Database Systems Concepts and Design

CSC201S2/G2

Chapter 2: Introduction to Relational Models

Outline

- Structure of Relational Databases
- Database Schema
- Keys
- Schema Diagrams
- Relational Query Languages
- The Relational Algebra

Example

attributes
(or columns)

ID	*name*	*dept_name*	*salary*
10101	Srinivasan	Comp. Sci.	65000
12121	Wu	Finance	90000
15151	Mozart	Music	40000
22222	Einstein	Physics	95000
32343	El Said	History	60000
33456	Gold	Physics	87000
45565	Katz	Comp. Sci.	75000
58583	Califieri	History	62000
76543	Singh	Finance	80000
76766	Crick	Biology	72000
83821	Brandt	Comp. Sci.	92000
98345	Kim	Elec. Eng.	80000

tuples
(or rows)

Relation Schema and Instance

- A_1, A_2, …, A_n are attributes
- R = (A_1, A_2, …, A_n) is a relation schema

  Eg: instructor = (ID, name, dept_name, salary)

- A relation instance r defined over schema R is denoted by r (R).
- The current values a relation are specified by a table
- An element t of relation r is called a **tuple** and is represented by a **row** in a table

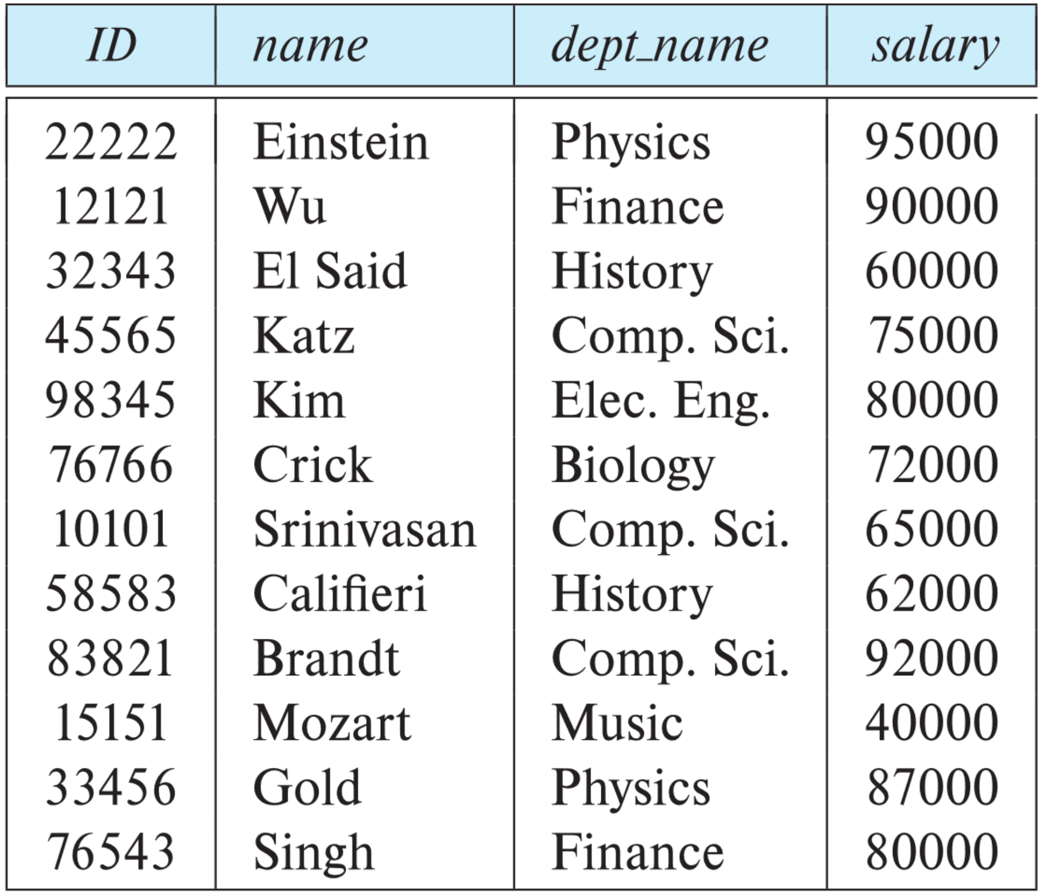
Attributes

- The set of allowed values for each attribute is called the domain of the attribute
- Attribute values are (normally) required to be atomic; that is, indivisible
- The special value **null** is a member of every domain

Relations are Unordered

- Order of tuples is irrelevant (tuples may be stored in an arbitrary order)

Eg: instructor relation with unordered tuples



ID	*name*	*dept_name*	*salary*
22222	Einstein	Physics	95000
12121	Wu	Finance	90000
32343	El Said	History	60000
45565	Katz	Comp. Sci.	75000
98345	Kim	Elec. Eng.	80000
76766	Crick	Biology	72000
10101	Srinivasan	Comp. Sci.	65000
58583	Califieri	History	62000
83821	Brandt	Comp. Sci.	92000
15151	Mozart	Music	40000
33456	Gold	Physics	87000
76543	Singh	Finance	80000

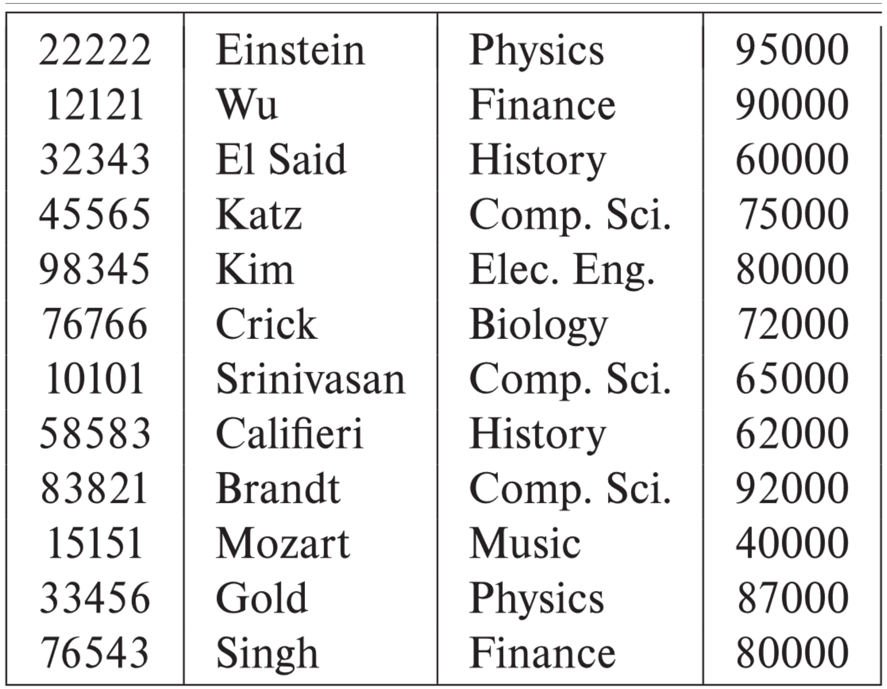
Database Schema

Database schema: is the logical structure of the database.

Database instance: is a snapshot of the data in the database at a given instant in time.

Eg: schema: instructor (ID, name, dept_name, salary)

Instance:



22222	Einstein	Physics	95000
12121	Wu	Finance	90000
32343	El Said	History	60000
45565	Katz	Comp. Sci.	75000
98345	Kim	Elec. Eng.	80000
76766	Crick	Biology	72000
10101	Srinivasan	Comp. Sci.	65000
58583	Califieri	History	62000
83821	Brandt	Comp. Sci.	92000
15151	Mozart	Music	40000
33456	Gold	Physics	87000
76543	Singh	Finance	80000

Keys

- Essential elements of any relational database.
- Identifies each tuple in a relation uniquely.
- Establish the relationship among the tables in a schema.

DBMS Keys

1. Super Key
2. Candidate Key
3. Primary Key
4. Foreign Key
5. Alternate Key
6. Composite Key

Super Key

- A single or a set of attributes can uniquely identify all attributes
- Can contains one or more than one attributes.

Employee
EmployeeID
EmployeeName
SSN
DeptID
DOB

Example

Super keys:

Emp_ID	Emp_Number	Emp_Name
E01	226	Steve
E02	227	Ajeet
E03	228	Chaitanya
E04	229	Robert

{Emp_ID}

{Emp_Number}

{Emp_ID, Emp_Number}

{Emp_ID, Emp_Name}

{Emp_ID, Emp_Number, Emp_Name}

{Emp_Number, Emp_Name}

Primary Key

- Special relational database table column (or combination of columns)
- Uniquely identify each table record.
- Main features are:
  - Contain a unique value for each row of data.
  - Cannot contain null values.
  - Every row must have a primary key value.

Employee
EmployeeID
EmployeeName
SSN
DeptID
DOB

Example

Stu_Id	**Stu_Name**	**Stu_Age**
101	Steve	23
102	John	24
103	Robert	28
104	Carl	22

Stu_Id column uniquely identifies each row of the table.

Identify Primary Key

- No two rows can have the same primary key value.
- Every row must have a primary key value.
- The primary key field cannot be null.
- Value in a primary key column can never be modified or updated.

Candidate Key

- A candidate key is a super key that contains no extra attribute.
- Selected from the set of super keys
- Should not have any redundant attributes.
- Also called minimal super key.

Employee
EmployeeID
EmployeeName
SSN
DeptID
DOB

Example

Emp_Id	Emp_Number	Emp_Name
E01	226	Steve
E12	227	Ajeet
E22	228	Chaitanya
E32	229	Robert

{Emp_Id}
{Emp_Number}

Note: A **primary key** is being selected from the group of candidate keys. That means we can either have Emp_Id or Emp_Number as primary key.

Foreign key

Columns of a table that points to the primary key of another table.

This concept is also know as Referential Integrity.

Employee
EmployeeID
EmployeeName
SSN
DeptID
DOB

Department
DeptID
DeptName

Example

Stu_Id in **Course_enrollment** is a foreign key as it points to the primary key of the **Student table**.

Course_enrollment table:

Course_Id	Stu_Id
C01	101
C02	102
C03	101
C05	102

Student Table

Stu_Id	Stu_Name	Stu_Age
101	Chaitanya	22
102	Arya	26
103	Bran	25
104	Jon	21

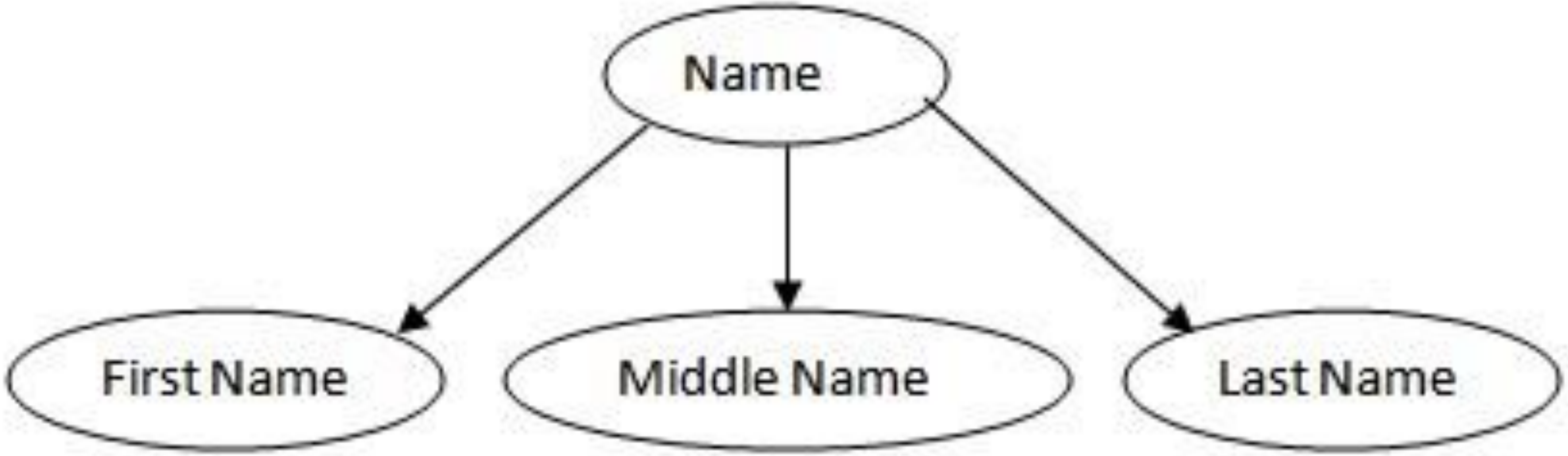
Alternate key

- The candidate key other than the primary key is called an alternate key.

Employee
EmployeeID
EmployeeName
SSN
DeptID
DOB

Composite key

- A key that has more than one attributes is known as composite key.
- It is also known as compound key.



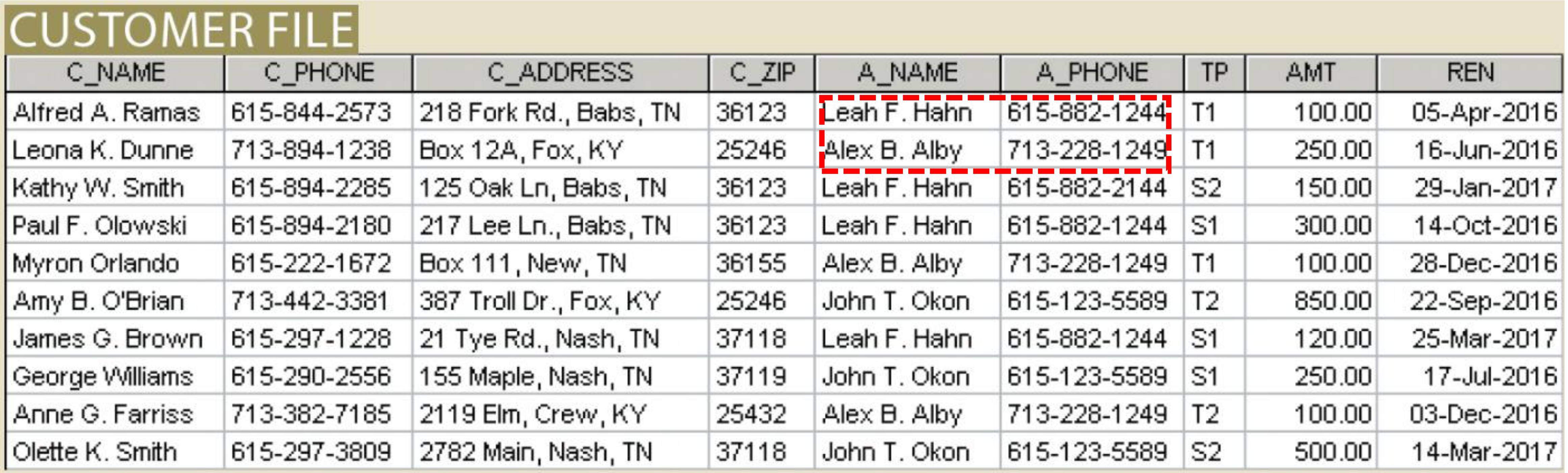
Q&A

Lab Activity 1

- Data Redundancy, Inconsistency & Security
- Relational model

Data Redundancy, Inconsistency & Security

- **Redundancy:** duplication of data, or storing the same information in multiple places.



CUSTOMER FILE

C_NAME	C_PHONE	C_ADDRESS	C_ZIP	A_NAME	A_PHONE	TP	AMT	REN
Alfred A. Ramas	615-844-2573	218 Fork Rd., Babs, TN	36123	Leah F. Hahn	615-882-1244	T1	100.00	05-Apr-2016
Leona K. Dunne	713-894-1238	Box 12A, Fox, KY	25246	Alex B. Alby	713-228-1249	T1	250.00	16-Jun-2016
Kathy W. Smith	615-894-2285	125 Oak Ln, Babs, TN	36123	Leah F. Hahn	615-882-2144	S2	150.00	29-Jan-2017
Paul F. Olowski	615-894-2180	217 Lee Ln., Babs, TN	36123	Leah F. Hahn	615-882-1244	S1	300.00	14-Oct-2016
Myron Orlando	615-222-1672	Box 111, New, TN	36155	Alex B. Alby	713-228-1249	T1	100.00	28-Dec-2016
Amy B. O'Brian	713-442-3381	387 Troll Dr., Fox, KY	25246	John T. Okon	615-123-5589	T2	850.00	22-Sep-2016
James G. Brown	615-297-1228	21 Tye Rd., Nash, TN	37118	Leah F. Hahn	615-882-1244	S1	120.00	25-Mar-2017
George Williams	615-290-2556	155 Maple, Nash, TN	37119	John T. Okon	615-123-5589	S1	250.00	17-Jul-2016
Anne G. Farriss	713-382-7185	2119 Elm, Crew, KY	25432	Alex B. Alby	713-228-1249	T2	100.00	03-Dec-2016
Olette K. Smith	615-297-3809	2782 Main, Nash, TN	37118	John T. Okon	615-123-5589	S2	500.00	14-Mar-2017

Data Redundancy, Inconsistency & Security

- Inconsistency: In spread sheet, update to data could use a search/replace; may not work if people are not consistent with data entry.

Mohammad or Muhammad or Mohammed or Mohamed? Which one is used? Typos may also happen!

- Security – an important consideration; who should have access to what information? Who should be able to make changes to data?

Data Redundancy, Inconsistency & Security

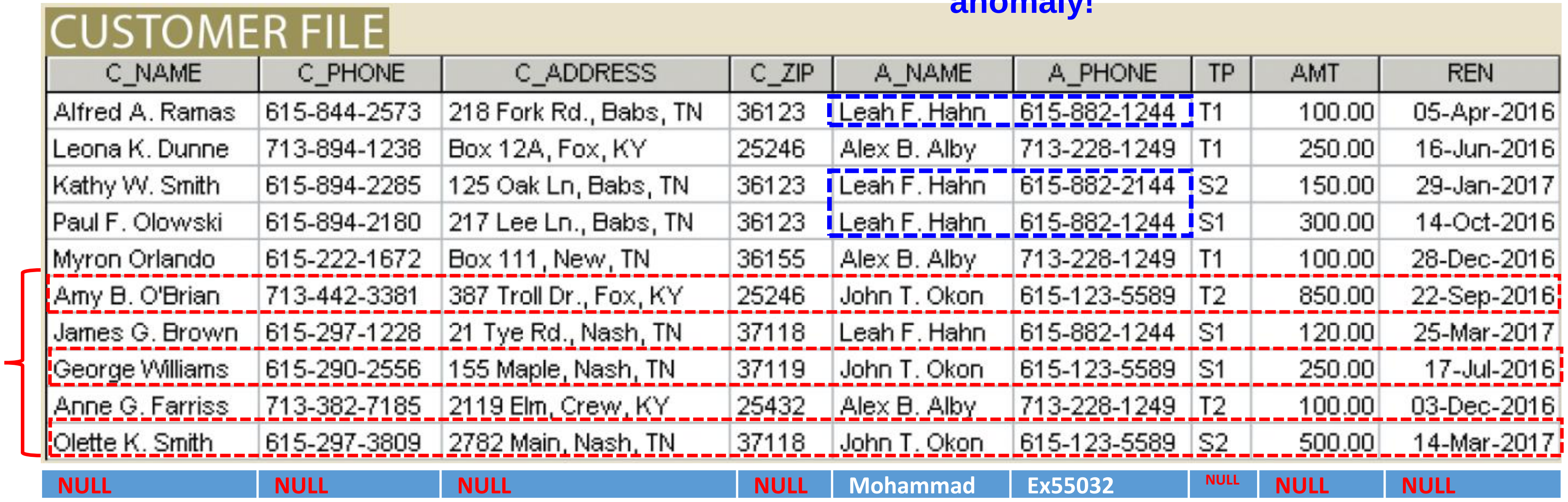
Anomalies introduced into the data when inserting new data, updating old data or deleting data.

Insertion anomaly – to add a new agent to customer table requires dummy customer entry customer table.

Deletion anomaly – Deletion of customer data require deletion of agent data too.

Update anomaly – to update an agent’s phone number requires update in more than one entries.

Update anomaly!



CUSTOMER FILE

C_NAME	C_PHONE	C_ADDRESS	C_ZIP	A_NAME	A_PHONE	TP	AMT	REN
Alfred A. Ramas	615-844-2573	218 Fork Rd., Babs, TN	36123	Leah F. Hahn	615-882-1244	T1	100.00	05-Apr-2016
Leona K. Dunne	713-894-1238	Box 12A, Fox, KY	25246	Alex B. Alby	713-228-1249	T1	250.00	16-Jun-2016
Kathy W. Smith	615-894-2285	125 Oak Ln, Babs, TN	36123	Leah F. Hahn	615-882-2144	S2	150.00	29-Jan-2017
Paul F. Olowski	615-894-2180	217 Lee Ln., Babs, TN	36123	Leah F. Hahn	615-882-1244	S1	300.00	14-Oct-2016
Myron Orlando	615-222-1672	Box 111, New, TN	36155	Alex B. Alby	713-228-1249	T1	100.00	28-Dec-2016
Amy B. O'Brian	713-442-3381	387 Troll Dr., Fox, KY	25246	John T. Okon	615-123-5589	T2	850.00	22-Sep-2016
James G. Brown	615-297-1228	21 Tye Rd., Nash, TN	37118	Leah F. Hahn	615-882-1244	S1	120.00	25-Mar-2017
George Williams	615-290-2556	155 Maple, Nash, TN	37119	John T. Okon	615-123-5589	S1	250.00	17-Jul-2016
Anne G. Farriss	713-382-7185	2119 Elm, Crew, KY	25432	Alex B. Alby	713-228-1249	T2	100.00	03-Dec-2016
Olette K. Smith	615-297-3809	2782 Main, Nash, TN	37118	John T. Okon	615-123-5589	S2	500.00	14-Mar-2017
NULL	NULL	NULL	NULL	Mohammad	Ex55032	NULL	NULL	NULL

Deletion anomaly!

Insertion anomaly!

Data Redundancy, Inconsistency & Security

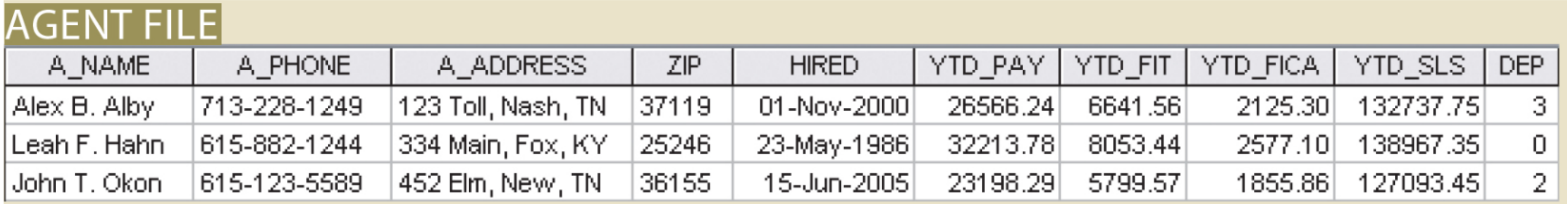
Solution:

What is the solution?

use a separate Agent file and

Customer file only have agent IDs

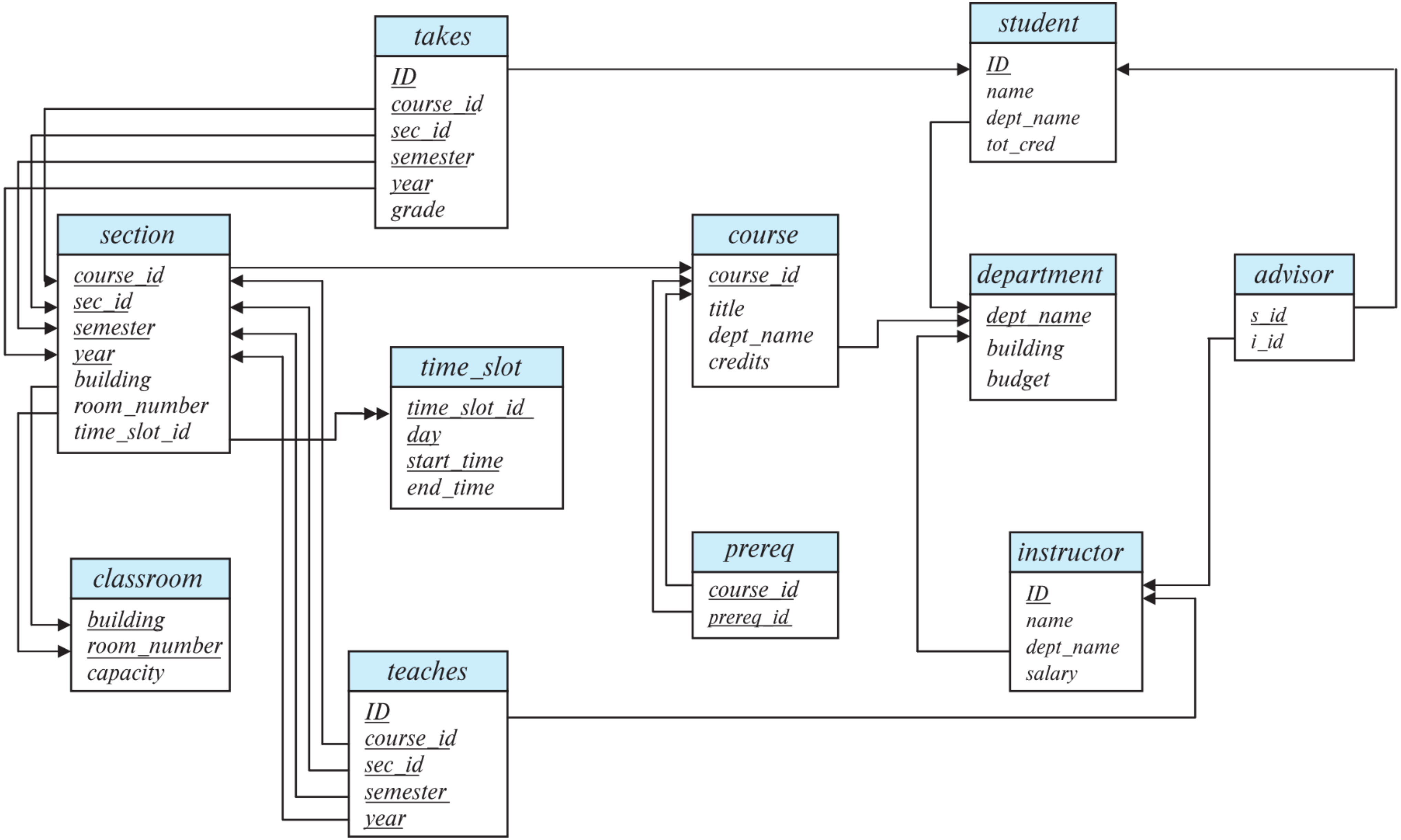
to refer to Agent table!



AGENT FILE

A_NAME	A_PHONE	A_ADDRESS	ZIP	HIRED	YTD_PAY	YTD_FIT	YTD_FICA	YTD_SLS	DEP
Alex B. Alby	713-228-1249	123 Toll, Nash, TN	37119	01-Nov-2000	26566.24	6641.56	2125.30	132737.75	3
Leah F. Hahn	615-882-1244	334 Main, Fox, KY	25246	23-May-1986	32213.78	8053.44	2577.10	138967.35	0
John T. Okon	615-123-5589	452 Elm, New, TN	36155	15-Jun-2005	23198.29	5799.57	1855.86	127093.45	2

Schema Diagram for University Database


takes
ID
course_id
sec_id
semester
year
grade
student
ID
name
dept_name
tot_cred
section
course_id
sec_id
semester
year
building
room_number
time_slot_id
course
course_id
title
dept_name
credits
department
dept_name
building
budget
advisor
s_id
i_id
time_slot
time_slot_id
day
start_time
end_time
prereq
course_id
prereq_id
instructor
ID
name
dept_name
salary
classroom
building
room_number
capacity
teaches
ID
course_id
sec_id
semester
year