# 11: Data Manipulation



\*\* Denotes the fundamental relational algebra operators

### an introduction to relational algebra



#### ∩ intersection

**T** projection



Uunion

★ natural join

### which colleges offer courses?

### π (Co\_college) (COURSE)

Co\_college

Arts and Sciences Business Education Engineering what is the name and college of each department? π (Dpt\_name, Dpt\_college) (DEPARTMENT) Dpt name Dpt college Economics Arts and Sciences QA/QM Business Economics Education Engineering Mathematics IS Business Philosophy Arts and Sciences

### set theoretic operators



# pg. 503

#### RELATION R

902 A

#### fall quarter classes

#### RELATION S

Se_section#	Se_qtr	Se_year	Se_time	Se_maxst	Se_room	Se_co_course#	Se_pr_profid
902	A	2006	H1700	25	Lindner 108	22IS270	SK85977
101	S	2006	т1045	29	Lindner 110	2215330	SK85977
102	S	2006	H1045	29	Lindner 110	2215330	CC49234
101	A	2007	н1700	29	Lindner 108	2215330	SK85977

Se\_section# Se\_qtr Se\_year Se\_time Se\_maxst Se\_room Se\_co\_course# Se\_pr\_profid 

901 A 2006 W1800 35 Rhodes 611 22IS270

 101 A
 2007 H1700
 29 Lindner 108
 22IS330

 101 A
 2007 T1015
 25
 22QA375

25 Lindner 108 22IS270

Baldwin 437 20ECES212

SK85977

SK85977

SK85977 HT54347

RR79345

#### Relational Algebra Syntax and Result: RUS

2006 H1700

101 A 2007 W1800

Se_section#	Se_qtr	Se_year	Se_time	Se_maxst	Se_room	Se_co_course#	Se_pr_profid
101	A	2007	H1700	29	Lindner 108	2215330	SK85977
101	A	2007	т1015	25		22QA375	HT54347
101	A	2007	W1800		Baldwin 437	20ECES212	RR79345
101	S	2006	т1045	29	Lindner 110	2215330	SK85977
102	S	2006	H1045	29	Lindner 110	2215330	CC49234
901	A	2006	W1800	35	Rhodes 611	22IS270	SK85977
902	А	2006	H1700	25	Lindner 108	22IS270	SK85977

#### Relational Algebra Syntax and Result: $R \cap S$

Se\_section# Se\_gtr Se\_year Se\_time Se\_maxst Se\_room Se\_co\_course# Se\_pr\_profid 101 A 2007 H1700 29 Lindner 108 22IS330 SK85977 25 Lindner 108 902 A 2006 H1700 22IS270 SK85977

#### classes in lindner

# pg. 503

#### **RELATION R**

Se_section#	Se_qtr	Se_year	Se_time	Se_maxst	Se_room	Se_co_course#	Se_pr_profid
902	A	2006	H1700	25	Lindner 108	22IS270	SK85977
901	A	2006	W1800	35	Rhodes 611	22IS270	SK85977
101	A	2007	H1700	29	Lindner 108	221S330	SK85977
101	A	2007	т1015	25		22QA375	HT54347
101	A	2007	W1800		Baldwin 437	20ECES212	RR79345

#### **RELATION S**

Se_section#	Se_qtr	Se_year	Se_time	Se_maxst	Se_room		Se_co_course#	Se_pr_profid
902	A	2006	н1700	25	Lindner	108	22IS270	SK85977
101	S	2006	т1045	29	Lindner	110	2215330	SK85977
102	S	2006	H1045	29	Lindner	110	2215330	CC49234
101	А	2007	H1700	29	Lindner	108	221S330	SK85977

#### Relational Algebra Syntax and Result: R - S

Se_section#	Se_qtr	Se_year	Se_time	Se_maxst	Se_room	Se_co_course#	Se_pr_profid
101	А	2007	т1015	25		22QA375	HT54347
101	А	2007	W1800		Baldwin 437	20ECES212	RR79345
901	А	2006	W1800	35	Rhodes 611	22IS270	SK85977
101	S	2006	т1045	29	Lindner 110	2215330	SK85977
102	S	2006	H1045	29	Lindner 110	22IS330	CC49234
901	A	2006	W1800	35	Rhodes 611	22IS270	SK85977
902	А	2006	н1700	25	Lindner 108	22IS270	SK85977

### let's add some more..

#### cartesian product

join operators

divide

### cartesian product - pg. 501



Figure 11.4 Cartesian Product operation

# relations derived from course and department

C_name C_course# C_cr	edit C_d_dcode
Intro to Economics 15ECON112 U	1
, Operations Research 22QA375 U	3
Intro to Economics 18ECON123 U	4
COURSES Supply Chain Analysis 22QA411 U	3
Principles of IS 22IS270 G	7
Programming in C++ 20ECES212 G	6
Optimization 22QA888 G	3
Financial Accounting 18ACCT801 G	4
Database Concepts 22IS330 U	7
Database Principles 22IS832 G	7
Systems Analysis 22IS430 G	7

D Relation

	D_name	D_dcode	D_college
/ departments	Economics QA/QM Economics Mathematics IS Philosophy	1 3 4 6 7 9	Arts and Sciences Business Education Engineering Business Arts and Sciences

### cartesian product - pg. 501 C X D ...first 18 rows below

D_name	D_dcode	D_college	C_name	C_course#	C_credit	C_d_dcode
				4.5-20114.0		
Economics	1	Arts and Sciences	Intro to Economics	15ECON112	U	1
QA/QM	3	Business	Intro to Economics	15ECON112	U	1
Economics	4	Education	Intro to Economics	15ECON112	U	1
Mathematics	6	Engineering	Intro to Economics	15ECON112	U	1
IS	7	Business	Intro to Economics	15ECON112	U	1
Philosophy	9	Arts and Sciences	Intro to Economics	15ECON112	U	1
Economics	1	Arts and Sciences	Operations Research	22QA375	U	3
QA/QM	3	Business	Operations Research	22QA375	U	3
Economics	4	Education	Operations Research	22QA375	U	3
Mathematics	6	Engineering	Operations Research	22QA375	U	3
IS	7	Business	Operations Research	22QA375	U	3
Philosophy	9	Arts and Sciences	Operations Research	22QA375	U	3
Economics	1	Arts and Sciences	Intro to Economics	18ECON123	U	4
QA/QM	3	Business	Intro to Economics	18ECON123	U	4
Economics	4	Education	Intro to Economics	18ECON123	U	4
Mathematics	6	Engineering	Intro to Economics	18ECON123	U	4
IS	7	Business	Intro to Economics	18ECON123	U	4
Philosophy	9	Arts and Sciences	Intro to Economics	18ECON123	U	4

### join operators

#### equijoin

natural join

theta join

outer joins

### equijoin



#### duplication of the joining attributes

# natural join



#### no duplication of joining attributes

# theta join



inequality (not common in practical applications)

### outer join

#### previous three joins were inner joins (only matching tuples were included in result)

outer joins keep the tuples in one or both relations regardless of whether they match

# left outer join

#### D ]X| D\_dcode = C\_d\_dcode C

D_name	D_dcode	D_college	C_name	C_course#	C_credit	C_d_dcode
Economics	1	Arts and Sciences	Intro to Economics	15ECON112	U	1
QA/QM	3	Business	Operations Research	22QA375	U	3
Economics	4	Education	Intro to Economics	18ECON123	U	4
QA/QM	3	Business	Supply Chain Analysis	22QA411	U	3
IS	7	Business	Principles of IS	2215270	G	7
Mathematics	6	Engineering	Programming in C++	20ECES212	G	6
QA/QM	3	Business	Optimization	22QA888	G	3
Economics	4	Education	Financial Accounting	18ACCT801	G	4
IS	7	Business	Database Concepts	2215330	U	7
IS	7	Business	Database Principles	2215832	G	7
IS	7	Business	Systems Analysis	2215430	G	7
Philosophy	9	Arts and Sciences				

if no matching tuple in right (second)relation, attributes in left (first) relation are "padded" with null values

### right outer join

C Relation			
C_name	C_course#	C_credit	C_d_dcode
Intro to Economics	15ECON112	U	1
Operations Research	22QA375	U	3
Intro to Economics	18ECON123	U	4
Supply Chain Analysis	22QA411	U	3
Principles of IS	22IS270	G	7
Programming in C++	20ECES212	G	6
Optimization	22QA888	G	3
Financial Accounting	18ACCT801	G	4
Database Concepts	22IS330	U	7
Database Principles	22IS832	G	7
Systems Analysis	22IS430	G	7

#### SS Relation

Ss_section#	Ss_qtr	Ss_year	Ss_c_course#
101	A	2007	22QA375
901	A	2006	22IS270
902	A	2006	22IS270
101	S	2006	22IS330
102	S	2006	22IS330
701	W	2007	22IS832
101	A	2007	20ECES212
101	U	2007	22QA375
101	A	2007	22IS330
101	S	2007	22QA375
101	W	2007	22QA375

# right outer join

#### SS |X[ Ss\_c\_course# = C\_course# C

Ss_section#	Ss_qtr	Ss_year	Ss_c_course#	C_name	C_course#	C_credit	C_d_dcode
101	A	2007	22QA375	Operations Research	22QA375	U	3
901	A	2006	22IS270	Principles of IS	22IS270	G	7
902	A	2006	22IS270	Principles of IS	22IS270	G	7
101	S	2006	221S330	Database Concepts	22IS330	U	7
102	S	2006	22IS330	Database Concepts	2215330	U	7
701	W	2007	22IS832	Database Principles	2215832	G	7
101	A	2007	20ECES212	Programming in C++	20ECES212	G	6
101	U	2007	22QA375	Operations Research	22QA375	U	3
101	A	2007	22IS330	Database Concepts	2215330	U	7
101	S	2007	22QA375	Operations Research	22QA375	U	3
101	W	2007	22QA375	Operations Research	22QA375	U	3
				Financial Accounting	18ACCT801	G	4
				Careers Colloquium	06US100	U	
				Systems Analysis	2215430	G	7
				Intro to Economics	18ECON123	U	4
				Supply Chain Analysis	22QA411	U	3
				Intro to Economics	15ECON112	U	1
				Optimization	22QA888	G	3

# full outer join

C_name	C_course#	C_credit	C_d_dcode
Intro to Economics	15ECON112	U	1
Operations Research	22QA375	U	3
Intro to Economics	18ECON123	U	4
Supply Chain Analysis	22QA411	U	3
Principles of IS	22IS270	G	7
Programming in C++	20ECES212	G	6
Optimization	22QA888	G	3
Financial Accounting	18ACCT801	G	4
Database Concepts	22IS330	U	7
Database Principles	22IS832	G	7
Systems Analysis	22IS430	G	7
Careers Colloquium	06US100	U	

# full outer join

#### D ]X[ D\_dcode = C\_d\_dcode C

D_name	D_dcode	D_college	C_name	C_course#	C_credit	C_d_dcode
Economics	1	Arts and Sciences	Intro to Economics	15ECON112	U	1
QA/QM	3	Business	Operations Research	22QA375	U	3
Economics	4	Education	Intro to Economics	18ECON123	U	4
QA/QM	3	Business	Supply Chain Analysis	22QA411	U	3
IS	7	Business	Principles of IS	22IS270	G	7
Mathematics	6	Engineering	Programming in C++	20ECES212	G	6
QA/QM	3	Business	Optimization	22QA888	G	3
Economics	4	Education	Financial Accounting	18ACCT801	G	4
IS	7	Business	Database Concepts	2215330	U	7
IS	7	Business	Database Principles	2215832	G	7
IS	7	Business	Systems Analysis	2215430	G	7
Philosophy	9	Arts and Sciences				
			Careers Colloquium	06US100	U	

# querying using SQL

# SELECT <column list> FROM WHERE <condition>

# additional query options

### GROUP BY <group\_by expression> HAVING <group\_condition> ORDER BY <column\_name(s)>

### course relation

Co_name	Co_course#	Co_credit	Co_college	Co_hrs	Co_dpt_code
Intro to Economics	15ECUNI12	U	Arts and Sciences	3	T
Operations Research	22QA375	U	Business	2	3
Intro to Economics	18ECON123	U	Education	4	4
Supply Chain Analysis	22QA411	U	Business	3	3
Principles of IS	22IS270	G	Business	3	7
Programming in C++	20ECES212	G	Engineering	3	6
Optimization	22QA888	G	Business	3	3
Financial Accounting	18ACCT801	G	Education	3	4
Database Concepts	22IS330	U	Business	4	7
Database Principles	2215832	G	Business	3	7
Systems Analysis	22IS430	G	Business	3	7

### section relation

Se_section# Se_qtr	Se_year Se_time Se	_maxst Se_room	Se_co_course#	Se_pr_profid
101 A	2007 T1015	25	22QA375	HT54347
901 A	2006 W1800	35 Rhodes 611	22IS270	SK85977
902 A	2006 H1700	25 Lindner 108	22IS270	SK85977
101 S	2006 T1045	29 Lindner 110	22IS330	SK85977
102 S	2006 H1045	29 Lindner 110	22IS330	CC49234
701 W	2007 M1000	33 Braunstien 2	211 22IS832	CC49234
101 A	2007 <b>W1</b> 800	Baldwin 437	20ECES212	RR79345
101 U	2007 T1015	33	22QA375	HT54347
101 A	2007 н1700	29 Lindner 108	22IS330	SK85977
101 S	2007 T1015	30	22QA375	HT54347
101 W	2007 T1015	20	22QA375	HT54347

### which courses are three-hour courses?

SELECT COURSE.CO\_NAME, COURSE.CO\_COURSE#, COURSE.CO\_CREDIT, COURSE.CO\_COLLEGE, COURSE.CO\_HRS, COURSE.CO\_DPT\_DCODE FROM COURSE WHERE COURSE.CO HRS = 3;

SELECT \* FROM COURSE WHERE COURSE.CO HRS = 3;

Co_name	Co_course#	Co_credit	Co_college	$Co_hrs$	Co_dpt_dcode
	1 5 5 6 0 1 1 0				
Intro to Economics	15ECON112	U	Arts and Sciences	3	L
Supply Chain Analysis	22QA411	U	Business	3	3
Principles of IS	221S270	G	Business	3	7
Programming in C++	20ECES212	G	Engineering	3	6
Optimization	22QA888	G	Business	3	3
Financial Accounting	18ACCT801	G	Education	3	4
Database Principles	2215832	G	Business	3	7
Systems Analysis	2215430	G	Business	3	7

# which courses offered by department 7 are three-hour courses?

```
SELECT *
FROM COURSE
WHERE COURSE.CO_DPT_DCODE = 7 AND
COURSE.CO_HRS = 3;
```

Co_name	Co_course	# Co_cre	dit Co_college	Co_hrs Co_dpt	_dcode
Principles of IS	22IS270	 G	Business	3	 7
Database Principles	22IS832	G	Business	3	7
Systems Analysis	<b>22IS430</b>	G	Business	3	7

which sections have a maximum number of students greater than 30 or are offered in Linder 110?

SELECT \*
FROM SECTION
WHERE SECTION.SE\_MAXST > 30 OR SECTION.SE\_ROOM =
'Lindner 110';

Se_section# Se_qtr	Se_year Se_time Se_m	axst Se_room	Se_co_course#	Se_pr_profid
901 A	2006 W1800	35 Rhodes 611	221S270	SK85977
101 S	2006 T1045	29 Lindner 110	22IS330	SK85977
102 S	2006 H1045	29 Lindner 110	22IS330	CC49234
701 W	2007 M1000	33 Braunstien 2	11 22IS832	CC49234
101 U	2007 T1015	33	22QA375	HT54347

### AND and OR

#### $\triangle$ AND is evaluated first

WHERE COURSE.CO\_CREDIT = 'U' and COURSE.CO\_COLLEGE = 'Business' or COURSE.CO\_COLLEGE = 'Engineering'

(undergraduate business students) OR (engineering students)

# IN and NOT IN

# comparison operators (is it in a set of values?)

SELECT \*
FROM COURSE
WHERE COURSE.CO\_CREDIT = 'U' AND
(COURSE.CO\_COLLEGE = 'Business' OR COURSE.CO\_COLLEGE =
'Engineering');

SELECT \*
FROM COURSE
WHERE COURSE.CO\_CREDIT = 'U' AND COURSE.CO\_COLLEGE IN
('Business', 'Engineering');

CO_NAME	CO_COURSE#	CO_CREDIT	CO_COLLEGE	CO_HRS	CO_DPT_DCODE
Operations Research	22QA375	U	Business	2	3
Supply Chain Analysis	2204411	υ	Business	3	3
Database Concepts	2215330	U	Business	4	7

# IN and NOT IN

SELECT \*
FROM COURSE
WHERE COURSE.CO\_CREDIT = 'U'
AND COURSE.CO\_COLLEGE NOT IN ('Business', 'Engineering');

Co_name	Co_course	Co_credit	Co_college	Co_hrs	Co_dpt_dcode
Intro to Econom	ics 15ECON112	U	Arts and Sciences	3	1
Intro to Econom	ics 18ECON123	U	Education	4	4

#### which professors earn more than \$6,000?

#### SELECT \* FROM PROFESSOR WHERE PROFESSOR.PR SALARY/12 > 6000;

#### also: TRUNC, ROUND

PR_NAME	PR_EMPID	PR_PHINE	PR_OFFICE	PR_BIRTHDATE	PR_DATEHIRED	PR_DPT_DCODE	PR_SALARY
Mike Faraday	FM49276	5235568492	249 McMicken	26-AUG-60	01-MAY-96	1	92000
Chelsea Bush	BC65437	5235567777	227 Lindner	03-SEP-46	01-MAY-93	3	77000
Tony Hopkins	HT54347	5235569977	324 Lindher	24-NOV-49	20-JAN-97	3	77000
Alan Brodie	BA54325	5235569876	238 Lindher	14-JAN-44	16 <b>-MAY-</b> 00	3	76000
Marie Curie	CM65436	5235569899	331 Dyer	29-FEB-72	22-0CT-99	4	99000
John Nicholson	NJ43728	5235569999	324 Dyer	01-MAY-66	22-JUN-03	4	99000

#### use of AS in column aliases

SELECT PROFESSOR.PR\_NAME, PROFESSOR.PR\_SALARY/12 AS "Monthly Salary" FROM PROFESSOR WHERE PROFESSOR.PR SALARY/12 > 6000;

PR_NAME	Monthly Salary
Mike Faraday	7666.66667
Chelsea Bush	6416.66667
Tony Hopkins	6416.66667
Alan Brodie	6333.33333
Marie Curie	8250
John Nicholson	8250

### use of BETWEEN and NOT BETWEEN

```
SELECT PROFESSOR.PR_NAME,
PROFESSOR.PR_SALARY/12 AS "Monthly Salary"
FROM PROFESSOR
WHERE PROFESSOR.PR_SALARY/12
BETWEEN 6000 AND 7000;
```

#### note: BETWEEN is inclusive

PR_NAME	Monthly Salary
Chelsea Bush	6416.66667
Tony Hopkins	6416.66667
Alan Brodie	6333.33333

#### which colleges offer courses?

### SELECT COURSE.CO\_COLLEGE FROM COURSE;

CO\_COLLEGE

Arts and Sciences Business Education Business Business Engineering Business Education Business Business

\_\_\_\_\_

### which colleges offer courses? (DISTINCT)

### SELECT DISTINCT COURSE.CO\_COLLEGE FROM COURSE;

CO\_COLLEGE

Arts and Sciences Business Education Engineering

\_\_\_\_\_

# what is the age (in years) of each professor in department 3 when hired?

SELECT PROFESSOR.PR\_NAME, TRUNC((PROFESSOR.PR\_DATEHIRED -PROFESSOR.PR\_BIRTHDATE)/365.25,0) "Age When Hired" FROM PROFESSOR WHERE PROFESSOR.PR DPT DCODE = 3;

365.25 accounts for leap years ,0 indicates the decimals AS is optional (not included here)

Pr_name	Age	When	Hired
Chelsea Bush			46
Tony Hopkins			47
Alan Brodie			56
Jessica Simpson			40
Laura Jackson			26