.
- \circ $% \left(Let's \right)$ Let's create a new query with a calculated expression and field to include in new table:

emplo emplo	yee			
ZIP				
DEPT				
MODE1	_то			
MODE2	_то			
T_COM	1 10000-0410			
D_COM	IN ALL IT MARKED AND A SUB-			
ID_CON				
Field: M		HH_EMPLOYM]/[HH_SIZE]*100	SURVEYNO	DOT_BRANCH
Field: M		HLEMPLOYM]/[HH_SIZE]*100	SURVEYNO employee	DOT_BRANCH employee
Field: M Table: Sort:			employee •	employee
Field: M Table: Sort: Show:			employee	
Field: M Table: Sort:			employee •	employee

Query1 : Select Query								
	My_New_Field	SURVEYNO	DOT_BRANCH	OFFICE	ZIP			
	66.666666667	1	FAA	AAP	21045	28		
	80	2	FAA	APR	21043	28		
	100	3	FAA	AEE	21044	28		
	66.666666667	4	FAA	AAD	21044	28		
	100	5	FAA	AAP	21044	28		
	50	6	FAA	AFS	21044	28		
	50	7	FAA	ASC	21045	29		
	100	8	FAA	AEE	21045	28		
Re	cord: 14 4	8 > > > of 4	1735					

			- -
0	At the Toolbar, c	lick on the Query select bu	Itton Itto and from the Query
		🗊 🔹 🚦 😋 🗵 🛛 All	
		🗊 Select Query	
		🔲 Crossta <u>b</u> Query 👘	
	Je I	🖬 Make-Table Query	
		🥂 Update Query	
		Append Query	
	options menu	XI Delete Query	

select **Make-Table Query**. At the **Make Table** window, we enter a new table name and choose to leave it in the current database. After clicking on the Query

now button . Microsoft Access prompts us with a window to confirm that we will create a new table. We click **Yes**. We now have a new table that we can perform queries on.

Relational Databases and Joins

More complex model is relational database where there are multiple tables that can be joined by common identifier (just as social security number can be used).

Relational joins link different tables with different types of data. This generates information otherwise not possible. For example, you will join tables of building owners, parcels, and fires and determine which owner sustained the most property damage due to fire.

Example of relational joins using tables: parcels, fires, tax, owners

Create database "Parcels_plus_other"

Get External Data... and then Import the following table: K:\11.208\Data\parcels.dbf and then repeat for the tables FIRES.DBF TAX.DBF and OWNERS.DBF. Data Dictionary for PARCELS, FIRES, TAX, OWNERS

PARCEL	S (Parcel information)	FIRE	S (Fire incidents)	TAX (Ta	x and value information	OWNE	RS (Owner informatio
id	unique parcel identifier	ld	Unique parcel identifier	id	unique parcel identifier	ownernum	owner identifier
wpb	ward/precinct/block	Wpb	Ward/precinct/block	wpb	ward/precinct/block	oname	owner name
parcel	parcel number in block	Parcel	Parcel number in block	parcel	parcel number in block	address	owner's street address
add1	street number	Fdate	Date of fire	prptype	property type	city	owner's city of residen
add2	street name	Ignfactor	Ignition factor	landval	value of land	state	owner's state of reside
zip	zip code	Estloss	Estimated loss	bidval	value of building	zip	owner's zip code of res
sqft	square footage			tax	property tax		
onum	owner identifier						
landuse	land use type						

Join tables with unique record identifiers: PARCELS.ID to FIRES.ID to TAX.ID. Then OWNERS.OWNERNUM to PARCELS.ONUM.

The same kinds of queries we ran on flat file we can run on joined tables. (Group by, Order by)

NOTE: Use of Criteria in joined tables:

• If your query includes linked tables, the values you specify in criteria on fields from the linked tables are *case-sensitive*. They must match the case of the values in the underlying table.

Relational joins can be helpful in other ways. Another example of adding a table and joining is to make analysis easier by providing lookup table to translate codes:

Open database myemployee.

Use **Get External Data**... and then **Import** and in **K:\11.208\Data** select dotmode1.dbf and open to see it's a simple look up table:

<pre>dotmode1 : Tab MODE1_COD</pre>		
1	drive alone	
10	walk	
11	dropped off	
12	other	
2	carpool	
3	vanpool	
4	metrobus	
5	metrorail	
6	commuterrail	
7	commuterbus	
8	motorcycle	
9	bicycle	
*		
Record: II I	7 > > > of 12	

Table contents of dotmode1

Solor *	Select Query	OFFICE	
•			
Field:	MODE1_TO	MODE_1	<u>+</u>
Table:	employee	dotmode1	
Total:	Group By	Group By	
Sort:			
Show:			
Criteria:			
OI.			_

Employee and dotmode1 tables JOINED on MODE1_CODE to MODE1_TO

	📰 Query1 : Select Query 📃 🗖 🗙					
	MODE1_TO	MODE_1				
115.11 115.11	1	drive alone				
	10	walk				
	11	dropped off				
	12	other				
10.50	2	carpool	矖			
	3	vanpool				
	4	metrobus				
1000	5	metrorail				
	6	commuterrail				
	7	commuterbus				
	8	motorcycle				
	9	bicycle				
Re	cord: 14 4	12 🔀 🖬 😕	c			

Result of above query